

Designing Auditory Misinformation Warnings to Foster Critical Engagement with Podcasts

by

Konstantin-Asen Yordanov

to obtain the degree of Master of Science
at the Delft University of Technology,
defended publicly on 24 June 2026 at 13:00.

Student number: 5529883
Project duration: November 10, 2025 – June 24, 2026
Thesis committee: Prof. dr. Ujwal Gadiraju, TU Delft, supervisor
Prof. dr. Pradeep K. Murukannaiah TU Delft, committee member

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Preface

This MSc thesis allowed me to experience the process of conducting empirical scientific research from start to finish – essentially turning a broad-spectrum brainstorming session dating back to the summer of 2025 into a cohesive and academically motivated user study over the span of seven months. I was given the opportunity to systematically reason about and arrange each component of the data collection and analysis methodologies in accordance with the contribution’s overarching scope and objectives – a skill developed through previous course projects in my studies and now put to the test in an exploratory work that bears conceptual overlap with multiple research areas, including human-computer interaction and responsible opinion formation.

Besides my own efforts, I owe the successful completion of this thesis to several vital individuals. First, I extend my deepest gratitude to my supervisor, Prof. Ujwal Gadiraju, for his time, approachability, insightful forward-looking guidance which was always tailored towards the next steps in my progress, and inherent excitement for the scientific unknown that in turn unfailingly boosted my motivation to aim further. I am also thankful to Alisa Rieger for her attentive feedback and experienced recommendations at the crux of this project – namely, devising the method, during which her course-correcting role proved invaluable. Additionally, I extend my appreciation to Prof. Pradeep K. Murukannaiah for taking the time to join the thesis committee and participate in my defense.

Finally, I want to thank my friends and colleagues whom I regularly interacted with throughout the project’s duration – Yiğit, Kaan, Codrin, Matej, Mihnea, Andreea, Vlad, Adina, Violeta-Mara: our conversations, get-togethers, and shared experiences helped me recharge and see this thesis through.

*Konstantin-Asen Yordanov
Delft, 2026*

Abstract

The spread of online misinformation undermines responsible opinion formation, this risk being exacerbated for heavily debated topics where emotionally-driven reasoning and the formation of echo chambers weaken critical scrutiny. Visual misinformation labels which nudge users towards more mindful information-seeking behaviors are common-place, yet, despite its increasing popularity, the podcast medium remains overlooked in these efforts compared to Web search. This thesis examines the design of auditory misinformation warnings: brief, non-verbal sound cues embedded in podcast audio to indicate that a nearby statement is misleading. To verify the relevant criteria for these signals' effectiveness and the factors influencing their reception by listeners, two exploratory crowdsourcing studies were conducted. The first assessed 15 auditory icons in terms of recognizability, consistency of conceptual mapping, and induced disruption, isolating the most favorably perceived candidates. The second embedded the latter as purposeful interventions within AI-generated podcast dialogues on three controversial scientific topics interspersed with myths and falsehoods, employing a between-subjects design across four warning-placement configurations (before, after, enclosing, and concurrent with a misinformation occurrence) and treating several psychometric traits as exploratory factors. The findings indicate salient auditory icons that alert listeners are a viable choice provided they do not severely depart from the podcast's setting or listeners' expectations of traditional podcast-editing effects; otherwise, they may trigger prolonged dissatisfaction and ultimately break immersion. No conclusive data on the cues' role in misinformation recognition emerged, although some listeners correctly inferred their intended function when these followed or enclosed a misleading claim. Warnings placed after a falsehood were least disruptive while concurrent placement was associated with the lowest content recall. Regarding contextual factors, listeners' prior stance on a debated topic and a podcast's perceived density and pace exhibited significant correlations with several Likert-scale sound-cue evaluations. Participant's reactions and follow-up attitudes towards the interventions diverged significantly – anticipation, habituation, and active resistance against perceived paternalism all being expressed. These results yield preliminary recommendations for auditory misinformation warnings and, given the many non-trivial trade-offs at play, outline a multitude of directions for future research.

Contents

1	Introduction	3
1.1	Project Scope and Research Questions	4
1.2	Study Layout and Key Contributions	4
1.3	Thesis Outline	5
2	Related Work	7
2.1	Responsible Opinion Formation	7
2.2	Interplay Between Misinformation and Cognition	8
2.3	Design Guidelines for Fact-Checking Interventions	8
2.4	Podcast Medium Characteristics	9
2.5	Sound Cues and Auditory Warnings	10
3	Preliminaries	13
3.1	Hypothesis Formulation	13
3.2	Materials	13
4	Method	17
4.1	Variables and Measures	17
4.2	Experimental Setup and Procedure	22
4.2.1	Study 1	22
4.2.2	Study 2	22
4.3	Cognitive Bias Checklists	24
4.3.1	Study 1	24
4.3.2	Study 2	25
4.4	Analyses	26
4.5	Participants	28
4.5.1	Study 1	29
4.5.2	Study 2	29
5	Results	31
5.1	Study 1	31
5.1.1	Quantitative Data (Descriptive Statistics)	32
5.1.2	Qualitative Findings	32
5.1.3	Selected Carryover Auditory Icons	34
5.2	Study 2	35
5.2.1	Quantitative Data (Descriptive Statistics)	35
5.2.2	Contextual Factors (Statistical Tests and Models)	37
5.2.3	Qualitative Findings	44
5.2.4	Exploratory Post-Debrief Listener Perceptions	49
6	Discussion and Implications	54
6.1	Impact and Reception of Podcast-Integrated Auditory Icons (RQ1)	54
6.2	Interplay Between Warnings and Misinformation Awareness (RQ2)	55
6.3	Influence of Placement Alternatives on Warning Performance (RQ3)	55
6.4	Influence of Contextual Factors on Warning Performance (RQ4)	56
6.5	Broad Research Implications	57
6.6	Limitations and Directions for Future Work	57

7 Conclusion	60
A Exhaustive Study Results	63
B NotebookLM Instructions	79
B.1 Podcast Generation	79
B.2 Recall Questionnaire Generation	80
C Podcast Materials	84
C.1 Layout Visualizations	84
C.2 Transcripts	85
C.2.1 Nuclear Energy	85
C.2.2 Cell Phone Radiation	90
C.2.3 GMO Foods	94
C.3 Content-Supplying Sources	99
C.3.1 Nuclear Energy (32 Sources)	99
C.3.2 Cell Phone Radiation (15 Sources)	100
C.3.3 GMO Foods (36 Sources)	101
D Recall Questionnaires	103
D.1 Nuclear Energy	103
D.2 Cell Phone Radiation	105
D.3 GMO Foods	107
E Informed Consent Materials	111
E.1 Study 1 Opening Statement	111
E.2 Study 2 Opening Statement (Partial Disclosure)	111
E.3 Study 2 Debriefing Form (Full Disclosure)	112
F Web Application Frontend Views	114

Introduction

Characterized by the lack of a satisfactory consensus-driven resolution, debated issues are often polarizing, emotionally charged, and conducive to a wide variety of cognitive biases [54]. These byproducts of a topic's controversiality in turn suppress the faculties for objective reasoning and skew information processing so it aligns with one's pre-existing attitudes [38]. From an information seeking perspective, such selective exposure to (and retention of) content is diametrically opposed to the central principles of responsible opinion formation – namely, a searcher's engagement with diverse viewpoints and their reliance on critical thought rather than emotion [54, 43]. Moreover, it enables the formation of tightly-knit echo chambers, diminishing one's tolerance for alternative outlooks and thus increasing their susceptibility to falsehoods [19, 40, 10], while the cognitive biases underpinning these behaviors actively induce belief in misinformation [3, 46, 24].

The latter's reinforcement and dissemination are further facilitated by the prominence of more accessible and convenient alternatives to traditional Web search, podcasts representing one such medium [49]. Relying on natural language to convey their intended messages and typically blending information and entertainment, such on-demand recordings are more engaging and require less cognitive effort to interact with [60]. Precisely these attributes, however, prompt searchers to prioritize the utility of podcasts over the critical assessment of their content's veracity: this shift has significant implications for the spread and durability of misleading narratives [49]. The recordings' fast-paced delivery of information and lower cognitive demands restrict the opportunity for listeners to scrutinize the statements they are presented with: this turns the interaction into a passive process [60]. Furthermore, certain key aspects of podcasting – namely, the formation of para-social relationships between listeners and hosts as well as the latter's frequent appeals to emotion and identity while offering elaboration or justification – are tailored towards persuasion, leading to misplaced trust and reduced misinformation resilience [59].

In the visual modality, fake news interventions like content flags and labels have become commonplace on popular media platforms, promoting mindful information consumption habits [47]. Serving as transparent automatic nudges, these preventative measures rely on the intentionally introduced user-experience "friction" to fulfill the dual objective of (i) raising misinformation awareness and (ii) prompting critical engagement with online publications [11, 13]. Inspired by this design, a recent contribution forming the basis for the present work shifts focus to the auditory space and suggests the integration of flag-like signals into podcasts – such cues indicating that an adjacent audio snippet contains misleading statements [49]. Sound interventions take advantage of hearing as one's natural "always-on" warning sense and act as a useful corollary to visual indicators in workplace settings, especially under high workload conditions [20]. Particular signal-carrier types, such as auditory icons with their intuitive associations to real-world entities [9], provide additional benefits for learnability and ease-of-recall, also avoiding the language-comprehension limitations of speech-based alternatives.

Considering podcasts' reliance on an uninterrupted and intelligible narrative for message conveyal and listener comprehension, however, incorporating such interventions into the medium is a non-trivial process. To fulfill the function of a warning, a sound cue should be consistently recognizable and salient, its perceptual prominence and assumed importance enabling it to instill a sense of urgency and reliably capture one's attention [20]. Meanwhile, to remain truly effective, it must trigger neither cognitive load interference nor prolonged user dissatisfaction, the latter two potentially causing a listener to lose track

of the discussion, mishear and misremember the statements presented, or exit the podcast prematurely [26]. Finding the right balance between misinformation awareness and warning-induced disruption in turn requires understanding how listeners perceive and interact with auditory interventions, as well as what factors influence the audience's receptiveness to such cues and should therefore be considered at an early design stage [55]. This understudied interplay is precisely the research gap that the present contribution aims to address in the pursuit of (i) guidelines for effective auditory misinformation warnings and (ii) their adoption by practitioners and podcast-hosting platforms seeking to curtail the dissemination of falsehoods in this medium.

1.1. Project Scope and Research Questions

An examination of auditory interventions is itself a multi-tiered undertaking. At the base level, podcast-integrated sound cues can be assessed in terms of their effectiveness and user reception, essentially verifying whether a given implementation is well-suited and minimally disruptive. For their interplay with cognition to be studied beyond simply immediate attention-capture, the extent to which they impact listeners' misinformation awareness and are consistently associated with this hazard can be investigated. Going further, these stimuli's influence on decision-making and the consumption behaviors they might induce form a separate cluster of analysis, user studies exploring whether a listener would pause the recording, carry out independent research, or outright close the audio player and unsubscribe from the podcast upon hearing a warning. From a common psychological lens, the capacity of such sound cues to trigger attitude changes or belief updates can also be examined, offering insights into their long-term cognitive effects.

Of these four plausible directions, this work primarily focuses on the first two – with inspiration from the preliminary research agenda outlined in [49] as well as the overarching objective of offering an early glimpse into the properties, perceptions, and caveats of podcast-integrated warnings. To that end, the following research questions are investigated:

RQ1 Which metaphorical auditory icons demonstrate high effectiveness and salience when integrated into podcast content?

RQ2 To what extent do auditory warnings affect listeners' recognition of misinformation?

RQ3 How does the placement of auditory warnings (w.r.t. a misinformation segment – (i) before; (ii) after; (iii) enclosing, i.e. before and after; (iv) concurrent) influence their effectiveness and salience?

RQ4 What other contextual factors influence auditory warnings' effectiveness and salience?

As delineated by these four questions, the project's scope encompasses both implementation-level attributes of auditory interventions and key facets of their potential cognitive impact (in particular, misinformation recognition and podcast content recall), extending the analysis by also considering listeners' psychometric characteristics as external factors.

1.2. Study Layout and Key Contributions

To address the RQs, a pair of exploratory crowdsourcing studies were carried out consecutively. In the first, a total of 15 auditory icons belonging to different sound classes were incorporated into short conversation excerpts and evaluated in terms of their recognizability, conceptual-mapping ambiguity, and induced disruption (RQ1). Obtained worker responses were used to identify the most consistently and favorably perceived sound cues, which then assumed the role of misinformation warnings in the second study. The latter was centered on an assessment of these interventions in the functional context of podcast listening, utilizing AI-generated podcast-dialogues tackling three controversial scientific topics and injected with misleading or inaccurate claims interspersed uniformly throughout their duration. Relying on both subjective participant impressions and objective measures like the assessment of podcast content recall as a basis for examination, this second study explored the indicative markers facilitating misinformation recognition (RQ2), used a between-subjects design to vary experimental conditions across four warning-placement configurations (RQ3), and incorporated several psychometric characteristics (including attitude strength, need for cognition, and intellectual humility) as exploratory variables (RQ4) to account for individual differences and provide a multi-faceted understanding of listeners' perceptions and engagement-patterns.

Findings from both studies confirm the benefits of high recognizability and conceptual-mapping consistency characteristic to auditory icons: nonetheless, the sound serving as the intervention should not exhibit a severe discrepancy with the podcast setting or listeners' appropriateness-expectations, otherwise perceived as a purposeless interruption and breaking their immersion. No influence of the sound cues on participants' misinformation recognition is observed; however, their function as misinformation warnings is recognized by some when *placed after* or *enclosing* the misleading claims. Interventions positioned *after* a falsehood also yield the lowest levels of induced distraction and interruption out of all placement configurations, significantly outperforming *concurrently placed* warnings – which are in turn associated with the overall lowest recall-assessment scores. Participants' subjective evaluations of the sound cues are influenced by their pre-existing personal stance on the topic of the podcast, as well as the latter's perceived information density and pace. Reactions and attitudes towards the interventions identified through qualitative data analysis vary substantially – some perceive an underlying purpose behind the signals and actively anticipate them while others discard them due to habituation or a sense of unsolicited paternalism. The many external factors and trade-offs relevant in the design of auditory misinformation warnings thus offer various directions for follow-up contributions on this topic.

1.3. Thesis Outline

This report is structured as follows: Chapter 2 first outlines the project's underpinning knowledge foundations, highlighting the insights from previous contributions and positioning the present work within the literature landscape. Chapter 3 offers key pre-study considerations, such as the exploration-driven approach to answering the RQs, and describes the use of external tools and materials throughout the project's implementation. Afterwards, Chapter 4 establishes the employed methodology and experimental setup, denoting the full list of variables, elaborating on the dual-study procedure and the associated analysis techniques, and addressing relevant crowdsourcing concerns. Chapter 5 then presents the results obtained from both studies and through any additional post-processing, whereas Chapter 6 interprets the findings in a broader context, addressing the RQs, identifying crucial research limitations, and offering recommendations for future work. A final summary of the project and its contributions is provided in Chapter 7.

2

Related Work

This chapter outlines the key insights from previous contributions on (i) responsible opinion formation, (ii) the implications of falsehood dissemination, (iii) established practices in the design and integration of misinformation warnings, (iv) the distinctive attributes of the podcast domain, and (v) the categorization of auditory cues based on their signal-types and applications. Relevant open questions and challenges are also identified, positioning the present work in the context of prior literature and demonstrating how it builds upon the existing knowledge base to bridge the research gap.

2.1. Responsible Opinion Formation

Due to their controversiality and the difficulty of finding an all-round satisfactory resolution, debated topics induce a reliance on emotion at the expense of critical thinking, thereby increasing the risk of polarization [54]. These effects are compounded by the prominence of different biases throughout the information seeking process, expressed on the level of searcher-system interactions, ranking and indexing algorithms, or the stored data themselves carrying some latent real-world skew [54, 2]. A number of user studies have investigated the interplay between such biases and debated topics in the context of Web search [68, 16, 53], demonstrating that users browse search results selectively based on their pre-existing attitudes while search engines in turn diminish viewpoint diversity through non-transparent and one-sided rankings of Web resources.

Individual searcher characteristics are another factor at play in the interpretation and reconciliation of findings on debated topics. Prior research in personality and social psychology points to need for cognition [30] and attitude importance [17, 32] as particularly influential attributes, the former facilitating the formation of durable stances resistant to change and the latter affecting the retention of topic-relevant knowledge by encouraging selective information exposure. In addition to elaboration (defined as “the amount of issue-relevant thinking” one carries out [51]), moral conviction in a given viewpoint also leads to distancing from alternative opinions and increases one’s resistance to compromise [58, 45]. Collectively, these four attributes act as predictors of attitude strength [44], searchers with strong prior stances having been found to prioritize attitude-consistent results and discount the remaining Web resources’ credibility *a priori* [62]. The influence of such psychological characteristics extends beyond information processing and Web search, also manifesting in conspiracy belief [14].

Furthermore, picking out information in support of a prior attitude and interpreting it accordingly are behaviors indicative of confirmation bias, where findings not conforming to that pre-existing stance are discarded: this behavioral pattern has been examined in user studies concerned with Web search on controversial topics [65]. In a broader sense, such phenomena fall under the category of cognitive biases (defined as “systematic deviations from rational thinking and decision-making” [1]), and they permeate all stages of the information seeking process [54].

In contrast, critical thought and engagement with diverse viewpoints serve as the foundation of responsible opinion formation: a cognitive process that inevitably involves interactions with digital media and underpins the welfare of democratic societies [54]. To that end, different strategies for fostering those two desirable faculties have been explored, including nudges (interventions “nudging people’s behavior in a predictable way”) and boosts (interventions “boosting people’s cognitive and motivational

competencies”) [38, 43]. From the psychological lens, other contributions have investigated the benefits of intellectual humility (“being open to the possibility that one may be misinformed or uninformed” [25]) for objective information seeking, in particular the ability to accurately examine one’s personal stance with respect to the available evidence [52].

The many caveats of debated topics, online search interactions, and interwoven biases (cognitive or otherwise) explored in the literature are instrumental in motivating this research, existing contributions asserting the significance of responsible opinion formation as well as identifying its driving and inhibiting factors. This work contributes to the body of knowledge on the subject by applying a well-known nudge (the misinformation warning) in a domain drastically different from Web search – podcasts, whose fast pace and inherently passive process of consumption suppress critical engagement with their content.

2.2. Interplay Between Misinformation and Cognition

Comprising different kinds of false, inaccurate, or misleading narratives that circulate in the public domain, misinformation can originate from rumors, statements made by figures of authority, or media reports [40]. Since the onset of the digital age and social media platforms’ rise to prominence, such falsehoods can be disseminated quickly and reach a worldwide audience, distorting facts, capturing people’s attention through continuous sharing, and being increasingly perceived as accurate due to repeated exposure – a phenomenon known as the illusory truth effect [10, 66]. While not all misinformation arises from the intention to deceive, purposefully generated misleading narratives are further categorized as disinformation, typically circulating “through adversarial identity-based grievances” that reinforce polarization and solidify opposing viewpoints into tightly-knit echo chambers [12].

In turn, misinformation belief is closely associated with unwarranted trust in institutions and a failure to rationally scrutinize the content’s credibility or veracity [28]. This affective shift in the object of trust from information itself to the entity broadcasting it (including social media platforms in recent crises like the COVID-19 pandemic [69]) is correlated both positively with overconfidence in one’s ability to detect misinformation and negatively with their motivation to verify what they are presented with – an indication that utility and convenience may outweigh the value of effortful critical thinking [63]. Changes in information seeking priorities have also been observed on a generational level, prior research pointing out Gen Z’s tendency to encounter narratives ad hoc, interpreting and authenticating them with an underlying awareness of their social group’s consensus on the subject (a phenomenon termed “information sensibility” [29]), rather than purposefully seek out insights grounded in truth.

Playing a critical role in information processing, individual characteristics further impact one’s perception of falsehoods. Domain knowledge, education, and personal conscientiousness have been found to improve fake news detection [70], whereas heightened emotionality and poor digital literacy both induce belief in misleading narratives [3, 46]. In addition to confirmation bias, several cognitive biases (including anchoring, bandwagon, and framing effects) were identified as influential to misinformation susceptibility in a recent qualitative study involving social media users [24].

Due to the individually and socially catastrophic consequences of misinformation dissemination left unchecked, different intervention designs have been considered, generally categorized as prebunking and debunking strategies [19, 55]. The former include preexposure warnings (akin to nudging) and preemptive inoculation (a form of boosting), whereas the latter address misleading narratives retroactively through refutations centered on an alternative factually-accurate explanation [19, 40]. Analyses of these techniques highlight the importance of explicitly referencing credible sources and adjusting the language used to communicate the presence of falsehoods – both practices helping reduce the risk of backfire effects brought on by information overload or a perceived threat to one’s worldview [19].

Understanding the means by which misinformation traverses the digital landscape as well as the cognitive traits facilitating its internalization or detection is a self-evident prerequisite to any research into the design of fake news interventions. In particular, the misconception-enabling trust dynamics and the creation of echo chambers examined in prior contributions are both deeply rooted in the podcast medium itself through the resilient para-social relationships formed between host and listener, and they will have implications for how the latter perceives the auditory warnings explored in this study.

2.3. Design Guidelines for Fact-Checking Interventions

Commonly integrated into social media platforms and news websites, visual misinformation warnings can be a boon in the struggle to curtail the spread of falsehoods online, transcending demographic

characteristics and pre-existing attitudes [47]. The intuition behind these interventions originates from the notion of design frictions: purposefully introduced moments of thoughtful interaction with a UI component that cause minor dissatisfaction yet trigger the change of some specific risk-taking behavior, yielding a net benefit for both system integrity and user well-being [11, 13]. Adhering to this paradigm, prior research has demonstrated that nudges prompting social media users to assess a post's accuracy via a checklist and a free-form rationale successfully reduce the sharing of misleading content [34].

For warnings to be effective interventions, however, their recipients' preferences should be accounted for at the design stage: a wide range of user studies have been carried out to that end. Compared to the easy-to-miss contextual flags that offer users the option to "find out more" when clicked, interstitial warnings entirely obscuring the inaccurate news claim prove successful in capturing users' attention and appropriately alerting them to the risk of misinformation, an accompanying explanation for why the item is flagged further improving their reception [26, 37]. Warnings should be visually consistent and accurately placed, their urgency corresponding to users' perceived risk level and factually correct content remaining unaffected (false alarms in turn associated with a "cry wolf effect") [26]. Other desirable qualities include transparency and credibility, the design featuring a salient veracity indicator, listing relevant fact-checking sources, and including details about their authors' background [37, 31]: otherwise, the interventions may be perceived as "biased and overly punitive" [57].

Users' cognitive traits also influence their attitude towards misinformation warnings. According to one study investigating this interplay, need for cognitive closure, trust in medical professionals, and a liberal political stance are all positively associated with a favorable outlook on fake news interventions, whereas trust in religious leaders has a negative correlation [36]. Another contribution shows the effects of uncertainty-aversion bias, observing that "claims tagged with the 'Lack of Evidence' label are recognized similarly as false information" [48]. This finding aligns with previous insights into how doubt obstructs decision-making and prompts users to more consistently take heed of warnings when judging an item's veracity – in such cases, the interventions do not even have to be explicitly interacted with to be considered effective, influencing information seekers simply through exposure [26].

Nonetheless, several limitations of visual misinformation warnings have been identified in the literature. When deemed too high compared to the perceived risk level, the cost of separately researching a given claim's veracity causes users to ignore such interventions [56]. The latter also fail to root out the continued influence of misleading information according to a psychology contribution on memory and cognition, suggesting that users comply with them due to the experienced friction rather than informed decision-making [18]. Furthermore, a negative spillover effect has been explored as a consequence of warning integration, the flags inadvertently "priming general distrust in authentic news" [61]. Another line of research has investigated the issue of habituation, caused by users unconsciously processing a frequently encountered visual stimulus and becoming desensitized to the warnings unless a polymorphic design is utilized [64, 7].

The implementation of visual misinformation labels brings forth a host of non-trivial choices and user experience considerations, making it necessary to reconcile conflicting objectives – alerting information seekers to the presence of falsehoods while also avoiding cognitive overload and prolonged dissatisfaction. Building on previous work in the area, this contribution explores the degree to which those insights translate onto auditory space, facing two challenges unique to this domain: the limited room and duration for accurately positioning the warning, as well as the inability to incorporate a more detailed fact-checking report justifying the flag in a well-timed and salient manner.

2.4. Podcast Medium Characteristics

Offering diverse on-demand content in the auditory modality, podcasts garner an increasing worldwide listenership and carry several benefits over traditional radio [49, 60]. Characterized by the amalgamation of information and entertainment (termed "infotainment" or "edutainment"), these recordings tackle popular and niche topics alike, appealing to audiences and sub-cultures of various sizes and backgrounds [60]. They are also inexpensive and easy to create: although production quality matters for a podcast to be well-received and gain popularity, tailored smartphone applications enable amateurs to effortlessly record and publish their own content [35]. In turn, listeners can personalize their experience by subscribing to and downloading their preferred podcasts, engaging with and replaying them at their leisure – often while completing mundane tasks or household chores, a multitasking behavior enabled by podcasts' audio-only nature and the eliminated need to interact with a visual interface [60, 59]. From

an information seeking perspective, podcasts are recognized for the use of natural language to convey intended messages and the reliance on narrative elements for contextualization: both these attributes reduce the time and cognitive effort users need to invest in the interaction, making the medium more engaging and accessible than the Web search alternative [49, 60].

Prior research has specifically investigated the motivations of podcast consumption, a US national survey indicating that learning new things, entertainment, and keeping up-to-date with trending topics are the three primary reasons for listening [35]. Furthermore, an international user study demonstrates the role of interest-based curiosity, need for cognition, and openness to experience as psychometric predictors of engagement with podcasts and identifies the latter's key outcomes as competence, personal growth, and relatedness – suggesting that a listener's information needs are accompanied by the desire for social involvement [59]. These findings align with previously established characteristics of the medium itself, particularly the significance of community-driven interactions for a podcast's popularity and the formation of para-social relationships between host and listener [60].

From a researcher's standpoint, however, many of these recordings' attributes become problematic. Systematic analyses of the "podosphere" are obstructed by its diversity and internal inconsistency, exemplified by incomplete or unreliable metadata, varying episode lengths and formats (e.g. monologues, dialogues, or group discussions), mixed media through the use of music and custom sound effects, and noisy transcriptions [35]. These challenges have resulted in the medium being overlooked and underexplored, inadvertently discounting its role in the dissemination of falsehoods. In the latter regard, podcasts' fast pace and the rapid-fire delivery of claims put forward by the hosts leave little time for listeners to scrutinize the statements they are presented with [60]. Furthermore, community engagement and the aforementioned para-social attachments can increase one's susceptibility to misinformation: the use of pathos (appeals to emotion) and ethos (appeals to identity) coupled with a host's voice, attitude, and personality becomes an instrument of persuasion to the listener's potential detriment [60]. Prior research on podcast listenership has also explored the dynamics of perceived risk, suggesting that the degree to which audiences take the presence of misleading claims seriously depends on their purpose for listening [59]. In turn, the accessibility and convenience podcasts are recognized for transform one's engagement with the medium into a passive process, restricting users' control over what content they are exposed to and causing them to discard a critical approach centered around veracity assessment in favor of utility: misplaced trust and higher proneness to misinformation are the immediate consequences of this shift [49].

Identifying the defining characteristics of the podcast domain and how they could reinforce belief in hoaxes or misconceptions is understandably a critical precursor to the present contribution. Based on this exploration of prior research, several subtleties of introducing fake news interventions into audio content have been brought to light – in particular, its fast pace, inherent suitability for off-screen hands-free consumption, and reliance on an ongoing narrative where ill-placed interruptions could severely impede listener comprehension and satisfaction.

2.5. Sound Cues and Auditory Warnings

The appropriate utilization of auditory signals has been a dedicated research area for more than five decades, prior contributions exploring the use of sounds for non-visual interfaces benefiting blind computer users [50] and as warnings in workplace settings [20]. In the latter context, sound cues have been identified as a useful corollary to visual indicators – especially "under conditions of high workload" – taking advantage of human hearing as a natural warning sense that cannot be "shut off" to avoid stimuli and, importantly, exhibiting higher levels of compliance compared to alerts in the visual modality [20]. Nonetheless, several key factors influencing the effectiveness of auditory signals have been established. First, the latter's learnability and long-term retention in memory significantly impact identification and response rates, related non-trivial considerations including the time-span of exposure and the employed training procedure [20]. Second, the concrete audio sample's recognizability and the perceived appropriateness of its mapping onto the domain entity it is intended to indicate are vital aspects to capture prior to the sound cue's incorporation – for example, using pilot studies and early-stage evaluations [50]. Third, the salience and assumed importance of a sample's overarching sound class further regulate people's reactions to auditory signals, one empirically-motivated perceptual ordering ranking alarms, speech, and animal sounds as the three categories most highly judged as salient and urgent by participants [42].

In the pursuit of a sound cue design that accounts for these elements, a wide range of signal-carrier options can be considered, including speech and several non-verbal alternatives typically classified into three types [49]. Auditory icons are “emulations of naturally occurring sounds” produced by real-world objects and events, this conveyed sound-source property essentially underpinning the way information is represented [9]. In turn, earcons constitute synthetic audio patterns that lack clear everyday-life associations yet combine into intricate hierarchical structures through the manipulation of sound parameters, making these cues suitable for menu-type interface components [49]. Finally, spearcons are created by augmenting and speeding up utterances to the point where they no longer resemble speech, subsequently forming sound-families akin to earcon hierarchies where relatedness is derived from the similarity of the spearcons’ underlying spoken phrases [49]. While all these approaches have distinct utility-benefits and practical applications, a recent contribution outlining the research agenda this work is based on offers warning-implementation suggestions specifically suited to the podcast medium: the authors point out earcons’ complexity and limited learnability and further discard both spearcons and speech-based interventions, asserting that the latter two’s interference with podcast audio makes them “not viable to sustain the train of thought of listeners and create good listening experiences” [49].

Under these constraints, auditory icons remain an appropriate choice for a non-verbal signal carrier. Prior research has emphasized how their reliance on real-world sound associations enables them to both utilize a listener’s prior knowledge and transcend cultural and linguistic barriers [9]. Furthermore, being natural and relatable to everyday life, these cues are easy to recognize, learn, and remember provided that a sensible conceptual mapping is established: such benefits have been clearly demonstrated in user studies concerning clinical scenarios and medical device standards, where alarms based on auditory icons outperform their traditional tonal alternative in terms of both localizability and response rates [21, 22]. Unlike earcons and spearcons, warnings implemented using auditory icons have also been uniquely identified to not cause resource interference in working memory under concurrent cognitive loads, the experiment featuring verbal and spatial recall tasks to investigate whether the interventions’ inattentive processing (as a consequence of bounded attention) is correlated with the underlying sound-cue type [39].

From a taxonomy perspective, auditory icons are divided into three groups depending on their relationship with the entity they are intended to indicate (i.e. the referent): nomic (the relation is based on a physical sound-producing source), symbolic (the relation is “essentially arbitrary and based on social convention”), and metaphorical (the relation is “based on some similarities” between sound and referent, making it “not completely arbitrary, yet also not fully dependent on physical causation”) [9]. Specific domain applications have prompted the careful evaluation of these categories and their viability with respect to the context, including in literature on the podcast medium: symbolic icons are typically considered “reserved sounds” fulfilling a predefined social function that cannot be overridden (for example, ambulance and police sirens or cell phone notification alerts) [9, 49], whereas interventions based on nomic icons face the risk of being obscured by other topically-relevant jingles and sound effects interspersed into a given podcast that in turn use the same referent, detracting from a misinformation warning’s salience and thus limiting its effectiveness [35, 49]. With these considerations in mind, judiciously chosen metaphorical icons could act as consistently identified and perceived auditory signals, their inherent arbitrariness and ad hoc semantics allowing them to simultaneously avoid the symbolic type’s overlapping-functionality restriction and capture listeners’ attention more reliably than the nomic variant [9].

Tackling the integration of auditory misinformation warnings into podcasts, this work necessitates an in-depth exploration of the available sound cues, their known applications, existing design guidelines for indicators outside the visual modality, and a plethora of relevant trade-offs concerning utility, salience, learnability, conceptual associativity, cognitive load, and perceived importance. Based on these literature findings, the category of metaphorical auditory icons and the particularly salient sound classes of alarms and animal sounds have been designated for further consideration. By operationalizing them in the novel functional context of podcasts, the present work will build upon this foundation, concerning itself with both warning urgency and listener satisfaction – attention-capture as well as podcast-content discernibility – as conflicting objectives.

3

Preliminaries

This chapter motivates the research approach adopted throughout the remainder of the contribution, laying the groundwork for the choice of analysis and interpretation techniques (described in later sections) and establishing any explicit hypotheses. It also outlines the key materials used in the project's implementation, elaborating on the process of their creation and/or modification where necessary.

3.1. Hypothesis Formulation

Due to the insufficient prior work on the mid-narrative incorporation of non-thematic audio signals and, consequently, this contribution's exploratory character, a predominant portion of the research scope is examined without any prior expectations from the findings. Specifically, no hypotheses are associated with RQs 1, 2, and 4 – instead, the results and identified patterns are reported as they appear from the experiment and the post-study analyses, offering the preliminary foundation for follow-up investigations to build upon as they continue to bridge the knowledge gap.

In turn, considering the differentiation of four icon-positioning configurations and the second study's between-groups design, the following generalized hypothesis pertaining to RQ3 is formulated:

H1 There is a relationship between the placement of auditory misinformation warnings and the overall podcast listening experience.

In line with the open-ended nature of this work, no expected differences among the four placement-alternatives are made explicit in the hypothesis specification, whose generality is only reduced by operationalizing the “podcast listening experience” notion. To capture the latter, six quantitative dependent variables are considered for hypothesis testing (podcast content recall, warning acclimatization, as well as four Likert-matrix subjective evaluations, later described in Section 4.1), and their correlations with warning placement are separately investigated.

All remaining statistical tests (aiming to address RQ4) are conducted without setting a significance threshold, the obtained coefficients, effect sizes, and p values presented in-full as indicators of latent interactions that warrant further investigation in follow-up confirmatory studies. Coupled with the insights obtained from qualitative data analysis, this exploratory approach allows for the desirable multi-faceted examination of misinformation warnings in the novel and largely unique context of podcasting.

3.2. Materials

There are two key components underpinning this contribution: the podcasts tackling controversial scientific topics and the sound cues serving as auditory interventions. For the former, three subjects were picked that differ from the common choices in previous works on responsible opinion formation and for which there is widespread discussion and information online: “Nuclear Energy for Power Generation”, “Health Effects of Cell Phone Radiation”, as well as “Applications and Safety of GMO Foods”. After assembling a collection of Web resources for each, NotebookLM¹ was used to generate medium-length

¹<https://notebooklm.google/>

podcasts formatted as dialogues between a speaker guiding the conversation and a knowledgeable co-host answering the questions put forward by the other.

For the purposes of the study, several modifications were applied by issuing explicit instructions to NotebookLM. To keep the podcasts accessible and easy-to-follow for participants regardless of their prior knowledge and personal stance, a prompt was included to approach the topics from a scientific as well as beginner-friendly perspective. Furthermore, each podcast was purposefully injected with misinformation around which the integration of auditory warnings would be centered. From the collected Web resources (among which academic and scientific publications, NGO platforms, and international government sites), a few discuss common myths and misconceptions for each of the debated subjects: these falsehoods were interspersed into the generated audio, ranging from 4 to 5 occurrences uniformly across a podcast's duration. Crucially, to avoid trivializing the study by revealing the presence of misinformation, a prompt was included to avoid referring to the false statements as such, the speaker introducing them using phrases such as "I have read [that]", "I have the impression [that]", or "Some sources point out [that]" instead. For reproducibility, the instructions issued to NotebookLM are provided in Appendix B. The generated "Deep Dive" podcasts have the following specifications: Nuclear Energy (17'00", 5 misinformation injections), Cell Phone Radiation (19'45", 4 misinformation injections), GMO Foods (19'45", 5 misinformation injections); Appendix C includes detailed visualizations of each podcast's structure and timeline, full transcripts (where the added misconceptions are highlighted in red), and URLs of the sources supplied to NotebookLM.

This model was also used to generate content-based questions for the recall assessment, aiming to examine the warnings' potential impact on post-listening information retention. The output was manually screened and modified, rephrasing or replacing answer-options to ensure they are comparable in length and detail, similarly formatted, and plausible. After excluding questions found to be ambiguous and easy-to-misinterpret in pilot testing, one multiple-choice questionnaire per topic containing 15 items (with 4 answer-options each) was obtained. The complete question-sets are provided in Appendix D, and the generation prompts issued to NotebookLM are included in Appendix B.

The other external resource relevant for this contribution are the metaphorical auditory icons integrated as misinformation warnings. In accordance with high-salience guidelines established from prior research [42], the sound classes of alarms, animal noises, and natural/domestic sounds were chosen given the need to invoke urgency and trigger attention-capture. Within each of these categories, five cues were selected from an applications-based taxonomy of auditory icons [9] such that they have a consistent conceptual mapping and high identifiability: these attributes are relevant for the intervention to remain learnable and minimally disruptive to the podcast's ongoing narrative (and thus the listening experience). Table 3.1 shows the full list of 15 sound samples included in the first study: they were obtained from the Freesound² collection (all of them freely downloadable and modifiable under a Creative Commons license), chosen based on audio clarity and the production equipment used, and edited with Audacity³ to adjust their decibel-level (so it matches that of the podcast they are embedded into, avoiding jarring audio spikes) and shorten their duration where necessary.

²<https://freesound.org/>

³<https://www.audacityteam.org/>

Table 3.1: The complete list of auditory icons used throughout Study 1 and their characteristics.

Category	Sound Cue	Dur.	Original Sample
Alarms	Referee Whistle	00:01	https://freesound.org/people/joedeshon/sounds/78508/
	Car Honk	00:02	https://freesound.org/people/wanaki/sounds/569613/
	Bike Bell	00:01	https://freesound.org/people/pallen33/sounds/405628/
	Factory Warning	00:02	https://freesound.org/people/qubodup/sounds/832414/
	Sonar Ping	00:02	https://freesound.org/people/LG/sounds/20223/
Animals	Dog Bark	00:01	https://freesound.org/people/dersuperanton/sounds/433698/
	Owl Hoot	00:02	https://freesound.org/people/faxfaxfax/sounds/655380/
	Seagull Screech	00:03	https://freesound.org/people/dobroide/sounds/317432/
	Frog Croak	00:02	https://freesound.org/people/TheKingOfGeeks360/sounds/744307/
	Horse Whinny	00:03	https://freesound.org/people/Kubuzz/sounds/347036/
Natural & Domestic	Whip Cracking	00:01	https://freesound.org/people/JayRom01/sounds/615761/
	Paper Ripping	00:02	https://freesound.org/people/bdunis4/sounds/271471/
	Scrap Metal Falling	00:02	https://freesound.org/people/SamsterBirdies/sounds/587443/
	Thunder Rumbling	00:04	https://freesound.org/people/csengeri/sounds/434359/
	Glass Breaking	00:03	https://freesound.org/people/Kneeling/sounds/448004/

4

Method

This chapter provides an in-depth overview of the research methodology, covering the design choices, variables, study procedure, and data analysis techniques. It also elaborates on the strategies employed for cognitive-bias management and the participant samples for both crowdsourcing studies. The design and workflow decisions were also made explicit in an OSF pre-registration⁴ prior to study execution.

4.1. Variables and Measures

This section outlines the variables captured or manipulated throughout this contribution, elaborating on their purpose and intended interpretation. Table 4.1 offers an exhaustive list of all these attributes and metrics, indicating their type (and, for quantitative data, their domain/scale).

Independent Variables

Warning Placement Per RQ3, four placement-configurations for the auditory intervention are investigated – before, after, enclosing (both before and after), or concurrent with a misinformation segment (Figure 4.1). In addition, a control group that does not encounter the warnings (the “without” condition) is considered for the recall and veracity assessments that are part of Study 2: for both components, this fifth category serves as a baseline to comparatively examine the impact of auditory-warning integration. Since the latter is a prerequisite to all subsequent data collection (such as the subjective sound cue evaluations, cf. Section 4.2), the “without” condition is excluded during the remaining findings’ analysis.

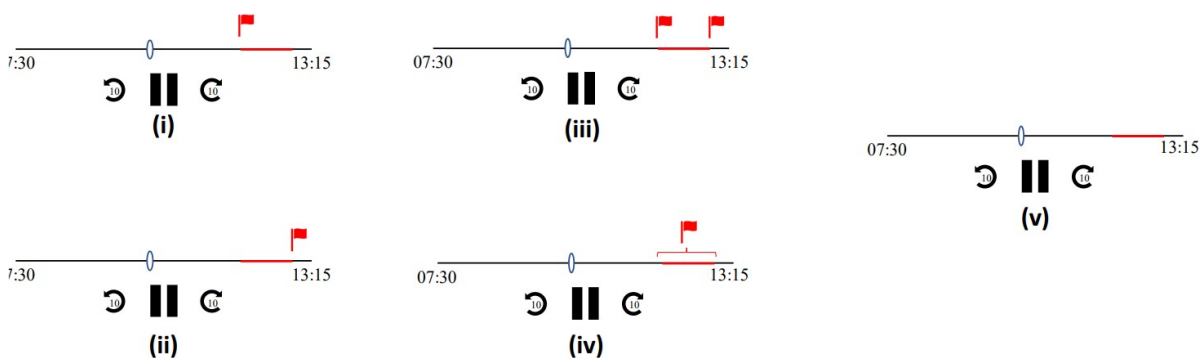


Figure 4.1: An illustration of the 4 warning-placement alternatives (red flags) with respect to the misinformation segment (red horizontal-bar portions) – before (i), after (ii), enclosing (iii), and concurrent (iv) – as well as the control-group condition without sound cues (v), this figure originally included in [49].

⁴<https://doi.org/10.17605/OSF.IO/2N4C8>

Table 4.1: The complete list of variables outlined in this contribution and their characteristics.

Category	Name	Type	Scale
Independent	Warning Placement	Nominal	W.r.t. a Misinformation Segment: Before / After / Enclosing / Concurrent (4-way split in Study 2)
Dependent	Podcast Content Recall	Count Proportion	$\llbracket 0; 15 \rrbracket$ (+) [0.0; 1.0]
	Presumed Sound Purpose	Qualitative	–
	Pausing/Backtracking Statistics	Count Qualitative	$\mathbb{N}_0$
	Misinformation Recognition	Binary Qualitative	Yes / No
	Conceptual Mapping	Qualitative	–
	Elicited First Reaction	Qualitative	–
	Warning Acclimatization	Ordinal Qualitative	Became less pronounced / Remained stable throughout / Became more pronounced (3-pt)
	Sound Recognizability	Ordinal	Not at all $\leftrightarrow$ Extremely (5-pt)
	Sound-Induced Distraction	Ordinal	Not at all $\leftrightarrow$ Extremely (5-pt)
	Sound Abruptness	Ordinal	Not at all $\leftrightarrow$ Extremely (5-pt)
	Sound-Induced Interruption	Ordinal	Not at all $\leftrightarrow$ Extremely (5-pt)
Exploratory	Topic Familiarity	Ordinal	Non-existent $\leftrightarrow$ Excellent (7-pt)
	Pre-Existing Personal Stance	Ordinal	Strongly (dis)agree (7-pt)
	Attitude Strength	Composite	Not important at all $\leftrightarrow$ Extremely important (+) Strongly (dis)agree (7-pt)
	Need for Cognition	Composite	Extremely (un)like me (5-pt)
	Intellectual Humility	Composite	Strongly (dis)agree (9-pt)
	Perceived Podcast Density	Ordinal	Too sparse $\leftrightarrow$ Too cluttered (5-pt)
	Perceived Podcast Pace	Ordinal	Very slow $\leftrightarrow$ Very fast (5-pt)
	Induced Warning-Aware Immediate Reaction	Qualitative	–
Control	Listener Category	Nominal	Selected: Attentive Listeners (no multitasking involved)
	Podcast Format	Nominal	Selected: Dialogue
	Debated Topic	Nominal	3 Selected: Nuclear Energy, Cell Phone Radiation, GMO Foods
Profiling	Frequency of Podcast Listening	Ordinal	Daily / Weekly / Monthly / Less than monthly (4-pt)
	Amount of Podcast Listening	Count	$\mathbb{N}_0$ (3 scales: years, hours/week, podcasts/month)

Dependent Variables

Podcast Content Recall Via the three NotebookLM-generated questionnaires, listeners' recall is assessed immediately after the podcast's duration has elapsed to examine how disruptive the differently-positioned warnings prove. Besides the number of correctly-answered questions (each questionnaire containing 15 in total), a more fine-grained breakdown of recall scores is also carried out by uniformly dividing the podcasts into three time intervals (each equal to one-third of the respective podcast's duration, cf. Appendix C) with the aim of capturing latent patterns that count data alone would not reveal (e.g. skew introduced by primacy or recency effects). These "early", "middle", and "late" recall-score proportions are based on the following partitions:

- Nuclear Energy – Q1-3 (early), Q4-11 (middle), Q12-15 (late);
- Cell Phone Radiation – Q1-5 (early), Q6-10 (middle), Q11-15 (late);
- GMO Foods – Q1-4 (early), Q5-11 (middle), Q12-15 (late).

Presumed Sound Purpose A brief-reflection prompt to inspect what participants assume as the cue's intended role(s) and to thereby determine how intuitive its function as a misinformation indicator is.

Pausing/Backtracking Statistics The number of times participants clicked the "Pause" / "Backtrack-by-10-Seconds" buttons while listening to the podcast, as well as a brief explanation of their reason(s) for doing so requested afterwards to contextualize those counts and confirm/discard the possibility that they occurred in reaction to encountering a misinformation warning mid-podcast.

Misinformation Recognition The number of times a piece of injected misinformation was recognized as such by participants across warning-placement conditions, alongside a free-text rationale for why they believed it was (not) a falsehood, these responses being used to determine whether the sound cues were heeded as "hazard" indicators. Across the 3 podcast topics, a pair of benign and misinformative excerpts were selected for this veracity assessment (Table 4.2), participants being asked to indicate whether the entire discussion presented within is factually accurate (separately for both snippets).

Conceptual Mapping An examination of the real-world entities participants associate with any given sound sample, included to ascertain the consistency and degree of ambiguity characterizing that sound cue and determine whether the desirable properties of auditory-icon-based interventions (i.e. learnability and pre-existing familiarity) are preserved.

Elicited First Reaction Participants' immediate perceptions of the auditory warnings upon encountering the latter for the first time (a source of must-have qualitative data crucial for contextualizing the numerical findings and observing the complete variety of first-time reactions without limiting the discussion to solely attention-capture and a heightened sense of urgency).

Warning Acclimatization An indicator of habituation potentially taking place as the misinformation warnings re-occur throughout the podcast (an ordinal scale and a free-text prompt being used to capture any changes in listeners' perceptions of the sound cues).

Likert-Matrix Subjective Evaluations A collection of four fine-grained subjective evaluations targeting specific sound-cue aspects via a 5-point scale (Not at all / Slightly / Moderately / Very / Extremely):

- Sound Recognizability – to what extent the real-world entity a listener associated with the sound was easy to recognize upon hearing it;
- Sound-Induced Distraction – how distracting a listener found the sound mid-podcast;
- Sound Abruptness – to what extent the sound seemed out of place whenever it occurred (seeking to capture how intuitive each warning-placement alternative appeared);
- Sound-Induced Interruption – how disruptive the sound was of the overall listening experience (including aspects like the podcast content's discernibility and the ongoing narrative's subsistence).

Table 4.2: Pairs of podcast excerpts used for the veracity assessment in Study 2, one factually accurate and another containing injected misinformation (highlighted in yellow) chosen for every debated topic. The misinformative snippets are **not** accompanied by any warnings during this stage to avoid priming participants' responses and revealing the sounds' purpose prematurely.

Topic		Podcast Excerpt
Nuclear	Benign	Co-host: "Think of it like this. You can light a bucket of high octane gasoline with a spark and get an explosion – that is the bomb. Reactor fuel is like a wet log. It creates heat, sure, but it simply doesn't have the concentration of energy to detonate. The physics just isn't there."
	Myth	Host: "Speaking of technology, I want to look forward. Because those big concrete domes, they feel very mid-century. I have read that nuclear energy is essentially outdated technology and that we don't really need it anymore because renewables can do the whole job. "
Cell Phone	Benign	Co-host: "Organizations like the American Academy of Pediatrics are actively demanding that governments overhaul their testing protocols. They argue that our current FCC guidelines, many of which were established back in 1996 during the absolute infancy of the modern telecom era, completely fail to account for the unique anatomical vulnerabilities of pregnant women and growing children."
	Myth	Co-host: "It proved that non-thermal biological effects like the development of tumors are theoretically possible under extreme prolonged conditions." Host: "That shift in understanding the biological boundaries brings up a concern I keep seeing as cellular networks upgrade. Some sources point out that 5G frequencies replicate inside the body and cause re-radiation. "
GMO Food	Benign	Co-host: "We are now moving into an era of targeted gene expression. It's incredibly helpful to contrast this modern targeted process with traditional selective breeding which humans have relied on for millennia, of course. But with traditional cross-breeding or even mutagenesis where seeds are exposed to radiation to force random mutations, you are mixing tens of thousands of genes at once. You're rolling the genetic dice."
	Myth	Co-host: "They immediately scrapped the entire project. It never reached the farm, let alone the grocery store. The testing mechanism worked exactly as designed." Host: "That's reassuring. But I have the impression that all research on GMOs has been funded by big agricultural companies. If the massive corporations making the seeds are the ones paying for the safety tests, the public trust is naturally going to be zero."

Exploratory Variables

Topic Familiarity A participant's self-assessed prior knowledge on the debated topic of their assigned podcast, serving as an additional predictor to capture latent interactions with the outcome variables.

Pre-Existing Personal Stance A participant's prior attitude towards their assigned controversial podcast topic, included to make the impact of relevant cognitive biases (such as confirmation bias / affect heuristic) on the response variables explicit and captured based on (dis)agreement with the following scientifically-backed statements (w.r.t. the podcast subject):

- Nuclear Energy for Power Generation – "Nuclear power is among the cleanest, safest, and most reliable energy sources supplying the electricity grid.";

- Health Effects of Cell Phone Radiation – “The energy emitted by cell phones has no known interactions with human anatomy besides causing tissue heating and occasional nerve stimulation.”;
- Applications and Safety of GMO Foods – “Genetically modified foods authorized for commercial distribution are safe for human consumption and do not pose a direct environmental hazard.”.

Attitude Strength A composite (approximately continuous) value calculated as the mean of 3 ordinal predictors captured using 7-point Likert scales based on the following prompts (from a prior user study concerned with opinion-change [67]):

- Attitude Importance – “How important is your attitude on this topic to you?”;
- Attitude Elaboration – “To what extent do you agree with the following statement: ‘Your attitude on this topic is a result of careful thinking about relevant information.’?”;
- Attitude Moralization – “To what extent do you agree with the following statement: ‘Your attitude on this topic is connected to your core moral values.’?”.

Need for Cognition A psychometric measure known to influence searchers’ information processing and interactions with sources [30] – computed as a composite (approximately continuous) score equal to the mean of 6 (occasionally reverse-scored) items composing the widely-used NfC-6 scale [41].

Intellectual Humility A psychometric measure indicative of an individual’s limitation self-awareness and receptiveness to counter-attitudinal information [52, 25, 33] (therefore, included to capture its potential influence on warning reception) that is treated as composite (approximately continuous) – equal to the mean of 12 (occasionally reverse-scored) items making up the Limitations-Owning scale [27].

Perceived Podcast Density A listener’s perception of density as the amount of information conveyed in the podcast (sparse vs. cluttered) – potentially informative for recall scores and warning reception.

Perceived Podcast Pace A listener’s perception of dialogue pace (slow vs. fast) – an extra variable with (possibly non-trivial) latent influences on any of the observed outcomes.

Induced Warning-Aware Immediate Reaction One of two free-text prompts included after Study 2’s debriefing to offer an early glimpse into participants’ potential interactions with misinformation warnings during regular podcast-listening sessions (outside experimental settings) – this one inquiring about their plausible immediate reaction upon encountering one of these interventions.

Induced Warning-Aware Downstream Behavior The second post-debrief exploratory prompt at the end of Study 2, asking participants what their follow-up actions would be (if any) in response to hearing a warning during their regular podcasts (this way, offering insights into how the intervention could affect a listener’s subsequent decision-making w.r.t. misinformation).

Control Variables

Listener Category The focus of this work is attentive podcast listening, which implies no multitasking is taking place, and listeners are processing solely the audio content, keeping track of the ongoing narrative and between-host discussions. Controlling for the listener category is a design choice made in line with the research objectives of this exploratory contribution: namely, identifying the factors that influence listeners’ perceptions of auditory misinformation warnings and informing first-iteration guidelines for the interventions’ effective integration into publicly accessible podcasts. These factors would be expressed the most clearly and noise-free – and listeners’ subjective impressions of the sound cues will be the most reliable – in a setting where the misinformation-injected podcast assumes a central role, and external distractions are minimized: this corresponds to the selected attentive-listening condition. Accordingly, specific measures were taken while designing the study – like the emphasis on desktop / headphone usage and certain audio-player restrictions (described in Section 4.2).

Podcast Format Based on the capabilities of NotebookLM for podcast generation, the dialogue format was chosen: this two-speaker discussion style is nevertheless predominant in the “podosphere”.

Debated Topic Three fundamentally-scientific controversial topics were chosen in order to randomize the assigned podcast and prevent the results from being fully attributable to the choice of one particular debated subject: the three being nuclear energy, cell phone radiation, and GMO foods. This selection was also made to deviate from commonly-picked topics in the literature on responsible opinion formation (abortion, obesity, gun control, etc.), introducing variety and making the study more engaging.

Participant Profiling

Frequency of Podcast Listening A measure included in Study 2 to characterize the participant sample’s overall listening habits (for reproducibility purposes) using the question-formulation from [59].

Amount of Podcast Listening 3 measures added to Study 2 for extra contextualization – listening trends in terms of years, hours per week, and podcasts per month (also based on [59]).

4.2. Experimental Setup and Procedure

This section describes the procedure and execution of the two consecutive crowdsourcing studies that underpin the research contribution. The web application created for the purpose of obtaining participant responses is presented via frontend views of the workflow in Appendix F, and its codebase is accessible online⁵. Considering the need for human research subjects, the studies were approved by the Human Research Ethics Committee at the Delft University of Technology prior to their deployment, all relevant informed consent materials included in Appendix E.

4.2.1. Study 1

The first study is centered around a multi-faceted evaluation of all 15 metaphorical auditory icons considered in this contribution (cf. Section 3.2). The criteria by which participants carry out this assessment are a sound cue’s recognizability, conceptual ambiguity, and degree of introduced disruption. Using the findings, the study’s objective is to identify the samples that are the most salient and consistently (as well as favorably) perceived by listeners: those auditory icons would then assume the role of misinformation warnings in Study 2, which in turn involves listening to a full medium-length podcast.

Study 1’s workflow consists of 3 stages and is unidirectional, participants being unable to return to a previous stage after progressing further. First, crowdworkers are presented with an opening statement that elaborates on the study’s structure and research objectives, offers work-environment recommendations, and specifies the collection of personal research data. After launching the study (thereby also granting consent), participants are asked to indicate their age bracket, gender, and native language (an opt-out choice included for each), and they listen to a mandatory volume calibration sample before proceeding to the main (final) stage. The latter contains three podcast excerpts (one per debated topic), into all of which one of the 15 samples is inserted: these snippets do not contain misinformation, instead simply used to re-create a podcast-listening scenario by wrapping a dialogue-style narrative around each sound-cue encounter. Due to the lack of falsehoods to serve as a vantage point, auditory-icon placement is also not considered as a governing independent variable: the sounds are interspersed between utterances, appearing twice in each podcast snippet (shown in Table 4.3). After listening to all three at least once, participants reflect on the sound’s connotations via two free-text responses (conceptual mapping & elicited first reaction) and evaluate its impact on the listening experience using the four Likert-matrix scales – this concludes Study 1.

4.2.2. Study 2

This study applies the salient and minimally disruptive sound cues from Study 1 to the context of podcast listening, where they serve the functional purpose of misinformation warnings. Using a between-subjects design based on the intervention’s placement (w.r.t. a misinformative segment), Study 2 aims to compare these warnings’ effectiveness, perceptual prominence, and induced misinformation awareness across the 4 configurations and identify any latent interactions or external factors at play.

⁵<https://github.com/palokando/cs5000-audio-project>

Table 4.3: Podcast excerpts used in Study 1, the sound samples' points of insertion also indicated.

Topic	Podcast Excerpt
Nuclear	Co-host: “In the US, we use light water reactors. There are two main flavors, and the names are refreshingly literal – pressurized water reactors and boiling water reactors.” Host: “I’m guessing the difference is whether the water boils.” Co-host: “You nailed it. In a boiling water reactor, the water surrounding the fuel boils directly into steam that heads off to the turbine. It is like a pot on the stove. [SOUND] The pressurized water reactor, which is more common, is more like a pressure cooker. They keep the water around the nuclear fuel under insane pressure so it gets to 600° F, but cannot turn into steam. [SOUND] Because of the pressure, it stays liquid. They pump that super hot water through a heat exchanger to boil a completely separate supply of water.”
Cell Phone	Co-host: “You really have to draw a hard line between two completely different categories, which are ionizing radiation and non-ionizing radiation.” Host: “Let’s break those down.” Co-host: “Ionizing radiation is the high frequency, high energy stuff. This is what you encounter with X-rays, gamma rays, or ultraviolet light from the sun.” Host: “We’re talking about energy waves that carry so much kinetic power they can physically alter matter at the atomic level.” Co-host: “Yes, ionizing radiation carries photons with enough energy to literally knock electrons out of their orbit around an atom. [SOUND] But non-ionizing radiation operates on the completely opposite end of the spectrum – it is low frequency and low energy. This category includes things like FM radio waves, microwaves, and the RF energy used by your cell phone, your home Wi-Fi router, and 5G networks. [SOUND] Because it cannot break chemical bonds, the primary biological mechanism it actually does have is tissue heating. It’s very similar to just stepping out into direct sunlight.”
GMO Food	Co-host: “We really have to look closely at the harsh biology of the digestive system. Because whether a cow, or a human for that matter, eats a protein from a completely natural wild plant or a genetically modified soybean, the stomach treats it the exact same way.” Host: “Which is by destroying it, right?” Co-host: “Utterly. [SOUND] The stomach environment is just a vat of hydrochloric acid. It’s got a pH level around 1.5, which is extremely acidic. And combined with aggressive digestive enzymes, this environment just breaks down complex biological structures. It unwinds the DNA and chops the proteins down into individual tiny building blocks called amino acids. [SOUND] The body doesn’t see a biological blueprint at that stage. It only sees basic building blocks.”

Unlike Study 1, the second relies on partial disclosure to obscure the genuine research objectives, the opening statement suggesting its focus is the immediate recall of podcast content measured using a multiple-choice questionnaire to identify patterns of listener engagement. This is done to preserve the authenticity of the obtained responses as going into the study with the mental preparation for encountering misinformation warnings – and the knowledge to interpret them as such – would skew the results and invalidate any conclusions. To justify the design choice and request consent under full disclosure,

a dedicated debriefing stage is included at the end (relevant materials presented in Appendix E).

Study 2 consists of 6 to 7 stages depending on whether a participant is assigned to the no-warning control group (in which case one stage is omitted). After proceeding past the opening statement (thus granting partial-disclosure consent), crowdworkers are asked to complete a profiling questionnaire that includes (optional) personal data (age, gender, and native language), podcast listening habits, and psychometric characteristics (ranging from topic familiarity to intellectual humility). Following mandatory volume calibration, they are then directed to the podcast listening stage, which consists of an audio player with a pre-loaded file (covering one of the three debated topics at random) that only exposes pausing and backtracking functionality: fast-forwarding and seeking are disabled to ensure listeners can only progress to the next stage once the podcast's full duration has elapsed. The recall assessment comes after, the topic-appropriate questionnaire (Appendix D) being pre-loaded and several extra variables being captured (presumed sound purpose, pausing/backtracking rationales, perceived density and pace). Afterwards, the veracity assessment (centered around the misinformation recognition variable) is carried out to observe whether the warnings are (i) associated with the hazard of misinformation as intended and (ii) heeded by participants when determining if a podcast excerpt contains falsehoods. Omitted for control-group crowdworkers, the next stage involves evaluating the sound cues based on the set of criteria from Study 1, now extended with the warning acclimatization variable: this redundancy in subjective-data collection is used to validate and expand upon the findings from the first study, considering the sounds' now-purposeful integration into podcast content. As discussed, Study 2 concludes with a debriefing form that offers full disclosure, requests informed consent, and is accompanied by two final exploratory variables (one's immediate reaction and downstream behavior potentially triggered by encountering a misinformation warning in non-experimental settings).

4.3. Cognitive Bias Checklists

When studies are deployed on a crowdsourcing platform for data collection, a wide range of cognitive biases manifesting as deviations from rational decision-making can affect participants' answers and judgments, these persisted results yielding a skewed and potentially unreliable representation of the real-world phenomenon being investigated [23]. In turn, this may have significant downstream implications for the accuracy and applicability of systems that are built, tuned, and verified based on such findings. In the context of crowdsourcing as a commonly-used tool to fulfill these data needs, several factors driving the prominence of cognitive biases are present, including a study's difficulty, level of clarity, duration, and reliance on memory or recall: the higher the cognitive demands, uncertainty, and time pressure a participant is exposed to while completing it, the more pronounced their tendency to abuse cognitive shortcuts and exhibit biases becomes [23].

To make these considerations explicit and assist requesters in devising a more robust study design, a 12-item checklist has been proposed [15], outlining a set of influential cognitive biases motivated by existing crowdsourcing literature. Per the authors' recommendations, the checklist was applied before data collection – as a step in the design process of the studies underpinning this contribution. In accordance with best requester practices, means of measuring and/or mitigating certain biases were added during implementation while those that could not be circumvented were nonetheless acknowledged. The following sub-sections elaborate on the outcomes of this checklist-driven analysis.

4.3.1. Study 1

The results of applying the checklist to Study 1 are described in Table 4.4. To prevent submissions by profit-maximizing and “satisficing” workers from introducing inaccuracies and data noise (self-interest bias), an attention-check question is included in the subjective-evaluations page as a common quality-control technique which filters out inattentive workers and bots. Mitigating the impact of affect heuristic and availability bias, meanwhile, requires explicitly motivating the research and emphasizing the role that the sound cues are intended to fulfill. Through precise question-formulation (i.e. inquiring about specific aspects, such as the sound's recognizability and abruptness, rather than a generalized notion of “perceived appropriateness”), workers are prompted to conduct a multi-faceted evaluation of each auditory icon at a lower level of granularity – this could help decouple (some of) their judgments from both pre-existing personal preferences and exposure to past examples.

Salience bias is the one item whose influence is largely the focus of the study, several dependent variables capturing the perceptual prominence of an auditory icon. To that end, the only (vital) mitigation

step was to regulate the volume of all 15 featured sound samples, adjusting their decibel-levels to match the surrounding podcast excerpts they were integrated into. Alongside the volume calibration step in the study introduction, this measure would prevent unrelated factors (like the sound being overbearing or, conversely, insufficiently audible) from skewing salience-related findings.

Table 4.4: Application of the Cognitive-Biases-in-Crowdsourcing Checklist to Study 1.

Cognitive Bias	Present?	Impact
Self-Interest Bias	Yes	Workers could attempt to speedrun the task, exerting the minimal required effort to reach the end, obtain the completion code, and receive payment in a short amount of time (“satisficing”).
Affect Heuristic	Yes	While subjectively evaluating a given sound cue, workers could be disproportionately lenient or strict in their judgment, primarily guided by their pre-existing personal preferences.
Bandwagon Effect (Groupthink)	No	N/A
Salience Bias	Yes	Workers’ evaluations of a given sound cue could be influenced by how perceptually prominent it appears in the audio snippets off which judgments are based.
Confirmation Bias	No	N/A
Availability Bias	Yes	A worker’s prior exposure to a given auditory cue (namely, how many examples of that sound they have heard previously – and how easily they are able to recall those) could skew the leniency or strictness of their judgments.
Anchoring Effect	No	N/A
Halo Effect	No	N/A
Sunk Cost Fallacy	No	N/A
Overconfidence (Optimism Bias)	No	N/A
Disaster Neglect	No	N/A
Loss Aversion	No	N/A

4.3.2. Study 2

The checklist-based analysis of Study 2 is presented in Table 4.5. Self-interest bias remains a concern in this longer procedure, and additional mitigation methods are employed: aside from attention-check questions, forward-skipping and seeking are disabled in the podcast listening stage to ensure participants can progress only after the full duration has elapsed, and submissions with anomalous response times are inspected to verify their integrity. Since this study is more difficult and time-consuming, the sunk cost fallacy could also manifest, urging workers to strive towards completion even if their interest has diminished and thereby causing performance to deteriorate. To mitigate this, the study’s expected duration of 45 minutes is made explicit upfront (in the brief synopsis posted on Prolific as well as the Opening Statement), and an option to exit prematurely is provided, meaningful responses being incentivized proportionately to workers’ time investment.

A collection of cognitive biases becomes prominent during both the objective recall assessment and the subjective warning-evaluation stage, driven by participants’ personal preferences (affect heuristic), pre-existing notions (confirmation bias), prior exposure (availability bias), and degree of issue-relevant reasoning (halo effect). Besides impacting the reception of podcast-integrated auditory warnings, these four attributes could affect a worker’s capacities for content retention and misinformation recognition by reinforcing cognitive “shortcuts” and introducing a viewpoint-aligned skew, which would remain latent in the resulting data. To manage the influence of these biases, three of the study’s features prove

instrumental: (i) the measurement of psychometric characteristics as dedicated exploratory variables (including topic familiarity, personal stance, and attitude strength) at the start, (ii) the formulation of precise and granular questions targeting one specific analysis-lens at a time (akin to the sound-judgment inquires from Study 1), and (iii) the inclusion of a standalone debriefing stage at the end, where a full justification of the conducted research and the auditory cues' integration is provided alongside two final questions aiming to capture participants' self-reported interactions with misinformation warnings – now in the circumstances of full disclosure.

Similarly to the first study, certain cognitive biases inadvertently overlap with the phenomena being investigated and therefore remain in-effect: here, these are the exhibited salience and anchoring bias. The former's dependence on how perceptually prominent an auditory warning appears will regulate the degree to which it influences the obtained results for sound recognizability, distraction, and interruption. Meanwhile, the latter's impact is two-fold: a listener's first encounter with a misinformation warning may affect their subsequent perceptions (primacy effect), and their performance on the recall assessment could be governed by how much time has passed since they heard a given piece of information (recency effect). Two components of the methodology for Study 2 were introduced specifically with these biases in mind: respectively, those are the warning acclimatization dependent variable (cf. Table 4.1) and the podcast segmentation scheme used in the interpretation of recall assessment scores.

4.4. Analyses

Considering the diverse range of variables examined in this exploratory work, employing suitable analysis techniques via across-method triangulation becomes key to processing and interpreting both quantitative and qualitative data. The former provide measurable evidence relevant to the RQs and help to identify relationships among the factors at play with a quantifiable degree of certainty – meanwhile, the latter enrich and contextualize these findings, revealing patterns that might otherwise be omitted.

Besides descriptive statistics, carrying out inferential analyses of quantitative results yields insights into the effect size and statistical significance of the observed interactions. To that end, a two-tiered approach was adopted: non-parametric statistical tests that serve as assumption-light preliminary indicators of latent relationships, followed by regression models that offer greater reliability by fitting the underlying distributions. In the first tier, the Kruskal-Wallis test⁶ was used to compare the five groups formed based on the warning-placement independent variable (the no-warning control group being considered alongside the 4-way split per RQ3), where the null hypothesis states that the population median across all groups is equal, and Kendall's Tau correlation⁷ (preferred due to the modest sample size of the study) was computed with the remaining explanatory variables to check for monotonic relationships between each of them and the quantitative dependent variables, treating both as ordered ranks. Each test calculates a statistic (H or τ) and a p value, Kruskal-Wallis also yielding ϵ^2 as a separate indicator of effect size.

In the second tier, three regression models were chosen based on the type of dependent variable under analysis: Poisson regression⁸ for count data (the number of correctly-answered questions from the recall assessment), Fractional Logistic regression⁹ for proportion data (the recall scores partitioned into early, middle, and late segments – this particular model capable of handling the inclusive interval $[0.0; 1.0]$), and Ordinal Logistic regression¹⁰ for ordinal data (the warning-acclimatization trend and the four Likert-matrix subjective evaluations). Due to the Poisson distribution's underlying characteristic of its mean and variance being equal, an extra measure is computed to verify that the obtained count data behaves accordingly: the overdispersion ratio R – the quotient of the Pearson's Chi-Squared statistic and the model's residual degrees of freedom – which is close to 1 for Poisson-distributed data. In case the recall-count measurements were overdispersed (i.e. a variance greater than the mean and $R > 1$), Negative Binomial regression would be a suitable alternative to the Poisson model.

Across the three regression implementations, four statistics are captured: the estimated coefficients β , each representing the logarithm of the difference in outcome, their exponentiations e^β , which make the magnitude of that change easier to determine by yielding the concrete ratios, the standard errors σ_β

⁶<https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kruskal.html>

⁷<https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kendalltau.html>

⁸https://www.statsmodels.org/stable/generated/statsmodels.discrete.discrete_model.Poisson.html

⁹https://www.statsmodels.org/stable/generated/statsmodels.genmod.generalized_linear_model.GLM.html

¹⁰https://www.statsmodels.org/stable/generated/statsmodels.miscmodels.ordinal_model.OrderedModel.html

Table 4.5: Application of the Cognitive-Biases-in-Crowdsourcing Checklist to Study 2.

Cognitive Bias	Present?	Impact
Self-Interest Bias	Yes	Workers could attempt to speedrun the task, exerting the minimal required effort to reach the end, obtain the completion code, and receive payment in a short amount of time (“satisficing”).
Affect Heuristic	Yes	A worker’s own partiality for and interest in a specific topic could influence their performance on the recall assessment, affecting content retention and perceived podcast density. This bias may also manifest in their subjective evaluations of the auditory misinformation warning.
Bandwagon Effect (Groupthink)	No	N/A
Salience Bias	Yes	The perceptual prominence of an auditory warning could impact its reception by listeners and, from an intervention-oriented perspective, its potential interplay with content retention and misinformation awareness.
Confirmation Bias	Yes	Workers’ preconceived notions concerning a given controversial topic could trigger selective viewpoint-aligned content retention, skewing results from the recall assessment, and affect the misinformation warning’s perceived appropriateness.
Availability Bias	Yes	A worker’s prior exposure to a podcast’s controversial topic, including stereotypical associations, personal elaboration, and familiarity with details, could influence both their content retention and misinformation recognition performance, potentially also affecting their reception of misinformation warnings.
Anchoring Effect	Yes	This bias could manifest in two ways: (i) a primacy effect may be observed in terms of a listener’s first exposure to a misinformation warning influencing subsequent encounters, & (ii) a recency effect w.r.t. a worker’s content retention being skewed towards information appearing later in (rather than uniformly across) the entire podcast may be captured.
Halo Effect	Yes	Irrelevant information not mentioned in the podcast (from one’s pre-existing knowledge on the subject instead, for instance) may affect workers’ performance on the recall assessment – causing them to misremember statements and arguments they heard.
Sunk Cost Fallacy	Yes	As this task is more cognitively involved (and longer) than Study 1, workers could be prompted to “persevere” and complete it the more time and effort they invest – even after losing their interest in it – which would cause their performance to deteriorate.
Overconfidence (Optimism Bias)	No	N/A
Disaster Neglect	No	N/A
Loss Aversion	No	N/A

(the standard deviations of the coefficient estimates), and the associated p values. The e^{β} multipliers themselves have different interpretations depending on the regression model (and, in turn, the dependent/outcome variable’s type): expected-count ratios (or incidence-rate ratios) for Poisson regression, recall-proportion odds ratios for Fractional Logistic regression, and higher-category odds ratios for Ordinal Logistic regression. To isolate an explanatory variable’s unique contribution and detect the influence

of predictor-confounding and latent correlations on the statistics obtained in univariable (raw) settings, multivariable (adjusted) regression models – where each predictor is estimated while holding the rest constant – were additionally carried out for greater observability and higher inferential robustness.

Handling categorical explanatory variables requires extra considerations. To avoid the dummy trap (where linearly dependent columns and perfect multicollinearity cause the regression model to try estimating infinite coefficients), nominal and ordinal predictors were dummy-coded, which involved dropping a sensibly-chosen reference category for each of them. In this work, all such explanatory variables have a “neutral” baseline that could naturally serve as a default: the “without” control-group condition for the recall assessment, the “concurrent” placement-configuration for the warnings’ subjective evaluations, and the “neither/nor” (or “moderately”) options in the middle of the psychometrics’ Likert scales. These reference levels are then implicitly coded whenever 0 appears across all columns of the remaining binary-coded categories, and the predictor itself is interpreted as factor: each category’s coefficients β and e^{β} quantify the observed change in outcome compared to the reference-group baseline, and the p values indicate the statistical significance of that category-specific difference. Besides dropping a reference level to address the dummy-trap problem, infrequent categories represented in less than 3% of the sample size were further excluded as they would introduce overhead to the regression model’s convergence, produce inflated standard errors, and lead to unstable (high-variance) coefficient estimates. A final measure to curtail near-collinearity, which also triggers these effects, was to calculate the variance inflation factor¹¹ for each predictor, verifying whether it significantly contributes to the variance of the model’s estimates ($VIF > 5.0$) and, if so, dropping its column from the design matrix.

To contextualize the descriptive and inferential statistics obtained from these methods, qualitative data were interpreted using the reflexive thematic analysis (RTA) technique [6, 4, 8]. In particular, an essentialist (realist) perspective was adopted, reporting the meanings and experiences of participants as articulated and reflected on through natural language [5]. The inductive (i.e. bottom-up or data-driven) RTA approach was employed, coding the data without trying to fit them into a pre-defined framework, and semantic themes were identified within the explicit meanings of the free-text responses [6, 8]. This entailed a two-step process of description (where the patterns were revealed and summarized) and interpretation (through which their significance and broader implications were elaborated further).

Lastly, it is relevant to outline the procedure in which auditory icons were selected to assume the role of misinformation warnings based on the findings from Study 1. In line with the benefits of learnability, minimal cognitive interference, and real-world familiarity offered by this sound cue type, a sample’s consistent conceptual mapping and high recognizability were the first two selection criteria. Furthermore, based on the notion of “design frictions” and the minor dissatisfaction associated with them, low reported levels of distraction and interruption were chosen as the remaining criteria: these would indicate that listeners experienced a brief narrative disruption yet were able to quickly recover from the warning encounter, refocusing on the podcast content without prolonged confusion. Abruptness was not taken into account for auditory-icon selection as placement and its effects were not investigated in Study 1, whose focus were the sound samples themselves: unlike Study 2, it followed a non-intentional between-utterance template to intersperse the sound cues into the three podcast excerpts rather than using misinformation occurrences as a vantage point. Specific thresholds for the three quantitative criteria were chosen off the “Very” Likert-scale option (recognizability $\geq 4 \wedge$ distraction $< 4 \wedge$ interruption < 4). In the event of ties, the elicited reactions (interpreted using RTA) were inspected, prioritizing the themes of attention-capture, minor (rather than prolonged) dissatisfaction, and podcast-content discernibility.

4.5. Participants

The two studies were published on the crowdsourcing platform Prolific¹² with a payment rate of 9 GBP per hour. In line with the targeted scenario of attentive podcast listening, participation was restricted to desktop usage only, and crowdworkers were encouraged to wear headphones and find a quiet environment where they could remain undisturbed. Due to the American-centered narration lens in the audio output produced by NotebookLM, a pre-screening criterion for the United States as a current country of residence was added to avoid skewing the obtained results; additional screeners concerned age (18-years-old or above) and language proficiency (fluent in English). The participant counts for each study were chosen with respect to the number of experimental arms, the contribution’s exploratory character

¹¹https://www.statsmodels.org/dev/generated/statsmodels.stats.outliers_influence.variance_inflation_factor.html

¹²<https://www.prolific.com/>

(where the emphasis is placed on subjective perceptions, shifting away from majority-vote aggregation methods and the notion of “ground-truth” data), and budget constraints.

4.5.1. Study 1

The first study involved 150 participants (5 of those submissions later discarded during post-processing due to attention-check failure), such that each sound cue would be evaluated by 10 crowdworkers. The median completion time was 6 minutes and 3 seconds, the initial estimate shown in the Prolific overview having been set to 8 minutes. The participant sample was diverse in terms of age (18-24: 18, 25-34: 41, 35-44: 48, 45-54: 24, 55-64: 8, 65+: 6) and gender (Female: 79, Male: 61, Non-binary: 4, Unspecified: 1), and the reported native language was primarily English (English: 135, Urdu: 2, Unspecified: 3, and five other languages indicated by a single crowdworker each).

4.5.2. Study 2

The second study featured 115 participants, who were evenly split across the five experimental arms (the control group without misinformation warnings, as well as the four warning-placement alternatives, each being assigned to 23 crowdworkers). 8 participants exited the study directly after the veracity assessment (using the Quit Study option explicitly provided): due to this, quantitative data from listeners' subjective impressions of the auditory cues (and the subsequent inferential analysis) were based on a slightly lower number of responses compared to the fully-obtained measures of podcast content recall and misinformation recognition (the concrete participant counts for subjective warning-evaluation were Before: 20, After: 20, Enclosing: 22, & Concurrent: 22). The median completion time was 40 minutes and 44 seconds, where the initial estimate specified on Prolific was set to 45 minutes. The participant sample for this study was also diverse with respect to age (18-24: 13, 25-34: 40, 35-44: 44, 45-54: 6, 55-64: 8, 65+: 4) and gender (Female: 58, Male: 56, Non-binary: 1), and English was once more the predominant native language (English: 110, Urdu: 1, Yoruba: 1, Unspecified: 3).

In terms of personal podcast listening habits, the majority of participants tune in frequently (Daily: 35, Weekly: 63, Monthly: 9, Less than monthly: 8) and have considerable prior exposure to this medium (listening amounts: 7.1 ± 5.2 years, 7.0 ± 7.2 hours per week, and 4.4 ± 3.7 podcasts per month).

5

Results

This chapter presents the results obtained from the 2 crowdsourcing studies and any subsequent data-processing methods – including inferential and reflexive thematic analyses. The quantitative evidence included here is demonstrated in tandem with qualitative findings to offer additional context and avoid misattributing the factors at play. The full results and all processing scripts are accessible online¹³.

5.1. Study 1

The findings yielded from listeners' subjective evaluations of the 15 auditory icons considered in Study 1 are described in this section, at the end of which the selection of samples carrying over to Study 2 as misinformation warnings is explained and justified.

Table 5.1: Subset of best-performing sound cues (highest recognizability, lowest distraction, and lowest interruption out of all 15 auditory icons included in Study 1). For each metric, μ values indicate the mean, σ values show the standard deviation, and the distribution of participant judgments across the 5 levels (Not at all / Slightly / Moderately / Very / Extremely) is reported based on the obtained Likert-scale input. See Table A.1 for the complete results.

Sound Cue	#evaluators	Recognizability		Distraction		Interruption	
		μ	σ	μ	σ	μ	σ
Referee Whistle	10	4.5	1.27	3.7	0.82	2.8	1.4
		1 / 0 / 0 / 1 / 8		0 / 0 / 5 / 3 / 2		2 / 3 / 1 / 3 / 1	
Bike Bell	10	4.8	0.42	2.6	1.51	2.4	1.35
		0 / 0 / 0 / 2 / 8		3 / 2 / 3 / 0 / 2		3 / 3 / 2 / 1 / 1	
Dog Bark	9	4.89	0.33	3.67	1.41	3.56	1.33
		0 / 0 / 0 / 1 / 8		0 / 3 / 1 / 1 / 4		0 / 3 / 1 / 2 / 3	
Owl Hoot	9	4.0	0.71	2.56	1.33	2.78	1.56
		0 / 0 / 2 / 5 / 2		2 / 3 / 2 / 1 / 1		2 / 3 / 1 / 1 / 2	
Paper Ripping	10	4.7	0.48	3.6	1.58	3.6	1.65
		0 / 0 / 0 / 3 / 7		2 / 0 / 2 / 2 / 4		2 / 0 / 3 / 0 / 5	
Scrap Metal Falling	10	4.6	0.7	3.5	1.27	3.4	1.26
		0 / 0 / 1 / 2 / 7		1 / 1 / 2 / 4 / 2		1 / 1 / 3 / 3 / 2	

¹³<https://github.com/palokando/cs5000-results-processing>

5.1.1. Quantitative Data (Descriptive Statistics)

In line with the criteria and desirable metric-thresholds described in Section 4.4, 6 auditory icons fulfilling those conditions were identified based on the quantitative findings from Study 1 (listed in Table 5.1).

All 6 sounds were considered representative of the real-world entities listeners associate with them, most judgments of each sample's recognizability falling within the Very / Extremely categories (and all six means for that metric being ≥ 4.0).

Two samples exhibited the lowest levels of induced distraction and interruption: the sounds of a bike bell and an owl hooting, for which the majority of evaluators gravitated towards the Not at all, Slightly, and Moderately options on the respective Likert scales (yielding means < 3.0).

Performing somewhat worse along these 2 criteria, the next-best pair consists of the referee whistle and dropped scrap metal sounds – neither of which fully dominates the other. The former has a lower interruption score ($\mu = 2.8$), yet the latter is perceived as less distracting by participants (1 Not at all and 1 Slightly, compared to 0 / 0 for the referee whistle).

Lastly, the sounds of a dog barking and a sheet of paper being ripped are both fully dominated by the rest in terms of introduced distraction and interruption, yielding comparable measurements that do not indicate one sample prevailing over the other within the pair.

5.1.2. Qualitative Findings

The samples included in Study 1 were also evaluated using free-text responses in order to capture the conceptual mappings, perceived functions, and elicited first-encounter reactions for each. The themes identified through RTA are presented below, accompanied by brief explanations where necessary and exemplified using selected excerpts from participants' submissions.

Conceptual Mapping

Listeners' real-world associations were highly accurate and consistent for nearly all samples considered in the study – one major exception was the factory warning sound. Its ambiguity becomes evident from the variety of obtained Conceptual Mapping responses, which include (i) a "basketball game buzzer" / "game show buzzer", (ii) a "forklift horn" / "warehouse machinery", and (iii) a "fire drill air horn" / "emergency door klaxon". Overall, the collected rationales also indicate some of the foundations underpinning participants' subjective sound interpretations, as demonstrated by the following selected excerpts.

Associations Based on Prior Personal Experience

- "I think of track and field days when our coaches would use whistles non-stop for various reasons."
- "I've heard owls late at night in my backyard, and it sounded a lot like that."
- "It sounds like crickets and frogs. Reminds me of camping or attending a cookout, or even swimming in the evening when the sun is just starting to go down."

Associations Based on Popular Culture

- "It reminds me of a radar/sonar sound effect you would hear in a movie or video game."
- "It could have been a sound effect from an Indiana Jones movie." (whip crack)

Presumed Sound Purpose

Several perceived functions fulfilled by the auditory icons were also denoted in participants' responses, the samples being grouped into the three functional categories below.

Attention Grabber The sound was aimed at making listeners more alert, returning and/or maintaining their focus on the conversation (referee whistle, bike bell, owl hooting, horse whinny, whip crack, paper ripping, scrap metal falling, glass breaking).

- "It was used as a way to punctuate the conversation."

Transition Marker The sound was used to partition the conversation into sections based on changing sub-topics (bike bell, factory warning, paper ripping).

- “It was meant to invoke the sense of a new topic’s being broached.”

Indicator of Something Amiss The sound was used to stress some negative aspect of the conversation (referee whistle, factory warning, and the samples below).

- “It reminded me of a fact-check.” (dog bark)
- “It sounds like something bad just happened.” (glass breaking)
- “It was used to make something seem more dramatic.” (glass breaking)

Elicited First Reaction

Each participant’s initial encounter with their assigned sound sample prompted a variety of immediate reactions, the following thematic groups being identified through RTA.

Surprise All 15 auditory icons were unexpected in their appearance, listeners taken aback at first.

- “I thought it was a mistake in the snippet.”
- “It caught me off guard.” / “It came out of nowhere.”

Fright Certain samples were considerably more shocking/startling than others at their first occurrence (referee whistle, car honk, factory warning, whip crack, thunder rumbling, glass breaking).

- “I was a bit startled, because I was actually getting invested in the conversation.”

Distraction and Interruption Some auditory icons were considered particularly jarring and annoying by listeners, throwing them off the conversation’s ongoing narrative (referee whistle, car honk, factory warning, sonar ping, frog croak, horse whinny, and the samples below).

- “The sound made me jump slightly and cringe at its unpleasantness. Once my brain had placed it a half second later, it wasn’t as bad, just moderately annoying.” (seagull screech)
- “I felt it ruined the flow of the conversation and contributed zero substance to the topic.” (seagull screech)
- “It disturbs the momentum of the conversation and shouldn’t be a part of the audio.” (whip crack)
- “It did make it difficult to concentrate on the educational content.” (thunder rumbling)
- “If I did not have my volume at a low, comfortable level, it would’ve caused some distress.” (glass breaking)

Blending with Perceived Podcast Environment Certain cues were contextually masked within the imaginary podcast setting perceived by listeners, appearing as background noise rather than an intentionally introduced effect.

- “It sounds like one of the hosts tearing a piece of paper from their notepad.” (paper ripping)

Overlapping with External Stimuli Samples may also be overridden by cognitive associations reinforced through prior exposure or by matching sounds from a listener’s surroundings.

- “I assumed my neighbors dogs were barking again and then heeded it no mind. It took me until the second time to realize that it wasn’t them but rather the distraction.” (dog bark)

Habituation Sounds may be passively processed either through repeated exposure, or when causing too much interference, listeners prioritizing the ongoing conversation instead.

- “It faded into background noise for the rest of the listening experience.” (car honk)

No Explicit Reaction

- “I just paused my thoughts.”
- “I didn’t have a reaction. I simply recognized the sound.”
- “I don’t think I really cared much.”

Mismatch with Pre-Existing Sound-Cue Expectations The sample is not typical for a sound cue.

- “I was expecting something more like static.” (car honk) / “I thought it was going to be like a loud high-pitched beep or harsh buzz.” (seagull screech)
- “I was expecting it to be much more subtle and difficult to notice.” (dog bark)
- “I was not expecting a normal environmental sound.” (dog bark)
- “I was surprised it was wildlife as the sound interrupting the conversation. I would have expected a sound more in-line with an indoor and human setting as the interruption.” (frog croak)

Initial Confusion The aforementioned unintuitive and overly distracting auditory icons, which are not expected or usually encountered in regular dialogue, exhibit no clear connection with the conversation topic or the surrounding context, thus creating confusion.

- “I found it odd and slightly annoying, because it had nothing to do with anything the subjects were discussing, nor did it seem to guide the audio in any way. It was really just an interruption.”
- “It felt out of place for me, as it had no business being there and was completely out of context from the subjects of these audio clips.”
- “I was confused about how it correlates to what the people in the audio sample were talking about, trying to see if there was any substantive reason for it to be included.”

Amusement The most contextually-mismatching and least expected sound cues were perceived as silly and ridiculous, breaking listeners’ immersion into the conversational narrative.

- “I actually laughed at it because of how comically out of place it was.” (car honk)
- “Sounds funny to me, because it had nothing to do with the conversation.” (seagull screech)
- “It was such a disassociation from the conversation that it made me chuckle a bit.” (frog croak)
- “It’s unprofessional, and it made me laugh.” (whip crack)

5.1.3. Selected Carryover Auditory Icons

In line with the quantitative data from Table 5.1, indicating the auditory icons that satisfy the threshold-criteria described in Section 4.4, and the insights obtained from participants’ free-text responses, three samples were selected to act as misinformation warnings in the second study: similarly to a participant’s assigned controversial podcast topic, one of these sounds is randomly determined while their instance of the study is being prepared – the purpose of this randomization being to introduce more variety and prevent the results from being attributable to the choice of one unique auditory icon.

The first two of these samples are the **bike bell** and **owl hoot** sounds: both exhibit the lowest levels of induced distraction and interruption, are identified as attention-capturing markers, and do not appear among the cues perceived as particularly disruptive by participants. The sound of **scrap metal falling** is chosen as the third auditory icon included in Study 2, essentially being a middle-ground selection out of the remaining 4 samples from Table 5.1 (that is nonetheless salient and induces a sense of urgency) based on the following considerations:

- **Referee Whistle** – associated with considerable listener dissatisfaction (namely, a “jarring” sound that “could cause a headache”) despite being perceived as prominent and important;

- **Dog Bark** – prone to overlapping with environmental stimuli or being considered ridiculous (“completely out of place”) due to an insurmountable contextual mismatch;
- **Paper Ripping** – prone to blending with the perceived podcast setting, thereby becoming masked as background noise rather than a salient intentionally-introduced marker.

5.2. Study 2

The findings obtained through the objective assessments, subjective evaluations, and free-text reflections included in Study 2 are presented in this section, demonstrating the impact of podcast-integrated misinformation warnings, their reception by listeners, and the latent influence of contextual factors that follow-up studies should account for.

5.2.1. Quantitative Data (Descriptive Statistics)

The quantitative results from Study 2 are grouped based on participants’ assigned study condition, podcast topic, and/or warning-sound sample to substantiate any differences and assist in drawing comparisons. The latter two grouping criteria are also considered to make the impact of the selected debated topics and sound cues more explicit – one of each being randomly assigned to every crowdworker to introduce variety and thus lower the likelihood of one-off (circumstantial) findings. Nonetheless, those groupings are included in the present discussion to account for any underlying interactions and improve analytical robustness. The sample sizes across all assignments are as follows:

- Study condition (relevant for the recall and veracity assessments): without – 23, before – 23, after – 23, enclosing – 23, concurrent – 23;
- Placement configuration (relevant for the subjective warning evaluations): before – 20, after – 20, enclosing – 22, concurrent – 22 (cf. Section 4.5 and the premature-exit occurrences);
- Sound samples (alongside their subjective-evaluation counts corrected for premature exits): bike bell – 33 (29), owl hoot – 35 (32), scrap metal falling – 24 (23);
- Debated topics: nuclear energy – 34, cell phone radiation – 39, GMO foods – 42.

The scores obtained from the recall assessment are presented in Table 5.2. Out of the three controversial topics, the subject of GMO foods is associated with the highest average content recall (for both count and segment-wise proportion data), nuclear energy and cell phone radiation yielding somewhat lower (and mutually comparable) scores across the four recall-metric types. Regarding the five study conditions, a consistent pattern may be observed in the count, middle-proportion, and (to a lesser extent) late-proportion results: concurrently-placed warnings exhibit the lowest recall scores, slightly outperformed by the *enclosing* configuration, which is in turn dominated by the remaining three conditions (whose values are internally similar). The only exception appears to be the early-segment recall proportion, where the *enclosing* placement performs the poorest ($\mu = 0.7$) while *concurrent*’s score is aligned with those of the *without* and *after* conditions. Despite these trends, the Kruskal-Wallis test executed based on the nominal warning-placement variable offers no preliminary indications of statistical significance across the 5 study-condition groups for any of the 4 recall-metric outcomes (Table 5.7).

Inspecting the measurements of perceived podcast density & pace, the 3 controversial topics yield similar Likert-scale distributions: for both metrics, judgments gravitate towards the 3rd and 4th category (Moderately and Cluttered / Fast, respectively) across all subjects, the central (neutral) level carrying the largest weight. Consequently, the three podcast-density averages are all approximately equal ($\mu \approx 3.5$) while those for podcast pace fall roughly within the range [3.1; 3.2] (Table 5.3).

The functionalities for pausing and backtracking 10 seconds were used quite infrequently across all 5 study conditions: the average pause-click counts range from 0.26 (*enclosing*) to 1.22 (*after*), whereas the backtrack-click averages vary between 0.78 (*before*) and 1.39 (*enclosing*). Having been added for the contextualization of these metrics, participants’ rationales for pressing either of the two buttons do not demonstrate any connection between interacting with the functionalities and encountering a misinformation warning: instead, the overarching reasons include (i) briefly double-checking segments that were misheard or inattentively processed, (ii) repeating the more scientifically-involved and dense discussions to help with concept clarification, and (iii) recovering from external interruptions and refocusing one’s attention onto the podcast.

Table 5.2: Recall assessment scores in terms of the overall count (up to 15) and the proportion (based on the early / middle / late partitions) of correctly-answered questions across the 3 debated topics and the 5 warning-placement alternatives (including the control-group condition). For each metric, μ values indicate the mean, and σ values show the standard deviation.

Topic / Condition	Count		Early		Middle		Late	
	μ	σ	μ	σ	μ	σ	μ	σ
Nuclear	11.65	3.5	0.8	0.2	0.78	0.26	0.74	0.3
Cell Phone	11.0	3.76	0.7	0.27	0.76	0.28	0.75	0.29
GMO Food	12.45	3.0	0.86	0.22	0.82	0.22	0.82	0.25
Without	12.2	3.25	0.8	0.25	0.85	0.22	0.79	0.27
Before	12.4	3.12	0.88	0.22	0.81	0.23	0.8	0.29
After	11.9	3.29	0.79	0.26	0.81	0.22	0.78	0.28
Enclosing	11.3	3.2	0.7	0.25	0.78	0.22	0.75	0.24
Concurrent	10.87	4.36	0.79	0.26	0.68	0.34	0.73	0.32

Table 5.3: Perceived density and pace across the 3 podcasts for Study 2. μ values indicate the mean, σ values show the standard deviation, and the distribution of participant judgments across the 5 levels (Too sparse/slow $\leftrightarrow$ Too cluttered/fast) is reported based on the obtained Likert-scale input.

Topic	Density		Pace	
	μ	σ	μ	σ
Nuclear	3.53	0.71	3.18	0.52
	0 / 0 / 20 / 10 / 4		0 / 2 / 24 / 8 / 0	
Cell Phone	3.46	0.68	3.05	0.60
	0 / 1 / 22 / 13 / 3		1 / 3 / 28 / 7 / 0	
GMO Food	3.5	0.71	3.21	0.68
	0 / 0 / 26 / 11 / 5		0 / 4 / 27 / 9 / 2	

Findings from the veracity assessment are presented in Table 5.4, which shows the relative number of correctly recognized (non-)misinformative excerpts across podcast topics and study conditions. The identification of factually accurate (benign) snippets, which were added to the study design to counter-balance the falsehoods and gauge participants' reasoning in a "non-hazardous" setting, is uniform for the three podcasts (roughly in the range [0.8; 0.88]) and consistently high in all five conditions. A significant difference can be observed in the recognition of misleading claims (myths) across debated topics: while approximately equal for nuclear energy and cell phone radiation ($27/34 \approx 31/39 \approx 0.794$), that proportion plummets to one-third ($14/42$) for GMO foods – the reasons for this become apparent from the free-text rationales, which are interpreted alongside other qualitative data in a later section. In turn, inspecting the same proportions across study conditions to determine what impact warnings have on misinformation recognition (per RQ2), the *without* control-group condition is dominated for all 3 topics (by *before* and *concurrent* for nuclear energy / by *after* for cell phone radiation / by *after* and *enclosing* for GMO foods), yet it nonetheless outperforms several remaining warning-placement configurations in each topic-assignment. Coupled with the lack of a clear and consistent top-scoring study condition, this makes the standalone quantitative results obtained from the veracity assessment inconclusive.

The patterns of warning acclimatization captured in the second study are shown in Table 5.5. Across all icon types and placement configurations, listeners' initial perceptions of the interventions remained mostly stable after repeatedly encountering the latter throughout the podcast's duration – as suggested by the neutral majority-category (and $\mu \approx 2.0$) in each row.

Table 5.4: The number of times participants correctly recognized the benign / misinformative excerpts during the veracity assessment across debated topics. The values in each cell are reported out of the overall number of participants assigned the corresponding (topic, study-condition) combination.

Topic		Total	Without	Before	After	Enclosing	Concurrent
Nuclear	Benign	30/34	4/4	6/7	4/4	9/11	7/8
	Myth	27/34	3/4	6/7	3/4	8/11	7/8
Cell Phone	Benign	31/39	7/8	3/5	9/11	6/6	6/9
	Myth	31/39	7/8	3/5	11/11	3/6	7/9
GMO Food	Benign	35/42	10/11	9/11	6/8	4/6	6/6
	Myth	14/42	3/11	3/11	4/8	3/6	1/6

Table 5.5: Warning-acclimatization trends from Study 2 across sound samples and warning-placement alternatives. μ values indicate the mean, σ values denote the standard deviation, and the distribution of participant judgments across the 3 levels (Became less pronounced / Remained stable throughout / Became more pronounced) is reported based on the obtained Likert-scale input.

Cue / Condition	μ	σ	Distribution
Bike Bell	2.03	0.63	5 / 18 / 6
Owl Hoot	2.09	0.47	2 / 25 / 5
Scrap Metal Falling	1.96	0.56	4 / 16 / 3
Before	1.9	0.45	3 / 16 / 1
After	2.1	0.45	1 / 16 / 3
Enclosing	2.09	0.61	3 / 14 / 5
Concurrent	2.05	0.65	4 / 13 / 5

The final quantitative outcome data from Study 2 are the 4 Likert-matrix subjective warning evaluations, the findings for which are grouped by placement configuration in Table 5.6. In terms of sound recognizability (where higher measures are preferable), *concurrent* performs the worst, having the most judgments in the Not at all / Slightly categories and the lowest $\mu = 3.86$ out of all conditions. The remaining three Likert-matrix criteria, which should be minimized for a placement configuration to be considered effective, adhere to the following common pattern. The *after* condition performs the best throughout, characterized by the highest concentration of low-category judgments (Not at all / Slightly) and thus the smallest averages out of all placement alternatives. In contrast, *enclosing* and *concurrent* exhibit the poorest performance, yielding mutually similar Likert-scale distributions – with the largest number of Very / Extremely judgments – and comparably high average scores for all three metrics. The *before* placement lies in-between these extremes, *after* nonetheless dominating it by a large margin as evidenced by the differences in μ values. In turn, the Kruskal-Wallis test based on the warning-placement independent variable gives preliminary indications of statistical significance for some of these group-wise differences: namely, distraction ($H = 7.583$, $p = 0.055$, small-to-medium effect size $\epsilon^2 = 0.057$) and interruption ($H = 12.821$, $p = 0.005$, moderate effect size $\epsilon^2 = 0.123$) as shown in Table 5.7.

5.2.2. Contextual Factors (Statistical Tests and Models)

Inferential-analysis findings are obtained from both the non-parametric Kendall's Tau correlation, which involves all quantitative exploratory variables from Table 4.1 as individual predictors, as well as the raw (univariable) and adjusted (multivariable) regression models tailored towards the 3 different dependent-variable types included in this research. The following discussion separately addresses each outcome type, first providing the preliminary indicators of statistical significance based on Kendall's Tau and then reinforcing those with regression-model statistics for improved robustness.

Table 5.6: Subjective warning evaluations from the 4 Likert-matrix metrics. For each, μ values indicate the mean, σ values show the standard deviation, and the distribution of judgments across the 5 levels (Not at all / Slightly / Moderately / Very / Extremely) is reported based on the obtained Likert-scale input.

Condition	Recognizability		Distraction		Abruptness		Interruption	
	μ	σ	μ	σ	μ	σ	μ	σ
Before	4.1	0.72	2.85	1.31	3.85	1.57	2.75	1.41
	0 / 0 / 4 / 10 / 6		4 / 4 / 5 / 5 / 2		3 / 2 / 1 / 3 / 11		5 / 4 / 5 / 3 / 3	
After	4.05	1.05	2.35	1.27	3.2	1.7	1.95	1.19
	0 / 2 / 4 / 5 / 9		6 / 6 / 5 / 1 / 2		5 / 4 / 0 / 4 / 7		10 / 4 / 4 / 1 / 1	
Enclosing	4.09	0.97	3.32	1.36	4.0	1.27	3.45	1.34
	1 / 0 / 3 / 10 / 8		2 / 6 / 2 / 7 / 5		1 / 3 / 2 / 5 / 11		2 / 4 / 4 / 6 / 6	
Concurrent	3.86	1.25	3.36	1.33	4.05	1.25	3.18	1.5
	2 / 1 / 3 / 8 / 8		2 / 4 / 6 / 4 / 6		2 / 0 / 4 / 5 / 11		4 / 4 / 4 / 4 / 6	

Table 5.7: Kruskal-Wallis test results indicating the median-rank group differences based on the nominal warning-placement variable (including the control-group condition) and their statistical significance. Test outputs include the H statistic, p value, and ϵ^2 effect size for each outcome variable.

Outcome Variable	H	p	ϵ^2
recall_count	2.641	0.620	0
recall_early	7.582	0.108	0.033
recall_middle	3.331	0.504	0
recall_late	1.437	0.838	0
acclimate_trend	1.763	0.623	0
recognizability	0.248	0.969	0
distraction	7.583	0.055	0.057
abruptness	3.241	0.356	0.003
interruption	12.821	0.005	0.123

In the examination of count data (i.e. the number of correctly answered recall questions), Kendall's Tau indicates a significant correlation with topic familiarity (Table 5.8: $\tau = 0.161$, $p = 0.027$). To confirm and/or extend these observations, Poisson regression models were fit to the data – for all of them, the overdispersion ratio R remained ≈ 1.0 , removing the need to consider Negative Binomial regression as a fallback alternative. The univariable models yield two significant findings (Table 5.9): first, compared to the neutral baseline (level 4), the topic-familiarity category of Non-existent (level 1) is associated with a 23% reduction in expected recall count ($e^\beta = 0.77$); second, a prior stance in support of the podcast's scientifically-backed message (Agree, level 6) multiplies the estimated recall count by 1.22 compared to the neutral stance (level 4). However, the statistical significance of both results is diminished in the adjusted Poisson models (Table A.3), suggesting a latent overlap in the apparent raw effects of the two predictors that was not captured by univariable regression.

The three segmented recall scores, whose values lie in the range $[0.0; 1.0]$, are the next outcome variables to undergo inferential analysis. For early-stage recall, two separate significantly correlated predictors are identified through Kendall's Tau (Table 5.8) – topic familiarity and attitude strength: only the latter is confirmed to have a significant unique contribution based on both raw and adjusted Fractional Logistic regression, each 1-point increase in attitude strength associated with 29% higher odds of recall (Table 5.10: $e^\beta = 1.29$, $p = 0.022$).

Table 5.8: Subset of results obtained from Kendall's Tau correlation (where $p \leq 0.05$) indicating monotonic relationships in pairs of outcome and explanatory variables. Outputs include the τ statistic and p value for each pair. See Table A.2 for the complete results.

Outcome Variable	Explanatory Variable	τ	p
recall_count	topic_familiarity	0.161	0.027
recall_early	topic_familiarity	0.195	0.010
recall_early	attitude_strength	0.169	0.018
recall_late	need_for_cognition	0.154	0.033
acclimate_trend	topic_familiarity	0.272	0.004
acclimate_trend	attitude_strength	0.241	0.007
recognizability	podcast_density	-0.256	0.010

Table 5.9: Subset of results obtained from Poisson regression for count data (rows where $p \leq 0.05$ highlighted in yellow). For categorical predictors, the specific level (appearing in parentheses) is distinguished from the reference category (indicated in superscript) to make the grouping explicit. Table entries containing more than one uniquely-named explanatory variable indicate multivariable-model statistics. Outputs include the overdispersion ratio R , log-difference in outcome β , expected-count ratio e^β , standard coefficient error σ_β , and p value. See Table A.3 for the complete results.

Outcome	R	Explanatory Variable(s)	β	e^β	σ_β	p
recall_count	1.01	topic_familiarity(1) $\rightarrow^4$	-0.2651	0.77	0.135	0.049
recall_count	1.04	prior_stance(6) $\rightarrow^4$	0.1950	1.22	0.085	0.022

In turn, significant explanatory-variable interactions with middle-stage recall are only detected by the stronger Fractional Logit models. A *concurrent* warning placement (compared to the *without* condition), a Non-existent topic-familiarity level (compared to Average, level 4), and a prior stance of Agreement (compared to Neutral, level 4) all influence the recall-proportion odds, multiplying the latter by 0.37, 0.32, or 2.71 from their respective baselines. Adjusted models confirm the significance of these differences by isolating the predictors' "true" direct effects (Table 5.10) – only topic familiarity experiences a loss of significance once attitude strength is added to the model, namely $p = 0.056$, which is the result of predictor confounding. In particular, the interaction between study condition and recall score captured through regression reinforces the descriptive patterns presented in Table 5.2, for which Kruskal-Wallis initially yielded no indication of statistical significance.

Lastly, for late-stage recall, Kendall's Tau suggests a significant correlation with need for cognition, which is confirmed by both raw and multivariable Fractional Logit models, the latter yielding a recall-proportion odds multiplier of 1.92 for each 1-point increase in the composite NfC score ($p = 0.003$).

The remaining ordinal-data outcome variables are also considered for inferential analysis. Inspecting the trend of warning acclimatization, Kendall's Tau yields significant correlations with topic familiarity and attitude strength (Table 5.8). Both re-occur through univariable Ordinal Logistic regression models: a Very Good topic familiarity (compared to Average, level 4), a 1-point increase in attitude strength, as well as a prior stance of Somewhat Disagree (compared to Neutral, level 4) are each associated with higher odds of reporting more pronounced perceptions of the warning upon repeated encounters (with off-baseline ratios of 6.31, 4.57, and 1.45, respectively). Adjusted models indicate a diminished significance of these effects, however, suggesting the latter overlap while instead revealing a unique partial contribution of podcast pace – compared to the Moderate baseline (level 3), the Fast level leads to a 75% reduction of those odds (Table 5.11: $e^\beta = 0.25$, $p = 0.033$).

In terms of sound recognizability, both Kendall's Tau and Ordinal Logit regression confirm a negative correlation with podcast density: the non-parametric test yields $\tau = -0.256$ as the magnitude of this monotonic relationship (Table 5.8). Evidence for the latter is reinforced through both raw and adjusted regression models, according to which the podcast-density category of Too cluttered (5) is associated

with 91% lower odds of reporting high recognizability levels compared to the Moderate baseline (Table 5.11: $e^\beta = 0.09$, $p = 0.010$).

Several predictors with significant correlations are identified for the sound-induced distraction outcome variable via Ordinal Logit regression. The *after* warning-placement configuration (compared to a baseline of *concurrent*) and a prior stance of Strong Agreement (compared to Neutral, level 4) are both directly associated with heavily reduced odds of reporting higher distraction levels, these effects also persisting in multivariable models (Table 5.11: $e^\beta = 0.21$, $p = 0.010$ and $e^\beta = 0.09$, $p = 0.018$). In turn, compared to its Moderate (3) baseline, a Slow (2) perceived podcast pace multiplies the odds of reporting greater sound-induced distraction by 14.76 per adjusted-regression findings.

Finally, Ordinal Logit models with sound-induced interruption as their outcome variable yield a significant correlation with warning placement: specifically, compared to *concurrent*, the *after* configuration is associated with 74% lower odds of reporting higher interruption levels, this unique partial effect confirmed in adjusted as well as raw settings (Table 5.11: $e^\beta = 0.26$, $p = 0.025$). One univariable regression model also indicates a significant increase in higher-level interruption odds based on a Fast (4) perceived podcast pace compared to the Moderate baseline ($e^\beta = 2.87$, $p = 0.023$), yet confounding appears to diminish that effect's significance in the adjusted model. Combined with the findings related to distraction, these latent interactions between warning placement and interruption captured by Ordinal Logit regression reinforce the Kruskal-Wallis test's preliminary indicators of statistical significance (Table 5.7), improving the reliability of the subjective-evaluation metrics reported in Table 5.6.

Table 5.10: Subset of results obtained from Fractional Logistic regression for proportion data (rows where $p \leq 0.05$ highlighted in yellow). For categorical predictors, the specific level (appearing in parentheses) is distinguished from the reference category (indicated in superscript) to make the grouping explicit. Table entries containing more than one uniquely-named explanatory variable indicate multivariable-model statistics. Outputs include the log-difference in outcome β , recall-proportion odds ratio e^β , standard coefficient error σ_β , and p value. See Table A.4 for the complete results.

Outcome	Explanatory Variable(s)	β	e^β	σ_β	p
recall_early	attitude_strength	0.2335	1.26	0.088	0.008
recall_early	cond(before) $\rightarrow$ without	0.6553	1.93	0.526	0.213
	cond(after) $\rightarrow$ without	-0.1325	0.88	0.433	0.760
	cond(enclosing) $\rightarrow$ without	-0.5187	0.60	0.412	0.208
	cond(concurrent) $\rightarrow$ without	-0.1485	0.86	0.441	0.736
	attitude_strength	0.2448	1.28	0.100	0.014
recall_early	cond(before) $\rightarrow$ without	0.7508	2.12	0.494	0.129
	cond(after) $\rightarrow$ without	-0.1016	0.90	0.431	0.814
	cond(enclosing) $\rightarrow$ without	-0.5971	0.55	0.413	0.148
	cond(concurrent) $\rightarrow$ without	0.0221	1.02	0.448	0.961
	prior_stance(3) $\rightarrow$ 4	0.1955	1.22	0.490	0.690
	prior_stance(5) $\rightarrow$ 4	-0.1263	0.88	0.415	0.761
	prior_stance(6) $\rightarrow$ 4	0.8023	2.23	0.503	0.111
	prior_stance(7) $\rightarrow$ 4	-0.5429	0.58	0.492	0.269
	attitude_strength	0.2516	1.29	0.110	0.022
recall_middle	cond(concurrent) $\rightarrow$ without	-0.9745	0.38	0.476	0.041
recall_middle	topic_familiarity(1) $\rightarrow$ 4	-1.0083	0.36	0.361	0.005
recall_middle	prior_stance(6) $\rightarrow$ 4	1.3466	3.84	0.441	0.002

continued on next page

Table 5.10 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
recall_middle	cond(before) $\rightarrow$ without	-0.2346	0.79	0.434	0.589
	cond(after) $\rightarrow$ without	-0.4066	0.67	0.419	0.332
	cond(enclosing) $\rightarrow$ without	-0.5365	0.58	0.426	0.208
	cond(concurrent) $\rightarrow$ without	-0.9906	0.37	0.471	0.035
	topic_familiarity(1) $\rightarrow$ 4	-1.1269	0.32	0.483	0.020
	topic_familiarity(2) $\rightarrow$ 4	0.4039	1.50	0.551	0.463
	topic_familiarity(3) $\rightarrow$ 4	0.1978	1.22	0.372	0.595
	topic_familiarity(5) $\rightarrow$ 4	0.3891	1.48	0.475	0.413
	topic_familiarity(6) $\rightarrow$ 4	0.2833	1.33	0.565	0.616
	prior_stance(3) $\rightarrow$ 4	0.4353	1.55	0.480	0.364
	prior_stance(5) $\rightarrow$ 4	0.4080	1.50	0.398	0.305
	prior_stance(6) $\rightarrow$ 4	0.9971	2.71	0.461	0.031
	prior_stance(7) $\rightarrow$ 4	0.7443	2.10	0.610	0.223
recall_middle	cond(before) $\rightarrow$ without	-0.2208	0.80	0.432	0.609
	cond(after) $\rightarrow$ without	-0.4015	0.67	0.416	0.334
	cond(enclosing) $\rightarrow$ without	-0.5246	0.59	0.423	0.215
	cond(concurrent) $\rightarrow$ without	-0.9849	0.37	0.466	0.035
	topic_familiarity(1) $\rightarrow$ 4	-1.0576	0.35	0.554	0.056
	topic_familiarity(2) $\rightarrow$ 4	0.4435	1.56	0.554	0.423
	topic_familiarity(3) $\rightarrow$ 4	0.2129	1.24	0.370	0.565
	topic_familiarity(5) $\rightarrow$ 4	0.3670	1.44	0.494	0.457
	topic_familiarity(6) $\rightarrow$ 4	0.2466	1.28	0.583	0.672
	prior_stance(2) $\rightarrow$ 4	-0.2193	0.80	0.439	0.617
	prior_stance(3) $\rightarrow$ 4	0.4206	1.52	0.496	0.397
	prior_stance(5) $\rightarrow$ 4	0.3936	1.48	0.403	0.328
	prior_stance(6) $\rightarrow$ 4	0.9978	2.71	0.464	0.032
	prior_stance(7) $\rightarrow$ 4	0.7098	2.03	0.609	0.244
	attitude_strength	0.0323	1.03	0.135	0.810
recall_late	prior_stance(6) $\rightarrow$ 4	0.8720	2.39	0.443	0.049
recall_late	need_for_cognition	0.4849	1.62	0.215	0.024
recall_late	cond(before) $\rightarrow$ without	0.0335	1.03	0.513	0.948
	cond(after) $\rightarrow$ without	-0.1253	0.88	0.457	0.784
	cond(enclosing) $\rightarrow$ without	-0.2564	0.77	0.412	0.534
	cond(concurrent) $\rightarrow$ without	-0.6204	0.54	0.446	0.164
	need_for_cognition	0.5671	1.76	0.219	0.010

continued on next page

Table 5.10 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
recall_late	cond(before) $\rightarrow$ without	0.0690	1.07	0.481	0.886
	cond(after) $\rightarrow$ without	-0.1288	0.88	0.498	0.796
	cond(enclosing) $\rightarrow$ without	-0.3834	0.68	0.436	0.380
	cond(concurrent) $\rightarrow$ without	-0.4955	0.61	0.477	0.299
	prior_stance(2) $\rightarrow$ 4	0.6192	1.86	0.635	0.330
	prior_stance(3) $\rightarrow$ 4	0.1377	1.15	0.403	0.733
	prior_stance(5) $\rightarrow$ 4	0.3407	1.41	0.415	0.412
	prior_stance(6) $\rightarrow$ 4	0.9153	2.50	0.469	0.051
	prior_stance(7) $\rightarrow$ 4	-0.7369	0.48	0.627	0.240
	need_for_cognition	0.6538	1.92	0.221	0.003

Table 5.11: Subset of results obtained from Ordinal Logistic regression for ordinal data (rows where $p \leq 0.05$ highlighted in yellow). For categorical predictors, the specific level (appearing in parentheses) is distinguished from the reference category (indicated in superscript) to make the grouping explicit. Table entries containing more than one uniquely-named explanatory variable indicate multivariable-model statistics. Outputs include the log-difference in outcome β , higher-category odds ratio e^{β} , standard coefficient error σ_{β} , and p value. See Table A.5 for the complete results.

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
acclimate_trend	topic_familiarity(6) $\rightarrow$ 4	1.8415	6.31	0.909	0.043
acclimate_trend	prior_stance(3) $\rightarrow$ 4	1.5201	4.57	0.740	0.040
acclimate_trend	attitude_strength	0.3732	1.45	0.184	0.042
acclimate_trend	cond(before) $\rightarrow$ concurrent	-0.5144	0.60	0.732	0.482
	cond(after) $\rightarrow$ concurrent	0.3598	1.43	0.719	0.617
	cond(enclosing) $\rightarrow$ concurrent	0.4562	1.58	0.741	0.538
	topic_familiarity(1) $\rightarrow$ 4	0.3904	1.48	1.336	0.770
	topic_familiarity(2) $\rightarrow$ 4	-0.6899	0.50	0.940	0.463
	topic_familiarity(3) $\rightarrow$ 4	-1.0278	0.36	0.710	0.148
	topic_familiarity(5) $\rightarrow$ 4	0.4712	1.60	0.736	0.522
	topic_familiarity(6) $\rightarrow$ 4	2.0296	7.61	1.034	0.050
	prior_stance(2) $\rightarrow$ 4	0.5979	1.82	0.970	0.538
	prior_stance(3) $\rightarrow$ 4	0.9186	2.51	0.854	0.282
	prior_stance(5) $\rightarrow$ 4	0.1211	1.13	0.720	0.866
	prior_stance(6) $\rightarrow$ 4	-0.0752	0.93	0.882	0.932
	prior_stance(7) $\rightarrow$ 4	-0.2801	0.76	1.092	0.797

continued on next page

Table 5.11 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
acclimate_trend	topic_familiarity(1) $\rightarrow$ ⁴	−0.0996	0.91	1.350	0.941
	topic_familiarity(2) $\rightarrow$ ⁴	−0.7702	0.46	0.939	0.412
	topic_familiarity(3) $\rightarrow$ ⁴	−1.1095	0.33	0.713	0.120
	topic_familiarity(5) $\rightarrow$ ⁴	0.7096	2.03	0.738	0.336
	topic_familiarity(6) $\rightarrow$ ⁴	1.8157	6.15	1.009	0.072
	prior_stance(2) $\rightarrow$ ⁴	0.3299	1.39	0.978	0.736
	prior_stance(3) $\rightarrow$ ⁴	0.9792	2.66	0.847	0.248
	prior_stance(5) $\rightarrow$ ⁴	0.0915	1.10	0.735	0.901
	prior_stance(6) $\rightarrow$ ⁴	0.2436	1.28	0.858	0.776
	prior_stance(7) $\rightarrow$ ⁴	−0.7648	0.47	1.106	0.489
	podcast_pace(2) $\rightarrow$ ³	0.5408	1.72	1.234	0.661
	podcast_pace(4) $\rightarrow$ ³	−1.3976	0.25	0.657	0.033
recognizability	podcast_density(5) $\rightarrow$ ³	−2.1688	0.11	0.778	0.005
recognizability	prior_stance(2) $\rightarrow$ ⁴	−0.4731	0.62	0.791	0.550
	prior_stance(3) $\rightarrow$ ⁴	−0.4052	0.67	0.667	0.544
	prior_stance(5) $\rightarrow$ ⁴	−0.3927	0.68	0.557	0.481
	prior_stance(6) $\rightarrow$ ⁴	−1.2897	0.28	0.753	0.087
	prior_stance(7) $\rightarrow$ ⁴	−0.5359	0.59	0.905	0.554
	podcast_density(4) $\rightarrow$ ³	−0.4345	0.65	0.468	0.353
	podcast_density(5) $\rightarrow$ ³	−2.1300	0.12	0.855	0.013
	intellectual_humility	0.2937	1.34	0.252	0.243
recognizability	prior_stance(2) $\rightarrow$ ⁴	−0.5418	0.58	0.788	0.492
	prior_stance(3) $\rightarrow$ ⁴	−0.2608	0.77	0.656	0.691
	prior_stance(5) $\rightarrow$ ⁴	−0.5414	0.58	0.542	0.318
	prior_stance(6) $\rightarrow$ ⁴	−1.3455	0.26	0.742	0.070
	prior_stance(7) $\rightarrow$ ⁴	−0.4423	0.64	0.893	0.620
	podcast_density(4) $\rightarrow$ ³	−0.5290	0.59	0.460	0.251
	podcast_density(5) $\rightarrow$ ³	−2.2984	0.10	0.838	0.006
recognizability	cond(before) $\rightarrow$ concurrent	−0.1244	0.88	0.602	0.836
	cond(after) $\rightarrow$ concurrent	−0.1665	0.85	0.651	0.798
	cond(enclosing) $\rightarrow$ concurrent	−0.0682	0.93	0.637	0.915
	prior_stance(2) $\rightarrow$ ⁴	−0.5099	0.60	0.799	0.523
	prior_stance(3) $\rightarrow$ ⁴	−0.2610	0.77	0.669	0.696
	prior_stance(5) $\rightarrow$ ⁴	−0.5374	0.58	0.546	0.325
	prior_stance(6) $\rightarrow$ ⁴	−1.3308	0.26	0.756	0.079
	prior_stance(7) $\rightarrow$ ⁴	−0.4432	0.64	0.900	0.622
	podcast_density(4) $\rightarrow$ ³	−0.5277	0.59	0.464	0.255
	podcast_density(5) $\rightarrow$ ³	−2.3682	0.09	0.915	0.010

continued on next page

Table 5.11 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
distraction	cond(after) $\rightarrow$ concurrent	-1.3896	0.25	0.565	0.014
distraction	prior_stance(7) $\rightarrow$ 4	-2.2374	0.11	0.956	0.019
distraction	podcast_pace(2) $\rightarrow$ 3	2.6731	14.48	1.261	0.034
	podcast_pace(4) $\rightarrow$ 3	0.9413	2.56	0.464	0.043
distraction	cond(before) $\rightarrow$ concurrent	-0.8095	0.45	0.593	0.172
	cond(after) $\rightarrow$ concurrent	-1.5719	0.21	0.607	0.010
	cond(enclosing) $\rightarrow$ concurrent	-0.2856	0.75	0.579	0.622
	prior_stance(2) $\rightarrow$ 4	-0.0688	0.93	0.727	0.925
	prior_stance(3) $\rightarrow$ 4	0.4419	1.56	0.638	0.489
	prior_stance(5) $\rightarrow$ 4	0.5514	1.74	0.530	0.299
	prior_stance(6) $\rightarrow$ 4	0.6463	1.91	0.704	0.359
	prior_stance(7) $\rightarrow$ 4	-2.4038	0.09	1.020	0.018
	podcast_pace(2) $\rightarrow$ 3	2.6917	14.76	1.336	0.044
	podcast_pace(4) $\rightarrow$ 3	0.4646	1.59	0.512	0.365
interruption	cond(after) $\rightarrow$ concurrent	-1.6635	0.19	0.586	0.005
interruption	podcast_pace(4) $\rightarrow$ 3	1.0544	2.87	0.464	0.023
interruption	cond(before) $\rightarrow$ concurrent	-0.5583	0.57	0.574	0.331
	cond(after) $\rightarrow$ concurrent	-1.3477	0.26	0.601	0.025
	cond(enclosing) $\rightarrow$ concurrent	0.5005	1.65	0.556	0.368
	podcast_pace(2) $\rightarrow$ 3	1.7827	5.95	1.108	0.108
	podcast_pace(4) $\rightarrow$ 3	0.8152	2.26	0.496	0.100

5.2.3. Qualitative Findings

Qualitative data obtained from participants' free-text rationales for (i) the indicators they relied on during the veracity assessment and (ii) their subjective perceptions of the auditory interventions – in terms of presumed function, conceptual mapping, elicited first-encounter reaction, and degree of acclimatization – were categorized using RTA. The identified themes for all five qualitative variables are outlined below, briefly summarized where additional details are necessary and exemplified with selected excerpts from crowdworkers' responses.

Misinformation-Recognition Rationales

The following thematic clusters include the various indicators considered by participants when judging a statement's veracity during the misinformation recognition check – as well as the prominent reasons for failing to do so – thereby enriching the quantitative data outlined in Table 5.4.

Opinion or Assumption (Not a Factually-Backed Statement) No supporting evidence is presented.

- “That line is presented as a concern for the public, not as the answer to the science question.”
- “I think it is an assumption and not presented with facts.”
- “There is speculation involved. While I agree with the speculation, and have a feeling it is accurate, speculation by nature is unsure.”

Parroting of Rumors The statement discusses hearsay rather than verified scientific findings.

- “The male co-host is bringing up a rumor or idea he heard online that has not been verified.”

- “It sounds like a fear-mongering approach to non-ionizing radiation with no backing or proof.”

Word-Choice Suggesting Uncertainty or Vagueness Certain descriptors used by the hosts themselves (e.g. “claim”) indicate the given statement is alleged and not reliably true.

- “I’m not sure, but I’m leaning towards ‘no’ because they just say ‘some sources’ instead of saying where these claims actually come from.”
- “It talks a lot about ‘claims’ that could be true – but are not proven to be factual. Even though he says ‘another claim I keep seeing’, that doesn’t make the following statement factual.”
- “No, as he stated they are just claims coming from several different sources, none of which he specifically cited.”

Recall of Corrective Information from Podcast Content Participants are also occasionally able to detect a contradiction with scientifically-backed information from other parts of the podcast, recognizing the falsehood presented in the statement.

- “The claim he presents about renewable being able to do the whole job is contradicted by the interviewee.”
- “I remember the podcast guest saying that it was the opposite, and that we are in a sort of nuclear renaissance.”
- “The guy talks about 5G internal replication, which is nonsense and debunked by the more femme-voiced speaker soon thereafter.”
- “While major agricultural corporations do finance safety testing for regulatory approvals, it is completely false to state that all research on GMOs has been funded by them.”

Recognition Failure 1 (Reliance on Intuition and Pre-Existing Notions) Some participants issue instinctive veracity judgments influenced by familiarity bias, not formulated via fact-driven reasoning.

- “I have heard the information before.”
- “The discussion in the snippet captures what people have said.” / “It is common consensus.”
- “It makes sense that the extreme exposure to radiation would cause more problems.”

Recognition Failure 2 (Overreliance on Podcast-Host Authority) Others place unwarranted trust in the podcast hosts’ perceived infallibility, believing the presented statements to be thoroughly verified and therefore convincing.

- “The speaker pointed out the fact that he had read about the information he spoke about.”
- “Honestly, I have no idea, I would assume they know what they’re talking about.”
- “The explanation was clear enough to understand and be convinced.”

Recognition Failure 3 (Shared Ungrounded Opinion) Particular to the topic of GMO foods, a large number of participants have personal pre-conceptions that are in line with the claims presented in the misinformative statement used for the veracity assessment, failing to accurately judge its authenticity.

- “I think some science is concealed from the public.”
- “I know the concern with the corporate seed owners is real.”
- “Seed companies secretly test GMOs without regulation.”
- “Because realistically, these researches are always funded by big companies, so that obviously causes distrust.”
- “I feel this discussion is factually accurate as, with many industries, a big cause of mistrust stems from the corporations having too much of a hand in things.”

Conceptual Mapping

The 3 samples carrying over from Study 1 and interspersed as warnings into the full-length podcasts for the second study exhibited consistent real-world associations (per the Conceptual Mapping responses obtained from crowdworkers outside the no-warning control-group condition). Nonetheless, each cue's attributed real-world entity deviated from the majority for some individuals, asserting the importance of subjective experience for audio interpretation.

The bike bell was additionally considered to be “a cash register ding”, “a desk bell”, and “an alarm clock”. In turn, the sound of an owl hooting was also perceived as the call of different birds (including hawk, dove, crow, and macaw), a whistle, or even “a kids’ train ride at an amusement park”. Lastly, the sound of scrap metal falling was associated with “pots clanging in a kitchen”, “silverware falling onto a hard surface”, and “a metal tray full of steel medical instruments being dropped”.

Presumed Sound Purpose

The podcast-integrated sounds were assumed to fulfill different functions, grouped into the categories below (as identified through RTA).

Attention-Check Mechanism Some participants believed the cues to be a non-functional component of the study aimed at filtering out responses from inattentive crowdworkers.

- “To prevent participants from losing concentration while listening to the podcast.” / “To make sure participants stay focused on the podcast.”
- “To grab my attention before an important point was made that could be asked later.” / “To serve as an attention cue for the recall test.” / “For this question-set obviously.”

Attention Grabber Others identified their role as salient attention-capturing effects added to enhance listening sessions in general (not as a custom testing criterion for the study).

- “To keep the listener alert.” / “To refocus the listener in case they were not paying attention.”

Emphasis Cue The cues were perceived to highlight particularly important information in the podcast.

- “Something like a hard hitting point or punchline.”
- “It was used to accentuate the most important points.”
- “To identify when an important question has been asked.”

Transition Marker The sounds were also thought to partition the podcast into topic-based sections.

- “It felt like a cue used to mark a transition between topics in the podcast.”
- “It was likely used to signal a shift between major sections of the podcast.”
- “I remember it happening during topic changes or when further clarification was requested.”

Misinformation Flag Some participants that were assigned the *after* and *enclosing* placement configurations would also (accurately) identify the sounds’ intended role as misinformation warnings.

- “It was meant to punctuate commonly-held misconceptions about nuclear power.”
- “It seemed to come on every time one of the hosts would introduce a common misconception about nuclear energy.”
- “It happened whenever the male co-host brought up a common belief or myth.”
- “It seemed like whatever was stated in between the sound was a conspiracy theory, or information that was unproven.”

Purposeful Disruption The cues were also believed to serve as intentionally-added distractors which prevent listeners from internalizing the information they heard (an opinion expressed by crowdworkers assigned to the *after* study condition).

- “To try to distract from what was being said. It also may have been to disrupt the memory of what was being said.”

Podcast-Enhancement Editing Effect The sound effects were potentially used to increase engagement with the content and/or reinforce the podcast’s entertainment factor.

- “To break up dense information.”
- “It reminded me of a brief notification sound, kind of like a subtle podcast editing cue.”
- “The sound maybe replicates a mic drop after a point has been made.” / “To add drama or impact, the host was just being fun and using his soundboard.”
- “To make it lively.” / “It was used as an extra conversational cue to make the discussion feel more natural and engaging, almost like a real podcast with pauses, reactions, or emphasis moments.”

Elicited First Reaction

Various immediate reactions were triggered by participants’ first encounter with the podcast-integrated sound cues, most reinforcing the findings originally obtained from Study 1 – in which the auditory icons did not serve a functional purpose. Using RTA, the following themes were formulated.

Surprise The interventions were unexpected (especially given the use of partial disclosure).

- “It honestly startled me. Wasn’t expecting it. Kind of an abrasive sound.”
- “The first time I heard the sound, it caught me off guard because I wasn’t expecting non-verbal effects in a science podcast.”

Heightened Attention and Concentration

- “It made me refocus and listen harder.” / “It made me pay closer attention.”
- “I was mildly startled but not surprised. It caught my attention.” / “It wakes you up and gets your attention as well.”
- “I figured I was about to hear something significant and tried to concentrate more on the topic of the conversation.”
- “The first time I heard the sound, it caught my attention for a moment because it stood out from the normal conversation, but it also felt like a natural podcast transition or editing effect rather than something distracting.”

Initial Confusion Upon the sound cue’s initial occurrence, participants were unable to identify a clear connection between its appearance and the overarching conversation topic, nor perceive its relevance given the surrounding context – both these challenges create a sense of confusion.

- “I was confused because it sounded completely out of place and disrupted the flow of the interview.” / “I was confused as it did not match the discussion or the overall topic of the podcast.”
- “I was confused for a second and tried to figure out how it fit in with the podcast. Then I figured it had some ulterior motive.”
- “I wasn’t sure if it was meant to be a transition, an attention check of sorts, or what the overall relevance was.”
- “My first reaction was really confusion, I thought some utensil had fallen.”
- “I was thrown off – I wondered if the sound was coming from the podcast or from somewhere else, and I also wondered briefly if it was intended to be in the podcast, or if it was added by researchers for this survey.”

Distraction and Interruption The apparent lack of purpose & intention behind the sound cues' occurrences also led to listener dissatisfaction, some participants believing them to be a mistake in editing.

- "I figured it was quite unprofessional to have this amount of noise interfere with such an interactive and educative session."
- "I thought it was jarring, a sensory nightmare."
- "I was briefly considering the origin of the noise rather than the words of the podcast."
- "No podcaster allows that to happen or stay in the audio after editing."
- "The first time I heard it, it stood out as quite noticeable and distracting because it sounded like a real-world accident rather than a typical podcast sound effect."
- "It made me pause my thoughts and, for a second, I thought an ad might play."

Amusement and Loss of Immersion Like in Study 1, whenever the interventions are perceived as drastically inappropriate (even ridiculous), they may also elicit laughter and break listeners' immersion.

- "I didn't expect to hear that sound, and afterwards thought it was bit funny and strange for it to be in the podcast."
- "It sounds goofy and dumb. It makes me not take the podcast seriously."
- "I thought it was kind of cringe-y, but I ignored it at first and didn't think much until I heard it again."

Overlapping with External Stimuli An intervention's first occurrence could also occasionally overlap with sounds from a listener's present (or regular) environment, some participants briefly wondering if it came from the podcast before confirming this was the case.

- "The first time I heard the sound, for a moment, I thought this was occurring outside my window. Many people ride bikes in my neighborhood, so that is not an uncommon noise."
- "I was confused and thought it was a noise outside. Once I heard it again, I realized it came from the audio." / "At first, I thought it came from somewhere in my environment until I heard it again later in the episode."
- "The first time I heard it, I thought a notification had gone off on my device, but then I realized it was part of the podcast."

Warning-Acclimatization Rationales

The following categories of perception-change (or lack thereof) in relation to the repeated occurrences of podcast-integrated sound cues throughout Study 2 were identified based on the free-text responses requested from participants.

No Change Some crowdworkers' perceptions of the sounds remained stable throughout the podcast.

- "It happened sparsely enough that I wasn't really expecting it to happen again."
- "Each sound happened far enough apart for my guard to lower in between them."
- "Once I recognized the sound's purpose as a deliberate cue, my perception of it stayed the same each time it played."
- "The sound stayed about the same for me throughout the podcast because it was brief and blended naturally into the conversation without becoming distracting."

Decline in Confusion Many participants also managed to identify either a latent purpose for the cues' integration into the podcast, or a pattern in their occurrences throughout the audio, their initial confusion dissipating in both cases.

- "The more I heard the sound, which was little more than 4 minutes apart, it confirmed my thought that it was there for a specific reason."
- "As it re-occurred, I understood that it's a pointer to take note of the next set of responses."
- "After hearing it multiple times, I felt maybe it was trying to highlight a set of information that I need to be aware of whether it is accurate or misinformation."
- "After I first heard it, I began to expect it whenever they introduced information that they were going to debunk."

Decline in Surprise and Distraction The sound's subsequent occurrences were less jarring/disruptive from some listeners' perspective, reducing their initial dissatisfaction.

- "In my opinion, it was some sort of deliberate sound intended to highlight something special, and when I encountered it again, it no longer seemed to distract me."
- "It became as expected predictable pattern instead of a surprising or jarring audio disruption."
- "My perceptions of the sound changed. As I become more involved and focused on the information in the podcast, the sound was less distracting."
- "At first, it seemed to stick out like a sore thumb, then it got less in your face."
- "I became less surprised by it." / "The subsequent times it wasn't as unexpected and shocking."

Habituation Other participants started to disregard the cues later in the podcast, growing accustomed to their presence and/or not attributing a particular meaning to their inclusion.

- "I don't know if it keeps your attention as well as it went on." / "The more I heard it, the less I paid attention to it." / "It faded away slowly."
- "I perhaps thought about its significance less each time, but it was equally prominent during each occurrence."
- "It seemed to catch my attention at first, but the more times I heard it, I tried to tune it out. It was a bit of a distraction while trying to focus and listen to the podcasters."

Reinforced Salience Meanwhile, a subset of responses point out the sound becoming more prominent throughout repeated occurrences, its memorability and salience reinforced.

- "The noise stood out from the conversation and became more noticeable every time it occurred as the podcast progressed."
- "The more I heard it, the more I remembered the sound."

Anticipation Some participants who perceived the cues as important indicators also began to actively anticipate them as the podcast progressed.

- "I found myself waiting for the sound."
- "The sound seemed to highlight an important question or claim made within the discussion. For that reason, I was attentive in listening to whether the sound would reappear later in the podcast."

5.2.4. Exploratory Post-Debrief Listener Perceptions

Finally, the themes formulated from participants' responses to the two post-debrief qualitative variables (in the circumstances of full disclosure) are shown via representative excerpts below, extra clarifications also included (as necessary). These findings offer an initial glimpse into the perceptions and behaviors potentially triggered by encountering a misinformation warning in regular podcast-listening sessions.

Potential Immediate Reactions

A wide variety of elicited reactions in response to hearing the intervention in non-experimental settings were supposed by participants, yielding the following categories.

Heightened Attention and Concentration

- “I’d pay extra attention to the context surrounding it.”
- “I would just take it as a quiet signal to pay closer attention to what is being said next.”

Heightened Falsehood Awareness Some listeners would more carefully assess the reliability of the information they are presented with, less inclined to take the hosts’ statements at face value.

- “It would make me think about if the podcast has accurate information.”
- “I would appreciate the alert because it instantly primes me to question the statement being made.”
- “It would help me passively sort information into science/factual and rumors/beliefs.”

Extra Podcast-Listening Precautions Others would interact with the podcast more closely, rewinding the relevant section to double-check what they heard or consulting the comment section.

- “I would definitely check the comments to see if anyone else had pointed it out. This sound would be very unusual in the podcasts that I typically listen to.”
- “I think I would go back to listen to it to see if I heard it or not.”
- “I would probably pause and pay more attention to what was being said instead of fully trusting it right away. It would make me want to double-check the information later.”

Closing the Audio Player Altogether

- “I would stop listening entirely – if one part is misleading/inaccurate, why would I listen to more?”
- “This could be handy for knowing which podcasts I no longer want to listen to.”

Concerns Regarding Flagging Accuracy Some participants expressed doubts regarding the reliability of the misinformation warning itself, stating that the flagged content sections might require a more nuanced interpretation (rather than as plain falsehoods) – and that erroneous misinformation detection is also a possibility.

- “If the podcast was in a similar format as this one – where a designated speaker is bringing up points for an expert guest to clarify – I think it would be great. However, if it was more of a debate style podcast, or one that verges into theoretical/unproven theories, then I would think of the ring as being more of someone’s judgment call than a factually proven misinformation sound.”
- “I wouldn’t really trust it if I heard it in a podcast though. Assuming the podcast itself is adding the sound, you’d have to trust they are accurate on what is labeled as misinformation.”

No Immediate Reaction Several responses instead suggested that the auditory interventions would be discarded, the ongoing narrative taking precedence in listeners’ priorities.

- “I’d probably just register it as a built-in signal and keep listening without reacting much.”
- “I wouldn’t care much and would look past it all the same if the conversation was good enough.”
- “I would briefly notice it, then continue focusing on the podcast content.”

Dissatisfaction Caused by Sound Interruption Even when the cues' intended function is transparently communicated to listeners, some would nonetheless consider them obnoxious and disruptive.

- "I'd react with mild annoyance if it was a regular thing."
- "It would be distracting and ultimately diminish the podcast experience for me."
- "I'd understand its purpose, but I'd be annoyed."
- "To be honest, I would not enjoy to keep hearing that sound during a podcast. It just drifts my mind elsewhere."

Deliberate Warning Opposition The possibility that the warnings are perceived as paternalistic and immersion-breaking cannot be ruled out, some listeners reacting with outright disapproval toward them.

- "I'd continue ignoring it unless it was persistent and layering over speech or otherwise affecting the audio for usability."

Potential Downstream Behaviors

Lastly, the follow-up actions potentially undertaken by participants after hearing a misinformation warning in their regular podcasts are grouped into the categories below.

Further Independent Research

- "I would look up the claims on my own to try to get accurate information, and I may stop listening to the podcast due to misinformation."
- "I would pause the episode to look up credible sources and check the facts myself."
- "I would likely fact-check the claims online or compare them with other reliable sources before believing or sharing the information. Depending on the topic, I might also continue listening more carefully to understand why the warning appeared."

Cautionary Continued Listening Participants were inclined to continue listening to the podcast more attentively – actively filtering the information they are presented with based on its veracity (indicated by the warning's presence or absence) and accompanying justification.

- "I would take everything with a grain of salt."
- "I would probably disregard the information or not hold as much weight to it."
- "I would pause for a moment to think about what was said and consider whether the information I just heard might be unreliable."
- "I would listen more closely to what came after the sounds, and if I believed that rumor, I would likely either change my beliefs or do more research into it."

Discontinued Podcast Subscription

- "I'd probably stop listening to that podcast and move on to something else. There are plenty of podcasts and audio-books to choose from. There's no reason to put up with that."
- "I honestly would be less interested in listening to the podcast as even if I do further research, misinformation can be difficult to unlearn, especially if given the comfort of phrases like *research shows* and *a reputable source states*."

Seeking Additional Confirmation Driven by concerns regarding the accuracy of the flagging procedure through which warnings are interspersed, some participants would undertake further measures to verify the content's status as misinformation (e.g. consulting additional indicators and sources).

- "I would replay it to confirm, then proceed to ask others about it."
- "A clue to misinformation can itself be misinformation. What I am not sure of I will look up."

No Downstream Behavior Depending on the perceived risk level and the amount of effort required, participants might be inclined to discard the warning, not taking up any follow-up actions.

- “I wouldn’t really care that much.”
- “I might click off of the video, but, honestly, that is pretty unlikely too, because when someone is presenting false information as fact, I like to listen to their whole speech.”

Deliberate Warning Opposition Such interventions could also be considered patronizing, their presence seemingly downplaying listeners’ own reasoning capabilities – thus diminishing their autonomous podcast-consumption experience.

- “I would try and keep my volume down.”
- “Probably try to find a way to prevent, disable, or turn off the nanny bots, because I am a full grown adult and don’t live in child-world anymore where I need someone interjecting their opinions into my media selections.”

Discussion and Implications

This chapter interprets and contextualizes the results obtained from the crowdsourcing studies, identifying any overarching trends and answering the four research questions. It also provides a reflection on the broad implications of the conducted research and offers a variety of suitable directions for future work on the implementation and empirical evaluation of auditory misinformation warnings.

6.1. Impact and Reception of Podcast-Integrated Auditory Icons (RQ1)

Based on Study 1's findings, all three sound classes considered in this work (alarms, animal noises, natural/domestic sounds) contain representative samples that would be suitable non-verbal conversational cues – namely, such that they are salient, instill a sense of urgency, and capture the attention of the listener without the latter reporting high distraction or interruption levels. To that end, the data are in accordance with prior literature – each sound category is perceptually prominent, and nearly all auditory icons included in Study 1 are highly recognizable and exhibit consistent conceptual mappings among participants.

The collected free-text responses offer additional nuances relevant for the selection of candidate auditory icons. A listener's first encounter with the sound cue is invariably surprising, yet certain samples can be overpowering (e.g. car honk, referee whistle, the sound of thunder), shocking or frightening listeners rather than triggering a minor disturbance. Additionally, while instances where the sound overlaps with external or familiar stimuli are recoverable, only affecting its first occurrence, auditory icons that could blend with the imaginary podcast environment perceived by the listener (i.e. a host tearing off a piece of paper close to the microphone rather than the sound being intentionally edited-in) should also be avoided.

The main driver of prolonged dissatisfaction is a sample's appropriateness with respect to (i) the podcast's ongoing narrative and (ii) the listener's pre-existing expectations of what a typical sound cue should be. The more severe the perceived mismatch along either of these two dimensions is, the more distracting and inappropriate an auditory icon is considered (e.g. car honk, seagull screech, frog croak, the sound of glass breaking). Such discrepancy also creates confusion on the listener's part – seeing as a given sample is not usually encountered in regular podcasts or conversation settings, its relevance in view of the surrounding context and the overarching topic is unclear, making the non-verbal cue appear as a purposeless interruption.

The reported confusion may also have been exacerbated as a consequence of Study 1's design. Since its focus was obtaining subjective judgments and impressions of the sound samples themselves, the cues did not fulfill the functional purpose of misinformation warnings, instead simply being inserted into the podcast excerpts at pre-defined between-utterance positions (unlike in Study 2). Given that the sounds appear midway through plain dialogue rather than accompany misinformative statements, therefore, participants are unable to identify an intentional pattern in their occurrences, which would in turn magnify the experienced confusion.

Regardless of its underlying factors, this puzzled state of mind and the apparent disconnect between the sound and the podcast's narrative then break the immersion and give way to genuine amusement,

the auditory icon being considered ridiculous and “comically out of place”. Given the objective of minimal listening-experience disruption, this outcome is highly undesirable and emphasizes the importance of a non-verbal cue’s perceived contextual appropriateness for effectiveness to be maintained.

To that end, while (certain) alarms, animal noises, and natural/domestic sounds can serve as salient and easily-recognizable interventions, their mismatch with the established “warning” archetype should not become too severe: to avoid the negative consequences of such a scenario, pilot studies could be carried out to verify a given design in practice before fully committing to it.

6.2. Interplay Between Warnings and Misinformation Awareness (RQ2)

Inspecting the quantitative results from Study 2’s veracity assessment, no definitive conclusions can be made regarding a change in misinformation-recognition consistency with or without podcast-integrated auditory warnings, nor can a universally best-performing placement configuration be determined out of the four implemented alternatives. The outlier observed for misinformation recognition on the topic of GMO foods (14/42) can be explained through the collected free-text rationales: that particular misinformative excerpt used in the assessment turns out to be an opinion shared by many of the participants, causing them to misjudge its veracity and leading to the comparatively low recognition scores for the GMO topic. This finding once again emphasizes the significant impact of prior exposure (and the associated availability bias) on misinformation awareness, indicating that the decades-long media circulation of hoaxes and rumors can reinforce the believability and long-term retention of false claims.

The obtained qualitative data reveal other prominent reasons for failing to recognize misinformation as well. One of them is relying on personal pre-conceptions and intuition rather than logical evidence-based deduction to judge a statement’s veracity, some participants referring to a “common consensus” off which they based their ruling. Another involves appealing to the podcast hosts’ authority as reliable arbiters of authentic information who could be trusted to have thoroughly researched the topic before commenting on it: this fallacy matches insights from prior literature indicating the negative implications of podcasting’s many elements tailored towards persuasion for critical engagement [60].

Crucially for the investigation of RQ2, the collected rationales offer no explicit mention of participants recalling the sound cue from the podcast and considering it while judging a statement’s veracity. Instead, listeners who successfully identified the misinformation point out flaws in the statements themselves or were able to recall corrective information from the podcast content – attributing their veracity ruling to those indicators.

Despite this, findings pertaining to the Presumed Sound Purpose dependent variable indicate that some participants were able to surmise the sound cues’ role as misinformation warnings – specifically, when the latter were *enclosing* or placed *after* the falsehood-containing segments. These data suggest that a cognitive association between the non-verbal sounds and the “hazard” of misinformation was made even in the circumstances of partial disclosure: the intended purpose behind the auditory icons’ integration can nonetheless be intuited by listeners.

Without full disclosure from the outset, however, RQ2 cannot be answered conclusively. From the crowdworkers’ perspective, the sounds’ intended function remains unknown (or at least uncertain) until the post-study debrief: no purpose behind their integration is immediately evident, which causes participants to not heed them explicitly during the veracity assessment (as evidenced by the missing mentions of the cues in the rationales). The only findings from the study indicative of the auditory icons’ interplay with misinformation awareness are the aforementioned presumed-purpose responses.

6.3. Influence of Placement Alternatives on Warning Performance (RQ3)

The placement of warnings with respect to the misinformative segment they are meant to indicate has implications for their effectiveness and reception by listeners. Overall, interventions that play concurrently with the podcast’s audio are associated with the least accurate retention of content, yielding the lowest recall-assessment scores out of all study conditions. Moreover, compared to listening without auditory interventions, a statistically significant reduction in the odds of middle-stage recall (i.e. in the second of each podcast’s three equal-length partitions) of > 60% is estimated for the *concurrent* condition. These results suggest the latter considerably disrupts the podcasting experience, diminishing

the content's discernibility and throwing listeners off the ongoing narrative.

In turn, based on participants' subjective evaluations of the sound cues, the *after* placement configuration outperforms all others in terms of perceived distraction, abruptness, and interruption, having the majority of judgments for these metrics in their lowest-rank categories. Two of these performance indicators are also determined to be statistically significant: compared to concurrently positioned warnings, ones appearing *after* a falsehood-containing segment are associated with reductions of > 70% in the odds of reporting higher distraction and interruption levels.

These findings therefore demonstrate that interventions positioned *after* occurrences of misinformation are the most favorably perceived and least disruptive to the overall listening experience. From a cognitive lens, this could be attributed to their unique property of enabling preliminary information processing to take place. Listeners first hear the misleading statement, all preceding context remaining intact and the narrative proceeding uninterrupted, after which they encounter the purposefully embedded warning. Provided a suitable auditory icon is used (per the guidelines formulated through RQ1), the intervention would (i) trigger a momentary surprise, (ii) prompt listeners to identify the reason for its inclusion, and thereby (iii) lead them to re-assess the statement they just heard. In the circumstances of attentive podcast listening – with a heightened misinformation awareness and having re-focused on the content – listeners could then behave according to their perceived risk level (e.g. pause/rewind the audio, approach subsequently presented information more cautiously, or resume the podcast without minding the warning).

6.4. Influence of Contextual Factors on Warning Performance (RQ4)

Several of the exploratory variables considered in this contribution exhibit non-trivial latent interactions with the outcome measures. Among the subjective warning evaluations, three statistically significant correlations are observed: (i) a high perceived podcast density (or, a sense of large information clutter) is associated with reduced odds of reporting high sound recognizability; (ii) a slow perceived podcast pace is associated with increased odds of reporting high warning-induced distraction levels; (iii) a strong pre-existing stance is associated with a reduction in the latter odds. Aside from participants' judgments of the interventions themselves, their recall-assessment performance is also influenced by external factors: one's pre-existing stance, attitude strength (and its underlying predictors), and need for cognition are each linked to a statistically significant increase in the odds of proportional recall. These findings suggest that follow-up contributions should model such explanatory metrics as independent variables in their methodologies, more closely examining and quantifying the latent interactions described here to confirm or reject their significance with greater certainty.

Additional factors influencing the auditory interventions' effectiveness can be identified through the analysis of qualitative data from participants' responses. In terms of the initial reactions prompted by a warning's occurrence, Study 2's overarching themes are consistent with those outlined in Study 1: the sound cues are surprising when encountered for the first time and instill a sense of urgency, yet they may lead to confusion, prolonged dissatisfaction, and a loss of immersion if they drastically mismatch the narrative context or listeners' own pre-conceptions of "traditional" podcast-editing effects.

Unique to Study 2, however, data regarding warning acclimatization suggest that some participants experience a decrease in the distraction and initial confusion triggered by the interventions, particularly after identifying an underlying pattern or purpose behind their presence in the podcast. From the second encounter onwards, these listeners perceive the sounds as intentional signals, attributing greater importance to them and finding them less disruptive over time. Still, this cluster of subjective impressions is only one of many diverse (and often conflicting) themes distinguished during analysis – other participants report being equally startled by the cues upon each occurrence (as the latter "were spaced out enough for [one's] guard to lower") or instead becoming indifferent towards them, the latter getting passively discarded through habituation or perceptually overruled by the podcast's narrative.

A similar lack of cohesion can be inferred from the post-debrief exploratory rationales: in response to hearing a misinformation warning during their regular podcasts, participants would react by (i) closing the audio player (as well as unsubscribing from the podcast), (ii) continuing to listen with a heightened awareness of the content's veracity and the presented evidence, (iii) individually verifying the claims to obtain additional confirmation (beyond blindly trusting the warning), (iv) discarding the signal and instead prioritizing the ongoing conversation, or (v) actively opposing the interventions, perceiving them

as paternalistic and authoritative.

Taken altogether, these qualitative findings highlight two key conclusions. First, there is no silver-bullet solution to implementing an effective, reliably salient, and minimally disruptive misinformation warning that can effortlessly be integrated into podcasts as the motivations, priorities, habits, and perceived stakes of podcast consumption differ for each individual listener. Second (and related to that), the overarching vision for such an intervention should remain that of a nudge fostering critical engagement with the presented information and thereby helping users avoid the negative effects of falsehood retention – rather than one of an embedded interaction mechanism that listeners are expected to heed as the final authority on the podcast information’s veracity. Without incurring the notion of paternalism associated with such a design, the warnings’ transparency and reliability could instead be improved by justifying why they were inserted at specific positions throughout the podcast – in turn, integrating such explanations without disrupting the overall listening experience remains an open challenge.

6.5. Broad Research Implications

Tackling different design considerations for a user-centered intervention intended to curtail the dissemination and reinforcement of misinformation, this contribution is naturally positioned at the intersection of (i) responsible opinion formation and (ii) human-computer interaction. Prior works in the field of computer science have also explored this topic, primarily concerned with visual interfaces and the use-case scenario of Web search. The present research extends the focus of these efforts by considering the domain of podcasts – an increasingly prominent medium, whose appealing inherent properties (accessibility, convenience, and an engaging “infotainment” approach) can inadvertently facilitate the spread of falsehoods online.

This exploratory contribution’s primary objective is to therefore inform a first iteration of guidelines for the implementation of misinformation warnings in the audio modality. Chosen as the focus of this investigation, the medium of podcasts already presents unique challenges in the way of attaining this goal, such as its reliance on an uninterrupted narrative for content discernibility and listener comprehension, as well as its fast pace of delivery. These aspects make the selection of sound types and samples, their pitch and duration, and the precise placement of such cues throughout a podcast’s audio track all non-trivial decisions to consider.

Furthermore, the effectiveness of misinformation interventions is not determined solely by their optimized technical parameters, human cognition and user perceptions also playing an equally critical role. Altogether, this becomes an expansive research topic that bridges several disciplines (among which computer science, psychology, and sociology) and involves a host of influential factors, including: individual psychometric characteristics; chosen controversial topics; relevant cognitive biases; different levels of misinformation severity; benefits and drawbacks of nudging as opposed to boosting interventions; personal podcast-listening motivations and preferences.

This widened research gap presents a large variety of insightful (and often conflicting) trade-offs to account for and examine. Consequently, many directions for follow-up contributions can be delineated, which is the focus of the next section.

6.6. Limitations and Directions for Future Work

The topic of designing auditory misinformation warnings leaves plenty of room for conducting additional research. Several directions for follow-up work can already be identified by addressing the limitations of the present contribution. One such constraint is the use of partial disclosure, which restricted the extent to which the sound cues’ downstream cognitive implications could be assessed. A differently designed study could explore these trends, capturing the impact of warning encounters on podcast-consumption behavior, attitude change, or belief update.

Misinformation detection methods are another aspect beyond the focus of this work, having been circumvented through the use of generative AI for the controlled injection of falsehoods based on which the sound cues were positioned. A follow-up contribution could instead integrate auditory interventions into podcasts in line with findings reported by state-of-the-art techniques for detecting inaccurate claims and investigate the resulting effects on the cues’ perceived transparency and reliability. Podcasts that are longer, readily available online, and differently formatted (such as multi-host and group discussions) could also be chosen for improved study realism.

In the latter regard, subsequent experiments on this topic could utilize an alternative approach to

crowdsourcing – for instance, a confirmatory lab study could be conducted to enable improved monitoring, a tighter control over experimental conditions, and a closer representation of real-world podcast-listening settings. Furthermore, such a design would facilitate an investigation of additional podcast-consumer categories besides attentive listeners, researchers being able to simulate menial-labor multitasking as well as inattentive listening that involves a concurrent yet also cognitively demanding task. Another potentially insightful option would be a longitudinal diary study for maximal realism, capturing participants' personal podcast preferences, actual listening patterns, and authentic experiences with auditory warnings throughout their regular podcast sessions.

Moving beyond these limitations, follow-up works could explore completely novel implementations and caveats of auditory interventions. For example, the impact of inaccurately-placed warnings on listeners' perceptions and misinformation awareness could be examined, determining whether engagement with the sound cues is reduced – and if a negative spillover effect occurs, lowering the believability of the podcast's factually accurate content as well. Other contributions could in turn investigate the disambiguation of multiple warning-types tailored towards the severity of a detected fallacy and listeners' perceived urgency: for instance, an individual might consider incomplete information to be less cognitively detrimental than intentional misinformation, in which case a more subtle intervention could be implemented. A similar multi-layered design could also be applied across different "confidence" levels of flagging accuracy, especially if an automated misinformation detection algorithm is used, in an effort to improve the sound cues' perceived transparency and reliability.

Furthermore, a series of contributions could examine opportunities for warning customization. Inspired by insights from the visual space [64], a polymorphic intervention-design could be implemented in order to combat habituation – reducing a listener's desensitization to one particular sound cue as a result of repeated exposure and familiarity. Other promising research avenues relevant for the warnings' adoption by podcast-hosting platforms include customizable designs that can be tailored towards an individual user's preferences or the creation of a personalized avatar in order to foster engagement with the interventions.

Finally, the integration of "boosting" measures within the overall podcast-listening experience could be explored as an alternative to nudges, considering the latter's shortcomings related to paternalism, perceived user autonomy, and the durability of their cognitive effects. In turn, the impact and potential long-term benefits of "boosting" interventions could be assessed in terms of misinformation correction and knowledge revision – in addition to capturing listeners' subjective impressions to verify whether the design is favorably perceived.

Conclusion

Disrupting the information-seeking process, the spread of online misinformation constitutes an ongoing social challenge and is especially prominent for controversial topics, which are associated with polarization, heightened emotionality, suppressed critical thought. In an effort to curtail it, visual misinformation flags have been implemented as nudges which prompt changes in user behavior by introducing minor “design frictions” (e.g. obscuring part of a page’s content). Beyond Web search, podcasts represent an increasingly popular medium that offers greater accessibility and engagement, yet their fast pace and passive mode of consumption make it difficult for the audience to identify falsehoods. An audio-based flag-like intervention is therefore necessary to facilitate listeners’ critical assessment of podcast-content veracity – the design of such sound cues being the focus of this contribution.

The overarching research objective is the development of guidelines for effective podcast-integrated misinformation warnings, which requires (i) analyzing the impact of these interventions on the podcast listening experience and (ii) examining the external factors that influence the way they are perceived by listeners. To do so, two crowdsourcing studies were carried out. The first involved collecting subjective judgments of 15 auditory icons (the chosen non-verbal sound-cue type for this research) interspersed into brief dialogue excerpts to identify the least disruptive and most favorably perceived samples among them. Integrating those as misinformation warnings into several AI-generated podcasts modified to contain misleading claims, the second study then investigated the interventions’ interplay with participants’ recall of podcast content, misinformation recognition, and subjective experiences during listening – its design using partial disclosure to not reveal the sound cues’ presence / purpose, featuring a between-subjects partition across 4 warning-placement configurations, and including several psychometric variables as exploratory predictors.

Through Study 1, salient samples instilling a sense of urgency without being overly distracting were identified across the categories of alarms, animal noises, and natural / domestic sounds, the qualitative findings also highlighting the importance of a sound’s listener-perceived appropriateness (based on pre-existing notions of “regular” podcast-editing effects) for it to be interpreted as an intentional signal rather than an out-of-context interruption. Serving as misinformation indicators throughout Study 2, the cues were not explicitly heeded by participants during the veracity assessment, yet those *enclosing* a falsehood or *appearing after* one were accurately presumed to fulfill that role by some evaluators. Additional results pertaining to warning placement were obtained – auditory signals *positioned after* a misleading statement were considered the least distracting and disruptive of the listening experience, both these metrics yielding a significant difference compared to *concurrently placed* interventions – which were in turn further associated with the lowest recall-assessment performance. Regarding listeners’ subjective evaluations and the influence of contextual factors, statistically significant correlations were observed with one’s pre-existing personal stance, as well as a podcast’s perceived density and pace.

The interpreted trends of warning acclimatization and participants’ follow-up actions in response to hearing such interventions in non-experimental settings, however, show a lack of cohesion: attributing greater importance to the cues from their second occurrence onwards, individually assessing a statement’s veracity without automatically trusting them, passively discarding them through habituation, and actively opposing them as paternalistic are all among the prominent themes. Such variety only widens the research gap, offering different suitable directions for future work on this topic.

Declaration on the Use of Generative AI

Throughout this research contribution, two generative AI models were used – Google’s **NotebookLM** and Anthropic’s **Claude Opus 4.7** – the present declaration elaborating on the manner and individual process stages in which either tool was involved.

NotebookLM was used for the purposes of generating the three podcasts which underpin the study and their associated recall-assessment questionnaires. For podcast generation, the *Audio Overview* feature was utilized with the *Deep Dive* and *Default Length* settings toggled on: the concrete prompts issued to the model are included in Appendix B to facilitate reproducibility. The outputs themselves are described in Section 3.2 of the report and fully documented in Appendices C and D (via full transcripts and exhaustive collections of relevant Web resources).

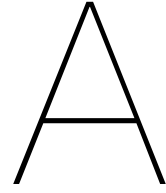
The podcasts were manually reviewed to verify their structure and the consistency of misinformation injection, as well as insert auditory interventions across the 4 warning-placement configurations (using each misinformation occurrence as a vantage point). In turn, the 3 recall questionnaires were carefully inspected to examine each question’s phrasing and degree of ambiguity, as well as to ensure all answer options are similarly formatted and equally plausible: manual modifications were carried out to resolve any such inconsistencies.

Claude Opus 4.7 was used during the development of the web application housing the logic for both crowdsourcing studies, as well as for the generation of automated data-processing scripts. The model was prompted via the *Claude Code* extension available for Visual Studio Code & utilized exclusively in *Plan Mode*, in which it adheres to detailed Markdown specification files modified and reviewed by the requester before any output generation takes place.

During web app development, the following four tasks were issued to Claude Opus: (i) the arrangement and connectivity of the HTML elements in each frontend page (buttons, answer options, free-text areas, and instruction-text containers); (ii) the dynamic routing of frontend pages based on the staging specification for each study; (iii) the implementation of audio-player templates with varying functionality restrictions (e.g. disabled fast-forwarding and seeking); (iv) the generation of a unified CSS file governing all frontend pages’ appearance. The final application underwent thorough manual testing, covering both good- and bad-weather scenarios to simulate different workflow branches and verify the correct persistence of response data.

In turn, the data-processing scripts generated by Opus (written in Python) were used to extract and aggregate study data from the raw CSVs – these files having been pulled directly from the database – into JSON-formatted summaries with the aim of improving readability and facilitating subsequent analyses (namely, descriptive and reflexive thematic). Both the automated scripts and the JSON outputs themselves were manually reviewed to verify the resulting files are structured as intended and the data remain correct regardless of file format.

Besides these models and use-cases, no additional generative AI tools were used at any other point in this contribution, including the exploration of prior literature, the findings’ analysis and interpretation, or the report-writing process.



Exhaustive Study Results

Table A.1: Complete quantitative findings for all 15 sound cues investigated in Study 1. For each metric, μ values indicate the mean, σ values show the standard deviation, and the distribution of participant judgments across the 5 levels (Not at all / Slightly / Moderately / Very / Extremely) is reported based on the obtained Likert-scale input.

Sound Cue	#evaluators	Recognizability		Distraction		Interruption	
		μ	σ	μ	σ	μ	σ
Referee Whistle	10	4.5	1.27	3.7	0.82	2.8	1.4
		1 / 0 / 0 / 1 / 8		0 / 0 / 5 / 3 / 2		2 / 3 / 1 / 3 / 1	
Car Honk	10	4.9	0.32	4.2	1.14	4.3	1.25
		0 / 0 / 0 / 1 / 9		0 / 1 / 2 / 1 / 6		0 / 2 / 0 / 1 / 7	
Bike Bell	10	4.8	0.42	2.6	1.51	2.4	1.35
		0 / 0 / 0 / 2 / 8		3 / 2 / 3 / 0 / 2		3 / 3 / 2 / 1 / 1	
Factory Warning	9	4.0	1.12	3.89	1.36	4.22	1.3
		0 / 1 / 2 / 2 / 4		1 / 0 / 2 / 2 / 4		1 / 0 / 0 / 3 / 5	
Sonar Ping	10	3.6	1.17	3.8	1.75	4.2	1.23
		0 / 2 / 3 / 2 / 3		2 / 1 / 0 / 1 / 6		1 / 0 / 0 / 4 / 5	
Dog Bark	9	4.89	0.33	3.67	1.41	3.56	1.33
		0 / 0 / 0 / 1 / 8		0 / 3 / 1 / 1 / 4		0 / 3 / 1 / 2 / 3	
Owl Hoot	9	4.0	0.71	2.56	1.33	2.78	1.56
		0 / 0 / 2 / 5 / 2		2 / 3 / 2 / 1 / 1		2 / 3 / 1 / 1 / 2	
Seagull Screech	10	4.8	0.42	3.8	1.48	4.0	1.33
		0 / 0 / 0 / 2 / 8		1 / 1 / 2 / 1 / 5		1 / 0 / 2 / 2 / 5	
Frog Croak	10	4.3	1.34	3.2	1.32	4.0	0.94
		1 / 0 / 1 / 1 / 7		1 / 2 / 3 / 2 / 2		0 / 1 / 1 / 5 / 3	
Horse Whinny	9	4.78	0.44	4.11	1.05	4.11	1.05
		0 / 0 / 0 / 2 / 7		0 / 1 / 1 / 3 / 4		0 / 1 / 1 / 3 / 4	

continued on next page

Table A.1 – continued from previous page

Sound Cue	#evaluators	Recognizability		Distraction		Interruption	
		μ	σ	μ	σ	μ	σ
Whip Cracking	10	3.9	1.1	3.8	1.4	4.0	1.33
		0 / 2 / 0 / 5 / 3		1 / 1 / 1 / 3 / 4		1 / 0 / 2 / 2 / 5	
Paper Ripping	10	4.7	0.48	3.6	1.58	3.6	1.65
		0 / 0 / 0 / 3 / 7		2 / 0 / 2 / 2 / 4		2 / 0 / 3 / 0 / 5	
Scrap Metal Falling	10	4.6	0.7	3.5	1.27	3.4	1.26
		0 / 0 / 1 / 2 / 7		1 / 1 / 2 / 4 / 2		1 / 1 / 3 / 3 / 2	
Thunder Rumbling	10	4.6	0.52	3.6	1.43	4.1	0.99
		0 / 0 / 0 / 4 / 6		1 / 1 / 3 / 1 / 4		0 / 1 / 1 / 4 / 4	
Glass Breaking	9	4.44	0.73	4.0	1.22	3.67	1.22
		0 / 0 / 1 / 3 / 5		1 / 0 / 0 / 5 / 3		1 / 0 / 2 / 4 / 2	

Table A.2: Complete results obtained from Kendall's Tau correlation indicating monotonic relationships in pairs of outcome and explanatory variables. Test outputs include the τ statistic and the p value.

Outcome Variable	Explanatory Variable	τ	p
recall_count	topic_familiarity	0.161	0.027
recall_count	prior_stance	0.078	0.285
recall_count	podcast_density	-0.050	0.528
recall_count	podcast_pace	0.115	0.143
recall_count	attitude_strength	0.116	0.094
recall_count	need_for_cognition	0.038	0.582
recall_count	intellectual_humility	0.013	0.853
recall_early	topic_familiarity	0.195	0.010
recall_early	prior_stance	0.042	0.574
recall_early	podcast_density	-0.103	0.205
recall_early	podcast_pace	0.021	0.794
recall_early	attitude_strength	0.169	0.018
recall_early	need_for_cognition	0.008	0.909
recall_early	intellectual_humility	0.042	0.554
recall_middle	topic_familiarity	0.139	0.062
recall_middle	prior_stance	0.134	0.070
recall_middle	podcast_density	-0.043	0.590
recall_middle	podcast_pace	0.113	0.159
recall_middle	attitude_strength	0.101	0.153
recall_middle	need_for_cognition	0.017	0.814
recall_middle	intellectual_humility	-0.056	0.414

continued on next page

Table A.2 – *continued from previous page*

Outcome Variable	Explanatory Variable	τ	p
recall_late	topic_familiarity	0.081	0.287
recall_late	prior_stance	0.002	0.983
recall_late	podcast_density	−0.040	0.625
recall_late	podcast_pace	0.125	0.127
recall_late	attitude_strength	0.012	0.871
recall_late	need_for_cognition	0.154	0.033
recall_late	intellectual_humility	0.105	0.135
acclimate_trend	topic_familiarity	0.272	0.004
acclimate_trend	prior_stance	−0.101	0.286
acclimate_trend	podcast_density	0.129	0.213
acclimate_trend	podcast_pace	−0.133	0.197
acclimate_trend	attitude_strength	0.241	0.007
acclimate_trend	need_for_cognition	0.050	0.581
acclimate_trend	intellectual_humility	0.017	0.847
recognizability	topic_familiarity	0.022	0.811
recognizability	prior_stance	−0.044	0.634
recognizability	podcast_density	−0.256	0.010
recognizability	podcast_pace	−0.086	0.390
recognizability	attitude_strength	−0.026	0.766
recognizability	need_for_cognition	−0.051	0.562
recognizability	intellectual_humility	0.138	0.105
distraction	topic_familiarity	0.001	0.993
distraction	prior_stance	−0.067	0.449
distraction	podcast_density	0.121	0.208
distraction	podcast_pace	0.106	0.267
distraction	attitude_strength	−0.005	0.949
distraction	need_for_cognition	−0.032	0.706
distraction	intellectual_humility	0.009	0.909
abruptness	topic_familiarity	−0.167	0.065
abruptness	prior_stance	−0.046	0.612
abruptness	podcast_density	−0.109	0.268
abruptness	podcast_pace	0.005	0.958
abruptness	attitude_strength	−0.109	0.202
abruptness	need_for_cognition	0.032	0.706
abruptness	intellectual_humility	0.105	0.212

continued on next page

Table A.2 – continued from previous page

Outcome Variable	Explanatory Variable	τ	p
interruption	topic_familiarity	−0.117	0.185
interruption	prior_stance	−0.009	0.916
interruption	podcast_density	0.021	0.824
interruption	podcast_pace	0.144	0.132
interruption	attitude_strength	−0.081	0.333
interruption	need_for_cognition	−0.082	0.326
interruption	intellectual_humility	−0.123	0.132

Table A.3: Complete results obtained from Poisson regression for count data (significant findings highlighted in yellow). For categorical predictors, the specific level (appearing in parentheses) is distinguished from the reference category (indicated in superscript) to make the grouping explicit. Table entries containing more than one uniquely-named explanatory variable indicate multivariable-model statistics. Outputs include the overdispersion ratio R , log-difference in outcome β , expected-count ratio e^β , standard coefficient error σ_β , and p value.

Outcome	R	Explanatory Variable(s)	β	e^β	σ_β	p
recall_count	1.04	cond(before) $\rightarrow$ without	0.0141	1.01	0.084	0.866
		cond(after) $\rightarrow$ without	−0.0252	0.98	0.085	0.766
		cond(enclosing) $\rightarrow$ without	−0.0777	0.93	0.086	0.367
		cond(concurrent) $\rightarrow$ without	−0.1169	0.89	0.087	0.179
recall_count	1.01	topic_familiarity(1) $\rightarrow$ ⁴	−0.2651	0.77	0.135	0.049
		topic_familiarity(2) $\rightarrow$ ⁴	0.0386	1.04	0.098	0.693
		topic_familiarity(3) $\rightarrow$ ⁴	0.0116	1.01	0.073	0.875
		topic_familiarity(5) $\rightarrow$ ⁴	0.1050	1.11	0.081	0.195
		topic_familiarity(6) $\rightarrow$ ⁴	0.0831	1.09	0.096	0.387
recall_count	1.04	prior_stance(2) $\rightarrow$ ⁴	0.0548	1.06	0.111	0.622
		prior_stance(3) $\rightarrow$ ⁴	0.0925	1.10	0.085	0.274
		prior_stance(5) $\rightarrow$ ⁴	0.0715	1.07	0.076	0.348
		prior_stance(6) $\rightarrow$ ⁴	0.1950	1.22	0.085	0.022
		prior_stance(7) $\rightarrow$ ⁴	0.0354	1.04	0.112	0.752
recall_count	1.04	podcast_density(4) $\rightarrow$ ³	−0.0337	0.97	0.061	0.582
		podcast_density(5) $\rightarrow$ ³	−0.1088	0.90	0.095	0.250
recall_count	1.04	podcast_pace(2) $\rightarrow$ ³	0.0977	1.10	0.099	0.325
		podcast_pace(4) $\rightarrow$ ³	0.0642	1.07	0.067	0.336
recall_count	1.01	attitude_strength	0.0292	1.03	0.021	0.166
recall_count	1.02	need_for_cognition	0.0455	1.05	0.041	0.263
recall_count	1.03	intellectual_humility	0.0197	1.02	0.028	0.489

continued on next page

Table A.3 – continued from previous page

Outcome	R	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
recall_count	1.05	topic_familiarity(1) $\rightarrow$ ⁴	−0.2305	0.79	0.137	0.092
		topic_familiarity(2) $\rightarrow$ ⁴	0.0737	1.08	0.104	0.480
		topic_familiarity(3) $\rightarrow$ ⁴	0.0258	1.03	0.074	0.728
		topic_familiarity(5) $\rightarrow$ ⁴	0.0865	1.09	0.083	0.299
		topic_familiarity(6) $\rightarrow$ ⁴	0.0628	1.06	0.098	0.523
		prior_stance(2) $\rightarrow$ ⁴	0.0117	1.01	0.115	0.919
		prior_stance(3) $\rightarrow$ ⁴	0.0659	1.07	0.090	0.465
		prior_stance(5) $\rightarrow$ ⁴	0.0564	1.06	0.080	0.480
		prior_stance(6) $\rightarrow$ ⁴	0.1570	1.17	0.090	0.081
		prior_stance(7) $\rightarrow$ ⁴	0.0212	1.02	0.114	0.852
recall_count	1.07	cond(<i>before</i>) $\rightarrow$ <i>without</i>	0.0310	1.03	0.087	0.721
		cond(<i>after</i>) $\rightarrow$ <i>without</i>	−0.0488	0.95	0.087	0.577
		cond(<i>enclosing</i>) $\rightarrow$ <i>without</i>	−0.1013	0.90	0.089	0.253
		cond(<i>concurrent</i>) $\rightarrow$ <i>without</i>	−0.1125	0.89	0.093	0.225
		topic_familiarity(1) $\rightarrow$ ⁴	−0.2618	0.77	0.139	0.060
		topic_familiarity(2) $\rightarrow$ ⁴	0.0630	1.07	0.106	0.554
		topic_familiarity(3) $\rightarrow$ ⁴	0.0134	1.01	0.076	0.859
		topic_familiarity(5) $\rightarrow$ ⁴	0.1015	1.11	0.084	0.226
		topic_familiarity(6) $\rightarrow$ ⁴	0.0277	1.03	0.100	0.783
		prior_stance(2) $\rightarrow$ ⁴	0.0247	1.03	0.115	0.831
		prior_stance(3) $\rightarrow$ ⁴	0.0768	1.08	0.092	0.402
		prior_stance(5) $\rightarrow$ ⁴	0.0641	1.07	0.080	0.424
		prior_stance(6) $\rightarrow$ ⁴	0.1409	1.15	0.093	0.131
		prior_stance(7) $\rightarrow$ ⁴	0.0058	1.01	0.115	0.960

Table A.4: Complete results obtained from Fractional Logistic regression for proportion data (significant findings highlighted in yellow). For categorical predictors, the specific level (appearing in parentheses) is distinguished from the reference category (indicated in superscript) to make the grouping explicit. Table entries containing more than one uniquely-named explanatory variable indicate multivariable-model statistics. Outputs include the log-difference in outcome β , recall-proportion odds ratio e^β , standard coefficient error σ_β , and p value.

Outcome	Explanatory Variable(s)	β	e^β	σ_β	p
recall_early	cond(<i>before</i>) $\rightarrow$ without	0.5900	1.80	0.523	0.259
	cond(<i>after</i>) $\rightarrow$ without	-0.0974	0.91	0.446	0.827
	cond(<i>enclosing</i>) $\rightarrow$ without	-0.5731	0.56	0.399	0.151
	cond(<i>concurrent</i>) $\rightarrow$ without	-0.0626	0.94	0.452	0.890
recall_early	topic_familiarity(1) $\rightarrow$ ⁴	-0.6804	0.51	0.416	0.102
	topic_familiarity(2) $\rightarrow$ ⁴	-0.2394	0.79	0.418	0.566
	topic_familiarity(3) $\rightarrow$ ⁴	-0.2924	0.75	0.385	0.447
	topic_familiarity(5) $\rightarrow$ ⁴	0.5612	1.75	0.449	0.211
	topic_familiarity(6) $\rightarrow$ ⁴	0.5221	1.69	0.575	0.364
recall_early	prior_stance(2) $\rightarrow$ ⁴	0.1647	1.18	0.496	0.740
	prior_stance(3) $\rightarrow$ ⁴	0.3058	1.36	0.444	0.491
	prior_stance(5) $\rightarrow$ ⁴	0.0663	1.07	0.386	0.864
	prior_stance(6) $\rightarrow$ ⁴	0.8334	2.30	0.488	0.088
	prior_stance(7) $\rightarrow$ ⁴	-0.0148	0.99	0.428	0.972
recall_early	podcast_density(4) $\rightarrow$ ³	-0.4644	0.63	0.310	0.134
	podcast_density(5) $\rightarrow$ ³	-0.3704	0.69	0.467	0.427
recall_early	podcast_pace(2) $\rightarrow$ ³	0.8707	2.39	0.612	0.155
	podcast_pace(4) $\rightarrow$ ³	0.0604	1.06	0.361	0.867
recall_early	attitude_strength	0.2335	1.26	0.088	0.008
recall_early	need_for_cognition	0.0722	1.07	0.204	0.723
recall_early	intellectual_humility	0.1370	1.15	0.147	0.350
recall_early	cond(<i>before</i>) $\rightarrow$ without	0.6553	1.93	0.526	0.213
	cond(<i>after</i>) $\rightarrow$ without	-0.1325	0.88	0.433	0.760
	cond(<i>enclosing</i>) $\rightarrow$ without	-0.5187	0.60	0.412	0.208
	cond(<i>concurrent</i>) $\rightarrow$ without	-0.1485	0.86	0.441	0.736
	attitude_strength	0.2448	1.28	0.100	0.014

continued on next page

Table A.4 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
recall_early	cond(before) $\rightarrow$ without	0.7508	2.12	0.494	0.129
	cond(after) $\rightarrow$ without	-0.1016	0.90	0.431	0.814
	cond(enclosing) $\rightarrow$ without	-0.5971	0.55	0.413	0.148
	cond(concurrent) $\rightarrow$ without	0.0221	1.02	0.448	0.961
	prior_stance(2) $\rightarrow$ 4	-0.0023	1.00	0.472	0.996
	prior_stance(3) $\rightarrow$ 4	0.1955	1.22	0.490	0.690
	prior_stance(5) $\rightarrow$ 4	-0.1263	0.88	0.415	0.761
	prior_stance(6) $\rightarrow$ 4	0.8023	2.23	0.503	0.111
	prior_stance(7) $\rightarrow$ 4	-0.5429	0.58	0.492	0.269
	attitude_strength	0.2516	1.29	0.110	0.022
recall_middle	cond(before) $\rightarrow$ without	-0.2918	0.75	0.469	0.534
	cond(after) $\rightarrow$ without	-0.2624	0.77	0.462	0.570
	cond(enclosing) $\rightarrow$ without	-0.4509	0.64	0.448	0.314
	cond(concurrent) $\rightarrow$ without	-0.9745	0.38	0.476	0.041
recall_middle	topic_familiarity(1) $\rightarrow$ 4	-1.0083	0.36	0.361	0.005
	topic_familiarity(2) $\rightarrow$ 4	0.2693	1.31	0.551	0.625
	topic_familiarity(3) $\rightarrow$ 4	0.2167	1.24	0.383	0.572
	topic_familiarity(5) $\rightarrow$ 4	0.5097	1.66	0.515	0.322
	topic_familiarity(6) $\rightarrow$ 4	0.4742	1.61	0.522	0.364
recall_middle	prior_stance(2) $\rightarrow$ 4	0.0016	1.00	0.400	0.997
	prior_stance(3) $\rightarrow$ 4	0.5505	1.73	0.499	0.270
	prior_stance(5) $\rightarrow$ 4	0.4441	1.56	0.370	0.230
	prior_stance(6) $\rightarrow$ 4	1.3466	3.84	0.441	0.002
	prior_stance(7) $\rightarrow$ 4	0.6611	1.94	0.545	0.225
recall_middle	podcast_density(4) $\rightarrow$ 3	-0.0262	0.97	0.327	0.936
	podcast_density(5) $\rightarrow$ 3	-0.4781	0.62	0.434	0.270
recall_middle	podcast_pace(2) $\rightarrow$ 3	0.6258	1.87	0.568	0.271
	podcast_pace(4) $\rightarrow$ 3	0.2933	1.34	0.432	0.497
recall_middle	attitude_strength	0.1562	1.17	0.089	0.079
recall_middle	need_for_cognition	0.1080	1.11	0.209	0.605
recall_middle	intellectual_humility	-0.0237	0.98	0.146	0.871

continued on next page

Table A.4 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
recall_middle	cond(before) $\rightarrow$ without	-0.2346	0.79	0.434	0.589
	cond(after) $\rightarrow$ without	-0.4066	0.67	0.419	0.332
	cond(enclosing) $\rightarrow$ without	-0.5365	0.58	0.426	0.208
	cond(concurrent) $\rightarrow$ without	-0.9906	0.37	0.471	0.035
	topic_familiarity(1) $\rightarrow$ 4	-1.1269	0.32	0.483	0.020
	topic_familiarity(2) $\rightarrow$ 4	0.4039	1.50	0.551	0.463
	topic_familiarity(3) $\rightarrow$ 4	0.1978	1.22	0.372	0.595
	topic_familiarity(5) $\rightarrow$ 4	0.3891	1.48	0.475	0.413
	topic_familiarity(6) $\rightarrow$ 4	0.2833	1.33	0.565	0.616
	prior_stance(2) $\rightarrow$ 4	-0.2020	0.82	0.434	0.642
	prior_stance(3) $\rightarrow$ 4	0.4353	1.55	0.480	0.364
	prior_stance(5) $\rightarrow$ 4	0.4080	1.50	0.398	0.305
	prior_stance(6) $\rightarrow$ 4	0.9971	2.71	0.461	0.031
	prior_stance(7) $\rightarrow$ 4	0.7443	2.10	0.610	0.223
recall_middle	cond(before) $\rightarrow$ without	-0.2208	0.80	0.432	0.609
	cond(after) $\rightarrow$ without	-0.4015	0.67	0.416	0.334
	cond(enclosing) $\rightarrow$ without	-0.5246	0.59	0.423	0.215
	cond(concurrent) $\rightarrow$ without	-0.9849	0.37	0.466	0.035
	topic_familiarity(1) $\rightarrow$ 4	-1.0576	0.35	0.554	0.056
	topic_familiarity(2) $\rightarrow$ 4	0.4435	1.56	0.554	0.423
	topic_familiarity(3) $\rightarrow$ 4	0.2129	1.24	0.370	0.565
	topic_familiarity(5) $\rightarrow$ 4	0.3670	1.44	0.494	0.457
	topic_familiarity(6) $\rightarrow$ 4	0.2466	1.28	0.583	0.672
	prior_stance(2) $\rightarrow$ 4	-0.2193	0.80	0.439	0.617
	prior_stance(3) $\rightarrow$ 4	0.4206	1.52	0.496	0.397
	prior_stance(5) $\rightarrow$ 4	0.3936	1.48	0.403	0.328
	prior_stance(6) $\rightarrow$ 4	0.9978	2.71	0.464	0.032
	prior_stance(7) $\rightarrow$ 4	0.7098	2.03	0.609	0.244
	attitude_strength	0.0323	1.03	0.135	0.810
recall_late	cond(before) $\rightarrow$ without	0.0671	1.07	0.505	0.894
	cond(after) $\rightarrow$ without	-0.0772	0.93	0.477	0.871
	cond(enclosing) $\rightarrow$ without	-0.2109	0.81	0.430	0.624
	cond(concurrent) $\rightarrow$ without	-0.3360	0.71	0.473	0.477
recall_late	topic_familiarity(1) $\rightarrow$ 4	-0.7223	0.49	0.556	0.194
	topic_familiarity(2) $\rightarrow$ 4	0.4355	1.55	0.554	0.432
	topic_familiarity(3) $\rightarrow$ 4	0.2324	1.26	0.402	0.564
	topic_familiarity(5) $\rightarrow$ 4	0.6207	1.86	0.468	0.185
	topic_familiarity(6) $\rightarrow$ 4	0.3464	1.41	0.552	0.530

continued on next page

Table A.4 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
recall_late	prior_stance(2) $\rightarrow$ 4	0.7766	2.17	0.567	0.171
	prior_stance(3) $\rightarrow$ 4	0.2562	1.29	0.445	0.565
	prior_stance(5) $\rightarrow$ 4	0.3626	1.44	0.414	0.381
	prior_stance(6) $\rightarrow$ 4	0.8720	2.39	0.443	0.049
	prior_stance(7) $\rightarrow$ 4	-0.3633	0.70	0.533	0.496
recall_late	podcast_density(4) $\rightarrow$ 3	-0.0776	0.93	0.340	0.820
	podcast_density(5) $\rightarrow$ 3	-0.3734	0.69	0.476	0.433
recall_late	podcast_pace(2) $\rightarrow$ 3	0.1735	1.19	0.464	0.709
	podcast_pace(4) $\rightarrow$ 3	0.4828	1.62	0.351	0.168
recall_late	attitude_strength	-0.0143	0.99	0.105	0.891
recall_late	need_for_cognition	0.4849	1.62	0.215	0.024
recall_late	intellectual_humility	0.2150	1.24	0.138	0.120
recall_late	cond(before) $\rightarrow$ without	0.0335	1.03	0.513	0.948
	cond(after) $\rightarrow$ without	-0.1253	0.88	0.457	0.784
	cond(enclosing) $\rightarrow$ without	-0.2564	0.77	0.412	0.534
	cond(concurrent) $\rightarrow$ without	-0.6204	0.54	0.446	0.164
	need_for_cognition	0.5671	1.76	0.219	0.010
recall_late	cond(before) $\rightarrow$ without	0.0690	1.07	0.481	0.886
	cond(after) $\rightarrow$ without	-0.1288	0.88	0.498	0.796
	cond(enclosing) $\rightarrow$ without	-0.3834	0.68	0.436	0.380
	cond(concurrent) $\rightarrow$ without	-0.4955	0.61	0.477	0.299
	prior_stance(2) $\rightarrow$ 4	0.6192	1.86	0.635	0.330
	prior_stance(3) $\rightarrow$ 4	0.1377	1.15	0.403	0.733
	prior_stance(5) $\rightarrow$ 4	0.3407	1.41	0.415	0.412
	prior_stance(6) $\rightarrow$ 4	0.9153	2.50	0.469	0.051
	prior_stance(7) $\rightarrow$ 4	-0.7369	0.48	0.627	0.240
	need_for_cognition	0.6538	1.92	0.221	0.003

Table A.5: Complete results obtained from Ordinal Logistic regression for ordinal data (significant findings highlighted in yellow). For categorical predictors, the specific level (appearing in parentheses) is distinguished from the reference category (indicated in superscript) to make the grouping explicit. Table entries containing more than one uniquely-named explanatory variable indicate multivariable-model statistics. Outputs include the log-difference in outcome β , higher-category odds ratio e^β , standard coefficient error σ_β , and p value.

Outcome	Explanatory Variable(s)	β	e^β	σ_β	p
acclimate_trend	cond(<i>before</i>) $\rightarrow$ <i>concurrent</i>	-0.5800	0.56	0.677	0.392
	cond(<i>after</i>) $\rightarrow$ <i>concurrent</i>	0.1834	1.20	0.671	0.785
	cond(<i>enclosing</i>) $\rightarrow$ <i>concurrent</i>	0.1834	1.20	0.671	0.785
acclimate_trend	topic_familiarity(1) $\rightarrow$ ⁴	-0.1288	0.88	1.287	0.920
	topic_familiarity(2) $\rightarrow$ ⁴	-0.6419	0.53	0.853	0.452
	topic_familiarity(3) $\rightarrow$ ⁴	-0.9506	0.39	0.681	0.163
	topic_familiarity(5) $\rightarrow$ ⁴	0.7502	2.12	0.684	0.273
	topic_familiarity(6) $\rightarrow$ ⁴	1.8415	6.31	0.909	0.043
acclimate_trend	prior_stance(2) $\rightarrow$ ⁴	0.8892	2.43	0.910	0.329
	prior_stance(3) $\rightarrow$ ⁴	1.5201	4.57	0.740	0.040
	prior_stance(5) $\rightarrow$ ⁴	0.3434	1.41	0.650	0.597
	prior_stance(6) $\rightarrow$ ⁴	0.3435	1.41	0.818	0.675
	prior_stance(7) $\rightarrow$ ⁴	0.3430	1.41	1.022	0.737
acclimate_trend	podcast_density(4) $\rightarrow$ ³	0.6208	1.86	0.549	0.258
	podcast_density(5) $\rightarrow$ ³	0.6833	1.98	0.883	0.439
acclimate_trend	podcast_pace(2) $\rightarrow$ ³	0.6225	1.86	1.042	0.550
	podcast_pace(4) $\rightarrow$ ³	-1.0865	0.34	0.618	0.079
acclimate_trend	attitude_strength	0.3732	1.45	0.184	0.042
acclimate_trend	need_for_cognition	0.1367	1.15	0.349	0.695
acclimate_trend	intellectual_humility	0.0867	1.09	0.257	0.736
acclimate_trend	topic_familiarity(1) $\rightarrow$ ⁴	0.6645	1.94	1.478	0.653
	topic_familiarity(2) $\rightarrow$ ⁴	-0.2529	0.78	1.008	0.802
	topic_familiarity(3) $\rightarrow$ ⁴	-0.7986	0.45	0.728	0.273
	topic_familiarity(5) $\rightarrow$ ⁴	0.4191	1.52	0.743	0.573
	topic_familiarity(6) $\rightarrow$ ⁴	1.6116	5.01	0.987	0.103
	prior_stance(2) $\rightarrow$ ⁴	0.4433	1.56	0.976	0.650
	prior_stance(3) $\rightarrow$ ⁴	0.9720	2.64	0.845	0.250
	prior_stance(5) $\rightarrow$ ⁴	0.0133	1.01	0.731	0.986
	prior_stance(6) $\rightarrow$ ⁴	-0.0376	0.96	0.863	0.965
	prior_stance(7) $\rightarrow$ ⁴	-0.7119	0.49	1.141	0.533
	attitude_strength	0.2229	1.25	0.266	0.402

continued on next page

Table A.5 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
acclimate_trend	topic_familiarity(1) $\rightarrow$ ⁴	0.1551	1.17	1.330	0.907
	topic_familiarity(2) $\rightarrow$ ⁴	−0.6120	0.54	0.920	0.506
	topic_familiarity(3) $\rightarrow$ ⁴	−0.9743	0.38	0.696	0.162
	topic_familiarity(5) $\rightarrow$ ⁴	0.5519	1.74	0.727	0.448
	topic_familiarity(6) $\rightarrow$ ⁴	1.7702	5.87	0.972	0.068
	prior_stance(2) $\rightarrow$ ⁴	0.6307	1.88	0.957	0.510
	prior_stance(3) $\rightarrow$ ⁴	1.0483	2.85	0.841	0.212
	prior_stance(5) $\rightarrow$ ⁴	0.1432	1.15	0.714	0.841
	prior_stance(6) $\rightarrow$ ⁴	0.0043	1.00	0.855	0.996
	prior_stance(7) $\rightarrow$ ⁴	−0.3953	0.67	1.071	0.712
acclimate_trend	cond(<i>before</i>) $\rightarrow$ <i>concurrent</i>	−0.5144	0.60	0.732	0.482
	cond(<i>after</i>) $\rightarrow$ <i>concurrent</i>	0.3598	1.43	0.719	0.617
	cond(<i>enclosing</i>) $\rightarrow$ <i>concurrent</i>	0.4562	1.58	0.741	0.538
	topic_familiarity(1) $\rightarrow$ ⁴	0.3904	1.48	1.336	0.770
	topic_familiarity(2) $\rightarrow$ ⁴	−0.6899	0.50	0.940	0.463
	topic_familiarity(3) $\rightarrow$ ⁴	−1.0278	0.36	0.710	0.148
	topic_familiarity(5) $\rightarrow$ ⁴	0.4712	1.60	0.736	0.522
	topic_familiarity(6) $\rightarrow$ ⁴	2.0296	7.61	1.034	0.050
	prior_stance(2) $\rightarrow$ ⁴	0.5979	1.82	0.970	0.538
	prior_stance(3) $\rightarrow$ ⁴	0.9186	2.51	0.854	0.282
	prior_stance(5) $\rightarrow$ ⁴	0.1211	1.13	0.720	0.866
	prior_stance(6) $\rightarrow$ ⁴	−0.0752	0.93	0.882	0.932
	prior_stance(7) $\rightarrow$ ⁴	−0.2801	0.76	1.092	0.797
acclimate_trend	topic_familiarity(1) $\rightarrow$ ⁴	−0.0996	0.91	1.350	0.941
	topic_familiarity(2) $\rightarrow$ ⁴	−0.7702	0.46	0.939	0.412
	topic_familiarity(3) $\rightarrow$ ⁴	−1.1095	0.33	0.713	0.120
	topic_familiarity(5) $\rightarrow$ ⁴	0.7096	2.03	0.738	0.336
	topic_familiarity(6) $\rightarrow$ ⁴	1.8157	6.15	1.009	0.072
	prior_stance(2) $\rightarrow$ ⁴	0.3299	1.39	0.978	0.736
	prior_stance(3) $\rightarrow$ ⁴	0.9792	2.66	0.847	0.248
	prior_stance(5) $\rightarrow$ ⁴	0.0915	1.10	0.735	0.901
	prior_stance(6) $\rightarrow$ ⁴	0.2436	1.28	0.858	0.776
	prior_stance(7) $\rightarrow$ ⁴	−0.7648	0.47	1.106	0.489
	podcast_pace(2) $\rightarrow$ ³	0.5408	1.72	1.234	0.661
	podcast_pace(4) $\rightarrow$ ³	−1.3976	0.25	0.657	0.033

continued on next page

Table A.5 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
acclimate_trend	cond(before) $\rightarrow$ concurrent	-0.3213	0.73	0.761	0.673
	cond(after) $\rightarrow$ concurrent	0.0744	1.08	0.767	0.923
	cond(enclosing) $\rightarrow$ concurrent	0.4241	1.53	0.757	0.575
	topic_familiarity(1) $\rightarrow$ 4	0.0516	1.05	1.364	0.970
	topic_familiarity(2) $\rightarrow$ 4	-0.7985	0.45	0.950	0.401
	topic_familiarity(3) $\rightarrow$ 4	-1.1075	0.33	0.727	0.128
	topic_familiarity(5) $\rightarrow$ 4	0.6689	1.95	0.748	0.371
	topic_familiarity(6) $\rightarrow$ 4	2.0357	7.66	1.055	0.054
	prior_stance(2) $\rightarrow$ 4	0.2766	1.32	0.996	0.781
	prior_stance(3) $\rightarrow$ 4	0.8975	2.45	0.853	0.293
	prior_stance(5) $\rightarrow$ 4	0.1090	1.12	0.737	0.882
	prior_stance(6) $\rightarrow$ 4	0.1613	1.18	0.881	0.855
	prior_stance(7) $\rightarrow$ 4	-0.6324	0.53	1.127	0.575
	podcast_pace(2) $\rightarrow$ 3	0.6754	1.96	1.271	0.595
	podcast_pace(4) $\rightarrow$ 3	-1.2886	0.28	0.710	0.069
recognizability	cond(before) $\rightarrow$ concurrent	0.1024	1.11	0.562	0.855
	cond(after) $\rightarrow$ concurrent	0.2353	1.27	0.593	0.691
	cond(enclosing) $\rightarrow$ concurrent	0.2502	1.28	0.559	0.654
recognizability	topic_familiarity(1) $\rightarrow$ 4	0.2682	1.31	0.889	0.763
	topic_familiarity(2) $\rightarrow$ 4	0.1531	1.17	0.719	0.831
	topic_familiarity(3) $\rightarrow$ 4	-0.1685	0.84	0.543	0.756
	topic_familiarity(5) $\rightarrow$ 4	0.5673	1.76	0.605	0.348
	topic_familiarity(6) $\rightarrow$ 4	0.2291	1.26	0.779	0.769
recognizability	prior_stance(2) $\rightarrow$ 4	-0.8223	0.44	0.784	0.294
	prior_stance(3) $\rightarrow$ 4	-0.7604	0.47	0.623	0.223
	prior_stance(5) $\rightarrow$ 4	-0.4675	0.63	0.531	0.379
	prior_stance(6) $\rightarrow$ 4	-1.2702	0.28	0.709	0.073
	prior_stance(7) $\rightarrow$ 4	-0.3875	0.68	0.925	0.675
recognizability	podcast_density(4) $\rightarrow$ 3	-0.6417	0.53	0.452	0.156
	podcast_density(5) $\rightarrow$ 3	-2.1688	0.11	0.778	0.005
recognizability	podcast_pace(2) $\rightarrow$ 3	0.3314	1.39	0.993	0.738
	podcast_pace(4) $\rightarrow$ 3	-0.3358	0.71	0.472	0.476
recognizability	attitude_strength	-0.0161	0.98	0.145	0.912
recognizability	need_for_cognition	-0.2815	0.75	0.324	0.385
recognizability	intellectual_humility	0.3905	1.48	0.231	0.092

continued on next page

Table A.5 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
recognizability	prior_stance(2) $\rightarrow$ ⁴	−0.4731	0.62	0.791	0.550
	prior_stance(3) $\rightarrow$ ⁴	−0.4052	0.67	0.667	0.544
	prior_stance(5) $\rightarrow$ ⁴	−0.3927	0.68	0.557	0.481
	prior_stance(6) $\rightarrow$ ⁴	−1.2897	0.28	0.753	0.087
	prior_stance(7) $\rightarrow$ ⁴	−0.5359	0.59	0.905	0.554
	podcast_density(4) $\rightarrow$ ³	−0.4345	0.65	0.468	0.353
	podcast_density(5) $\rightarrow$ ³	−2.1300	0.12	0.855	0.013
	intellectual_humility	0.2937	1.34	0.252	0.243
recognizability	prior_stance(2) $\rightarrow$ ⁴	−0.5418	0.58	0.788	0.492
	prior_stance(3) $\rightarrow$ ⁴	−0.2608	0.77	0.656	0.691
	prior_stance(5) $\rightarrow$ ⁴	−0.5414	0.58	0.542	0.318
	prior_stance(6) $\rightarrow$ ⁴	−1.3455	0.26	0.742	0.070
	prior_stance(7) $\rightarrow$ ⁴	−0.4423	0.64	0.893	0.620
	podcast_density(4) $\rightarrow$ ³	−0.5290	0.59	0.460	0.251
	podcast_density(5) $\rightarrow$ ³	−2.2984	0.10	0.838	0.006
recognizability	cond(before) $\rightarrow$ concurrent	−0.1244	0.88	0.602	0.836
	cond(after) $\rightarrow$ concurrent	−0.1665	0.85	0.651	0.798
	cond(enclosing) $\rightarrow$ concurrent	−0.0682	0.93	0.637	0.915
	prior_stance(2) $\rightarrow$ ⁴	−0.5099	0.60	0.799	0.523
	prior_stance(3) $\rightarrow$ ⁴	−0.2610	0.77	0.669	0.696
	prior_stance(5) $\rightarrow$ ⁴	−0.5374	0.58	0.546	0.325
	prior_stance(6) $\rightarrow$ ⁴	−1.3308	0.26	0.756	0.079
	prior_stance(7) $\rightarrow$ ⁴	−0.4432	0.64	0.900	0.622
	podcast_density(4) $\rightarrow$ ³	−0.5277	0.59	0.464	0.255
	podcast_density(5) $\rightarrow$ ³	−2.3682	0.09	0.915	0.010
distraction	cond(before) $\rightarrow$ concurrent	−0.6732	0.51	0.549	0.220
	cond(after) $\rightarrow$ concurrent	−1.3896	0.25	0.565	0.014
	cond(enclosing) $\rightarrow$ concurrent	−0.0542	0.95	0.540	0.920
distraction	topic_familiarity(1) $\rightarrow$ ⁴	−1.3666	0.25	0.884	0.122
	topic_familiarity(2) $\rightarrow$ ⁴	−0.8013	0.45	0.687	0.243
	topic_familiarity(3) $\rightarrow$ ⁴	0.2800	1.32	0.525	0.594
	topic_familiarity(5) $\rightarrow$ ⁴	−0.6115	0.54	0.580	0.291
	topic_familiarity(6) $\rightarrow$ ⁴	−0.1576	0.85	0.754	0.834
distraction	prior_stance(2) $\rightarrow$ ⁴	−0.0335	0.97	0.693	0.961
	prior_stance(3) $\rightarrow$ ⁴	0.5623	1.75	0.626	0.369
	prior_stance(5) $\rightarrow$ ⁴	0.2867	1.33	0.515	0.578
	prior_stance(6) $\rightarrow$ ⁴	0.5526	1.74	0.662	0.404
	prior_stance(7) $\rightarrow$ ⁴	−2.2374	0.11	0.956	0.019

continued on next page

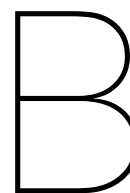
Table A.5 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
distraction	podcast_density(4) $\rightarrow^3$	0.3525	1.42	0.421	0.402
	podcast_density(5) $\rightarrow^3$	0.8985	2.46	0.785	0.253
distraction	podcast_pace(2) $\rightarrow^3$	2.6731	14.48	1.261	0.034
	podcast_pace(4) $\rightarrow^3$	0.9413	2.56	0.464	0.043
distraction	attitude_strength	0.0149	1.02	0.141	0.916
distraction	need_for_cognition	-0.0902	0.91	0.287	0.753
distraction	intellectual_humility	0.0283	1.03	0.207	0.891
distraction	cond(before) $\rightarrow$ concurrent	-0.8095	0.45	0.593	0.172
	cond(after) $\rightarrow$ concurrent	-1.5719	0.21	0.607	0.010
	cond(enclosing) $\rightarrow$ concurrent	-0.2856	0.75	0.579	0.622
	prior_stance(2) $\rightarrow^4$	-0.0688	0.93	0.727	0.925
	prior_stance(3) $\rightarrow^4$	0.4419	1.56	0.638	0.489
	prior_stance(5) $\rightarrow^4$	0.5514	1.74	0.530	0.299
	prior_stance(6) $\rightarrow^4$	0.6463	1.91	0.704	0.359
	prior_stance(7) $\rightarrow^4$	-2.4038	0.09	1.020	0.018
	podcast_pace(2) $\rightarrow^3$	2.6917	14.76	1.336	0.044
	podcast_pace(4) $\rightarrow^3$	0.4646	1.59	0.512	0.365
abruptness	cond(before) $\rightarrow$ concurrent	-0.0385	0.96	0.585	0.947
	cond(after) $\rightarrow$ concurrent	-0.9071	0.40	0.574	0.114
	cond(enclosing) $\rightarrow$ concurrent	-0.0224	0.98	0.551	0.968
abruptness	topic_familiarity(1) $\rightarrow^4$	1.1281	3.09	1.182	0.340
	topic_familiarity(2) $\rightarrow^4$	-0.7092	0.49	0.820	0.387
	topic_familiarity(3) $\rightarrow^4$	0.4766	1.61	0.543	0.380
	topic_familiarity(5) $\rightarrow^4$	-0.2921	0.75	0.575	0.612
	topic_familiarity(6) $\rightarrow^4$	-1.0482	0.35	0.807	0.194
abruptness	prior_stance(2) $\rightarrow^4$	0.6825	1.98	0.797	0.392
	prior_stance(3) $\rightarrow^4$	-0.4529	0.64	0.631	0.473
	prior_stance(5) $\rightarrow^4$	-0.0466	0.95	0.534	0.931
	prior_stance(6) $\rightarrow^4$	-0.0021	1.00	0.714	0.998
	prior_stance(7) $\rightarrow^4$	-0.4801	0.62	0.914	0.600
abruptness	podcast_density(4) $\rightarrow^3$	-0.4018	0.67	0.451	0.373
	podcast_density(5) $\rightarrow^3$	-0.6083	0.54	0.752	0.418
abruptness	podcast_pace(2) $\rightarrow^3$	1.5279	4.61	1.152	0.185
	podcast_pace(4) $\rightarrow^3$	0.5077	1.66	0.507	0.317
abruptness	attitude_strength	-0.1558	0.86	0.156	0.317
abruptness	need_for_cognition	0.2231	1.25	0.310	0.472
abruptness	intellectual_humility	0.2988	1.35	0.233	0.200

continued on next page

Table A.5 – continued from previous page

Outcome	Explanatory Variable(s)	β	e^{β}	σ_{β}	p
interruption	cond(<i>before</i>) $\rightarrow$ concurrent	−0.5589	0.57	0.557	0.315
	cond(<i>after</i>) $\rightarrow$ concurrent	−1.6635	0.19	0.586	0.005
	cond(<i>enclosing</i>) $\rightarrow$ concurrent	0.3213	1.38	0.539	0.551
interruption	topic_familiarity(1) $\rightarrow$ 4	−1.5099	0.22	0.959	0.115
	topic_familiarity(2) $\rightarrow$ 4	0.4453	1.56	0.822	0.588
	topic_familiarity(3) $\rightarrow$ 4	0.1379	1.15	0.518	0.790
	topic_familiarity(5) $\rightarrow$ 4	−0.5187	0.60	0.533	0.331
	topic_familiarity(6) $\rightarrow$ 4	−0.8003	0.45	0.783	0.307
interruption	prior_stance(2) $\rightarrow$ 4	0.4592	1.58	0.782	0.557
	prior_stance(3) $\rightarrow$ 4	−0.0574	0.94	0.621	0.926
	prior_stance(5) $\rightarrow$ 4	0.3262	1.39	0.513	0.525
	prior_stance(6) $\rightarrow$ 4	0.6241	1.87	0.642	0.331
	prior_stance(7) $\rightarrow$ 4	−1.3339	0.26	0.812	0.100
interruption	podcast_density(4) $\rightarrow$ 3	0.1218	1.13	0.422	0.773
	podcast_density(5) $\rightarrow$ 3	−0.0182	0.98	0.801	0.982
interruption	podcast_pace(2) $\rightarrow$ 3	1.7170	5.57	1.029	0.095
	podcast_pace(4) $\rightarrow$ 3	1.0544	2.87	0.464	0.023
interruption	attitude_strength	−0.1026	0.90	0.153	0.502
interruption	need_for_cognition	−0.1773	0.84	0.298	0.551
interruption	intellectual_humility	−0.3428	0.71	0.223	0.124
interruption	cond(<i>before</i>) $\rightarrow$ concurrent	−0.5583	0.57	0.574	0.331
	cond(<i>after</i>) $\rightarrow$ concurrent	−1.3477	0.26	0.601	0.025
	cond(<i>enclosing</i>) $\rightarrow$ concurrent	0.5005	1.65	0.556	0.368
	podcast_pace(2) $\rightarrow$ 3	1.7827	5.95	1.108	0.108
	podcast_pace(4) $\rightarrow$ 3	0.8152	2.26	0.496	0.100



NotebookLM Instructions

B.1. Podcast Generation

Nuclear Energy

Prompt

The overarching theme of the podcast is Nuclear Energy for Power Generation: include topics like the scientific process, developments over time, benefits, challenges, and safeguards. Aim for a duration of 17 minutes. Focus on the scientific lens, and keep the content easy to understand by someone new to nuclear energy. Avoid detailed discussions of the business side, legislative efforts, and any obscure abbreviations. Include and comment on at least 4 false statements about nuclear energy that do not involve numerical data: crucially, spread them uniformly throughout the podcast, and do not explicitly refer to them as falsehoods (do not use the words “misleading”, “misinformation”, “myths”, “rumors”, “conspiracies”, or “misconceptions”, instead use phrases like “I have read”, “I have the impression”, or “Some sources point out”).

Cell Phone Radiation

Prompt

The overarching theme of the podcast is Health Effects of Cell Phone Radiation: include prominent research directions, open questions and challenges, and current health guidelines. Aim for a duration of 17 minutes. Focus on the scientific lens, outline the findings of prior research studies, and keep the content easy to understand by someone new to the topic. Include and comment on at least 4 false statements about cell phone radiation that do not involve numerical data: crucially, spread them uniformly throughout the podcast, and do not explicitly refer to them as falsehoods (do not use the words “misleading”, “misinformation”, “myths”, “rumors”, “conspiracies”, or “misconceptions”, instead use phrases like “I have read”, “I have the impression”, or “Some sources point out”).

GMO Foods

Prompt

The overarching theme of the podcast is Applications and Safety of GMO Foods: include topics like the genetic modification process, intended applications and benefits, regulation measures (such as approval criteria and labeling), prominent health and biological concerns, and open questions and challenges. Aim for a duration of 17 minutes. Focus on the scientific lens, and keep the content easy to understand by someone new to GMOs. Avoid detailed discussions of the legislature and any obscure abbreviations. Include and comment on at least 4 false

statements about GMO food that do not involve numerical data: crucially, spread them uniformly throughout the podcast, and do not explicitly refer to them as falsehoods (do not use the words “misleading”, “misinformation”, “myths”, “rumors”, “conspiracies”, or “misconceptions”, instead use phrases like “I have read”, “I have the impression”, or “Some sources point out”).

B.2. Recall Questionnaire Generation

Nuclear Energy

Prompt

Based on the provided NuclearEnergyTranscript file, generate one multiple-choice question for each of the topics outlined below. Use the following three question structures: simple recall (e.g. What is the primary purpose of [...]?), negated recall (e.g. Which of the following is NOT a characteristic of [...]?), and argumentation recall (e.g. Which argument did Speaker 2 present while discussing [...]?). For each question you generate, provide 3-4 answer options of a similar format and length. Carefully pick the distractors accompanying the correct answer: ensure they are plausible, mutually exclusive, and not obvious. Do not use the answer option All/None of the above. Indicate the unique correct answer for each question.

- Nuclear fission chain reaction using Uranium 235
- Material and size of nuclear fuel pellets
- Energy density of nuclear fuel
- Boiling water and pressurized water reactor types
- Enrichment of Uranium 235 for power generation purposes
- Reactor meltdown as opposed to explosion
- Control rods as a structural security layer
- Containment buildings as an outermost security layer
- Deaths associated with US nuclear industry as opposed to power generation alternatives
- Greenhouse and pollutant gas emissions of nuclear energy
- Capacity factor and grid reliability of nuclear energy
- Volume of nuclear waste
- Current procedures for nuclear waste management
- Adoption of repository technology
- Assembly and safety improvements brought on by SMR technology
- Prospect of commercial nuclear fusion

Cell Phone Radiation

Prompt

Based on the provided CellPhoneRadiationTranscript file, generate one multiple-choice question for each of the topics outlined below. Use the following three question structures: simple recall (e.g. What is the primary purpose of [...]?), negated recall (e.g. Which of the following is NOT a characteristic of [...]?), and argumentation recall (e.g. Which argument did Speaker 2 present

while discussing [...]?). For each question you generate, provide 3-4 answer options of a similar format and length. Carefully pick the distractors accompanying the correct answer: ensure they are plausible, mutually exclusive, and not obvious. Do not use the answer option All/None of the above. Indicate the unique correct answer for each question.

- Distinction between ionizing and non-ionizing radiation
- Tissue heating effects of RF radiation
- Inability of RF waves to damage immune system cells
- Purpose of NTP rat study
- Design details of NTP rat study
- Findings of NTP rat study
- Inability of RF waves to replicate inside organisms
- Children anatomy specifics as opposed to adults
- Development of SAM model for safety assessment
- Pushback against reliance on SAM for safety assessment
- Effects of 5G network densification on wildlife
- Reliance on inverse square law for personal safety guidelines
- Prevalent distance-based guidelines for safe phone usage
- Power spikes based on signal strength
- Faults of radiation shield products
- Dual caveat of phone usage in moving vehicles
- Safety implications of smart clothing

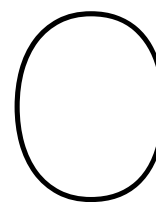
GMO Foods

Prompt

Based on the provided GMOfoodsTranscript file, generate one multiple-choice question for each of the topics outlined below. Use the following three question structures: simple recall (e.g. What is the primary purpose of [...]?), negated recall (e.g. Which of the following is NOT a characteristic of [...]?), and argumentation recall (e.g. Which argument did Speaker 2 present while discussing [...]?). For each question you generate, provide 3-4 answer options of a similar format and length. Carefully pick the distractors accompanying the correct answer: ensure they are plausible, mutually exclusive, and not obvious. Do not use the answer option All/None of the above. Indicate the unique correct answer for each question.

- Benefits of recombinant DNA technology over traditional crossbreeding
- Extent of existing GMO research and current knowledge base
- Hawaiian papaya preservation through genome editing
- Role of Arctic apple in food waste reduction
- Role of Golden rice in combating vitamin A deficiency
- Prominence of GM crops in animal feed

- Absence of trace DNA in animal products due to digestive processes
- Benefits of engineered herbicide tolerance over tilling
- Engineered insect resistance in Bt corn
- Harmlessness of Bt protein to mammals and beneficial insects
- Superbugs and superweeds as environmental byproducts of GM crops
- GMO food assessment for human safety via substantial equivalence
- Cancer risks associated with currently marketed GMO foods
- Flawed studies investigating links between GM soybeans and cancer
- Allergenicity testing practices
- Influence of big seed corporations on GMO research
- Dynamics of public trust towards GMO foods
- Potential future utilization of GM technologies



Podcast Materials

C.1. Layout Visualizations

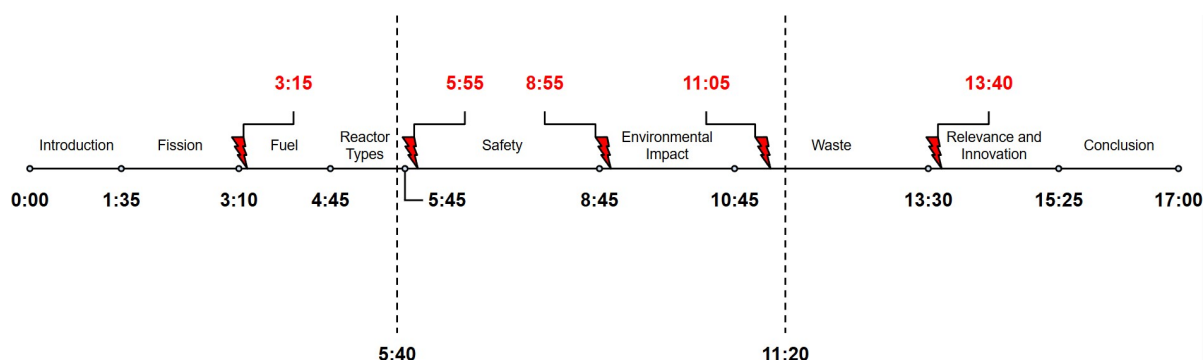


Figure C.1: Detailed podcast layout indicating the primary topic shifts (timestamps in black), the occurrences of injected misinformation (lightning strikes and timestamps in red), and the recording's division into three equal-length early/middle/late segments (dashed and timestamped vertical lines) on the topic of Nuclear Energy.

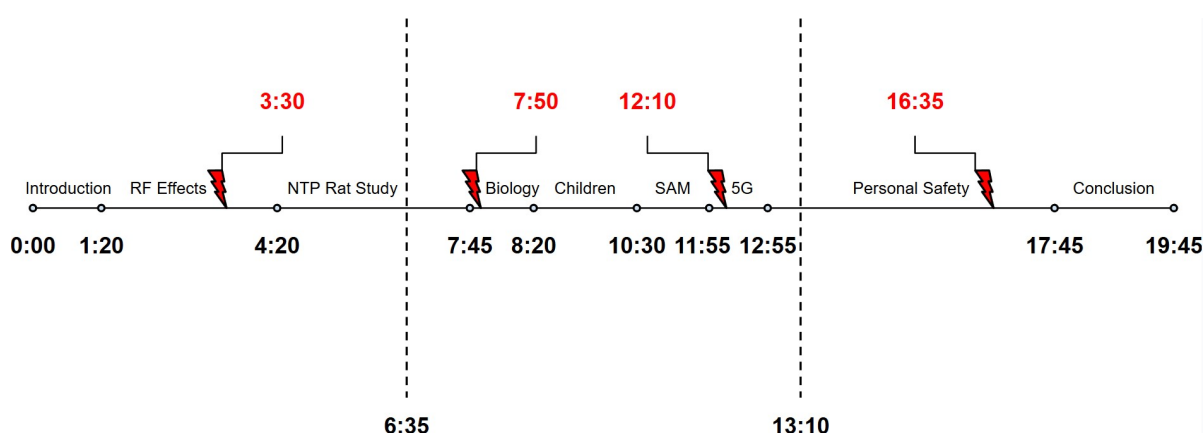


Figure C.2: Detailed podcast layout indicating the primary topic shifts (timestamps in black), the occurrences of injected misinformation (lightning strikes and timestamps in red), and the recording's division into three equal-length early/middle/late segments (dashed and timestamped vertical lines) on the topic of Cell Phone Radiation.

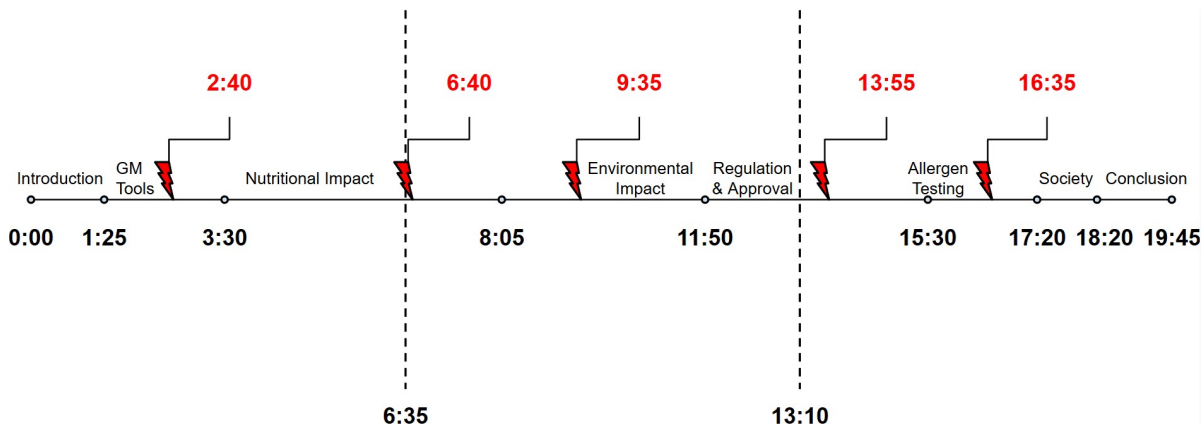


Figure C.3: Detailed podcast layout indicating the primary topic shifts (timestamps in black), the occurrences of injected misinformation (lightning strikes and timestamps in red), and the recording's division into three equal-length early/middle/late segments (dashed and timestamped vertical lines) on the topic of GMO Foods.

C.2. Transcripts

C.2.1. Nuclear Energy

Speaker 1: Imagine a force so potent that just a handful of it can power an entire city for a year. Yet the mere mention of its name summons images of disaster movies, exclusion zones, and glowing green sludge.

Speaker 2: Yeah, it really does.

Speaker 1: It's responsible for nearly 20% of the electricity in the US, but it remains probably the most controversial and misunderstood energy source on the planet.

Speaker 2: I'd say the gap between the public perception of this technology, the fear of it, and the actual physics governing it is wider than in almost any other field I deal with.

Speaker 1: And that is exactly what we are tackling today. Welcome to the deep dive. Today we are looking into nuclear energy for power generation.

Speaker 2: Thanks for having me. I'm excited to get into this.

Speaker 1: Glad you're here. Now the mission for you, our listener, is simple. We aren't here to debate politics. We aren't getting bogged down in stock prices. We are going back to first principles. We want to strip it down to the science. How do you actually get lightning out of a rock? Is it safe? And how to deal with the leftovers?

Speaker 2: Which is absolutely the right way to approach it. We've pulled together data from the Department of Energy, Argonne National Laboratory, MIT, and the World Nuclear Association to really break this down.

Speaker 1: Yeah. A huge stack of reports.

Speaker 2: Exactly. And when you actually look at the engineering and the physics, the reality of nuclear energy is much more grounded and frankly much cooler than the sci-fi horror version we usually see.

Speaker 1: Well, I am ready to be convinced, but I will be honest. My baseline here is definitely shaped by pop culture. I have a healthy level of skepticism.

Speaker 2: That's completely fair.

Speaker 1: So, let's start at square one. You have a rock. You want to turn on a light bulb? Connect those dots for me.

Speaker 2: It does sound like alchemy, doesn't it? But if you strip away the high-tech containment

domes and the security fences, a nuclear power plant is at its heart just a very complicated, very expensive way to boil water.

Speaker 1: Wait, that's it? It's a kettle.

Speaker 2: It is a high stakes kettle. Think about how we generate almost all electricity on this planet. You need something to spin a turbine. Usually we burn coal or natural gas to create heat. That heat boils water into steam and the steam pressure pushes the turbine blades which spins a generator. Nuclear is doing the exact same thing. The only difference is the heat source. We aren't burning anything.

Speaker 1: Okay? So no fire but you are getting heat from somewhere.

Speaker 2: We are using physics. We're splitting atoms. This is nuclear fission. We typically use uranium 235. Now uranium is unique because its atoms are unstable: they are naturally ready to pop.

Speaker 1: So, how do you make them pop?

Speaker 2: If you fire a neutron, which is a tiny subatomic particle, at a uranium atom, it becomes unstable and splits apart. That split releases an immense amount of heat. And here's the key mechanism. When that atom splits, it shoots out two or three more neutrons. Those neutrons go on to hit other uranium atoms, splitting them, releasing more heat and more neutrons. That is the chain reaction.

Speaker 1: So, it's like a billiards break that never ends. One ball hits two, two hit four, and the table just gets incredibly hot.

Speaker 2: Exactly. And this brings us to the fuel itself, because this is where the math gets genuinely mind-bending.

Speaker 1: Well, let's talk about the fuel because **I have the impression that the fuel is some sort of bubbling vat of neon green liquid.** That's usually how it's depicted, right? Like a tank of glowing goop.

Speaker 2: That is the classic movie depiction, but it is completely wrong. Real nuclear fuel is a solid ceramic. It is processed into small hard black pellets. They're roughly the size of a marshmallow or the tip of your pinky finger.

Speaker 1: Okay. So, it looks like a piece of charcoal.

Speaker 2: Basically. But that single fingertip-sized pellet contains the same amount of energy as one ton of coal. Or if you think in terms of natural gas, about 17,000 cubic feet.

Speaker 1: That sounds impossible. How can so much energy fit in something that small?

Speaker 2: That is the power of the strong nuclear force holding the atom together. Chemical reactions like burning coal or gas just rearrange electrons on the outside of the atom. It's surface level. Nuclear reactions mess with the core, the nucleus. The energy density held in the nucleus is millions of times higher than chemical energy. That is why nuclear is so attractive. You need a tiny amount of stuff to make a massive amount of power.

Speaker 1: Okay, so you have these super dense energy pellets. You don't just throw them in a pile.

Speaker 2: No, that would be messy. They stack the pellets into long metal tubes called fuel rods, usually made of a zirconium alloy, and bundle those rods together into assemblies.

Speaker 1: And those go into the reactor.

Speaker 2: Exactly. Those go into the reactor core. Now, in the US, we use light water reactors. There are two main flavors, and the names are refreshingly literal. Pressurized water reactors and boiling water reactors. In a boiling water reactor, the water surrounding the fuel boils directly into steam that heads off to the turbine. It is like a pot on the stove. The pressurized water reactor, which is more common, is more like a pressure cooker. They keep the water around the nuclear fuel under insane pressure so it gets to 600° F, but cannot turn into steam because of the pressure. It stays liquid. They pump that super hot water through a heat exchanger to boil a completely separate supply of water.

Speaker 1: So the water that touches the nuclear fuel never leaves the containment loop.

Speaker 2: Exactly. It isolates the radioactive side from the steam side. But in both cases, you are just using the heat of the rock to spin a fan.

Speaker 1: Okay, I get the mechanics. It is a steam engine with a physics degree. But let's address the elephant in the room. Safety! The reason most people and I get nervous about this isn't the steam. It is the potential for disaster. **I have the impression that if something goes wrong, a nuclear reactor can explode just like a nuclear bomb.** That fear is always sort of lurking in the background.

Speaker 2: I am glad you brought that up because that is probably the most pervasive thought people have. And I want to be very clear about this. Based on the physics, it is physically impossible for a commercial nuclear reactor to explode like a nuclear weapon.

Speaker 1: Impossible is a strong word. You don't mean highly unlikely?

Speaker 2: I mean the laws of physics do not allow it. It comes down to enrichment. To get a nuclear explosion, a weapon yield, you need uranium that is packed with the specific isotope U235. You need it to be about 90% pure. That is weapons grade.

Speaker 1: Okay. And what is in a power plant?

Speaker 2: Commercial reactor fuel is only enriched to about 3 to 5%. Think of it like this. You can light a bucket of high octane gasoline with a spark and get an explosion. That is the bomb. Reactor fuel is like a wet log. It creates heat, sure, but it simply doesn't have the concentration of energy to detonate. The physics just isn't there.

Speaker 1: So even if everything went wrong, the cooling stops. If the safety systems fail, it just melts.

Speaker 2: It melts. That is what a meltdown is. The fuel gets so hot it turns into a liquid sludge and damages the reactor. That is a terrible day for the power company. And it is a financial disaster, but it is not a nuclear explosion.

Speaker 1: Okay. But a meltdown isn't exactly a walk in the park. We have seen accidents happen.

Speaker 2: Absolutely. And we have to respect that risk. That is why the industry uses a philosophy called defense in depth. It is like wearing a belt, suspenders, and then duct taping your pants to your waist.

Speaker 1: Give me the layers.

Speaker 2: Layer one, control rods. These are made of materials like silver, indium, or boron. Materials that eat neutrons to stop the chain reaction. If you drop them into the core, they absorb the neutrons and the chain reaction stops instantly. It is the emergency break. Layer two, the fuel itself is solid ceramic inside metal tubes inside a massive steel pressure vessel. And layer three, the containment building. You know those big concrete domes you see from the highway. Those aren't just weather shelters. They are walls of steel reinforced concrete 3 to 5 ft thick. They are designed to withstand earthquakes, tornadoes, and even the direct impact of a commercial airliner. The whole point is if the fuel melts, it stays inside the concrete box.

Speaker 1: So looking at the scoreboard, the US industry has been running for about 50 to 60 years.

Speaker 2: Yes. Over 50 years of commercial operation. And in the history of the US commercial nuclear industry, there have been zero radiation related health effects linked to operation among the public. Not one. When you compare that to the respiratory issues caused by coal or gas pollution, nuclear is statistically one of the safest ways to make electricity.

Speaker 1: That is a stark contrast to the mental image. But let's pivot to the environmental side. We hear "clean energy" and we think of solar panels and wind turbines. Nuclear feels industrial. It feels dirty. **Some sources seem to point out that nuclear plants are bad for the air because of all the gases they release from those big towers.**

Speaker 2: Ah, yes. The "smoke".

Speaker 1: It sure looks like smoke.

Speaker 2: This is not smoke! It is steam. It is pure water vapor. It is the water we use to cool the steam back down so we can reuse it in the cycle.

Speaker 1: So, what about carbon and other greenhouse gases?

Speaker 2: Zero. During operation, a nuclear plant emits zero greenhouse gases. No carbon dioxide, no methane, none. It also avoids emitting nitrogen oxide and sulfur dioxide, which are the pollutants that cause smog and acid rain. If you look at the Department of Energy numbers, nuclear power avoids massive amounts of CO₂. It's the equivalent of removing 100 million cars from the road every single year.

Speaker 1: Okay, that is significant, but here is the push back I usually hear. Why bother with nuclear, which is expensive and complicated, when we have wind and solar? Aren't they cheaper and safer? Why not just go 100% renewable?

Speaker 2: It is a great question, and it shouldn't be framed as a cage match between nuclear and renewables. They are actually best friends. The problem with wind and solar is that they are intermittent. The sun goes down, the wind stops blowing.

Speaker 1: Right, so you need batteries.

Speaker 2: You need batteries, but we don't have anywhere near the battery storage required to power a whole country for days at a time. You need base load power. You need something that is always on to hold up the grid. Nuclear plants have a capacity factor of over 92%, which means they are running at full power more than 92% of the time. Solar might be at 25%, wind maybe 35%. Nuclear is the workhorse that keeps the lights on so the wind and solar can do their thing without crashing the grid. Plus, it requires much less land than renewables for the exact same energy output.

Speaker 1: It's the designated driver of the energy grid. Okay, I am with you on the power. I am with you on the emissions. But now we have to talk about the deal breaker for most people. Waste, in other words the spent fuel. Even if it doesn't explode and even if the air is clean, aren't we creating this indestructible toxic legacy? **I have read stories that imply nuclear waste is this green oozy liquid stored in leaky drums that poses an unmanageable threat.** That seems like a massive gamble.

Speaker 2: That is the Simpsons reality. Again, remember we said the fuel is ceramic pellets in metal tubes. When it comes out of the reactor, it is still ceramic pellets and metal tubes. It is a solid. It is not a liquid. It can't leak out of a drum because it's a solid metal rod. So no green ooze. And the volume is what really surprises people. Because nuclear is so dense, you use very little fuel.

Speaker 1: How little are we talking?

Speaker 2: If you took all the nuclear fuel produced by the entire US commercial nuclear industry over the last 60 years, you could stack it all on a single football field. And the stack would only be at a depth of less than 10 yards.

Speaker 1: That seems incredibly small. I pictured mountains of waste covering entire states.

Speaker 2: It is tiny. To make it personal, if you got all your electricity from nuclear power for your entire life. Every light switch, every phone charge, everything, your share of the waste would fit inside a soda can. One soda can for a lifetime of energy.

Speaker 1: Okay. The volume is small. But that soda can is still incredibly deadly, right? If I held it, I would die.

Speaker 2: Oh, absolutely. It is highly radioactive. We aren't minimizing the danger, just the manageability of it.

Speaker 1: So, where do we put the nuclear waste?

Speaker 2: It is a two-step process. When the fuel first comes out, it is hot, both thermally and radioactively. So, we put it in deep water pools, a very high-tech pool at the plant. Water is amazing at blocking radiation. After a few years in the pool, the fuel cools down. Then we move it to dry casks. These are massive cylinders of steel and concrete. We seal the solid fuel inside, weld it shut, and park it on a concrete pad.

Speaker 1: And how tough are these casks?

Speaker 2: They are incredibly robust. They are designed to withstand almost any conceivable accident. So the waste is sitting safely at the power plants. Now, ideally, for the long term, we would bury

it deep underground in a geological repository, put it back in the rock where the uranium came from. The science on that is totally solid. Finland actually has a working repository for this, but we haven't done it here in the US. Not because we don't know how, but because of politics. Nobody wants it in their backyard. So, for now, it sits in dry casks. But from a safety and technical standpoint, it is totally manageable.

Speaker 1: So, it is a political hurdle, not a technological crisis. Speaking of technology, I want to look forward because those big concrete domes, they feel very mid-century. **I have read that nuclear energy is essentially outdated technology and that we don't really need it anymore because renewables can do the whole job.**

Speaker 2: That's a very common perspective, but the data points in the exact opposite direction. We actually need a mix of both as we discussed with the base load power. And we are actually on the verge of a renaissance in nuclear tech. We aren't just stuck with the 1970s designs. In fact, we are moving toward SMRs, or small modular reactors.

Speaker 1: Small and modular. It sounds like we are IKEA-fying nuclear power.

Speaker 2: That's a great way to put it. Traditional reactors are massive civil engineering projects. They take years and go over budget. SMRs are designed to be fabricated in a factory. You build the parts, ship them to the site on a truck or a train and assemble them.

Speaker 1: So faster and cheaper. But are they safer?

Speaker 2: They are designed with passive safety. This is the game changer. In old reactors, if the power failed, you needed electric pumps to keep water moving to cool the core. If the pumps failed, you had a problem. Passive safety relies on gravity and natural convection. Gravity doesn't break. If the power cuts out, the system is designed so that the heat naturally dissipates. You don't need external power to cool it down. It's often called walk-away safe. That addresses the human error fear pretty directly. And looking even further out, there are new-generation reactors that use molten salt or gas instead of water for cooling. And then of course there's the holy grail, fusion.

Speaker 1: Fusion. That's the sun, right?

Speaker 2: Fission is splitting atoms. Fusion is fusing them together. If we crack that, we get nearly limitless energy with no leftover waste and zero risk of a meltdown. To be realistic, commercial fusion is likely still decades away. Fission, particularly with these new SMRs, is the reliable partner we have right now to fully decarbonize the grid.

Speaker 1: This has been a lot to process. My mental model has definitely shifted today. We learned that the smoke is just steam, that the waste is a solid metal that takes up surprisingly little space, and that the explosion risk is physically impossible because the fuel isn't weapons grade.

Speaker 2: And crucially, that renewables and nuclear aren't enemies. They have to work together in a carbon constrained world. Understanding the scientific reality of nuclear energy and getting past the fear are crucial for making informed choices about our climate future.

Speaker 1: It feels like we are standing at a crossroads. We have this tool that is incredibly powerful, but it carries so much baggage.

Speaker 2: It raises a massive question for you to think about as a listener. If nuclear energy is statistically one of the safest and cleanest ways to generate power, yet we fear it more than sources that kill millions annually through pollution, is our biggest energy hurdle actually technological? Or is it psychological? What would the world look like if we treated nuclear physics as a tool rather than a threat?

Speaker 1: That is a provocative thought to end on. I'm definitely going to look at those cooling towers differently next time I drive by. Thank you for guiding us through the atom today. And to you listening, thank you for jumping into this deep dive with us. Challenge your assumptions, check the data, and stay curious. We will see you on the next deep dive.

C.2.2. Cell Phone Radiation

Speaker 1: Welcome to this custom-tailored deep dive. We are unpacking a topic that is literally sitting in your pocket right now, or maybe you're even holding it while listening. Today's mission is to sift through this massive stack of sources to separate the scientific reality of cell phone radiation from all the noise out there. We are diving into everything from the national toxicology program's \$30 million animal studies to the current guidelines from the FDA and the FCC and even pressing policy warnings from places like the American Academy of Pediatrics and the Cyprus Medical Association.

Speaker 2: And we are going to approach this entirely through a scientific lens. The goal here is to lay out the actual mechanisms of how this energy works, unpack the prominent research and the major open questions, and finally define some practical health guidelines for you because it's so easy to just feel overwhelmed.

Speaker 1: Oh, absolutely. By the end of this, you should be able to make informed choices about your daily technology habits without feeling panicked by conflicting headlines.

Speaker 2: Which can easily happen when we're dealing with an entirely invisible environment. We are surrounded by this constant silent exchange of energy everywhere we go.

Speaker 1: But before we can even begin to understand how a smartphone might interact with human biology, we have to establish a baseline. What kind of energy is a cell phone actually emitting into the space around us?

Speaker 2: It all begins with the electromagnetic spectrum. Cell phones send and receive data using radio frequency energy, or RF energy for short. And to understand how that interacts with your body, you really have to draw a hard line between two completely different categories, which are ionizing radiation and non-ionizing radiation.

Speaker 1: Let's break those down.

Speaker 2: Ionizing radiation is the high frequency, high energy stuff. This is what you encounter with X-rays, gamma rays, or ultraviolet light from the sun. We're talking about energy waves that carry so much kinetic power they can physically alter matter at the atomic level. Ionizing radiation carries photons with enough energy to literally knock electrons out of their orbit around an atom. That process creates ions and is highly disruptive to biological tissue. But non-ionizing radiation operates on the completely opposite end of the spectrum. It is low frequency and low energy. This category includes things like FM radio waves, microwaves, and the RF energy used by your cell phone, your home Wi-Fi router, and 5G networks. Because it cannot break chemical bonds, the primary biological mechanism it actually does have is tissue heating, essentially warming you up. When these rapid waves pass through human tissue, they cause the molecules, particularly your water molecules, to vibrate slightly. That vibration creates friction, and the friction creates heat. It's actually very similar to just stepping out into direct sunlight. At the legally permitted levels for cell phones, the only widely agreed upon biological effects are a microscopic warming of the skin and occasionally some mild nerve stimulation which might cause a slight tingling sensation.

Speaker 1: Okay. But because this energy is completely invisible, it definitely breeds a lot of anxiety for people. **Just scrolling online, I have the impression that 5G radiation affects the human immune system.**

Speaker 2: Yeah, you see that a lot. But we should look at the data from radiation watch dogs and agencies like Arpana, the Australian Radiation Protection and Nuclear Safety Agency. They've heavily investigated how RF energy interacts with our biology. And because this energy is strictly non-ionizing and the power levels are legally capped to keep that tissue heating well below any harmful threshold, there simply is no biological mechanism for these waves to interact with your immune cells.

Speaker 1: So, they literally can't do anything to your immunity.

Speaker 2: No, they cannot structurally damage your white blood cells or alter your T cells because they physically cannot break the atomic bonds that make up those structures. The physical properties of the waves restrict them to basically just causing minor molecular vibration.

Speaker 1: Okay, let me stop you right there because that raises a massive question. If the baseline physics proves that RF energy only causes minor heating, why did the US government just spend

\$30 million and 10 entire years studying it? And why did the World Health Organization's International Agency for Research on Cancer classify radio frequency electromagnetic fields as a group 2B possible carcinogen? Scientists don't spend a decade and tens of millions of dollars just to confirm that a phone makes your ear warm.

Speaker 2: They certainly don't. The scientific debate has shifted: it moved away from direct DNA damage to the much more complex question of long-term low-level exposure, like what happens over a lifetime. Even if it doesn't break bonds, does a lifetime of this vibrating thermal energy trigger other biological cascades? To investigate that, the National Toxicology Program, or NTP, ran one of the most comprehensive toxicology studies in history. They built these massive specially-designed underground reverberation chambers. These were highly shielded rooms with large antennas and paddle wheels that constantly scattered the radio waves to ensure that the rats and mice inside received a perfectly uniform dose of 2G and 3G cell phone radiation no matter where they moved in the room.

Speaker 1: And they kept them in these reverberation chambers for how long?

Speaker 2: Over 9 hours a day, every day for up to 2 years. And the exposure was delivered in intervals, too. 10 minutes on and 10 minutes off to mimic an incredibly heavy continuous usage pattern.

Speaker 1: That is an intense environment. So, what did a \$30 million 2-year rodent study actually uncover?

Speaker 2: Well, the findings were highly significant for the scientific community. They actually observed a link between this continuous RF exposure and the development of relatively rare heart tumors. These malignant tumors targeted cells that are responsible for producing myelin, which is the protective insulation wrapped around your nerves. And the researchers found these tumors on the nerves of the heart in about 5 to 6% of the male rats. But there is a crucial caveat here: this biological effect was only observed at extremely high power levels. The rats were subjected to whole body exposure at levels four times higher than the maximum permitted human limit.

Speaker 1: Wait, if a rat is subjected to whole body radiation for 9 hours a day, at four times the human limit, how does that translate to me just holding a phone to my ear for a 10-minute phone call? That seems like a massive leap to make.

Speaker 2: You cannot draw a straight line from a rodent living inside a high radiation reverberation chamber to a human just checking emails on a smartphone. The study was never designed to perfectly mimic everyday human use.

Speaker 1: Then what was the point of it?

Speaker 2: Its purpose was to test a biological boundary. It proved that non-thermal biological effects like the development of these tumors are theoretically possible under extreme prolonged conditions. That single finding shattered the old assumption that if it doesn't heat you, it can't affect you, which is exactly why the WHO and others maintain that more research is urgently needed.

Speaker 1: That shift in understanding the biological boundaries brings up another claim I keep seeing as cellular networks upgrade. **Some sources point out that 5G frequencies replicate inside the human body and cause reradiation.**

Speaker 2: Let's connect this to the bigger picture. To understand why that claim breaks down, we have to basically revisit the definition of what a radio wave actually is. Radiation from a cell phone is just an oscillating field of electric and magnetic energy. It is not a biological agent. Your body's cells just don't have the biological machinery or the antennas to emit radio waves on their own. RF energy dissipates as heat as it travels through tissue.

Speaker 1: So, it's not like a virus.

Speaker 2: Exactly. It does not behave like a pathogen that infects a host, hijacks its cells, and replicates. Energy simply cannot replicate itself.

Speaker 1: So it is an external force acting upon the body, not something growing inside it, okay. But that brings us back to the core challenge of those animal studies because extrapolating data from a rat to a fully grown human adult is already incredibly difficult. It makes me wonder about populations that aren't fully grown adults. Like how does this external energy interact with the physically different

bodies of children?

Speaker 2: This is where modern research is really focusing its most urgent attention because a child's body is not just a scaled down version of an adult. It possesses entirely different physical, anatomical, and electrical properties. Their heads are literally smaller. Groundbreaking dosimetric research, particularly the modeling done by Dr. Gandhi has completely changed the conversation here. Dr. Gandhi used these high-resolution anatomical models to simulate exactly how RF energy penetrates human tissue at different stages of human development.

Speaker 1: And the sources show the results of his modeling are pretty striking.

Speaker 2: Indeed. Because children have significantly thinner skulls and their ears are smaller, the antenna of a smartphone sits much closer to their actual brain tissue when they hold it to their head. Furthermore, a child's brain tissue has a higher water and ion content, which actually makes it more conductive to radio waves. And Dr. Gandhi's modeling demonstrated that a 10-year-old child absorbs up to 153% more radiation than an adult. This measurement is known as the specific absorption rate, or SAR. It tracks how much energy is absorbed per kilogram of tissue. And even more concerning, the absorption in the bone marrow of a child's skull can be 10 times greater than in an adult.

Speaker 1: Which brings us to a major point of contention in the medical community right now: the way these devices are actually tested for safety before they are legally allowed to hit the market. Because when manufacturers test a phone to make sure it falls within the government's legal radiation limits, they don't test it on a variety of diverse body types. Instead, they use a plastic model called the specific anthropomorphic mannequin, or SAM. And this mannequin is modeled after the top 10% of US military recruits from the year 1989. A very specific, very large physical demographic. How can regulatory agencies possibly say a device is fully safe for a toddler watching a cartoon on an iPad when the safety standard was calibrated for a fully grown elite soldier from over three decades ago?

Speaker 2: And that precise discrepancy is exactly why the medical community is raising alarms. Organizations like the American Academy of Pediatrics are actively demanding that governments completely overhaul their testing protocols because they're outdated. They argue that our current FCC guidelines, many of which were established back in 1996 during the absolute infancy of the modern telecom era, completely fail to account for the unique anatomical vulnerabilities of pregnant women and growing children. They're testing for a world that literally no longer exists.

Speaker 1: Yeah. And as we transition from older networks to 5G, the infrastructure is changing, too. We're seeing thousands of new smaller cell towers being built much closer to our homes to support these higher frequencies. And with that visual change, new concerns pop up. **Some sources point out that these denser 5G and mobile networks kill birds and bees.**

Speaker 2: That's another claim you see a lot. But environmental and ecological monitoring agencies, including the Audubon Society, have heavily investigated the impacts of radio waves on wildlife.

Speaker 1: And what did they find?

Speaker 2: Their extensive data track confirms that the radio waves emitted by these cellular networks do not kill birds or bees. The frequencies in use simply do not disrupt their biology or their navigational systems. The ecological data just does not support those claims. Therefore, the real focus of concern within the scientific community remains firmly on human developmental anatomy and chronic long-term exposure.

Speaker 1: So, given that our current government safety guidelines are based on an outdated military mannequin from the 90s and medical councils all around the world from the California Department of Health to the Cyprus Medical Association are actively advising caution, what can you, the listener, actually do? What are the practical science-backed steps you can take to protect yourself right now without throwing your phone in a lake?

Speaker 2: Well, the most empowering part of this research is that the actionable guidelines from these global health organizations are incredibly consistent, and they all rely on one fundamental rule of physics which is the inverse square law.

Speaker 1: Which means the intensity of the energy drops off a cliff the further it gets from the source.

Speaker 2: Exactly, if you double the distance between your body and the phone, you only receive a

quarter of the radiation. It drops fast. Distance is your absolute best defense. The primary recommendation across the board is to use the speaker-phone function or a wired headset whenever possible. Just keep the phone away from your head.

Speaker 1: The California Department of Health advisory was very specific about where we carry these devices when we aren't even using them. They explicitly advise keeping phones out of your pockets and out of bras. Carry them in a backpack or a purse instead.

Speaker 2: And there's a good reason for that. This is because a cell phone is never truly off unless it is fully powered down or in airplane mode. The phone's band processor is periodically pinging the nearest cell tower to announce its location, constantly checking in, which emits spikes of RF energy. And another critical factor that dictates exposure is signal strength. You should strictly avoid making voice calls when your screen only shows one or two bars of service. Because when the connection is weak, the phone has to compensate.

Speaker 1: It's kind of like trying to talk to a friend at a really loud, crowded concert. Because the ambient noise is so high and the connection is bad, your phone essentially starts screaming to reach the cell tower. The internal software commands the power amplifier to ramp up its output to the absolute maximum, which means it is emitting significantly more energy right next to your brain just to maintain that fragile connection.

Speaker 2: That mechanism is precisely why weak signals are so problematic. And you see the exact same power spike when you are in a moving vehicle like a car, a bus, or a train.

Speaker 1: Oh, really? Why is that?

Speaker 2: First, the metal frame of the vehicle acts like a Faraday cage. Basically blocking the signal and forcing the phone to work harder to pierce through. And second, because you are moving rapidly, your phone has to constantly drop the connection with one cell tower and perform a digital handshake to connect with the next one. That continuous rapid-fire handshaking process requires the phone to operate at peak power levels.

Speaker 1: So, if you are riding a train or sitting in the back of an Uber, you should probably just download your music beforehand and switch the device into airplane mode.

Speaker 2: That is a perfect habit. Simple adjustments to how you use the device drastically reduce your cumulative exposure. Texting instead of calling keeps the transmitter away from your skull. Putting your phone in airplane mode when you go to sleep, or at the very least, keeping it off your mattress and across the room gives your body an 8-hour break from those background RF pings.

Speaker 1: And it probably helps you sleep better anyway without the screen time.

Speaker 2: Definitely. Furthermore, pediatric medical groups strongly suggest that families use hardwired internet connections like Ethernet cables over Wi-Fi routers in homes. And they are particularly advocating for hardwired connections in school classrooms where kids spend hours every day.

Speaker 1: You know, with people actively looking for solutions to mitigate this, **I have read that applying a special radiation shield or sticker to your phone significantly reduces your exposure.**

Speaker 2: Yeah, it is a massive consumer market right now. But if we apply the physics we just discussed, it is actually completely counterproductive. The Federal Trade Commission has even issued warnings about these products. When you place a physical shield or blocking sticker over the phone's antenna, the device's internal sensors immediately detect a drop in the signal to noise ratio. The phone assumes it is in a terrible coverage area and automatically increases its power output to blast through the blockage. By trying to shield yourself, you basically trap the phone in a feedback loop where it continuously screams at maximum power. Thereby increasing the amount of RF energy radiating into your hand and head.

Speaker 1: That is so wild. The very tool you buy to protect yourself tricks the device into exposing you to more energy. So, the only true scientifically sound way to reduce your exposure is creating physical distance and being mindful of the phone's transmission habits.

Speaker 2: Distance and signal management are the only proven methods.

Speaker 1: Okay, let's bring this all together. We set out today to understand the invisible ocean of

energy we live in. We learned that cell phone radiation is non-ionizing, so it won't shadow your DNA like an X-ray, and its primary function is molecular vibration or heat. We also saw that massive decades-long toxicology studies found that extreme non-thermal doses can cause rare biological changes in rats, even if those scenarios don't perfectly map onto our daily habits.

Speaker 2: But we also learned that we really cannot treat the entire human population as a monolith. The unique anatomical vulnerabilities of children, specifically the higher fluid content in their developing brains, combined with the fact that our regulatory testing relies on an outdated adult-sized military mannequin from 1989, means there are still massive open questions in the science.

Speaker 1: Definitely, we are regulating for a demographic that no longer represents the primary user base. And as 5G rolls out, bringing higher frequencies and millions of new small cell towers into our neighborhoods, taking simple distance-based precautions isn't about panicking. It's just smart science. Moving the phone a few inches away, using speaker-phone, and keeping the device out of your pocket is simply honoring that precautionary principle.

Speaker 2: And that leaves us with a final thought to ponder as we look toward the future. As technology evolves, we're moving rapidly past the smartphone and into an era of the internet of things and specifically smart clothing. We're talking about biometric wearables, continuous health monitors, sensors that are literally woven into the fabrics we wear every single day. As these emitting devices become completely indistinguishable from our apparel, how will we define a safe distance from a device when the device is the very shirt on our back?

Speaker 1: That is wild to think about. If you can see those invisible threads connecting us right now, I think they're only going to get brighter, denser, and far more tangled. Well, thanks for joining us on this deep dive into the source material. Stay curious, keep your distance, and we'll catch you next time.

C.2.3. GMO Foods

Speaker 1: Welcome to the show. It is fantastic to have you with us today for this custom-tailored deep dive.

Speaker 2: Yeah, thanks for having me.

Speaker 1: We've built this entirely around your interests, and today we are tackling a subject that practically everyone interacts with, yet very few people understand at a fundamental scientific level. We are looking at the applications and the safety of GMO foods. Our mission today is straightforward. We want to cut through the noise. We're going to explore the actual biology of how genetic modification works, look at the real-world agricultural and nutritional applications, and we'll understand exactly how these products are regulated and examine the genuine biological considerations. And we are going to do all of this by keeping the concepts accessible whether you have been following agricultural biotechnology for years or you're just starting to look into what is on your plate.

Speaker 2: To accomplish that, we are pulling from a remarkably diverse stack of sources for you today. We have reviewed comprehensive documentation from national health, agricultural, and environmental authorities across the globe. We are also diving into peer-reviewed scientific journals, widespread media analyses, and extensive consumer guides. So, bringing all these different data points together, we can give you a clear grounded picture of what the empirical evidence actually demonstrates.

Speaker 1: Okay, let's unpack this from the very beginning. What exactly are we talking about when we say GMO?

Speaker 2: Well, the term stands for genetically modified organism. In agricultural science, this usually refers to recombinant DNA technology. So instead of broad unpredictable crosses, scientists are altering an organism's DNA by introducing a specific gene to express a highly desirable targeted trait.

Speaker 1: We are moving away from the era of hoping for a lucky combination of traits.

Speaker 2: Exactly, moving into an era of targeted gene expression. It's incredibly helpful to contrast this modern targeted process with traditional selective breeding which humans have relied on for millennia, of course. With traditional breeding or even mutagenesis where seeds are exposed to radiation to force random mutations, you are mixing tens of thousands of genes at once. You're rolling the genetic dice and hoping the resulting plant has the traits you want. Modern genetic engineering is simply

a much more precise, predictable, and controllable way to transfer a desired trait.

Speaker 1: Because you know exactly which sequence is being moved.

Speaker 2: Yes, and you know exactly what biological function that sequence performs.

Speaker 1: Now, hearing about all these high-tech methods, **I have the impression that these modifications are so new that we really don't know if they are dangerous yet.**

Speaker 2: It's really common to feel that way because the tech sounds so futuristic. But looking at the timeline in our sources, genetically engineered plants actually first appeared in labs about 30 years ago. And they hit the commercial market back in 1994, so it's been decades. And since then, there have been over 1,700 peer-reviewed safety studies.

Speaker 1: That is a massive amount of literature.

Speaker 2: Plus decades of real-world consumption by hundreds of millions of people without any verified negative health effects. What's fascinating here is that modern techniques are actually way more precise than the random mutations we relied on during 10,000 years of traditional crossbreeding.

Speaker 1: So why do we do this? What are the actual applications and intended benefits?

Speaker 2: The overarching goals generally come down to higher crop yields, lower food prices, longer shelf life, and improved nutritional value.

Speaker 1: Okay, give me an example.

Speaker 2: To see the actual mechanism at work, you have to look at the Hawaiian papaya. In the 1990s, the entire Hawaiian papaya industry was on the brink of total collapse because of a devastating pathogen. It's called the ring spot virus.

Speaker 1: And I'm guessing there was no conventional cure for it?

Speaker 2: None. They were running out of options, so scientists took a harmless tiny fragment of the virus's own genetic code and inserted it into the papaya's genome.

Speaker 1: Wait, they put the virus into the plant?

Speaker 2: Just a fragment. When the plant grows, it produces a small amount of the viral protein. This essentially acts like a vaccine. The plant's cellular immune system recognizes the protein. So, when the actual live ringspot virus attacks the field, the papaya already has the defenses built up to fight it off. It literally saved the industry.

Speaker 1: See, that makes biological sense, using a fragment of the enemy against it. But what about something like the Arctic apple? I know it was designed so it doesn't turn brown when you cut it, but honestly, that sounds like a purely cosmetic trick, just to make fruit look prettier in a lunchbox.

Speaker 2: It definitely looks cosmetic, but the mechanism solves a massive logistical and environmental problem, which is food waste. When you cut an apple, an enzyme called polyphenol oxidase reacts with oxygen in the air causing the browning. And consumers throw away millions of pounds of perfectly good apples simply because they look a little oxidized. So scientists use genetic engineering to silence the specific gene responsible for producing that browning enzyme. It just doesn't oxidize.

Speaker 1: Okay, that makes a lot of sense. And on the nutritional side, we have things like pink pineapples which were modified to produce higher levels of lycopene, right?

Speaker 2: Yes, that's the pigment that makes tomatoes red.

Speaker 1: Gotcha. But beyond boutique fruits, there's golden rice, which seems to be a huge deal in the sources.

Speaker 2: In many developing nations, severe vitamin A deficiency causes childhood blindness, and it fatally compromises immune systems. And the primary diet in these regions is heavily dependent on white rice, which completely lacks certain micronutrients. So, they edited the rice to produce vitamins. Specifically, they engineered it to produce beta carotene right in the edible part of the grain.

Speaker 1: Beta carotene. That's what makes carrots orange!

Speaker 2: Exactly. Hence the name golden rice. So, when a person eats it, the human body naturally converts that beta carotene directly into vitamin A. It addresses the public health crisis right at the source of the diet.

Speaker 1: Here's where it gets really interesting. Because despite the Hawaiian papayas and golden rice, the reality is that the vast majority of our genetically modified crops are not sitting in the produce aisle. Over 90% of all the corn and soy planted in the United States is actually used for animal feed. Which brings up a major physiological concern. **Some sources point out that the DNA from genetically modified animal food makes its way into the meat, eggs, or milk from the animal that we eventually consume.**

Speaker 2: To address that, we really have to look closely at the harsh biology of the digestive system. Because whether a cow, or a human for that matter, eats a protein from a completely natural wild plant or a genetically modified soybean, the stomach treats it the exact same way, which is by destroying it. The stomach environment is just a vat of hydrochloric acid. It's got a pH level around 1.5, which is extremely acidic. And combined with aggressive digestive enzymes, this environment just breaks down complex biological structures. It unwinds the DNA and chops the proteins down into individual tiny building blocks called amino acids.

Speaker 1: So the cow's stomach doesn't look at a protein and say, "Ah, this is from a modified soybean. I need to treat it differently."

Speaker 2: Not at all. The body doesn't see a biological blueprint at that stage. It only sees basic building blocks because the DNA is completely degraded by the acid and enzymes. The modified genetic material cannot and does not transfer into the meat, the milk, or the eggs.

Speaker 1: And the research backs that up!

Speaker 2: Completely: research consistently shows that dairy and meat from animals fed edited crops are nutritionally indistinguishable from those fed conventional crops.

Speaker 1: Okay. So, if the animal's stomach destroys the DNA anyway, then the most significant impacts aren't happening inside the cow. We need to look at the environment where the plant is actually grown.

Speaker 2: Yes. The fields, the soil, the bugs, and the weeds: the environmental equation.

Speaker 1: Because this is where the heaviest agricultural focus lies. From the sources, it looks like many of these crops are engineered for herbicide tolerance. Meaning a farmer can spray a field with weed killer and the weeds die, but the crop survives, which honestly sounds chemically intense.

Speaker 2: It does sound intense, but the alternative mechanism is often much worse for the soil, namely tilling. Without herbicide tolerant crops, farmers have to control weeds by tilling the land.

Speaker 1: Like plowing?

Speaker 2: Yeah. They drive massive tractors back and forth across the field, physically churning and plowing the earth just to rip up weeds. Tilling destroys the soil structure, causes massive top soil erosion, releases trapped carbon into the atmosphere, and burns incredible amounts of diesel fuel. Meanwhile, herbicide tolerance allows for no till farming, which actually preserves the soil ecosystem.

Speaker 1: That makes sense. The other major environmental application I saw was insect resistance primarily with something called Bt corn.

Speaker 2: Yes. Bt stands for *Bacillus thuringiensis*. It's a naturally occurring soil bacterium.

Speaker 1: So, scientists took the gene from that bacteria and put it into the corn. And the corn essentially acts as its own insecticide. **But I have read that these pest resistant crops harm beneficial insect species in the environment.**

Speaker 2: This is a perfect example of why the underlying mechanism really matters because the Bt protein produced by the corn is crystalline. In its natural state, it is completely harmless. It only becomes toxic when it dissolves and binds to specific cellular receptors in a highly alkaline environment.

Speaker 1: Alkaline meaning the opposite of acidic.

Speaker 2: Yes, exactly. The specific target pests like the European cornborer caterpillar happen to have highly alkaline guts with those exact cellular receptors. The protein binds to their gut, and they die.

Speaker 1: But mammals have acidic stomachs, right?

Speaker 2: As we just discussed, mammals have highly acidic stomachs. So, the protein simply breaks down.

Speaker 1: And what about the beneficial bugs?

Speaker 2: Crucially, beneficial insects like ladybugs or pollinators like bees do not have those specific receptors in their biology. The protein just passes through them completely harmlessly.

Speaker 1: So compared to a farmer indiscriminately crop dusting a field with a broad-spectrum chemical insecticide that kills everything it touches, a field of Bt corn is actually a much safer neighborhood for a ladybug.

Speaker 2: However, we cannot ignore the larger biological challenge here, which is the evolutionary arms race.

Speaker 1: If we blanket millions of acres of farmland with the exact same Bt corn or the exact same herbicide, don't the bugs and weeds eventually adapt?

Speaker 2: They do. Heavy continuous reliance on a single agricultural tool places immense selective pressure on the ecosystem. So, out of a billion bugs, one might just have a rare genetic mutation that allows it to survive the protein, and then it reproduces. With all its competition wiped out, that surviving bug reproduces rapidly leading to the evolution of resistant superbugs. And the same thing happens with weeds, creating superweeds.

Speaker 1: So, the microscopic editing solves one problem, but creates a much tougher one down the road.

Speaker 2: It does. Farmers then have to rotate crops, use different chemicals, or even plant refugee zones of unedited crops just to keep vulnerable non-resistant bugs in the mating pool.

Speaker 1: It really complicates the environmental benefits over the long term. So, if these superweeds and superbugs are an inevitable evolutionary byproduct, someone has to be calculating that risk before a seed is even planted. Who is actually making that call?

Speaker 2: The regulatory frameworks are rigorous, but they vary dramatically depending on global geography. In the United States, it's a three-pronged coordinated effort. The health agency ensures it is safe for human digestion. The agricultural agency ensures it won't crosspollinate out of control and become a noxious weed. And the environmental agency measures the impact on the surrounding ecosystem.

Speaker 1: Okay. But I have to imagine Europe, which is notoriously stricter about food standards, doesn't just wave these crops through using the same process.

Speaker 2: No, they don't. The European Union approach leans heavily on what's called a precautionary principle. They require an exhaustive 10-year renewable authorization process. It involves incredibly strict risk assessments by the European Food Safety Authority. If there is even theoretical uncertainty about a risk, the EU will often just halt approval entirely.

Speaker 1: But how do any of these regulators actually test a new crop? They don't just stare at a soybean and guess if it's safe, right?

Speaker 2: No, they rely on a foundational scientific mechanism called substantial equivalence. Regulators conduct deep mass spectrometry and compositional analyses. They basically map out the new food against its conventional counterpart down to the molecule. They measure every single amino acid, vitamin profile, and all known naturally occurring toxins.

Speaker 1: So, it's kind of like taking two cars apart down to the very last screw. And if every single screw, wire, and piston is identical, and the biochemical profile matches perfectly, except one car has a slightly upgraded spark plug, regulators consider them the exact same car.

Speaker 2: That is the perfect analogy. If the modified soybean has the exact same nutritional profile and naturally occurring toxin levels as a conventional soybean, it passes the substantial equivalence test.

Speaker 1: That makes sense. **Though I have read that eating genetically modified soybeans is highly correlated with developing tumors and cancer.**

Speaker 2: That claim has a very specific origin story in the media analyses we reviewed. It largely traces back to a massive cyber panic in China around 2013 sparked by claims that cited two specific studies. A 2010 study on hamsters by a Russian scientist named Alexei Surov and a 2012 study on mice by the Seralini team.

Speaker 1: Okay. So, what did the broader scientific community find when they looked at those studies?

Speaker 2: The research institute that Surov worked for stated they actually had no record of that hamster study ever existing. And as for the 2012 Seralini study, it was heavily investigated and ultimately retracted by the scientific journal that published it.

Speaker 1: Why was it retracted?

Speaker 2: Severe methodological flaws. For example, they used a specific strain of lab rat that is naturally highly prone to developing tumors regardless of what it eats. So the results were completely skewed. And beyond those two studies, when global health authorities actually looked for an epidemiological basis, meaning looking at large populations over time, they found absolutely zero evidence that GM crops cause cancer. The scientific consensus from major bodies like the World Health Organization and the American Association for the Advancement of Science is completely unified. They all state that currently marketed GM foods pose no greater risk to human health than conventional foods.

Speaker 1: But if the cars are identical down to the screws and the science proves substantial equivalence, why is the public so terrified of the upgraded spark plug? There's a glaring disconnect between the scientific consensus and human psychology here.

Speaker 2: There is, and part of it stems from highly valid initial questions about cross species biology, specifically allergenicity. For example, if scientists take a gene from a Brazil nut to improve the nutritional profile of a soybean, could someone with a tree nut allergy go into anaphylactic shock from eating soy?

Speaker 1: That sounds terrifying. You wouldn't know to look out for it.

Speaker 2: It is a serious concern, which is precisely why regulators require extensive allergenicity testing.

Speaker 1: Have they ever caught something like that?

Speaker 2: Actually, yes. That exact Brazil nut scenario happened during development in the 90s. During the mandatory safety testing protocols, researchers realized the newly synthesized protein in the soy was an active allergen.

Speaker 1: Did it slip through?

Speaker 2: No, not at all. They immediately scrapped the entire project. It never reached the farm, let alone the grocery store. The testing mechanism worked exactly as designed.

Speaker 1: That's reassuring. **But I have the impression that all research on GMOs has been funded by big agricultural companies.** If the massive corporations making the seeds are the ones paying for the safety tests, the public trust is naturally going to be zero.

Speaker 2: That is a very common concern actually. But when you look at the global data, the research landscape is far more robust than that. Over the past few decades, hundreds of independent researchers have published peer-reviewed safety studies independent of the big seed companies. The European Union alone, which as we discussed is highly skeptical of this technology, funded over 130 separate research projects. This involved 500 independent scientific groups over 25 years.

Speaker 1: Why doesn't the public know that? If the independent science says it's safe, why does the fear persist so aggressively?

Speaker 2: Well, because information moves much faster than scientific peer review. Public panic is often fueled by a combination of digital anxiety and highly publicized flawed studies.

Speaker 1: On social media, a scary image with a sensational headline travels a million times faster than peer review. It just creates an echo chamber of fear that no amount of World Health Organization data can easily penetrate.

Speaker 2: The lack of trust isn't really about the cellular biology at the end of the day. It is about a feeling of losing control over our food supply combined with a deep distrust of the massive corporations that hold the patents and intellectual property to these genetically engineered seeds. The molecular science is incredibly clear, but bridging that sociological divide remains an ongoing complex challenge. If we connect this to the bigger picture, genetic engineering is a powerful tool. It's essential for addressing global challenges like climate change and food security.

Speaker 1: But it needs public trust to work.

Speaker 2: It requires careful, transparent management.

Speaker 1: So, what does this all mean for you, the listener, navigating the grocery store aisles? GMOs represent a highly precise evolution of agriculture. They bring very tangible benefits, but they also require rigorous ongoing scientific oversight. It's not magic. It's a managed science.

Speaker 2: Exactly. And as we look toward the future, I want to leave you with a final thought to ponder stepping just beyond what we have discussed today. If human ingenuity now possesses the biological precision to engineer crops simply to survive a changing harsher climate, could we eventually design plants that actively reverse climate change? Imagine engineering vast tracks of agricultural crops or even entire massive forests to absorb vastly more carbon dioxide from the atmosphere than any natural plant ever could. It is a frontier that could redefine our relationship with nature entirely, moving from adapting to our environment to actively healing it.

Speaker 1: That is an incredible concept to think about, and it really reframes what this technology could be capable of in our lifetimes. We want to warmly thank you for joining us on this deep dive. We hope it has been as eye-opening and empowering for you as it was for us to put together. Until next time; keep asking questions and keep exploring.

C.3. Content-Supplying Sources

C.3.1. Nuclear Energy (32 Sources)

- <https://www.anl.gov/article/10-myths-about-nuclear-energy>
- <https://www.energy.gov/ne/articles/5-fast-facts-about-nuclear-energy>
- <https://www.energy.gov/ne/articles/5-fast-facts-about-spent-nuclear-fuel>
- <https://www.energy.gov/ne/articles/advantages-and-challenges-nuclear-energy>
- https://energy.ec.europa.eu/topics/nuclear-energy/decommissioning-nuclear-facilities_en
- https://energy.ec.europa.eu/topics/nuclear-energy/emergency-preparedness-and-response_en
- https://energy.ec.europa.eu/topics/nuclear-energy/euratom-safeguards_en
- <https://www.energy.gov/ne/articles/nuclear-101-how-does-nuclear-reactor-work>
- <https://www.nei.org/news/2019/nuclear-101-an-introduction-to-nuclear-energy>
- <https://ourworldindata.org/nuclear-energy>
- <https://climate.mit.edu/explainers/nuclear-energy>
- <https://world-nuclear.org/nuclear-essentials/how-is-uranium-made-into-nuclear-fuel>
- <https://world-nuclear.org/nuclear-essentials/are-there-different-types-of-reactor>
- <https://world-nuclear.org/nuclear-essentials/what-are-the-effects-of-nuclear-accidents>

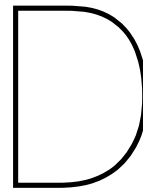
- <https://world-nuclear.org/nuclear-essentials/is-radiation-safe>
- <https://world-nuclear.org/nuclear-essentials/where-does-our-electricity-come-from>
- <https://world-nuclear.org/nuclear-essentials/how-does-a-nuclear-reactor-work>
- <https://world-nuclear.org/nuclear-essentials/what-is-nuclear-waste-and-what-do-we-do-with-it>
- <https://world-nuclear.org/nuclear-essentials/how-can-nuclear-combat-climate-change>
- <https://www.ucs.org/energy/nuclear-power>
- https://energy.ec.europa.eu/topics/nuclear-energy/nuclear-safety_en
- <https://funginstitute.berkeley.edu/news/op-ed-lets-dispel-the-myths-about-nuclear-energy/>
- https://energy.ec.europa.eu/topics/nuclear-energy/radiation-protection/radiation-protection-legislation_en
- <https://world-nuclear.org/information-library/nuclear-fuel-cycle/nuclear-waste/radioactive-wastes-myths-and-realities>
- https://energy.ec.europa.eu/topics/nuclear-energy/radioactive-waste-and-spent-fuel_en
- https://energy.ec.europa.eu/topics/nuclear-energy/radiation-protection/radioactivity-environment_en
- https://energy.ec.europa.eu/topics/nuclear-energy/radiation-protection/radiological-impacts_en
- https://energy.ec.europa.eu/topics/nuclear-energy/small-modular-reactors_en
- <https://www.orau.org/blog/programs/top-10-nuclear-energy-myths.html>
- <https://energy.mit.edu/wp-content/uploads/2018/09/The-Future-of-Nuclear-Energy-in-a-Carbon-Constrained-World.pdf>
- https://www.nrc.gov/sites/default/files/doc_library/cdn/legacy/reading-rm/basic-ref/students/what-is-nuclear-energy.pdf
- <https://www.sciencedirect.com/science/article/pii/S2772783122000085>

C.3.2. Cell Phone Radiation (15 Sources)

- https://en.wikipedia.org/wiki/5G_misinformation
- <https://www.canada.ca/en/health-canada/services/health-risks-safety/radiation/everyday-things-emit-radiation/cell-phones-towers.html>
- <https://www.healthychildren.org/English/safety-prevention/all-around/Pages/Cell-Phone-Radiation-Childrens-Health.aspx>
- <https://ntp.niehs.nih.gov/research/topics/cellphones>
- <https://www.cancer.gov/about-cancer/causes-prevention/risk/radiation/cell-phones-fact-sheet>
- <https://www.cancer.org/cancer/risk-prevention/radiation-exposure/cellular-phones.html>
- <https://www.cancerresearchuk.org/about-cancer/causes-of-cancer/cancer-myths-questions/do-mobile-phones-cause-cancer>
- <https://ehsciences.org/doctors-and-scientists-on-cell-phone-radiation-health-effects/>
- <https://www.cdc.gov/radiation-health/data-research/facts-stats/cell-phones.html>
- <https://www.yalescientific.org/2014/07/mythbusters-do-cell-phones-really-cause-cancer/>
- <https://ehtrust.org/science/scientific-imaging-cell-phone-wi-fi-radiation-exposures-human-body/>
- <https://www.cdph.ca.gov/Programs/CCDC/DEOD/CEHIB/CEPH%20document%20Library/Cell-Phone-Guidance.pdf>
- https://www.ewg.org/sites/default/files/2022-05/EWG_EMF-05.22_C01.pdf
- https://www.niehs.nih.gov/sites/default/files/2025-08/cellphonerfr_long_508.pdf
- <https://www.frontiersin.org/journals/communications-and-networks/articles/10.3389/frcmn.2021.635716/full>

C.3.3. GMO Foods (36 Sources)

- <https://www.fda.gov/food/consumers/agricultural-biotechnology>
- <https://www.foodstandards.gov.au/consumer/gmfood/current-status-genetically-modified-foods-applications>
- <https://www.cancerresearchuk.org/about-cancer/causes-of-cancer/cancer-myths-questions/do-genetically-modified-foods-cause-cancer>
- <https://www.who.int/news-room/questions-and-answers/item/food-genetically-modified>
- <https://www.fda.gov/food/agricultural-biotechnology/gmo-crops-animal-food-and-beyond>
- <https://www.popsci.com/article/science/core-truths-10-common-gmo-claims-debunked/>
- <https://www.foodstandards.gov.au/consumer/gmfood/gmoverview>
- <https://education.nationalgeographic.org/resource/genetically-modified-organisms/>
- <https://www.fsai.ie/enforcement-and-legislation/legislation/food-legislation/genetically-modified-organisms-%28gmos%29>
 - [/eu-authorisation-of-gm-food](#)
 - [/labelling-of-gm-food](#)
 - [/regulating-gm-food-in-the-eu-and-ireland](#)
 - [/safeguard-measures](#)
 - [/traceability-for-gm-food](#)
 - [/transboundary-movement-of-gmos](#)
- <https://www.nature.com/scitable/topicpage/genetically-modified-organisms-gmos-transgenic-crops-and-732/>
- https://food.ec.europa.eu/plants/genetically-modified-organisms_en
- https://en.wikipedia.org/wiki/Genetically_modified_food_controversies
- <https://www.food.gov.uk/safety-hygiene/genetically-modified-foods>
- <https://medlineplus.gov/ency/article/002432.htm>
- <https://www.efsa.europa.eu/en/topics/genetically-modified-organisms>
- <https://www.fda.gov/food/agricultural-biotechnology/how-gmo-crops-impact-our-world>
- <https://www.fda.gov/food/agricultural-biotechnology/how-gmos-are-regulated-united-states>
- <https://www.ams.usda.gov/rules-regulations/be/bioengineered-foods-list>
- <https://www.foodstandards.gov.au/food-standards-code/proposals/p1055-definitions-for-gene-technology-and-new-breeding-techniques>
- <https://www.foodstandards.gov.au/consumer/gmfood/safety>
- <https://www.fda.gov/food/agricultural-biotechnology/science-and-history-gmos-and-other-food-modification-processes>
- <https://www.unlockfood.ca/en/Articles/Food-technology/Understanding-Genetically-Modified-Foods.aspx>
- <https://ohioline.osu.edu/factsheet/HYG-5058>
- <https://www.hse.gov.uk/biosafety/gmo/whatare.htm>
- <https://pmc.ncbi.nlm.nih.gov/articles/PMC3791249/>
- <https://www.sciencedirect.com/science/article/pii/S2213453016300295>
- <https://pmc.ncbi.nlm.nih.gov/articles/PMC11702960/>
- <https://journals.sagepub.com/doi/10.1177/209660831900200407>
- <https://www.sciencedirect.com/science/article/pii/S0950329324000508>
- <https://www.scientificamerican.com/article/why-people-oppose-gmos-even-though-science-says-they-are-safe/>



Recall Questionnaires

D.1. Nuclear Energy

1. Which of the following is **NOT** a step in the nuclear fission chain reaction described in the podcast?
 - A) An unstable atom splits and releases an immense amount of heat.
 - B) Splitting atoms shoot out two or three additional neutrons.
 - C) Newly released neutrons strike other nearby uranium atoms.
 - D) Two uranium atoms merge together to form a heavier nucleus.
2. Why is nuclear fuel considered more attractive than burning fossil fuels according to the hosts?
 - A) Because chemical reactions like burning are very difficult to control safely.
 - B) Because nuclear reactions only rearrange surface-level electrons.
 - C) Because only a small amount of fuel is needed to make a lot of power.
 - D) Because fossil fuels require more expensive turbine technology.
3. Which of the following is **NOT** a characteristic of commercial nuclear fuel pellets?
 - A) They are produced from a malleable metallic ore.
 - B) They are processed into small, hard, black units.
 - C) They are roughly the size of a marshmallow.
 - D) They are composed of a solid ceramic material.
4. How did the hosts characterize the scale of a reactor meltdown and a nuclear explosion?
 - A) A meltdown is more dangerous because it cannot be seen.
 - B) A meltdown is a financial disaster but not a nuclear explosion.
 - C) Meltdowns are only a potential risk in older 1970s reactor designs.
 - D) An explosion is a mechanical failure, while a meltdown is chemical.
5. Which of the following is **NOT** true regarding weapons-grade uranium compared to reactor fuel?
 - A) Weapons-grade uranium is enriched with about 90% U235.
 - B) Reactor fuel lacks the concentration of energy required to detonate.
 - C) High enrichment is a physical requirement for a nuclear explosion.
 - D) Reactor fuel and weapons-grade uranium are of comparable purity.
6. What is the primary safety function of control rods made from materials like silver, indium, or boron?

- **A)** To absorb neutrons and instantly stop the nuclear chain reaction.
 - B) To provide structural support for the heavy steel pressure vessel.
 - C) To filter out radioactive isotopes from the steam before it hits the turbine.
 - D) To act as a heat sink that melts before the uranium fuel itself would.
7. What is the primary safety function of the steel-reinforced containment building in the “defense in depth” philosophy?
- A) To serve as a weather shelter for the turbines and other sensitive equipment.
 - **B)** To ensure that even if fuel melts, it remains inside a secure concrete box.
 - C) To enclose a vacuum-sealed chamber that increases electrical output.
 - D) To provide the insulation needed to maintain a consistent reactor temperature.
8. What comparison did the hosts make regarding the safety of the US nuclear industry versus fossil fuels?
- A) Nuclear energy causes fewer workplace accidents than wind turbine maintenance.
 - B) Fossil fuels are far safer because they do not produce long-term radioactive waste.
 - **C)** Nuclear is safer as it avoids the respiratory issues caused by fossil fuel pollution.
 - D) The safety of nuclear energy is currently identical to that of natural gas.
9. What substance is primarily emitted from the large towers of a nuclear power plant during standard operation?
- A) Carbon dioxide.
 - **B)** Pure water vapor.
 - C) Nitrogen oxide.
 - D) Sulfur dioxide.
10. According to the podcast, what is the key reason nuclear energy is described as the “workhorse” of the energy grid?
- **A)** It has a high capacity factor of over 92%, providing constant baseload power.
 - B) It uses a simplified turbine system that requires less maintenance than solar.
 - C) It produces more energy per square mile than all other resources in the US.
 - D) It can be turned on and off more quickly than wind or solar power.
11. Which of the following statements was **NOT** made in the podcast regarding the volume of nuclear waste produced?
- A) One person’s lifetime of nuclear energy waste would fit in a soda can.
 - B) It is small enough that the entire US industry’s waste fits on one football field.
 - C) It is significantly smaller than the volume of coal ash produced in a year.
 - **D)** It would take up the full depth of two Olympic-size swimming pools.
12. What is the standard two-step process for managing spent fuel once it is removed from a nuclear reactor?
- **A)** Storage in deep water pools for cooling, followed by transfer to dry casks.
 - B) Immediate burial in salt mines, followed by surface-level encasement.
 - C) Grinding into solid powder, followed by storage in reinforced metal drums.
 - D) Chemical neutralization in specialized vats, followed by deep sea disposal.
13. What is the primary hurdle preventing the US from following Finland’s lead on geological repositories?

- A) The lack of scientific knowledge on how to build one.
 - B) The lower funding compared to European nations.
 - C) The lack of political consensus on a suitable location.
 - D) The higher radioactivity of US-designed fuel pellets.
14. Which of the following is **NOT** a characteristic of the “passive safety” systems in modern SMR (Small Modular Reactor) designs?
- A) They rely on gravity and natural convection for cooling.
 - B) They are designed to be “walk-away safe” to reduce human error.
 - C) They do not face the risk of electric pump failure if power is cut.
 - D) They are outfitted with a dome that automatically covers the core.
15. Which of the following is **NOT** a benefit of nuclear fusion that the hosts highlighted while contrasting it with fission?
- A) Nearly limitless energy production.
 - B) Zero risk of a reactor meltdown.
 - C) Present-day commercial availability.
 - D) No long-term leftover radioactive waste.

D.2. Cell Phone Radiation

1. What is the primary physical difference between ionizing and non-ionizing radiation?
- A) Ionizing radiation is used by home Wi-Fi routers, while non-ionizing is used for medical X-rays.
 - B) Ionizing radiation carries enough energy to knock electrons out of orbit, while non-ionizing does not.
 - C) Ionizing radiation generates surface heat, while non-ionizing radiation alters matter at the atomic level.
 - D) Ionizing radiation is low frequency and low energy, while non-ionizing radiation is high frequency.
2. As stated in the podcast, how does radio frequency (RF) energy cause the biological effect of tissue heating?
- A) It physically breaks chemical bonds, which releases thermal energy into the bloodstream.
 - B) It triggers a chemical reaction between the skin’s surface and the phone’s metallic casing.
 - C) It causes nearby water molecules to vibrate, which creates friction that in turn generates heat.
 - D) It interacts with the immune system, which reacts to the waves by raising internal temperature.
3. What physical constraint of RF waves did the hosts highlight regarding immune system interaction?
- A) The waves are blocked by the outermost layer of skin.
 - B) RF energy is absorbed by bone marrow before hitting T cells.
 - C) 5G wavelengths are too short to reach and alter immune cells.
 - D) The properties of the waves inherently restrict them to vibration.

4. Which of the following was **NOT** a purpose of the \$30 million National Toxicology Program (NTP) rodent study?
 - A) To test the biological boundaries of long-term, low-level RF exposure.
 - B) To see if non-thermal biological effects were theoretically possible.
 - C) To simulate the daily usage levels of current cellular technology.
 - D) To move the scientific debate beyond just immediate DNA damage.
5. Which of the following was **NOT** a characteristic of the rodent exposure in the NTP study?
 - A) It lasted for up to 9 hours every day for 2 years.
 - B) It utilized intervals of 10 minutes on and 10 minutes off.
 - C) The rats and mice were kept in glass cages in an open laboratory.
 - D) The rodents received whole-body exposure to 2G and 3G radiation.
6. What crucial detail did the hosts point out when discussing the significance of the NTP study's findings?
 - A) The rodents were exposed to levels four times the human limit.
 - B) The tumors were found to be benign rather than malignant.
 - C) The animals that were affected only developed cancer of the brain.
 - D) The RF exposure equally impacted rats and mice of either sex.
7. Which of the following is **NOT** a reason given for why RF energy cannot replicate inside the body?
 - A) The energy simply dissipates as heat as it travels through tissue.
 - B) Human cells lack the biological machinery required to emit RF waves.
 - C) The energy does not behave like a biological pathogen, such as a virus.
 - D) The human body's high salt content ensures external waves are reflected.
8. Why does a 10-year-old child absorb up to 153% more radiation than an adult when holding a phone to their head?
 - A) Their brain tissue has a higher water and ion content, making it more conductive.
 - B) Their ears are larger and thus serve as more efficient collectors for radio frequency energy.
 - C) Their more abundant fatty tissue acts as an insulator, trapping the energy beneath the skin.
 - D) Their higher metabolism accelerates the rate at which incoming waves are absorbed.
9. Which of the following is **NOT** a criticism leveled against current FCC safety guidelines for cellular technology?
 - A) They represent a world that no longer exists based on the infancy of the modern telecom era.
 - B) They rely on tools outright incapable of measuring the molecular vibration that causes heating.
 - C) They fail to capture the unique anatomical vulnerabilities of children and pregnant women.
 - D) They were derived using a model of a fully-grown elite soldier from more than three decades ago.
10. According to environmental agencies like the Audubon Society, what is the observed impact of cellular networks on birds and bees?
 - A) 5G frequencies are the primary cause of migratory bird confusion.
 - B) Waves from 5G networks do not disrupt the biology of birds or bees.
 - C) Modern cell towers cause significant declines in bee population fertility.

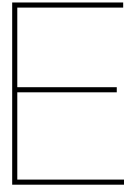
- D) The denser 5G towers attract more predatory birds to residential areas.
11. Regarding the discussion of the inverse square law, which of the following is **NOT** true?
- A) Energy intensity drops off significantly the further a wave travels from its source.
 - B) Doubling the distance from an emitter results in receiving only a quarter of the radiation.
 - C) Moving the phone next to the head minimizes the measured Specific Absorption Rate.
 - D) Distance is the best available scientifically-backed defense against RF exposure.
12. Which of the following is **NOT** a prevalent distance-based safety guideline mentioned in the podcast?
- A) Carry the device in a backpack or purse instead of against the body.
 - B) Leave the phone off the mattress and across the room while sleeping.
 - C) Keep the phone at least five inches away from the head even with Airplane Mode on.
 - D) Use the speaker-phone function or a wired headset during a call whenever possible.
13. As stated in the podcast, what mechanism triggered by weak signals (1-2 bars) makes them problematic for radiation exposure?
- A) The phone amplifier ramps up its power output to the allowed maximum to reach the tower.
 - B) Ambient signal noise at the user's location forces the phone to use more battery voltage.
 - C) The cell tower itself sends a high-voltage pulse to the device to force a one-way digital pairing.
 - D) The phone antenna extends its channel bandwidth to capture incoming waves that overlap.
14. Which of the following is **NOT** a reason cited for the failure of radiation shield and sticker products?
- A) The radiation shield tricks the device into assuming it is in a bad coverage area.
 - B) Internal sensors detect a drop in the signal-to-noise ratio due to the blockage.
 - C) The phone is trapped in a maximal-power feedback loop, increasing RF exposure.
 - D) Stickers interfere with GPS sensors, which start emitting high-frequency pings.
15. Which of the following was **NOT** considered by the hosts as a cause of power spikes during phone usage in a moving vehicle?
- A) The car's metal frame acts as a Faraday cage, blocking the outgoing signal.
 - B) A Bluetooth connection to the car itself creates interference, draining the signal.
 - C) The phone must constantly perform digital handshakes with new towers.
 - D) The device operates at peak power levels to maintain the shifting connection.

D.3. GMO Foods

1. How does recombinant DNA technology differ from traditional crossbreeding in its approach to transferring desirable traits?
- A) It uses radiation to force random mutations in the entire plant genome.
 - B) It involves the precise and targeted transfer of a specific beneficial gene.
 - C) It relies on mixing tens of thousands of genes to find a lucky combination.
 - D) It uses broad and unpredictable crosses between two unrelated species.
2. Which of the following was **NOT** part of the genetic solution used to save the Hawaiian papaya?
- A) Inserting a harmless fragment of the virus's own genetic code.

- B) Engineering the plant to produce a small amount of viral protein.
 - C) Modifying the plant to thrive in alkaline soil where the virus cannot survive.
 - D) Building up the plant's defenses similarly to an immune response from a vaccine.
3. Which of the following is **NOT** a characteristic of the Arctic apple's genetic modification?
- A) It prevents the fruit from turning brown after it has been sliced or bruised.
 - B) It aims to solve environmental and logistical problems related to food waste.
 - C) It works by silencing the specific gene responsible for the oxidative enzyme.
 - D) It introduces a gene from a tomato to increase the apple's lycopene levels.
4. Why was rice chosen as the target for the nutritional modification that yielded Golden rice?
- A) Rice is the only grain that can naturally hold beta carotene.
 - B) It is the primary diet in regions suffering from vitamin A deficiency.
 - C) Its genome is easiest to alter with sequences targeting childhood blindness.
 - D) Its color change was the side-effect with the least pushback from the general public.
5. Which popular physiological concern did one of the hosts raise regarding the prominence of GM crops in animal feed?
- A) Whether livestock eating edited crops might reproduce at a much faster rate.
 - B) Whether the pH level of cows' stomachs might drop further due to the GM diet.
 - C) Whether modified DNA might transfer into the meat, eggs, or milk consumers eat.
 - D) Whether cows might develop cellular receptors that make them immune to antibiotics.
6. Which of the following does **NOT** happen during the digestion of GM proteins in an animal's body?
- A) Hydrochloric acid in the stomach environment breaks down complex biological structures.
 - B) Aggressive digestive enzymes unwind the DNA and chop proteins into building blocks.
 - C) The animal's stomach recognizes the protein as a special-purpose biological blueprint.
 - D) The modified genetic material is completely degraded in the process of digestion.
7. According to the podcast, what is **NOT** an environmental disadvantage of traditional farmland tilling?
- A) It increases the nutritional toxicity of the resulting crop harvest.
 - B) It releases trapped carbon from the earth into the atmosphere.
 - C) It destroys the soil structure and causes massive topsoil erosion.
 - D) It requires the consumption of significant amounts of diesel fuel.
8. The crystalline protein produced by Bt corn (pronounced as "boot" in the podcast) becomes toxic only when it dissolves in what specific type of environment?
- A) A highly acidic environment like a mammal's stomach.
 - B) A highly alkaline environment like an insect pest's gut.
 - C) A carbon-rich environment like the soil of a tilled field.
 - D) A tailored environment like soil treated with insecticide.
9. Which argument did the hosts present to explain why Bt protein is safe for beneficial insects like ladybugs and bees?
- A) The protein only activates when it enters a pest's open circulatory system.
 - B) Ladybugs and bees have highly acidic guts that neutralize the crystalline protein akin to mammals' stomachs.

- C) The protein is only harmful to the European cornborer caterpillar as a pest.
→ **D)** Ladybugs and bees lack the specific cellular receptors necessary for the Bt protein to bind and become toxic.
10. What is the primary cause for the emergence of superbugs and superweeds in the environment?
- A) The accidental cross-pollination of Bt corn with local wildflower species.
→ **B)** The immense selective pressure from the continuous reliance on a single agricultural tool.
C) The accidental transfer of Bt genes into the soil's native bacterial population.
D) The lack of mass spectrometry analysis during the initial stages of GM seed development.
11. Which measurement is **NOT** the focus of a GM crop's substantial equivalence analysis (against its conventional counterpart)?
- A) The levels of every known amino acid.
B) The presence of naturally occurring toxins.
→ **C)** The speed and durability of gene insertion.
D) The full vitamin and biochemical profile.
12. What is **NOT** part of the scientific consensus from global health authorities regarding GMOs and cancer?
- A) Currently marketed GM foods pose no greater risk to health than conventional foods.
B) There is zero epidemiological evidence that GM crops cause cancer in humans.
C) Authoritative bodies, such as the WHO, are unified in their rulings on GMO safety.
→ **D)** Long-term population studies have linked GM soy to increased tumor growth.
13. Why did the hosts argue that the 2012 Seralini rodent study on GM corn and cancer was heavily flawed?
- **A)** The study used a strain of rat naturally prone to tumors regardless of diet.
B) The researchers cited a prior study of theirs that never actually existed.
C) The experiment did not use a control group to isolate the GM diet's effects.
D) The experimentation data was corrupted through a cyber panic in China.
14. What action did researchers take in the 1990s when testing showed that a new GM soybean contained an active allergen from a Brazil nut?
- A) They labeled the GM soy as a nut product and sold it only for animal feed.
→ **B)** They scrapped the product immediately before it could reach farms or markets.
C) They edited the soybean once again to remove the allergenic part of the protein.
D) They sought emergency approval to sell the product in developing nations.
15. According to the podcast, which of the following is **NOT** a driver of the sociological divide and lack of trust in GM foods?
- A) A feeling of losing control over the food supply.
B) Distrust towards the seed corporations that hold the patents.
→ **C)** The low volume of data released by the World Health Organization.
D) Sensational headlines on social media that create an echo chamber of fear.



Informed Consent Materials

E.1. Study 1 Opening Statement

You have been invited to participate in this research study on the salience of non-verbal sound cues by Konstantin-Asen Yordanov (TU Delft – Electrical Engineering, Mathematics, and Computer Science).

In this study, you will be asked to listen to short excerpts from a spoken conversation (e.g. a podcast) which contain one of the candidate sound cues we are investigating. Afterwards, you will provide your subjective impressions of the sound along different criteria using both judgment scales and free-form textual answers. Please take your participation seriously: responses that contain gibberish or are otherwise incoherent will be flagged during review.

The study will take **approximately 8 minutes**. If possible, **(i) connect and put on a pair of headphones** for improved audio clarity and **(ii) find a quiet environment where you will not be disturbed** before proceeding with the study. From this point on, **(iii) avoid closing the browser tab or refreshing the webpage** as that would reset your progress and redirect you back to the start of the survey.

You will also be asked for personal data, specifically your age, gender, and native language. The reason for requesting these details is to verify how balanced and representative the sample of participants is: the data will be presented and discussed in an aggregated and anonymized form alongside the study's main findings in a final report, which will be available in the public domain via TU Delft's open repository. All personal data will be deleted from our database at the end of the research project. If you are uncomfortable sharing any of this information, you may indicate the provided 'Prefer not to say' option. To minimize the risk of re-identification, we are going to store personal research data securely (deleting them when they are no longer needed), and only the Principal Investigator will have access to them to maintain privacy and confidentiality throughout the course of the research.

Your participation in this study is entirely voluntary, and you can withdraw at any time without adverse consequences by closing the survey. Partial responses that are meaningful and clearly show evidence of effortful participation will be reimbursed proportionally to the time spent. If you have questions or concerns, please contact the Principal Investigator (Konstantin-Asen Yordanov).

By clicking "Continue" below, you agree to the following: **I acknowledge that I have read and understood the information presented above, am aware of the research objectives and the use of the data I submit, and consent to be a participant in this study.**

E.2. Study 2 Opening Statement (Partial Disclosure)

You have been invited to participate in this research study on the immediate recall of podcast content by Konstantin-Asen Yordanov (TU Delft – Electrical Engineering, Mathematics, and Computer Science).

In this study, you will be asked to listen to a dialogue-formatted podcast (18-19 minutes) which tackles a scientific topic of high importance to society (e.g. fossil fuel exploitation). Beforehand, you will be presented with a few general questions that capture your prior knowledge on the topic and your podcast listening habits. After listening to the podcast in its entirety at least once, you will have access to the recall questionnaire that you are expected to complete as your main contribution. To keep the assessment authentic, you will not be able to return to the podcast audio player after transitioning to

the questionnaire. Please listen attentively, pause/backtrack if needed, and take your participation seriously: responses that contain gibberish or are otherwise incoherent will be flagged during review.

The study will take **approximately 45 minutes**. If possible, **(i) connect and put on a pair of headphones** for improved audio clarity and **(ii) find a quiet environment where you will not be disturbed** before proceeding with the study. From this point on, **(iii) avoid closing the browser tab or refreshing the webpage** as that would reset your answers in the page you were currently completing.

You will also be asked for personal data, specifically your age, gender, and native language. The reason for requesting these details is to verify how balanced and representative the sample of participants is: the data will be presented and discussed in an aggregated and anonymized form alongside the study's main findings in a final report, which will be available in the public domain via TU Delft's open repository. All personal data will be deleted from our database at the end of the research project. If you are uncomfortable sharing any of this information, you may indicate the provided 'Prefer not to say' option. To minimize the risk of re-identification, we are going to store personal research data securely (deleting them when they are no longer needed), and only the Principal Investigator will have access to them to maintain privacy and confidentiality throughout the course of the research.

Your participation in this study is entirely voluntary, and you can withdraw at any time without adverse consequences by clicking the Quit Study button. Partial responses that are meaningful and clearly show evidence of effortful participation will be reimbursed proportionally to the time spent. If you have questions or concerns, please contact the Principal Investigator (Konstantin-Asen Yordanov).

By clicking "Continue" below, you agree to the following: **I acknowledge that I have read and understood the information presented above, am aware of the research objectives and the use of the data I submit, and consent to be a participant in this study.**

E.3. Study 2 Debriefing Form (Full Disclosure)

We appreciate your participation in our study and the time you spent contributing to this research. The purpose of this post-study debrief is to provide crucial information about the task you just completed.

You were originally told that the aim of this study was to identify and analyze patterns of content recall measured immediately after participants listen to a dialogue-style podcast. While the recall assessment was indeed a significant task component, the study's actual objectives were more involved than what we explained to you.

The true purpose of this research is to compare different implementations of non-verbal auditory cues, which indicate that an adjacent podcast segment contains misleading statements. By examining how listeners perceive and react to such auditory warnings, this study informs the guidelines for integrating them into podcasts.

Under partial disclosure, the Opening Statement instead emphasized the interplay between attentive listening and memory as the research focus. It was necessary to obscure the remaining features and keep participants unaware of the auditory cues' presence and purpose to ensure the reactions we capture are authentic. Going into the survey with the knowledge of and mental preparation for encountering warnings would have otherwise skewed the results and detracted from the study's realism, invalidating our conclusions.

The process of handling personal data (i.e. age, gender, native language) remains as it was described in the Opening Statement. All personal information you shared (if any) is only featured as descriptive statistics of the participant sample in the final published report (presented in an aggregated and anonymized form). For the duration of the project, it is stored securely and accessible solely to the Principal Investigator, afterwards being deleted to minimize any risks to privacy and confidentiality.

We apologize for not offering you accurate information about the motivation, objectives, and procedure for this study, but we hope you understand why this was necessary. Due to the use of partial disclosure, we ask that you once again indicate whether you agree to the use and analysis of your submitted work via the checkboxes below. For any additional concerns or questions, please contact the Principal Investigator via Prolific or email.

- ☐ **I have read and understood the information presented above and am aware of the true research objectives.**
- ☐ **I approve the use and analysis of my submission.**



Web Application Frontend Views

Opening Statement

You have been invited to participate in this research study on the immediate recall of podcast content by Konstantin-Asen Yordanov ([email](#)) from TU Delft's Electrical Engineering, Mathematics, and Computer Science faculty.

In this study, you will be asked to listen to a dialogue-formatted podcast (18-19 minutes) which tackles a scientific topic of high importance to society (e.g. fossil fuel exploitation). Beforehand, you will be presented with a few general questions that capture your prior knowledge on the topic and your podcast listening habits. After listening to the podcast in its entirety at least once, you will have access to the recall questionnaire that you are expected to complete as your main contribution. To keep the assessment authentic, you will not be able to return to the podcast audio player after transitioning to the questionnaire. Please listen attentively, pause/backtrack if needed, and take your participation seriously: responses that contain gibberish or are otherwise incoherent will be flagged during review.

The study will take **approximately 45 minutes** of your time. If possible, **(i) connect and put on a pair of headphones** for improved audio clarity and **(ii) find a quiet environment where you will not be disturbed** before proceeding with the study. From this point on, **(iii) avoid closing the browser tab or refreshing the webpage** as that would reset your answers in the page you were currently completing.

You will also be asked for personal data, specifically your age, gender, and native language. The reason for requesting these details is to verify how balanced and representative the sample of participants is: the data will be presented and discussed in an aggregated and anonymized form alongside the study's main findings in a final report, which will be available in the public domain via TU Delft's open repository. All personal data will be deleted from our database at the end of the research project. If you are uncomfortable sharing any of this information, you may indicate the provided 'Prefer not to say' option. To minimize the risk of re-identification, we are going to store personal research data securely (deleting them when they are no longer needed), and only the Principal Investigator will have access to them to maintain privacy and confidentiality throughout the course of the research.

Your participation in this study is entirely voluntary, and you can withdraw at any time without adverse consequences by clicking the Quit Study button. Partial responses that are meaningful and clearly show evidence of effortful participation will be reimbursed proportionally to the time spent. If you have questions or concerns, please contact the Principal Investigator (Konstantin-Asen Yordanov).

By clicking 'Continue' below, you agree to the following: **I acknowledge that I have read and understood the information presented above, am aware of the research objectives and the use of the data I submit, and consent to be a participant in this study.**

Continue

Page 2: Podcast Listening

The podcast underpinning this study is accessible via the audio player below. Please listen attentively and freely use the Pause/Play and Backtrack buttons as needed. Note that the player will automatically pause if you switch focus to another window; after the full duration of the podcast has elapsed, the Continue button will become active.

Deep Dive: Nuclear Energy for Power Generation

Play

Back 10s

00:00

17:29

Quit Study

Continue

(a) Opening statement (Study 2)

Page 3: Recall Assessment

Based on what you remember from the podcast you just listened to, complete the questionnaire below. Each of the 15 content-related multiple-choice questions has one unique correct answer: please read both the question-text and the answer-options carefully before making your selection.

1. Which of the following is **NOT** a step in the nuclear fission chain reaction described in the podcast?

- ☐ A. An unstable atom splits and releases an immense amount of heat.
- ☐ B. Splitting atoms shoot out two or three additional neutrons.
- ☐ C. Newly released neutrons strike other nearby uranium atoms.
- ☐ D. Two uranium atoms merge together to form a heavier nucleus.

2. Why is nuclear fuel considered more attractive than burning fossil fuels according to the hosts?

- ☐ A. Because chemical reactions like burning are very difficult to control safely.
- ☐ B. Because nuclear reactions only rearrange surface-level electrons.
- ☐ C. Because only a small amount of fuel is needed to make a lot of power.
- ☐ D. Because fossil fuels require more expensive turbine technology.

3. Which of the following is **NOT** a characteristic of commercial nuclear fuel pellets?

- ☐ A. They are produced from a malleable metallic ore.
- ☐ B. They are processed into small, hard, black units.
- ☐ C. They are roughly the size of a marshmallow.
- ☐ D. They are composed of a solid ceramic material.

4. How did the hosts characterize the scale of a reactor meltdown and a nuclear explosion?

- ☐ A. A meltdown is more dangerous because it cannot be seen.
- ☐ B. A meltdown is a financial disaster but not a nuclear explosion.
- ☐ C. Meltdowns are only a potential risk in older 1970s reactor designs.
- ☐ D. An explosion is a mechanical failure, while a meltdown is chemical.

(b) Audio player for podcast listening (Study 2)

Can you recall hearing a distinctive non-verbal sound that might have occurred while the two hosts were talking?

- ☐ Yes
- ☐ No

How many times can you recall hearing the sound throughout the podcast?

- ☐ Once
- ☐ Twice
- ☐ Three times or more

What do you think the purpose of the sound was?

How dense did the podcast seem in terms of the amount of information it conveyed?

- ☐ 1 Too sparse
- ☐ 2 Sparse
- ☐ 3 Moderate
- ☐ 4 Cluttered
- ☐ 5 Too cluttered

Please rate the podcast's pace using the scale below.

- ☐ 1 Very slow
- ☐ 2 Slow
- ☐ 3 Moderate
- ☐ 4 Fast
- ☐ 5 Very fast

Quit Study

Continue

(c) Recall assessment questionnaire (Study 2)

(d) Warning reception under partial disclosure (Study 2)

Page 4: Veracity Assessment

Two short excerpts from the podcast are reproduced below. For each snippet, first listen to it all the way through and then indicate whether the discussion within is factually correct or contains misinformation.

Podcast Snippet 1

[Play](#) [Back 10s](#) 00:00 00:15

Is the entire discussion from the snippet factually accurate?

- ☐ Yes
☐ No

What makes you think so? (min. 1 sentence)

Podcast Snippet 2

[Play](#) [Back 10s](#) 00:00 00:15

Is the entire discussion from the snippet factually accurate?

- ☐ Yes
☐ No

What makes you think so? (min. 1 sentence)

[Quit Study](#)

[Continue](#)

Page 5: Sound Evaluation

This section of the study focuses on the non-verbal sound you encountered while listening to the podcast. Answer the questions below based on your subjective impressions, and (optionally) listen to the provided excerpts to remind yourself of the sound.

Podcast Snippet 1 (optional reminder)

▶ 0:00 / 0:26 [Speaker](#) [More](#)

Podcast Snippet 2 (optional reminder)

▶ 0:00 / 0:21 [Speaker](#) [More](#)

Podcast Snippet 3 (optional reminder)

▶ 0:00 / 0:28 [Speaker](#) [More](#)

What does the sound remind you of — does it match any real-world object or situation? Please type in what associations you make, and feel free to replay the provided audio snippets if you need to hear the sound again.

What was your reaction the first time you heard the sound while listening to the podcast? Please discuss the experience and your thoughts in the text box below (min. 1 sentence).

(e) Veracity assessment excerpts (Study 2)

Did your initial perceptions of the sound change as it re-occurred throughout the podcast? Please select an option...

- ☐ Became less pronounced
☐ Remained stable throughout
☐ Became more pronounced

...and elaborate in the text box below (min. 1 sentence).

Using the respective scale below, please indicate the extent to which:

(1 – Not at all, 2 – Slightly, 3 – Moderately, 4 – Very, 5 – Extremely)

The real-world entity you associate with the sound was easy to recognize upon hearing it.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
You found the sound distracting while listening.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
The sound seemed out of place whenever it occurred.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
The sound interrupted your overall listening experience.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
To verify you are completing the task attentively, please select "Slightly" for this question.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

[Quit Study](#)

[Continue](#)

Final Page: Post-Study Debrief

We appreciate your participation in our study and the time you spent contributing to this research. The purpose of this post-study debrief is to provide crucial information about the task you just completed.

You were originally told that the aim of this study was to identify and analyze patterns of content recall measured immediately after participants listen to a dialogue-style podcast. While the recall assessment was indeed a significant task component, the study's actual objectives were more involved than what we explained to you.

The true purpose of this research is to compare different implementations of non-verbal auditory cues, which indicate that an adjacent podcast segment contains misleading statements. By examining how listeners perceive and react to such auditory warnings, this study informs the guidelines for integrating them into podcasts.

Under partial disclosure, the Opening Statement instead emphasized the interplay between attentive listening and memory as the research focus. It was necessary to obscure the remaining features and keep participants unaware of the auditory cues' presence and purpose to ensure the reactions we capture are authentic. Going into the survey with the knowledge of and mental preparation for encountering warnings would have otherwise skewed the results and detracted from the study's realism, invalidating our conclusions.

The process of handling personal data (i.e. age, gender, native language) remains as it was described in the Opening Statement. All personal information you shared (if any) is only featured as descriptive statistics of the participant sample in the final published report (presented in an aggregated and anonymized form). For the duration of the project, it is stored securely and accessible solely to the Principal Investigator, afterwards being deleted to minimize any risks to privacy and confidentiality.

We apologize for not offering you accurate information about the motivation, objectives, and procedure for this study, but we hope you understand why this was necessary. Due to the use of partial disclosure, we ask that you once again indicate whether you agree to the use and analysis of your submitted work via the checkboxes below. For any additional concerns or questions, please contact the Principal Investigator via Prolific or [email](#).

☐ I have read and understood the information presented above and am aware of the true research objectives.

☐ I approve the use and analysis of my submission.

(g) Likert-matrix subjective evaluations (Studies 1 & 2) (h) Post-study debrief form (Study 2)

Figure F.1: Frontend views of the web application deployed for the two crowdsourcing studies.

Bibliography

- [1] Leif Azzopardi. “Cognitive Biases in Search: A Review and Reflection of Cognitive Biases in Information Retrieval”. In: *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*. CHIIR '21. Canberra ACT, Australia: Association for Computing Machinery, 2021, pp. 27–37. ISBN: 9781450380553. DOI: 10.1145/3406522.3446023. URL: <https://doi.org/10.1145/3406522.3446023>.
- [2] Ricardo Baeza-Yates. “Bias on the web”. In: *Commun. ACM* 61.6 (May 2018), pp. 54–61. ISSN: 0001-0782. DOI: 10.1145/3209581. URL: <https://doi.org/10.1145/3209581>.
- [3] Catherine Beauvais. “Fake news: Why do we believe it?”. In: *Joint Bone Spine* 89.4 (July 2022). Article 105371. ISSN: 1297-319X. DOI: 10.1016/j.jbspin.2022.105371. URL: <https://www.sciencedirect.com/science/article/pii/S1297319X22000306>.
- [4] Virginia Braun and Victoria Clarke. “Reflecting on reflexive thematic analysis”. In: *Qualitative Research in Sport, Exercise and Health* 11.4 (Aug. 2019), pp. 589–597. ISSN: 2159-676X. DOI: 10.1080/2159676X.2019.1628806. URL: <https://doi.org/10.1080/2159676X.2019.1628806>.
- [5] Virginia Braun and Victoria Clarke. “To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales”. In: *Qualitative Research in Sport, Exercise and Health* 13.2 (Mar. 2021), pp. 201–216. ISSN: 2159-676X. DOI: 10.1080/2159676X.2019.1704846. URL: <https://doi.org/10.1080/2159676X.2019.1704846>.
- [6] Virginia Braun and Victoria Clarke. “Using thematic analysis in psychology”. In: *Qualitative Research in Psychology* 3.2 (Jan. 2006), pp. 77–101. ISSN: 1478-0887. URL: <https://doi.org/10.1191/1478088706qp063oa>.
- [7] Bonnie Brinton Anderson et al. “How users perceive and respond to security messages: a NeurolS research agenda and empirical study”. In: *European Journal of Information Systems* 25.4 (July 2016), pp. 364–390. ISSN: 0960-085X. DOI: 10.1057/ejis.2015.21. URL: <https://doi.org/10.1057/ejis.2015.21>.
- [8] David Byrne. “A worked example of Braun and Clarke’s approach to reflexive thematic analysis”. In: *Quality & Quantity* 56.3 (June 2022), pp. 1391–1412. ISSN: 1573-7845. DOI: 10.1007/s11135-021-01182-y. URL: <https://doi.org/10.1007/s11135-021-01182-y>.
- [9] João Paulo Cabral and Gerard Bastiaan Remijn. “Auditory icons: Design and physical characteristics”. In: *Applied Ergonomics* 78 (July 2019), pp. 224–239. ISSN: 0003-6870. URL: <https://www.sciencedirect.com/science/article/pii/S0003687019300511>.
- [10] Danielle Caled and Mário J. Silva. “Digital media and misinformation: An outlook on multidisciplinary strategies against manipulation”. In: *Journal of Computational Social Science* 5.1 (May 2022), pp. 123–159. ISSN: 2432-2725. DOI: 10.1007/s42001-021-00118-8. URL: <https://doi.org/10.1007/s42001-021-00118-8>.
- [11] Anna L. Cox et al. “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. San Jose, California, USA: Association for Computing Machinery, 2016, pp. 1389–1397. ISBN: 9781450340823. DOI: 10.1145/2851581.2892410. URL: <https://doi.org/10.1145/2851581.2892410>.
- [12] Carlos Díaz Ruiz and Tomas Nilsson. “Disinformation and Echo Chambers: How Disinformation Circulates on Social Media Through Identity-Driven Controversies”. In: *Journal of Public Policy & Marketing* 42 (Jan. 2023), pp. 18–35. ISSN: 0743-9156. DOI: 10.1177/07439156221103852. URL: <https://doi.org/10.1177/07439156221103852>.

- [13] Verena Distler et al. "The Framework of Security-Enhancing Friction: How UX Can Help Users Behave More Securely". In: *Proceedings of the New Security Paradigms Workshop 2020*. Online, USA: Association for Computing Machinery, 2021, pp. 45–58. ISBN: 9781450389952. DOI: 10.1145/3442167.3442173. URL: <https://doi.org/10.1145/3442167.3442173>.
- [14] Karen M. Douglas, Robbie M. Sutton, and Aleksandra Cichocka. "The Psychology of Conspiracy Theories". In: *Current Directions in Psychological Science* 26.6 (2017), pp. 538–542. DOI: 10.1177/0963721417718261. eprint: <https://doi.org/10.1177/0963721417718261>. URL: <https://doi.org/10.1177/0963721417718261>.
- [15] Tim Draws et al. "A Checklist to Combat Cognitive Biases in Crowdsourcing". In: *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* 9.1 (Oct. 2021), pp. 48–59. DOI: 10.1609/hcomp.v9i1.18939. URL: <https://doi.org/10.1609/hcomp.v9i1.18939>.
- [16] Tim Draws et al. "This Is Not What We Ordered: Exploring Why Biased Search Result Rankings Affect User Attitudes on Debated Topics". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR '21. Virtual Event, Canada: Association for Computing Machinery, 2021, pp. 295–305. ISBN: 9781450380379. DOI: 10.1145/3404835.3462851. URL: <https://doi.org/10.1145/3404835.3462851>.
- [17] Asia A. Eaton and Penny S. Visser. "Attitude Importance: Understanding the Causes and Consequences of Passionately Held Views". In: *Social and Personality Psychology Compass* 2.4 (July 2008), pp. 1719–1736. ISSN: 1751-9004. DOI: 10.1111/j.1751-9004.2008.00125.x. URL: <https://doi.org/10.1111/j.1751-9004.2008.00125.x>.
- [18] Ullrich K. H. Ecker, Stephan Lewandowsky, and David T. W. Tang. "Explicit warnings reduce but do not eliminate the continued influence of misinformation". In: *Memory & Cognition* 38.8 (Dec. 2010), pp. 1087–1100. ISSN: 1532-5946. DOI: 10.3758/MC.38.8.1087. URL: <https://doi.org/10.3758/MC.38.8.1087>.
- [19] Ullrich K. H. Ecker et al. "The psychological drivers of misinformation belief and its resistance to correction". In: *Nature Reviews Psychology* 1.1 (Jan. 2022), pp. 13–29. ISSN: 2731-0574. DOI: 10.1038/s44159-021-00006-y. URL: <https://doi.org/10.1038/s44159-021-00006-y>.
- [20] Judy Edworthy and Rachael Hards. "Learning auditory warnings: The effects of sound type, verbal labelling and imagery on the identification of alarm sounds". In: *International Journal of Industrial Ergonomics* 24.6 (Oct. 1999), pp. 603–618. ISSN: 0169-8141. DOI: 10.1016/S0169-8141(98)00066-3. URL: <https://www.sciencedirect.com/science/article/pii/S0169814198000663>.
- [21] Judy Edworthy et al. "The Recognizability and Localizability of Auditory Alarms: Setting Global Medical Device Standards". In: *Human Factors* 59.7 (Nov. 2017), pp. 1108–1127. ISSN: 0018-7208. DOI: 10.1177/0018720817712004.
- [22] Judy Reed Edworthy, Natasha Talbot, and Nicole Martin. "Responding to clinical alarms in unfolding simulated clinical scenarios: auditory icons perform better than tonal alarms". In: *British Journal of Anaesthesia* 134.6 (June 2025), pp. 1773–1778. ISSN: 0007-0912. DOI: 10.1016/j.bja.2024.12.047. URL: <https://www.sciencedirect.com/science/article/pii/S0007091225001357>.
- [23] Carsten Eickhoff. "Cognitive Biases in Crowdsourcing". In: *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. WSDM '18. Marina Del Rey, CA, USA: Association for Computing Machinery, 2018, pp. 162–170. ISBN: 9781450355810. DOI: 10.1145/3159652.3159654. URL: <https://doi.org/10.1145/3159652.3159654>.
- [24] Aaron M. French, Veda C. Storey, and Linda Wallace. "The impact of cognitive biases on the believability of fake news". In: *European Journal of Information Systems* 34.1 (Jan. 2025), pp. 72–93. ISSN: 0960-085X. DOI: 10.1080/0960085X.2023.2272608. URL: <https://doi.org/10.1080/0960085X.2023.2272608>.
- [25] Tim Gorichanaz. "Relating information seeking and use to intellectual humility". In: *J. Assoc. Inf. Sci. Technol.* 73.5 (Apr. 2022), pp. 643–654. ISSN: 2330-1635. DOI: 10.1002/asi.24567. URL: <https://doi.org/10.1002/asi.24567>.

- [26] Chen Guo, Nan Zheng, and Chengqi (John) Guo. "Seeing is Not Believing: A Nuanced View of Misinformation Warning Efficacy on Video-Sharing Social Media Platforms". In: *Proc. ACM Hum.-Comput. Interact.* 7.CSCW2 (Oct. 2023). DOI: 10.1145/3610085. URL: <https://doi.org/10.1145/3610085>.
- [27] Megan Haggard et al. "Finding middle ground between intellectual arrogance and intellectual servility: Development and assessment of the limitations-owning intellectual humility scale". In: *Personality and Individual Differences* 124 (2018), pp. 184–193. ISSN: 0191-8869. DOI: 10.1016/j.paid.2017.12.014. URL: <https://www.sciencedirect.com/science/article/pii/S0191886917307286>.
- [28] Jutta Haider and Olof Sundin. "Information literacy challenges in digital culture: conflicting engagements of trust and doubt". In: *Information, Communication & Society* 25.8 (June 2022), pp. 1176–1191. ISSN: 1369-118X. DOI: 10.1080/1369118X.2020.1851389. URL: <https://doi.org/10.1080/1369118X.2020.1851389>.
- [29] Amelia Hassoun et al. "Practicing Information Sensibility: How Gen Z Engages with Online Information". In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Hamburg, Germany: Association for Computing Machinery, 2023. ISBN: 9781450394215. DOI: 10.1145/3544548.3581328. URL: <https://doi.org/10.1145/3544548.3581328>.
- [30] Curtis P. Haugtvedt and Richard E. Petty. "Personality and persuasion: Need for cognition moderates the persistence and resistance of attitude changes". In: *Journal of Personality and Social Psychology* 63.2 (1992), pp. 308–319. DOI: 10.1037/0022-3514.63.2.308. URL: <https://doi.org/10.1037/0022-3514.63.2.308>.
- [31] Danula Hettiachchi et al. "Designing and Evaluating Presentation Strategies for Fact-Checked Content". In: *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*. CIKM '23. Birmingham, United Kingdom: Association for Computing Machinery, 2023, pp. 751–761. ISBN: 9798400701245. DOI: 10.1145/3583780.3614841. URL: <https://doi.org/10.1145/3583780.3614841>.
- [32] Allyson L. Holbrook et al. "Attitude Importance and the Accumulation of Attitude-Relevant Knowledge in Memory". In: *Journal of Personality and Social Psychology* 88.5 (2005), pp. 749–769. DOI: 10.1037/0022-3514.88.5.749. URL: <https://doi.org/10.1037/0022-3514.88.5.749>.
- [33] Rick H. Hoyle et al. "Holding specific views with humility: Conceptualization and measurement of specific intellectual humility". In: *Personality and Individual Differences* 97 (2016), pp. 165–172. ISSN: 0191-8869. DOI: 10.1016/j.paid.2016.03.043. URL: <https://www.sciencedirect.com/science/article/pii/S0191886916301970>.
- [34] Farnaz Jahanbakhsh et al. "Exploring Lightweight Interventions at Posting Time to Reduce the Sharing of Misinformation on Social Media". In: *Proc. ACM Hum.-Comput. Interact.* 5.CSCW1 (Apr. 2021). DOI: 10.1145/3449092. URL: <https://doi.org/10.1145/3449092>.
- [35] Rosie Jones et al. "Current Challenges and Future Directions in Podcast Information Access". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR '21. Virtual Event, Canada: Association for Computing Machinery, 2021, pp. 1554–1565. ISBN: 9781450380379. DOI: 10.1145/3404835.3462805. URL: <https://doi.org/10.1145/3404835.3462805>.
- [36] Robert Kaufman, Aaron Broukhim, and Michael Haupt. "WARNING This Contains Misinformation: The Effect of Cognitive Factors, Beliefs, and Personality on Misinformation Warning Tag Attitudes". In: *Proc. ACM Hum.-Comput. Interact.* 9.7 (Oct. 2025). DOI: 10.1145/3757521. URL: <https://doi.org/10.1145/3757521>.
- [37] Jan Kirchner and Christian Reuter. "Countering Fake News: A Comparison of Possible Solutions Regarding User Acceptance and Effectiveness". In: *Proc. ACM Hum.-Comput. Interact.* 4.CSCW2 (Oct. 2020). DOI: 10.1145/3415211. URL: <https://doi.org/10.1145/3415211>.

- [38] Anastasia Kozyreva, Stephan Lewandowsky, and Ralph Hertwig. "Citizens Versus the Internet: Confronting Digital Challenges With Cognitive Tools". In: *Psychological Science in the Public Interest* 21.3 (2020), pp. 103–156. DOI: 10.1177/1529100620946707. eprint: <https://doi.org/10.1177/1529100620946707>. URL: <https://doi.org/10.1177/1529100620946707>.
- [39] Zhaoli Lei et al. "The Impact of Different Types of Auditory Warnings on Working Memory". In: *Frontiers in Psychology* 13 (2022). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2022.780657. URL: <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.780657>.
- [40] Stephan Lewandowsky et al. "Misinformation and Its Correction: Continued Influence and Successful Debiasing". In: *Psychological Science in the Public Interest* 13.3 (Dec. 2012), pp. 106–131. ISSN: 1529-1006. DOI: 10.1177/1529100612451018. URL: <https://doi.org/10.1177/1529100612451018>.
- [41] Gabriel Lins de Holanda Coelho, Paul H. P. Hanel, and Lukas J. Wolf. "The Very Efficient Assessment of Need for Cognition: Developing a Six-Item Version". In: *Assessment* 27.8 (Dec. 2020), pp. 1870–1885. ISSN: 1073-1911. DOI: 10.1177/1073191118793208. URL: <https://doi.org/10.1177/1073191118793208>.
- [42] Yu-Jeh Liu and Mounya Elhilali. "Sound identity, salience, and perceived importance in complex auditory environments". In: *The Journal of the Acoustical Society of America* 158.4 (Oct. 2025), pp. 3489–3502. ISSN: 0001-4966. DOI: 10.1121/10.0039710. URL: <https://doi.org/10.1121/10.0039710>.
- [43] Philipp Lorenz-Spreen et al. "How behavioural sciences can promote truth, autonomy and democratic discourse online". In: *Nature Human Behaviour* 4.11 (Nov. 2020), pp. 1102–1109. ISSN: 2397-3374. DOI: 10.1038/s41562-020-0889-7. URL: <https://doi.org/10.1038/s41562-020-0889-7>.
- [44] Andrew Luttrell and Vanessa Sawicki. "Attitude strength: Distinguishing predictors versus defining features". In: *Social and Personality Psychology Compass* 14.8 (Aug. 2020). ISSN: 1751-9004. DOI: 10.1111/spc3.12555. URL: <https://doi.org/10.1111/spc3.12555>.
- [45] Andrew Luttrell et al. "Making it moral: Merely labeling an attitude as moral increases its strength". In: *Journal of Experimental Social Psychology* 65 (July 2016), pp. 82–93. ISSN: 0022-1031. DOI: 10.1016/j.jesp.2016.04.003. URL: <https://www.sciencedirect.com/science/article/pii/S0022103116302372>.
- [46] Cameron Martel, Gordon Pennycook, and David G. Rand. "Reliance on emotion promotes belief in fake news". In: *Cognitive Research: Principles and Implications* 5.47 (Oct. 2020). ISSN: 2365-7464. DOI: 10.1186/s41235-020-00252-3. URL: <https://doi.org/10.1186/s41235-020-00252-3>.
- [47] Cameron Martel and David G. Rand. "Misinformation warning labels are widely effective: A review of warning effects and their moderating features". In: *Current Opinion in Psychology* 54 (2023). ISSN: 2352-250X. DOI: 10.1016/j.copsy.2023.101710. URL: <https://www.sciencedirect.com/science/article/pii/S2352250X23001550>.
- [48] Sungkyu Park et al. "An Experimental Study to Understand User Experience and Perception Bias Occurred by Fact-checking Messages". In: *Proceedings of the Web Conference 2021*. WWW '21. Ljubljana, Slovenia: Association for Computing Machinery, 2021, pp. 2769–2780. ISBN: 9781450383127. DOI: 10.1145/3442381.3450121. URL: <https://doi.org/10.1145/3442381.3450121>.
- [49] Sachin Pathiyen Cherumanal, Ujwal Gadiraju, and Damiano Spina. "Everything We Hear: Towards Tackling Misinformation in Podcasts". In: *Proceedings of the 26th International Conference on Multimodal Interaction*. ICMI '24. San Jose, Costa Rica: Association for Computing Machinery, 2024, pp. 596–601. ISBN: 9798400704628. DOI: 10.1145/3678957.3678959. URL: <https://doi.org/10.1145/3678957.3678959>.
- [50] Helen Petrie and Sarah Morley. "The use of non-speech sounds in non-visual interfaces to the MS-Windows GUI for blind computer users". In: *Proceedings of the 1998 International Conference on Auditory Display*. ICAD'98. UK: BCS Learning & Development Ltd., 1998.

- [51] Richard E. Petty, Curtis P. Haugtvedt, and Stephen M. Smith. "Elaboration as a determinant of attitude strength: Creating attitudes that are persistent, resistant, and predictive of behavior". In: Ohio State University series on attitudes and persuasion, Vol. 4. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc, 1995, pp. 93–130. ISBN: 0-8058-1086-2.
- [52] Tenelle Porter et al. "Predictors and consequences of intellectual humility". In: *Nature Reviews Psychology* 1.9 (Sept. 2022), pp. 524–536. ISSN: 2731-0574. DOI: 10.1038/s44159-022-00081-9. URL: <https://doi.org/10.1038/s44159-022-00081-9>.
- [53] Alisa Rieger et al. "Disentangling Web Search on Debated Topics: A User-Centered Exploration". In: *Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*. UMAP '24. Cagliari, Italy: Association for Computing Machinery, 2024, pp. 24–35. ISBN: 9798400704338. DOI: 10.1145/3627043.3659559. URL: <https://doi.org/10.1145/3627043.3659559>.
- [54] Alisa Rieger et al. "Responsible Opinion Formation on Debated Topics in Web Search". In: *Advances in Information Retrieval: 46th European Conference on Information Retrieval, ECIR 2024, Glasgow, UK, March 24–28, 2024, Proceedings, Part IV*. Glasgow, United Kingdom: Springer-Verlag, 2024, pp. 437–465. ISBN: 978-3-031-56065-1. DOI: 10.1007/978-3-031-56066-8_32. URL: https://doi.org/10.1007/978-3-031-56066-8_32.
- [55] Jon Roozenbeek, Miriam Remshard, and Yara Kyrychenko. "Beyond the headlines: On the efficacy and effectiveness of misinformation interventions". In: *advances.in/psychology* 2 (July 2024). DOI: 10.56296/aip00019.
- [56] Bjorn Ross et al. "Fake News on Social Media: The (In)Effectiveness of Warning Messages". English. In: *Proceedings of the 39th International Conference on Information Systems (ICIS 2018)*. United States: Association for Information Systems, Dec. 2018. URL: <https://www.research.ed.ac.uk/en/publications/fake-news-on-social-media-the-ineffectiveness-of-warning-messages/>.
- [57] Emily Saltz, Claire R Leibowicz, and Claire Wardle. "Encounters with Visual Misinformation and Labels Across Platforms: An Interview and Diary Study to Inform Ecosystem Approaches to Misinformation Interventions". In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI EA '21. Yokohama, Japan: Association for Computing Machinery, 2021. ISBN: 9781450380959. DOI: 10.1145/3411763.3451807. URL: <https://doi.org/10.1145/3411763.3451807>.
- [58] Linda J. Skitka, Anthony N. Washburn, and Timothy S. Carsel. "The psychological foundations and consequences of moral conviction". In: *Current Opinion in Psychology* 6 (Dec. 2015), pp. 41–44. ISSN: 2352-250X. DOI: 10.1016/j.copsyc.2015.03.025. URL: <https://www.sciencedirect.com/science/article/pii/S2352250X1500130X>.
- [59] Stephanie J. Tobin and Rosanna E. Guadagno. "Why people listen: Motivations and outcomes of podcast listening". In: *PLOS ONE* 17.4 (Apr. 2022), pp. 1–16. DOI: 10.1371/journal.pone.0265806. URL: <https://doi.org/10.1371/journal.pone.0265806>.
- [60] Manos Tsagkias et al. "PodCred: a framework for analyzing podcast preference". In: *Proceedings of the 2nd ACM Workshop on Information Credibility on the Web*. WICOW '08. Napa Valley, California, USA: Association for Computing Machinery, 2008, pp. 67–74. ISBN: 9781605582597. DOI: 10.1145/1458527.1458545. URL: <https://doi.org/10.1145/1458527.1458545>.
- [61] Toni G. L. A. van der Meer, Michael Hameleers, and Jakob Ohme. "Can Fighting Misinformation Have a Negative Spillover Effect? How Warnings for the Threat of Misinformation Can Decrease General News Credibility". In: *Journalism Studies* 24.6 (Apr. 2023), pp. 803–823. ISSN: 1461-670X. DOI: 10.1080/1461670X.2023.2187652. URL: <https://doi.org/10.1080/1461670X.2023.2187652>.
- [62] Johan L.H. van Strien et al. "How attitude strength biases information processing and evaluation on the web". In: *Computers in Human Behavior* 60 (2016), pp. 245–252. ISSN: 0747-5632. DOI: 10.1016/j.chb.2016.02.057. URL: <https://www.sciencedirect.com/science/article/pii/S0747563216301157>.

- [63] Ward van Zoonen, Vilma Luoma-aho, and Matias Lievonon. "Trust but verify? Examining the role of trust in institutions in the spread of unverified information on social media". In: *Computers in Human Behavior* 150 (Jan. 2024), pp. 1–9. ISSN: 0747-5632. DOI: 10.1016/j.chb.2023.107992. URL: <https://www.sciencedirect.com/science/article/pii/S0747563223003436>.
- [64] Anthony Vance et al. "What Do We Really Know about How Habituation to Warnings Occurs Over Time? A Longitudinal fMRI Study of Habituation and Polymorphic Warnings". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, pp. 2215–2227. ISBN: 9781450346559. DOI: 10.1145/3025453.3025896. URL: <https://doi.org/10.1145/3025453.3025896>.
- [65] Dáša Vedejová and Vladimíra Čavojová. "Confirmation bias in information search, interpretation, and memory recall: evidence from reasoning about four controversial topics". In: *Thinking & Reasoning* 28.1 (Jan. 2022), pp. 1–28. ISSN: 1354-6783. DOI: 10.1080/13546783.2021.1891967. URL: <https://doi.org/10.1080/13546783.2021.1891967>.
- [66] Valentina Vellani et al. "The illusory truth effect leads to the spread of misinformation". In: *Cognition* 236 (July 2023). ISSN: 0010-0277. DOI: 10.1016/j.cognition.2023.105421. URL: <https://www.sciencedirect.com/science/article/pii/S0010027723000550>.
- [67] Junjie Wang et al. "From SERPs to Sound: How Search Engine Result Pages and AI-generated Podcasts Interact to Influence User Attitudes on Controversial Topics". In: *Proceedings of the 2026 Conference on Human Information Interaction and Retrieval*. CHIIR '26. New York, NY, USA: Association for Computing Machinery, 2026, pp. 51–61. ISBN: 9798400724145. DOI: 10.1145/3786304.3787942. URL: <https://doi.org/10.1145/3786304.3787942>.
- [68] Luyan Xu, Mengdie Zhuang, and Ujwal Gadiraju. "How Do User Opinions Influence Their Interaction With Web Search Results?" In: *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*. UMAP '21. Utrecht, Netherlands: Association for Computing Machinery, 2021, pp. 240–244. ISBN: 9781450383660. DOI: 10.1145/3450613.3456824. URL: <https://doi.org/10.1145/3450613.3456824>.
- [69] Yixuan Zhang et al. "Shifting Trust: Examining How Trust and Distrust Emerge, Transform, and Collapse in COVID-19 Information Seeking". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: Association for Computing Machinery, 2022. ISBN: 9781450391573. DOI: 10.1145/3491102.3501889. URL: <https://doi.org/10.1145/3491102.3501889>.
- [70] Aljaž Zrnc, Marko Poženel, and Dejan Lavbič. "Users' ability to perceive misinformation: An information quality assessment approach". In: *Inf. Process. Manage.* 59.1 (Jan. 2022). ISSN: 0306-4573. DOI: 10.1016/j.ipm.2021.102739. URL: <https://doi.org/10.1016/j.ipm.2021.102739>.