Regular paper

Journal of Cognitive Engineering
and Decision Making
2026, Vol. 0(0) 1–39

DOI: 10.1177/15553434261472960
journals.sagepub.com/home/edm


Susceptibility to Misinformation: Exploring Persuasion and Prior Orientation in Recognition-Primed Decision Pathways

Boshra T. Haghighi[1] **and B. L. William Wong**[1]

Abstract

Misinformation poses growing risks to public well-being, contributing to vaccine refusal, climate inaction, and erosion of democratic institutions. A body of research has examined who is vulnerable to misinformation and the psychological and social mechanisms shaping belief, including analytic reasoning, motivated cognition, identity, and social and source cues. However, less attention has been given to how individuals evaluate misinformation in context and how this evaluation interacts with their prior topic-specific associative orientations. This study investigates how these orientations and persuasive framing interact within recognition-primed decision pathways when individuals encounter narratives. Using a qualitative Cognitive Task Analysis guided by the Recognition-Primed Decision framework, six participants evaluated misinformation narratives across three domains: vaccines, climate change, and bioweapons, presented using systematically varied rhetorical framings. Implicit Association Tests were used descriptively to indicate topic-specific associative patterns. Reasoning processes were examined through think-aloud protocols and Critical Decision Method interviews, tracing cue recognition, interpretation, and judgment. The findings illustrate how alignment between narrative cues and participants' associative orientations can reinforce confirmation tendencies, whereas mismatches may trigger skepticism and reconsideration. The study provides an exploratory process-level account of how persuasion cues interact with prior orientations during recognition-based decision pathways, illustrating susceptibility as a context-sensitive evaluation process.

[1]Auckland University of Technology (AUT), Auckland, New Zealand

Corresponding Authors:
Boshra T. Haghighi, Auckland University of Technology (AUT), Bay View, Auckland, 0629, New Zealand.
Email: boshra.talebihaghighi@aut.ac.nz

B. L. William Wong, Auckland University of Technology (AUT), Bay View, Auckland, 0629, New Zealand.
Email: william.wong@aut.ac.nz

Keywords
misinformation, susceptibility, persuasion models, implicit association, cognitive task analysis, recognition-primed decision model, decision-making processes

Introduction

Misinformation is defined as false or misleading information shared unknowingly without intent to deceive or cause harm (Wardle, 2018a). It has become a persistent challenge in contemporary information environments, influencing public health behavior, political attitudes, and responses to scientific evidence (Lewandowsky et al., 2017; van der Linden, 2022). Susceptibility reflects the degree to which someone is prone to be affected by personal traits, cognitive, and situational variables (Wardle & Singerman, 2021). In the context of misinformation, it refers to an individual's vulnerability to believing misleading or persuasive narratives.

A substantial body of research has examined factors associated with belief in misinformation, including analytic reasoning ability, motivated cognition, social identity dynamics, and the role of social and source cues in shaping credibility judgments. These studies have advanced the understanding of who may be vulnerable to misinformation and which psychological mechanisms are associated with belief (Ecker et al., 2022; Pennycook & Rand, 2021).

However, less is known about how individuals actually evaluate misinformation when encountering persuasive narratives in real time. Much of the existing literature relies on survey-based measures, accuracy judgments, or experimental manipulations of message features (Pennycook & Rand, 2020; Traberg et al., 2024). While these approaches have identified important psychological correlates of misinformation belief (Ecker et al., 2022; Hubeny, Nahon, Ng, & Gawronski, 2025; Pennycook & Rand, 2021), they offer more limited insight into the real-time cognitive processes through which individuals interpret, assess, and respond to persuasive claims during the moment of evaluation.

Decision-making research within the naturalistic decision-making (NDM) paradigm (Klein & Zsambok, 1997) offers a useful perspective for examining these processes. In complex and uncertain environments, individuals often rely on recognition-based judgments, drawing on prior knowledge and experience to interpret cues and generate plausible interpretations of unfolding situations. The recognition-primed decision (RPD) model (Klein, 1993) describes how decision makers identify cues, match them to stored patterns, and mentally simulate potential consequences to guide action. Although this framework has been widely applied to domains such as emergency response, medicine, and military decision-making, it has rarely been applied to the evaluation of persuasive information narratives.

At the same time, research on persuasive communication design shows that narratives can be framed to emphasize particular cues, values, or interpretive pathways (Entman, 1993). These cues may interact with individuals' prior topic-specific associative orientations, guiding how narratives are interpreted and whether they appear plausible or suspicious (Kunda, 1990; Roozenbeek et al., 2022). Understanding how such cues are recognized and interpreted during evaluation may help explain why similar messages lead to different judgments across individuals and contexts.

To explore these processes, this study adopts cognitive task analysis (CTA) (Crandall et al., 2006). CTA methods are widely used in human factors research to investigate cognitive processes underlying performance in complex environments, particularly behaviors that are not directly observable, such as how individuals recognize

cues, interpret situations, and make judgments in context. Consistent with HCI discussions on improving analytic sensitivity through repeated observations within participants (Dix, 2020), the study triangulates three sources of evidence: think-aloud protocols, Critical Decision Method interviews (Klein et al., 1989), and Implicit Association Tests (IAT) (Greenwald et al., 1998) used descriptively to indicate topic-specific associative patterns. The study adopts a qualitative CTA approach with structured stimulus variation, using controlled contrasts to examine how decision processes unfold within participants.

Participants evaluated misinformation narratives across three domains: vaccines, climate change, and bioweapons, presented using systematically varied rhetorical framings. These variations were used as structured contrasts to probe how cue–predisposition alignment influences interpretation and plausibility judgments within decision episodes. The analysis focuses on how participants recognized cues, formed interpretations, and reached plausibility judgments while interacting with these narratives.

Accordingly, the study investigates how prior topic-specific associative orientations and persuasive narrative cues interact during recognition-based evaluation of misinformation. Rather than seeking to explain stable differences in susceptibility across populations or estimate population-level effects, the study provides an exploratory, process-level account of how individuals evaluate misinformation narratives in context, highlighting the recognition dynamics that shape plausibility judgments. For readability, a glossary of abbreviations used throughout the paper is provided in Table 1.

Table 1. Glossary of Abbreviations Used in the Paper

Full term	Acronym
Aggregated cognitive strategies	ACSs
Cognitive reliability analysis	CRA
Cognitive strategies	CSs
Cognitive task analysis	CTA
Critical decision method	CDM
Elaboration likelihood model	ELM
Emergent themes approach	ETA
Heuristic systematic model	HSM
Hovland's attitude change approach	HACA
Human—computer interaction	HCI
Implicit association test	IAT
Naturalistic decision-making	NDM
Performance shaping factors	PSF
Persuasion knowledge model	PKM
Recognition-primed decision model	RPD
Theory of reasoned action	TRA

Background and Related Works

This section develops the conceptual basis for the study in four steps: it first defines misinformation and susceptibility as a receiver-side evaluation problem; it then examines how narratives and persuasion models structure persuasive cues; next, it considers prior topic-specific associative orientations and relevant decision-making models; finally, it brings these strands together through the RPD lens to motivate the study's analytical framework.

Misinformation and Characteristics

Misinformation is defined as incorrect or misleading information presented as authentic and shared unknowingly without intent to deceive, but often with harmful consequences (Wardle & Derakhshan, 2017). In contrast, disinformation is the deliberate and intentional use of information to mislead people and cause harm (Wardle, 2018b). When orchestrated and coordinated, disinformation can become a powerful and non-obvious instrument of propaganda, used to cause confusion, anger, and disagreement, and to sway opinions in order to polarize communities, undermine trust in governments, and interfere with elections (EUCommission, 2018; Roozenbeek, Schneider et al., 2020; Roozenbeek & van der Linden, 2019b, 2020). However, this paper does

not address this aspect of disinformation, as sender intent is outside the scope of analysis (Freelon & Wells, 2020; Tucker et al., 2018; Wardle, 2018b).

In this paper, misinformation is understood broadly as false, misleading, or deceptively framed information that may contain fabricated, distorted, or selectively presented elements. The focus of this study is therefore not limited to factual inaccuracy but includes how individuals interpret and evaluate persuasive narratives when the reliability of the information is uncertain.

Misinformation can take various forms, ranging from satire and parody to manipulated or fabricated content, with misleading framing often making factually accurate elements appear deceptive (Wardle & Derakhshan, 2017). Socio-psychological factors drive the spread of misinformation, with individuals sharing content to connect with like-minded groups, which can be amplified by the ubiquity of digital media (Roozenbeek & van der Linden, 2019a, 2022; Roozenbeek, van der Linden et al., 2020). The danger of misinformation lies not only in its falsity but in its ability to appear authentic, be presented through misleading framing techniques, and circulate rapidly in ways that undermine reliability analysis and decision-making (Roozenbeek & van der Linden, 2022).

Susceptibility to Misinformation

Susceptibility refers to the degree of vulnerability individuals have when influenced by particular stimuli or contexts, which in this research is the encounter with misinformation (Wardle & Singerman, 2021). Susceptibility is not a fixed trait but reflects the interaction of multiple factors, including cognitive biases, uncertainty, time pressure, prior knowledge, emotional and motivational states, and group identity commitments (Maertens et al., 2023; Nan et al., 2022). Different levels of susceptibility produce varied responses (Calvillo, 2014). Understanding susceptibility is crucial because it links human factors with the rhetorical and decision-making processes that may contribute to vulnerability to misinformation (Downs et al., 2006; van der Linden, 2022). Thus, susceptibility serves as the conceptual bridge between the cognitive variables individuals bring to an encounter with misinformation and the persuasion and decision pathways investigated in this study.

How Narratives and Framing Organize Information Into Persuasive Structures

Misinformation rarely appears as isolated claims. It is embedded within narratives that organize facts, values, and emotions into persuasive structures. Narratives provide coherence by linking events to causes and consequences, making even misleading information feel credible (Entman, 1993). Within these narratives, framing determines which aspects of reality are emphasized, downplayed, or omitted, thereby guiding interpretation and evaluation (Entman, 1993, 2004; Goffman, 1974).

Framing theory, developed by Goffman (1974), conceptualizes frames as mental structures that shape how people perceive and interpret information. In media research, framing has been shown to influence public perception by highlighting particular features of an issue while obscuring others (Entman, 1993; Entman & Rojecki, 1993). Frames are not neutral; they are produced through cultural norms, organizational routines, and ideological inclinations, a process known as frame-building (Gamson & Modigliani, 1989). Once established, frames channel attention and interpretation, shaping not only opinions but also behavioral responses (Scheufele, 1999).

The persuasive power of framing extends to decision-making, where it alters evaluative criteria. Classic studies by Tversky and Kahneman (1986) demonstrated that identical outcomes, such as survival rates versus mortality rates, produce divergent judgments when framed differently. In the context of misinformation, this means that misleading claims can gain traction not because of their factual accuracy, but because of how they are framed to resonate with prior beliefs,

emotional triggers, or cognitive shortcuts (Lewandowsky et al., 2017).

Thus, narratives and framing embed cues that interact with individuals' prior orientations during evaluation. Depending on how these cues are interpreted, they may reinforce existing beliefs, prompt skepticism, or encourage further sensemaking. Understanding these mechanisms is essential to explaining how misinformation appears persuasive, why its effects vary across individuals, and how persuasion models can be used to analyze its impact.

Persuasion Models

Having discussed how misinformation narratives can be shaped through framing, the next step is to consider the persuasion mechanisms that may make such frames effective. The literature identifies a set of recurring susceptibility indicators: trust, emotion, values, group identity, thinking style, and knowledge that are frequently associated with vulnerability to misinformation (Nan et al., 2022). These indicators mark points of susceptibility and may provide conditions under which persuasive techniques become more influential during evaluation. To understand how such techniques work, we draw on established persuasion models commonly used and classical rhetorical strategies that explain message effectiveness.

Several persuasion models have been discussed in the literature, including the Persuasion Knowledge Model (PKM) by Friestad and Wright (1994), the Elaboration Likelihood Model (ELM) by Petty and Cacioppo (1986), the Theory of Reasoned Action (TRA) by Fishbein and Ajzen (2011), the Heuristic Systematic Model (HSM) by Chaiken (1980), and Hovland's Attitude Change Approach (HACA) (Hovland et al., 1953). Three models, as reviewed below, were selected and used in this study to create different narrative versions used as stimuli, each representing a distinct but complementary mechanism of persuasion.

Persuasion Knowledge Model (PKM) originally developed in marketing and consumer research by Friestad and Wright (1994). PKM explains the dynamic between a persuader and a target, focusing on how people learn to recognize persuasion attempts and how that recognition affects persuasiveness. When persuasive intent is obvious (e.g., a pushy salesperson), resistance is more likely; when it is subtle or embedded (e.g., product placement in a movie), persuasion is more likely to succeed because the consumer's "persuasion knowledge" is not triggered. In misinformation contexts, PKM helps explain why resistance may fail when cues are unfamiliar or disguised.

The Elaboration Likelihood Model (ELM) was developed by Petty and Cacioppo (1986) to explain how people process persuasive messages in advertising, political campaigns, and public health. ELM proposes two routes: a central route based on careful, logical evaluation and a peripheral route relying on quick, surface-level cues. For example, a well-reasoned campaign on the benefits of seatbelt uses appeals to the central route, while a celebrity-endorsed soft drink ad uses the peripheral route. In misinformation contexts, emotional triggers and credibility cues may encourage greater reliance on peripheral-route processing, particularly when individuals are under cognitive load or lack motivation to process information deeply.

The Theory of Reasoned Action (TRA) was introduced by Fishbein and Ajzen (2011) in health, consumer, and environmental behavior research. TRA links attitudes toward behavior and perceived social norms to behavioral intentions, which predict actual behavior. For example, a person may decide to recycle because they believe it benefits the planet (attitude) and because most people they respect recycle (social norm). Persuasion operates by changing beliefs about the outcomes of behavior or by influencing perceived norms. In misinformation contexts, altering these determinants may influence public

behavior by reshaping attitudes or changing perceptions of what others think or do.

This study focuses on these three models because they represent analytically distinct and complementary mechanisms while jointly covering the core characteristics relevant to misinformation contexts. ELM explains variation in how people process persuasive information under cognitive constraint. HSM and HACA were excluded because they address similar heuristic and cue-based processes and therefore added limited explanatory value beyond ELM. In contrast, PKM addresses persuasion awareness, while TRA links beliefs and perceived norms to behavioral intention. Together, these models provided a complementary basis for structuring the narrative stimuli.

Aristotle's Rhetorical Communication Strategies

Persuasive design often draws on Aristotle's classical rhetorical strategies: ethos (credibility appeal), logos (logical appeal), and pathos (emotional appeal), which together form the rhetorical triangle (Haskins, 2004; Rapp, 2002). Ethos builds trust through perceived integrity, expertise, or relatability (e.g., endorsements from respected figures). Pathos evokes emotion through imagery, metaphor, or narratives to align with the audience's values (e.g., fear appeals or inspirational stories). Logos uses reasoning, evidence, or data to make arguments sound analytical (e.g., statistics or product comparisons). When balanced, these appeals enhance persuasiveness across contexts, including misinformation (Higgins & Walker, 2012). These appeals directly intersect with susceptibility indicators and may increase the perceived plausibility of persuasive claims when they resonate with existing beliefs, values, emotions, or social identities, potentially reducing reliance on reflective, analytic evaluation. In contemporary practice, persuasive design uses narrative framing through rhetorical strategies such as fear appeals, undermining trust, and moral contrast, which may increase reliance on cognitive shortcuts, including confirmation bias, group loyalty, and emotion-driven reasoning. Together, these models' mechanisms suggest that facts become persuasive within misinformation not because of their accuracy but because of how they are framed.

While these models have been broadly applied in advertising, health communication, and political campaigns, their integration into misinformation research, particularly under naturalistic decision-making conditions involving uncertainty, time pressure, and emotional intensity, remains underdeveloped. This gap motivates the current study's application of persuasion models within the RPD framework.

Prior Topic Orientations and Implicit Associations

Underlying these processes are prior topic orientations, the internalized assumptions, beliefs, and cognitive patterns that incline individuals to respond in particular ways (Dzvonik, 2008). Prior orientations are shaped by prior experiences, cultural norms, and socialization, and they frame what is perceived as plausible or acceptable in thought, feeling, and behavior. The Implicit Association Test (IAT), developed by Greenwald et al. (1998), assesses the strength of implicit associations between concepts and attributes, revealing patterns that may operate outside conscious awareness and influence evaluative judgments, such as the perceived reliability of misinformation.

While widely used, the IAT remains subject to debate regarding its predictive validity, contextual sensitivity, and the extent to which it reflects stable individual differences rather than situationally activated associations (Blanton et al., 2009; Oswald et al., 2013). In this study, the IAT is therefore treated cautiously as a descriptive indicator of topic-specific associative orientation, used to

provide contextual insight rather than to support predictive or causal claims.

Decision-Making Models Relevant to Belief Formation

To understand how individuals evaluate and interpret potentially misleading information under uncertainty, it is necessary to understand models of decision-making that account for the interaction between prior orientations, cues, and reasoning styles. These models provide the foundation for linking persuasion to susceptibility in ways that clarify why misinformation can appear convincing in some contexts and questionable in others.

Cognitive Processing Pathways and Dual-Process Models. Nan et al. (2022) propose that cognitive processing during persuasive encounters can follow four pathways: (i) accuracy-motivated reasoning, (ii) directionally-motivated reasoning, (iii) identity-motivated reasoning, and (iv) accurate reasoning. Accuracy-motivated reasoning reflects the intention to reach correct conclusions through reflective analysis. By contrast, directional and identity-protective reasoning can distort judgment by aligning it with pre-existing beliefs or group identity. Accurate reasoning, unlike the first three, refers to the outcome of reaching correct conclusions, regardless of the underlying motivation. Dual-process frameworks from different disciplines describe similar distinctions in human reasoning. In cognitive psychology, System 1 and System 2 thinking (Kahneman, 2011) differentiate between fast, intuitive, automatic processes (System 1) and slow, deliberate, analytical processes (System 2). In persuasion research, the heuristic–systematic model (Chaiken, 1980) makes a similar distinction between heuristic processing and systematic processing. Heuristic processing relies on cues such as source credibility, message fluency, or emotional salience, whereas systematic processing involves careful evaluation of message content. Heuristic or System 1 processing can often yield accurate conclusions when supported by extensive domain experience. For instance, experts such as firefighters, surgeons, or pilots may draw on recognition developed through repeated exposure to make rapid decisions under pressure (Klein, 1993). However, heuristic processing is not uniformly beneficial across all contexts. In misinformation contexts, judgments may rely more on salient cues, familiarity, or information consistent with existing beliefs, potentially reinforcing existing biases and increasing susceptibility to misinformation when persuasive cues align with individuals' prior topic-specific associative orientations, values, or identities (Pennycook & Rand, 2021; Roozenbeek et al., 2022; Taber & Lodge, 2006). These models collectively underscore that belief in misinformation is not a simple matter of knowledge deficiency but a consequence of which reasoning pathway is engaged in a given context. In misinformation contexts, these pathways show how predispositions and persuasion cues shape the balance between System 1 and System 2 processing in judgment.

Naturalistic Decision Making and the Recognition-Primed Decision Model. This study applies the NDM paradigm (Klein & Zsambok, 1997), specifically the RPD model (Klein, 1993), to illustrate how individuals make decisions when encountering persuasive misinformation. The RPD model was developed to describe how experienced decision-makers operate in dynamic, time-pressured, and uncertain situations. Rather than systematically comparing multiple options, people often recognize a situation as familiar based on key cues, retrieve relevant experiences, and run a mental simulation to predict likely outcomes before acting. In high-stakes domains, such as public health or national security, this process can be efficient and accurate for experts, whose prior experience supports rapid and appropriate action. However, in misinformation contexts, recognition can be misleading: persuasive cues may mimic the patterns of familiar, trustworthy information, activating implicit associations and heuristic shortcuts. This can trigger System 1–style processing (Kahneman, 2003) and reinforce confirmation bias, particularly when cues align

with pre-existing beliefs, values, or emotions. Conversely, when cues violate expectations, the decision-making process may shift toward more analytical System 2 evaluation, engaging fact-checking or dissonance reduction strategies. By mapping onto susceptibility to misinformation through the RPD framework, we can link situation assessment, information retrieval, and mental simulation to the influence of persuasion models that shape reliability judgments. This makes RPD particularly suited to illustrating how persuasion models and prior orientations converge as individuals assess the plausibility and reliability of a claim.

Together, these decision-making models show that belief in misinformation is shaped by the reasoning pathway engaged (System 1 vs. System 2), the recognition of cues under uncertainty, and the predispositions that bias interpretation. What they do not provide, however, is an integrated account that brings predispositions, persuasion, and recognition together in explaining susceptibility. This gap motivates the present study, which examines how prior topic-specific orientations and persuasive cues interact within recognition-based evaluation processes when individuals encounter misinformation.

Susceptibility as an Integrated Process

In this literature, susceptibility is better understood as a dynamic process shaped by the interaction of prior orientations, persuasive framing, and decision-making pathways. The literature identifies recurring factors associated with susceptibility, including trust, emotion, values, group identity, and knowledge (Nan et al., 2022), which persuasive framings may activate or exploit under different conditions. Framing and rhetorical strategies embed cues that may resonate with these factors, while recognition-based decision-making accounts help explain how such cues can guide rapid judgments under uncertainty.

To situate this account within a human factors perspective, we draw on cognitive reliability analysis (CRA). CRA treats human reliability as context-dependent cognitive performance in dynamic systems and shifts the analysis from individual error to the systemic conditions that shape judgment and action under uncertainty (Woods et al., 1990). CRA emphasizes the role of cognitive, social, and contextual conditions, often operationalized as performance shaping factors (PSFs), in influencing how people perform under such conditions. These include conditions such as time pressure, workload, and incomplete knowledge, which shape how individuals attend to cues, interpret situations, and evaluate information under constraints. In the context of misinformation, CRA offers a useful perspective for understanding how contextual and cognitive conditions shape attention, interpretation, and evaluation during exposure to persuasive messages.

From this perspective, susceptibility can be understood not as a stable trait but as a context-sensitive outcome. When persuasive cues align with prior orientations and are encountered under supportive contextual conditions, they may be perceived as familiar or plausible, increasing the likelihood of rapid, low-effort acceptance. When cues conflict with prior orientations or expectancies, individuals may shift toward analytic evaluation, dissonance reduction, or skepticism. This integrated perspective provides the conceptual basis for the present study, which examines how prior orientations and persuasive cues interact during recognition-based evaluation of misinformation. The following section outlines how this perspective is operationalized in the study design.

Methodology

This study adopts an exploratory CTA (Crandall et al., 2006) design to investigate how individuals evaluate the plausibility of misinformation narratives at the point of encounter. The study design focuses on making recognition and evaluation processes observable through repeated within-participant assessments under structured variation (Dix, 2020). In this study,

structured variation involves the deliberate manipulation of narrative topics and persuasion models across repeated exposures to generate diverse evaluation episodes within each participant. This design generated multiple evaluation episodes per participant under structured variation, enabling detailed tracing of recognition and evaluation processes within individuals.

Study Design

The study consisted of repeated narrative evaluation tasks combining three data sources: think-aloud protocols (Ericsson & Simon, 1998), CDM interviews (Klein et al., 1989), and an IAT (Greenwald et al., 1998). Each participant evaluated three model-aligned variations of narratives derived from three topical domains: vaccination, climate change, and Ukraine biolabs. These topics were selected because they represent widely recognized misinformation contexts that have generated substantial public debate and have been frequently examined in misinformation research. They were not selected to represent or critique any particular political or ideological position. Rather, they provided diverse contexts in which participants could evaluate contested claims relating to public health, environmental science, and geopolitical security. The purpose of the study was not to assess the validity of viewpoints associated with these topics, but to investigate the cognitive processes through which participants evaluated persuasive misinformation narratives.

Narrative Construction

Narratives were constructed to vary persuasive framing while holding underlying factual content constant. For each topic domain: anti-vaccine perspectives on COVID, natural-cause explanations of climate change, and Ukraine biolab conspiracy narratives, a neutral base message was developed from publicly available information. From this base, three narrative variants were created, each corresponding to a different selected persuasion model: PKM (Friestad & Wright, 1994), ELM (Petty & Cacioppo, 1986), and TRA (Fishbein & Ajzen, 2011), following a systematic approach to narrative variation in misinformation research described in detail elsewhere (Haghighi & Wong, inpress). They are not treated as competing frameworks; rather, they are used to systematically frame narrative variations for observing participants' evaluation processes. The analysis of the data collected is grounded in the NDM/RPD framework, which focuses on how such cues are evaluated during decision-making.

Each variant differed in how the same information was presented. In the PKM-aligned variant, the message frame drew attention to the source's intent and credibility by suggesting possible bias or hidden motives behind official statements; for instance, by referring to "dissenting voices" or claims that authorities "ignore evidence." This framing emphasized source intent and credibility cues, rather than focusing solely on whether the claim itself was true. In the ELM-aligned variant, the message was framed to encourage rapid, low-effort evaluation by emphasizing urgency and potential risk, such as "serious risks to public safety." This framing emphasized urgency and potential consequences, encouraging attention to salient emotional cues rather than detailed analytical evaluation. In the TRA-aligned variant, the message frame emphasized social norms and group alignment by suggesting what others believe or how "the public has been told" to interpret the issue. This emphasizes acceptance or skepticism regarding perceived group expectations and shared values.

Across all variants, the same factual content was retained, while wording and emphasis were varied. This allowed differences in participant responses to be examined in relation to how information was presented.

Participants

Six adult participants were recruited through public advertisements using a volunteer self-

selection approach. Interested individuals contacted the researcher and were screened for basic eligibility criteria, such as age, English proficiency, and general engagement with online information. Individuals with direct academic relationships to researchers were excluded. Participants had backgrounds in interior design, elite sport coaching and judging, pharmacology, computer science, transport and aviation, and AI-computation expertise. These characteristics are reported for context only and were not treated as analytical variables. Each participant contributed multiple evaluation episodes across topics and framing conditions. Consistent with the exploratory, process-oriented design, the sample is not intended to be representative, but to support detailed analysis of how evaluation unfolds under structured variation (Dix, 2020).

Implicit Association Test

The IAT required participants to rapidly sort topic-relevant words into paired categories. Response latencies and errors were used as descriptive indicators of relative associative strength between concepts (Greenwald et al., 1998; Greenwald et al., 2003).

Category sets were tailored to each topic domain to align the task with the key concepts used in the corresponding narratives. For example, the biolab task contrasted trust- and threat-related categories, the vaccine task contrasted vaccine- and natural-immunity-related categories, and the climate task contrasted natural-cause and human-cause framings. Participants sorted narrative-relevant words across practice and critical blocks. Pairings contrasted alternative associative mappings within each topic domain. Congruent and incongruent pairings were treated as relative indicators of the defined associative patterns within each task.

Response latencies from the critical blocks were converted to milliseconds and used to calculate D-scores using the improved IAT scoring algorithm (Greenwald et al., 2003). D-scores provide a standardized index of associative contrast between paired concepts within each task. Scores further from zero indicate stronger associative contrast in the direction defined by the category pairings, but are interpreted only in relative, task-specific terms. In this study, IAT results are interpreted only as descriptive indicators of topic-specific associative orientation and are not treated as measures of belief, endorsement, or stable individual differences.

Procedure

The study was conducted individually with each participant. Each session followed five steps: (1) Participants completed an IAT prior to encountering the message. The IAT was used descriptively to indicate relative associative orientation toward each topic based on response latency, providing contextual information for interpreting later qualitative responses rather than a measure of belief, reasoning ability, or susceptibility. (2) Participants then encountered a series of narratives across three topics (vaccines, climate change, and bioweapons), each presented using systematically varied model-aligned framings. (3) During each narrative encounter, participants were asked to think-aloud while reading and evaluating the narrative, verbalizing their thoughts as they occurred. The think-aloud procedure was used to capture concurrent reactions, recognition, interpretation, and emerging judgments during narrative evaluation. No follow-up prompts or interventions were used during the think-aloud task, including during periods of silence, in order to minimize researcher influence on participants' ongoing reasoning processes. (4) After each narrative encounter, a CDM interview was conducted. Participants were asked to reflect on how the message made them feel; whether those feelings influenced the choices or interpretations presented by the message; whether it prompted particular imagined outcomes or scenarios; how they resolved information that challenged or reinforced their expectations; and whether

it affected their confidence in the message's validity. The full set of CDM probes is provided in Appendix 3. (5) The IAT was then repeated to obtain a second descriptive indication of associative orientation after each narrative evaluation. The post-IAT data were collected to explore whether associative responses might vary following exposure. However, given the small sample size and the well-known sensitivity of IAT scores to context, priming, and task repetition, these patterns are interpreted only descriptively and are not treated as evidence of systematic change. Each of the six participants evaluated nine narratives (three topics × three model-aligned framings), resulting in 54 evaluation episodes across the dataset. The order of narrative topics and persuasion models was randomized across participants to reduce ordering effects. In keeping with the exploratory design, IAT results were used only to contextualize patterns in the think-aloud and interview data, not for statistical comparison or causal inference. Table 2 presents an overview of the study design.

Analysis

Think-aloud and CDM transcripts were analyzed using the emergent themes approach (ETA) (Wong, 2003), a structured variant of thematic analysis informed by CDM. This approach combines a priori coding based on recognition-based decision processes (RPD) with inductive identification of emergent themes. The analysis focused on identifying patterns in how participants evaluated narratives within each evaluation episode. Data was organized accordingly, with each episode corresponding to one participant's response to a single narrative. Initial coding was conducted inductively, with attention to decision-process constructs. This led to the identification of cognitive strategies.

Codes captured behaviors such as cue recognition, questioning content, reference to prior knowledge and experience, emotional reactions, resistance, mental simulation of consequences, and resolution of uncertainty. These were grouped into higher-level categories reflecting evaluation themes:

cue recognition,
interpretation,
response patterns,
RPD pathways.

These represent thematic groupings of coded behaviors. They were developed

Table 2. Experiment Design Structure to Elicit Cognitive Strategies (CSs)

	Story 1 (COVID vaccine)	Story 2 (climate change)	Story 3 (Ukraine biolab)
	Study Procedure	Study Procedure	Study Procedure
Persuasion model applied	Initial IAT	Initial IAT	Initial IAT
PKM	Think-aloud	Think-aloud	Think-aloud
	CDM interview	CDM interview	CDM interview
	Post-IAT	Post-IAT	Post-IAT
	Derived CSs	Derived CSs	Derived CSs
TRA	Think-aloud	Think-aloud	Think-aloud
	CDM interview	CDM interview	CDM interview
	Post-IAT	Post-IAT	Post-IAT
	Derived CSs	Derived CSs	Derived CSs
ELM	Think-aloud	Think-aloud	Think-aloud
	CDM interview	CDM interview	CDM interview
	Post-IAT	Post-IAT	Post-IAT
	Derived CSs	Derived CSs	Derived CSs

Table 3. Overview of Codes and Themes Emerged Through Thematic Analysis of Think-Aloud and CDM

	Excerpt sample from transcript	Codes	Themes	Description
Think-aloud and CDM interviews	"Parents like you, and if, you know, if you're not choosing the safer, then maybe, then you are less than these other parents, yeah." [P2, TRA, COVID, Think-aloud, 1/17]	Cue recognition	Cue recognition	How cues were recognized
	"I usually go straightly to their research source, but here's no source, so I do not think that it can be true, and they refer to traditional treatment, traditional medicine, which can be controversial sometimes. Hmm… so, there is no proof in this paragraph related to what was stated at the beginning." [P1, TRA, COVID, think-aloud, 1/26]	Questioning	Interpretation	How they were interpreted in relation to prior knowledge and experience
	"That's a red flag… Scientific claims without evidence don't hold weight, especially when they go directly against the broad scientific consensus" [P5, PKM, ClimateChange, Think-aloud, 1/9]	Prior knowledge		
	"It didn't really make me feel much… maybe just cautious. It's trying so hard to make fear, but I don't mind that… It made me more critical of the claims than accepting of the choices it suggests… yeah…" [P4, ELM, BioLab, CDM, 1/19]	Emotional reactions	Response patterns	How they reacted and responded to the cues and interpretations
	"If it is really a natural cycle and about Geology and earth science, so, we cannot do anything about it. But if it is based on human activities, then the international organizations and governments [may need to] align with each other. My only responsibility will be like, making less waste and garbage, or saving some energy and fuel just from my side, which may not be even considered serious or impactful." [P5, ELM, ClimateChange, CDM, 2/30]	Resistance to contradiction		
	"Like if green energy really caused economic problems… like people losing jobs or energy prices going up… That worried me… I want solutions that don't hurt anyone… But then I also wondered if ignoring human activities on climate might make worse environmental damage later… it's a complicated issue." [P6, PKM, ClimateChange, CDM, 1/28]	Simulation	RPD pathways	How judgments are formed or shifted during evaluation
	"I'd say I'm middle… moderate… confident. Like, I didn't accept it… I still want to see evidence, not claims from a single source. But also… yeah, the message did make me wait a bit… it didn't change my opinion completely, but it made me less sure that ignoring these stories is good [idea]. I Think maybe reconsider… like… if I'm maybe too quick to trust the official narrative sometimes. Not changed my mind, but it's a bit at how sure I am in not listening to these claims at all. I Still want to search more. That pulling and stuff… it does get a bit." [P3, ELM, BioLab, CDM, 2/26]	Resolution of uncertainty		

iteratively through repeated reading of the transcripts and constant comparison across episodes. Table 3 presents the codes and emerging themes identified during analysis. These themes were then used to identify recurring patterns of reasoning across episodes, referred to as cognitive strategies (CSs).

Think-aloud and CDM data were analyzed using a common coding scheme, as both data sources captured different aspects of the same evaluation process within an episode. Think-aloud data provided immediate, in-the-moment interpretation and judgment, while CDM data enabled retrospective probing into how participants explained and reflected on these decisions. Coding was applied across both data sources to capture consistent evaluation behaviors within each episode. Distinctions between concurrent (think-aloud) and retrospective (CDM) accounts were logged to support analyses of the findings.

Building on the initial inductive coding of evaluation episodes, similar codes were grouped and refined to identify recurring CSs. The same coding scheme was applied across think-aloud and CDM data, allowing these strategies to reflect consistent reasoning patterns across concurrent and retrospective accounts. Each CS represents a higher-level pattern of reasoning derived from coded segments across multiple evaluation episodes. These strategies capture how participants recognized, interpreted, and responded to persuasive cues. In this sense, CSs are second-order abstractions derived from 200 coded segments, rather than independent categories introduced at a later stage. These strategies were then organized into broader abstractions reflecting stages of evaluation, including:

recognition,
processing,
response,
action,

Table 4. Mapping of selected CSs to Aggregated Cognitive Strategies (ACSs)

	Theme	CS cluster (intermediate)	CS IDs	Description of grouping	Resulting ACS
PKM	Recognition	Trust and manipulation cue detection	CS1, CS35	Recognition of credibility cues (authority, trust) combined with detection of manipulative framing or propaganda	ACS1
	Action	Evidence-based rejection	CS2, CS36, CS37	Analytical evaluation and emotional resistance leading to rejection due to insufficient evidence	ACS4
	Processing	Biased mental simulation	CS6, CS42, CS26	Construction of plausible scenarios guided by prior beliefs and past experiences	ACS6
TRA	Processing	Analytical and heuristic evaluation	CS21, CS2, CS5b	Initial scrutiny of authority cues followed by analytical reasoning and contextual evaluation	ACS8
	Dissonance reduction	Motivated reasoning and rationalization	CS11, CS27	Resolving conflict by selectively prioritizing belief-consistent interpretations	ACS11
ELM	Consideration	Provisional consideration without belief shift	CS5b, CS42	Temporary engagement with alternative possibilities without commitment or belief revision	ACS14
	Processing	Heuristic–analytical balancing	CS30, CS36, CS2	Simultaneous use of emotional cues, authority heuristics, and partial analytical reasoning	ACS16

Table 5. Aggregated Cognitive Strategies

Stimuli	Aggregated CSs	Theme	Definition
PKM	ACS 1	Recognition	- Recognizing trust or safety cues and treating them with skepticism, while staying aware of how they are framed and not letting emotional intensity influence judgment.
	ACS 2		- Being more receptive to narratives that invoke distrust, while still resisting generic fear-based messages, shows that sensitivity to cues can vary.
	ACS 3	Processing	- Prioritizing analytical reasoning and external validation, with a preference for evidence over emotional or affective responses.
	ACS 4	Action	- Rejection of messages based on a lack of sufficient evidence, emphasizing a skeptical, deliberative approach over reliance on trust or heuristics.
	ACS 5	Recognition	- Implicit association between fear and mistrust, signaling heightened sensitivity to these cues.
	ACS 6	Processing	- Using mental simulation, where past experiences and prior beliefs guide assessment of plausibility, showing how directionally motivated reasoning works and how skepticism interacts with familiarity-based validation.
	ACS 7		- Verifying information before belief formation, with reasoning shaped by context and prior motivations, makes evaluation adaptive but potentially biased.
TRA	ACS 8	Processing	- Starting with critical recognition of authority cues, then relying on familiar cues such as source credibility and prior beliefs, showing a shift from skepticism to heuristic processing.
	ACS 9	Dissonance reduction	- Selectively prioritizing information that supports existing beliefs or identities, reinforcing comfort by accepting confirmatory evidence, and dismissing contradictions.
	ACS 10	Consideration	- Engaging selectively with information based on personal relevance and prior knowledge, using mental simulations of outcomes to respond to messages that fit existing beliefs and expectancies.
	ACS 11	Dissonance reduction	- Reducing discomfort with conflicting information by selectively prioritizing sources that resonate emotionally and align with pre-existing beliefs or identities, using cognitive shortcuts to maintain consistency.
ELM	ACS 12	Processing	- Systematic reasoning shaped by implicit biases and mental simulations, where recognition of negative implications interacts with prior beliefs to preserve or gradually adjust them.
	ACS 13		- Engaging in cognitive flexibility by considering different possibilities and broader contexts, but still maintaining alignment with core beliefs guided by both implicit associations and motivated reasoning.
	ACS 14	Consideration	- Message framing can recalibrate implicit associations by reinforcing prior but non-dominant beliefs.
	ACS 15	Processing	- Recognizing persuasion techniques while still leaning on authority cues, showing directionally motivated reasoning; critical engagement that evaluates arguments but ultimately preserves core beliefs.
	ACS 16		- Balancing expert validation with emotional and heuristic cues, where implicit associations support acceptance of predisposed messages, while analytical evaluation protects core beliefs from change.

representing a higher-level structuring of the evaluation process. These categories structure the presentation of findings in Section 4.

To ensure transparency in the analytic process, Table 6 (see Appendix 1) presents a sample of coded evaluation segments and their associated CSs assigned to each, illustrating how second-order abstractions were derived from the think-aloud and CDM data. Building on this, Table 4 shows how these CSs were grouped into aggregated cognitive strategies (ACSs) through constant comparison across episodes. CSs capture recurring patterns of reasoning within individual evaluation episodes. ACSs represent higher-level abstractions that organize these strategies into broader patterns of recognition, processing, dissonance management, and action. Table 5 presents the final set of ACSs used to structure the findings.

Where coding decisions were uncertain, interpretations were revisited and refined through comparison with other episodes and the emerging coding structure. The analysis was conducted by a single researcher and is interpretive and exploratory; it does not rely on frequency counts or statistical inference.

Results and Data Analysis

This section reports findings from the think-aloud and CDM analyses in two parts: think-aloud data captures immediate evaluation processes during exposure, while CDM interviews provide retrospective accounts that clarify and extend these processes. IAT results are used only as descriptive contextual information regarding topic-specific associative orientations, with detailed scores reported in Appendix 2.

Think-Aloud During Narrative Encounter

Think-aloud data captured participants' immediate reactions, interpretations, and plausibility judgments as they encountered the narratives. Analysis of the evaluation episodes identified themes derived from recurring patterns in: (1) how cues were recognized, (2) how they were interpreted in relation to prior orientations, including knowledge and experience, (3) how responses were formed, and (4) how judgments formed or shifted during evaluation.

Theme 1—Cue Recognition: Framing Directs Attention to Different Cue Types. Participants' immediate reasoning varied across the three persuasion framings, particularly in the kinds of cues that were noticed and how those cues were evaluated. PKM-framed narratives often prompted attention to persuasive intent, especially when participants questioned the motives or credibility of the source. P6, for instance, treated authority names as a momentary credibility boost but quickly defaulted to verification:

> "the names… make it more trustworthy, but I would search more to trust 100%" [P6, PKM, Biolab, 1/16].

In other cases, the same cues appeared to support plausibility when they matched participants' prior suspicions or expectations. For instance, P2, mildly misinformation-aligned on Biolab, admitted that:

> "with these consequences at stake, the world deserves to know the truth" [P2, PKM, Biolab, 1/14]

was sufficient to support their initial judgment.

ELM-framed narratives showed a contrast between emotionally charged and more neutral presentation styles. Dramatic phrasing often triggered suspicion, while more neutral or balanced wording was less likely to prompt immediate resistance. P6 dismissed the vaccine message's invitation as:

> "that's trying to create team and make me feel special… But it feels sly rather than smart" [P6, ELM, COVID, 1/12].

Yet calmer, familiar tones sometimes lowered resistance. P1, although strongly

evidence-aligned on climate, acknowledged the balanced narrative as:

> "it sounds fair… Sounds clear… And there is some evidence here" [P1, ELM, ClimateChange, 1/14].

TRA-framed narratives most consistently drew attention to normative and identity appeals, although participants often reassessed these against prior knowledge or lived experience. For example, P2 explicitly recognized the parental identity framing:

> "parents like you… you don't want to be a bad parent" [P2, TRA, COVID, 3/6]

but reported quickly that they were not persuaded and:

> "moved on" [P2, TRA, COVID, 3/6].

Similarly, P5 weighed the familiar appeal to natural immunity against their own experience of illness and recovery, concluding that:

> "natural immunity may fight it but is risky" [P5, TRA, COVID, 2/14]

and reinforcing their support for vaccination. P6 dismissed an exaggerated claim outright:

> "So just one case doesn't have to scare me" [P6, TRA, COVID, 1/18].

Even P1, who noticed the collective pressure implied in vaccine framings:

> "have no proofs … except for providing this information about the person who had some side effects of vaccination. But this can happen to anyone" [P1, TRA, COVID, 1/37],

framed this as recognition of a tactic rather than an influence on their decision. Taken together, TRA cues were consistently recognized, but assessment against prior knowledge or experience moderated their persuasive impact.

In this dataset, the three framings drew attention to different cue types, but their influence on participants' judgments depended less on the framing itself than on how those cues aligned with participants' prior knowledge and expectations. These differences reflect variation in how cues were recognized and evaluated, rather than consistent effects of the framing itself.

Theme 2—Interpretation and Assessment: Prior Orientations Filter Cue Evaluation. Participants' immediate responses suggested that prior orientations shaped how persuasive cues were interpreted in the moment. When cues resonated with prior expectations or existing leanings, judgments often form quickly. When cues conflicted with prior knowledge or expectations, they were more likely to trigger skepticism, irritation, or resistance. For instance, P4, who showed mild alignment on the Climate-Change topic, appeared to rely on a single expert citation as a basis for agreement:

> "because one… scientist has this opinion" [P4, ELM, Climate-Change, 2/12][1].

Similarly, participant P2's conspiracy-aligned predisposition made whistleblower claims on Biolab feel plausible:

> "I would assume perhaps there is truth in these leaked documents" [P2, PKM, Biolab, 1/23].

In both cases, recognition of authority or secrecy cues appeared to fit participants' existing expectations, allowing judgments to form with limited further scrutiny.

By contrast, when cues clashed with evidence-aligned predispositions, they were recognized as manipulative or exaggerated and prompted immediate resistance. P5 dismissed vaccine fear appeals as:

> "emotionally manipulative right away" [P5, ELM, COVID, 1/4],

while P1 described the Biolab's dramatic rhetoric as:

> "That kind of language really annoys me… it thinks I'm asleep or blind" [P1, ELM, Biolab, 1/11].

These participants did not ignore the cues; instead, recognition triggered defensive filtering through prior knowledge and trust in science, which limited further consideration of the message.

Taken together, the think-aloud data suggest that prior orientations acted as a selective filter: cues gain traction when they feel familiar or valued, but the same cues prompted resistance when they contradicted established knowledge.

Theme 3—Response Patterns: Emotional Appeals Often Elicit Resistance, Analytical Cues Prompt Verification. Emotional tones in the narratives often triggered irritation, anger, or rejection rather than persuasion, especially among participants who drew on prior scientific knowledge or verification-oriented reasoning. P4, when confronted with the vaccine narrative's rhetoric, stated directly:

> "People like you opting out! classic trick to belonging" [P4, ELM, COVID, 1/10]

and dismissed similar phrasing as:

> "Feels manipulative" [P4, ELM, COVID, 1/11].

Similarly, dramatic appeals and accusatory rhetoric elicited frustration and resistance from P1, who described the Ukraine Biolab framing as condescending and pressuring:

> "wake up… that's insulting" [P1, ELM, Biolab, 1/20].

Even when emotional appeals briefly captured attention, as in P2's reaction to vaccine claims framed as a bold revelation, the result was unease rather than persuasion:

> "many people are starting to realize… bold claim" [P2, PKM, COVID, 1/4].

These reactions show how strong emotional appeals can backfire, provoking defensive resistance rather than compliance.

By contrast, analytical framings or authority references more often led to fact-checking requests and further evaluation rather than uncritical acceptance. P4 pressed for alternatives in climate claims:

> "If not green energy, then what?" [P4, ELM, ClimateChange, 1/8].

P5 challenged the lack of specificity in climate solutions and flagged unsupported expertise by pointing out:

> "it's easy to throw a name and years of experience… that's a red flag" [P5, PKM, ClimateChange, 1/6],

illustrating a preference for concrete evidence over appeal to authority. P2 took a similar stance when presented with an authority name in the vaccine framing and admitted they would only be persuaded if journal sources were verifiable, remarking:

> "unless I was gonna look into them… I don't know them" [P2, TRA, COVID, 2/19]

or P6 emphasized the need for sources before trusting authority names:

> "I'd Google him first" [P6, ELM, COVID, 1/8].

These findings show that emotional appeals often triggered irritation and resistance, while analytical or authority-based cues prompted verification. Analytical responses did not necessarily lead to acceptance; instead, they often guide evaluation through questioning, comparison, and verification. Participants were often not passive recipients of the narratives; they actively questioned claims, assessed cue credibility, and challenged unsupported assertions in real time.

Theme 4—Evaluation Pathways (RPD): Cue–Prior Alignment Shapes Evaluation Trajectories. Consistent with the RPD framework, participants' immediate reasoning often moved through recognizable sequences of cue noticing, situation assessment, mental simulation, and judgment. The cues recognized included authority names, secrecy claims, emotional tone, identity positioning, and missing or weak evidence. Whether the cycle closed in rapid acceptance or rejection, or extended into resistance, depended on how well cues fit prior knowledge and experience.

Participants rapidly recognized surface cues such as authority names, emotional tone, or group-based appeals. For instance, P3 noticed the biolab narrative's secrecy pitch but doubted its veracity:

> "they always say secret documents… but I only trust maybe 40%" [P3, PKM, Biolab, 2/21].

In climate narratives, participant P6 was similarly quick to interrogate new authority cues:

> "who is this Dr. Greer? it needs a bit of checking" [P6, ELM, ClimateChange, 1/11].

When cues created uncertainty, participants assessed them against familiar knowledge or imagined outcomes. P2, reacting to the climate message, drew on memories of ozone debates:

> "this big hole in the ozone layer… we didn't change too much in the world" [P2, ELM, ClimateChange, 1/25].

Similarly, P3 expressed uncertainty about the statistical report in the vaccine message and drew on his knowledge of the prevalence of side effects within his network:

> "Some of the numbers, I'm not that, believe [e.g., just 0.01% of the people got the side effects] across the just the people around me. I heard a lot of times… Just feel not good because they just got vaccines, so they think maybe that's in fact" [P3, TRA, COVID, 3/7].

P5, when hearing natural cycle arguments, simulated possible consequences but dismissed the selective framing:

> "completely rejecting green energy seems silly… it only highlights one side" [P5, TRA, ClimateChange, 1/11].

Some participants extended the cycle into explicit verification attempts. For example, P3, when confronted with vaccine narratives, directly rejected vague natural-remedy claims, noting:

> "supplements… are not enough" [P3, ELM, COVID, 1/10]

and highlighting:

> "no specifics, no data, no sources" [P3, ELM, COVID, 1/6].

Similarly, P1, while irritated by conspiratorial tones, admitted that curiosity could drive further checking:

> "maybe if I had free time I would like to read more" [P1, PKM, Biolab, 2/28].

These statements show how recognition of weak or missing evidence moved the cycle beyond initial assessment, triggering fact-checking and critical scrutiny rather than acceptance.

Across episodes, cue recognition was associated with either provisional acceptance or active resistance. Provisional acceptance occurred when cues aligned with participants' prior knowledge or expectations. For example, P2 showed initial openness to leaked-document claims:

> "I would assume perhaps there is truth in these documents" [P2, PKM, Biolab, 1/23].

Resistance followed when cues clashed with prior knowledge or belief. For instance, P5 rejected dramatic biolab phrasing as manipulative:

> "sounds like a movie trailer, not something factual" [P5, ELM, Biolab, 1/5].

Taken together, these findings show that cue recognition set a trajectory: when cues conflicted with participants' prior knowledge or experience, participants extended evaluation through mental simulation or fact-checking before rejecting the claim. Where cues aligned with prior knowledge or experience, participants often considered the message plausible with minimal further scrutiny.

Summary of Think-Aloud Findings: Immediate Reasoning Patterns. Across the 54 participant–narrative evaluation episodes, plausibility judgments did not arise from message content alone, but from the interaction of prior orientations, persuasive cues, and recognition-based evaluation. Participants actively attended to cues such as authority names, secrecy claims, emotional tone, and identity appeals, assessing them against prior knowledge and expectations. When cues aligned with these expectancies, judgments formed quickly; when they conflicted, they triggered skepticism, irritation, or further verification. Across the dataset, PKM framings often drew attention to persuasive intent, ELM framings frequently prompted responses to tone and emotional intensity, and TRA framings highlighted norm and identity appeals, though their influence was moderated by prior knowledge and experience. These patterns reflect how recognition-based evaluation shaped immediate plausibility judgments during exposure.

The next section turns to the CDM interviews, which complement the Think-Aloud data by providing retrospective accounts of why participants judged messages as they did.

CDM Interviews: Retrospective Reasoning and Decision Pathways

The CDM interviews probed how participants retrospectively explained the judgments they had made during narrative encounters, focusing on emotional responses, imagined consequences and mental simulation, contradiction handling, and confidence in judgment.

These patterns reflect the same evaluation process identified in the think-aloud analysis: what appeared as cue recognition, interpretation and assessment, response patterns, and evaluation pathways is reflected here through how participants later articulated their reasoning, rather than repeating what they noticed first during the encounter or how evaluation unfolded in real time. The subsections are organized around these retrospective accounts to reflect how participants explained their reasoning.

These retrospective accounts clarify: (1) how participants explained the influence of emotional triggers, (2) how they connected imagined consequences to message plausibility, (3) how they resolved tension when a message challenged or reinforced expectations, and (4) whether they reconsidered their stance or confidence. The following subsections report these patterns across these aspects of retrospective reasoning.

Emotional Responses and Reasoning: Emotional Tone Prompted Caution, Resistance, or Brief Engagement. Across the dataset, participants often recalled strong emotional tones as prompting caution rather than persuasion. Fear and accusatory language often provoked irritation, while calmer framing occasionally encouraged momentary openness. For example, P4, evidence-aligned on vaccines, rejected fear appeals outright:

> "message is clearly trying to push fear… making distrust. I'm not buying it." [P4, PKM, COVID, 1/8].

Similarly, P5 dismissed accusatory language in biolab stories:

> "Right away, I don't like being told I'm blindly following if I don't agree… That's insulting… Too emotional. Too manipulative" [P5, TRA, Biolab, 1/4].

These responses suggest that emotional intensity often clashed with participants' prior knowledge or expectations, reinforcing resistance.

Emotional cues could also prompt curiosity when they resonate with participants' prior suspicions or concerns. P3, misinformation-aligned on biolab, acknowledged:

> "feel uneasy… scary… made me wait and think the possibility, even I was not sure it is true" [P3, ELM, Biolab, 1/29].

Here, fear opened space for contemplation without full acceptance. P2 likewise admitted being unsettled:

> "dangerous experiments… kind of scary" [P2, ELM, Biolab, 1/7],

though this unease eventually gave way to skepticism.

Emotions influenced how participants engaged with the message rather than simply increasing persuasion. In some cases, they triggered rejection by making the message appear manipulative; in others, they prompted further consideration of the scenario without leading to full acceptance.

Imagined Consequences and Mental Simulation: Mental Simulation Tested Plausibility. When asked to picture consequences, participants often constructed "what-if" scenarios, testing plausibility against prior knowledge and expectation. These simulations showed how imagined outcomes either supported or challenged emerging judgments, depending on how well they aligned with participants' prior knowledge or expectations.

P4, mildly misinformation-aligned on biolab, entertained geopolitical scenarios:

> "I imagined a scenario like geopolitical actors are running covert operations under the radar… not impossible, given history" [P4, TRA, Biolab, 1/30].

In this case, the framing opened space for a suspicious or conspiratorial interpretation of possible events. By contrast, P6, strongly evidence-aligned on climate, rejected selective framings through mental simulation, concluding that such claims exaggerated or distorted reality:

> "…to push the public to do something in their favour… the industries, the economy, whatever… Isn't it just exaggerating or changing things?" [P6, ELM, ClimateChange, 1/24].

Their simulation reinforced prior belief in anthropogenic causes. P3's simulations reflected doubt. On vaccines, they drew from personal experience:

> "people around me… feel not good because they got vaccines… but because this kind of illness is very serious, it will hurt our body a lot, so we may need that help" [P3, TRA, COVID, 3/8].

This pattern highlights how simulations introduced competing considerations that kept judgments provisional rather than fully resolved. P2 simulated worst-case biolab scenarios:

> "something leaking out of a lab and spreading" [P2, ELM, Biolab, 1/29].

Yet in both cases, the simulations did not produce settled judgments; instead, they remained provisional, shaped by competing knowledge or uncertainty.

Across participants, mental simulation operated as a secondary stage of assessment. It often supported plausibility when imagined outcomes fit prior expectations, reinforced skepticism when they did not, and occasionally moderated extreme positions, as in P3 case, by bringing competing considerations into view.

Navigating Contradictions and Resolving Tension: Contradictions Were Reinterpreted to Maintain Coherent Judgments. Participants frequently identified contradictions in the narratives and described how they resolved them. This dissonance reduction was often accomplished by reframing intent or evidence, and predispositions shaped whether

contradictions reinforced or weakened resistance. P4, evidence-aligned with vaccine, reflecting on the message, described:

> "It is correct claim but also contradictory to itself… correct facts but with wrong conclusion" [P4, TRA, COVID, 1/20].

Rather than producing confusion, this contradiction reinforced their existing judgment in favor of vaccination:

> "reinforced my decision… made me certain to get the vaccine" [P4, TRA, COVID, 1/20].

P6, misinformation-aligned with conspiracy, confronted with accusatory biolab framings, normalized them as typical of conspiracies:

> "These kinds of messages always come with guilt and drama… I already expect them to be tricky and accusing, and this one totally was" [P6, TRA, Biolab, 2/5].

Here, recognition of contradiction was defused by attributing it to a familiar manipulative style. P3, misinformation-aligned, highlighted tension in climate narratives:

> "I agree Earth has cycles. But the speed of warming now is different from before. So maybe natural and human both?" [P3, TRA, ClimateChange, 1/15].

On vaccines, they flagged inconsistencies:

> "like someone close to me getting sick and thinking, 'I'll just drink tea and take vitamins,' and then ending up in the hospital because they underestimated the severity of the virus…" [P3, PKM, COVID, 1/37].

Rather than resolving toward belief, these contradictions strengthened cautious skepticism.

These cases suggest that contradictions rarely disrupt judgments for long. Instead, they were reinterpreted to maintain coherence in participants' reasoning, reinforcing their current line of judgment rather than overturning it: in some cases, strengthening rejection by being read as signs of manipulation, while in others being acknowledged but leaving room for continued suspicion or doubt.

Confidence and Stance Reconsideration: Confidence Was Calibrated Rather Than Fully Revised. CDM interviews showed that confidence was not static, but calibrated in relation to perceived evidence, prior expectations, and the credibility of the message. Three recurring patterns were apparent in how participants described their confidence. First, some participants anchored confidence in evidence, although the specific percentage expressions reported below are individual examples rather than frequency-based indicators. P4, which showed more leaning toward knowledge and evidence, when discussing vaccines, expressed relatively high certainty:

> "70% I am certain that I am making a right decision [to take the vaccine]" [P4, TRA, COVID, 2/20],

explaining that statistics boosted reliability:

> "once the numbers are specific then we can be certain that they are reliable" [P4, TRA, COVID, 2/20].

By contrast, on biolab narratives, P4 showed mild conspiracy alignment. Confidence dropped:

> "about 45–50% confident in the message's validity" [P4, TRA, Biolab, 2/11].

Similarly, P1, though more aligned with the vaccine's topic, consistently anchored retrospective confidence in pro-science reasoning:

> "too many hidden facts and no real proofs" [P1, PKM, Biolab];

> "in our country it is important to make all the vaccine… it gives immunity for lifetime" [P1, TRA, COVID, 1/29].

Second, several participants reflected their confidence cautiously when evidence was missing or inconclusive. P5, with more alignment in relation to this topic, stated:

> "40% is correct… Statistical data shows it more. They can be the best evidence" [P5, TRA, COVID, 2/28].

On vaccines, missing evidence again shaped confidence:

> "I'm quite confident the message is misleading… no references instead of evidence… is not like what I read in science sources" [P3, PKM, COVID, 2/17].

P6, likewise, settled at "50–50" confidence:

> "I don't trust it, but I can't say it's totally false… It made me want to check and google and move on" [P6, TRA, Biolab, 2/15].

Finally, some participants reported momentary hesitation but reaffirmed prior stances. P5, with more alignment with the vaccine's topic, reflected:

> "I am hesitating on this but… I am not sure if I have enough courage to trust my immunity system or not" [P5, TRA, COVID, 1/32],

but quickly clarified:

> "No… I just had a pause and hesitated a little but no change… I still believe in vaccines" [P5, TRA, COVID, 2/21].

Taken together, these patterns show that confidence was not fixed but adjusted in relation to evidence, prior understanding, and perceived credibility. When narratives aligned with prior understanding and expectations, confidence was expressed more strongly; when evidence was weak or missing, participants expressed uncertainty in their judgments without fully accepting the message; and when messages challenged prior views, participants briefly hesitated but typically returned to their initial stance.

Summary of CDM Findings. The CDM interviews showed that participants did not simply report fixed reactions to the narratives; they retrospectively worked through how and why particular judgments had felt plausible, doubtful, or misleading. Across the accounts, emotional tone, imagined consequences, contradictions, and confidence were integrated within the same interpretive process through which participants explained and justified their stance. What mattered was not the message content alone, but how participants connected message cues to their prior knowledge, expectations and existing understanding, and the credibility they assigned to the account.

These retrospective explanations show that plausibility judgments were actively constructed and recalibrated rather than passively absorbed, often maintaining coherence with participants' prior understanding: contradictions were reinterpreted, simulations kept judgments provisional, and confidence was adjusted without necessarily leading to revision of stance. In this way, the CDM findings extend the think-aloud analysis by showing how participants later made sense of the judgments they had reached in the moment.

Pre- and post-IAT comparisons were examined descriptively and are reported in Appendix 2. Given the small sample size and the sensitivity of IAT scores to context and task repetition, these patterns are not interpreted as evidence of systematic change. The main findings of the study, therefore, rely on the think-aloud and CDM analyses.

Discussion

This study examined how prior topic-specific associative orientations and

persuasive narrative cues interact during recognition-based evaluation of misinformation. The findings from think-aloud and CDM analyses suggest that plausibility judgments were not attributable to message content alone, but were shaped through the interaction between prior orientations and persuasive framing as participants moved through recognition-primed evaluation. Evaluation trajectories differed depending on whether cues aligned or clashed with participants' prior orientations, and on how specific framings shaped subsequent responses.

To interpret these patterns, we draw on the thematic analysis (Blandford & Wong, 2004; Wong, 2003) to identify a set of CSs that describe recurring ways participants made sense of persuasive cues; for example, flagging manipulative tone, constructing worst-case scenarios, or quantifying confidence numerically. Across 54 participant–narrative evaluation episodes, more than 200 reasoning segments were identified and iteratively organized into 16 cognitive strategies.

We treat these CSs as observable patterns of reasoning that make visible how evaluation processes unfold in practice. They show how participants recognized cues, assessed them in relation to prior knowledge and expectations, simulated outcomes, and resolved contradictions. Rather than treating susceptibility as an abstract "bias," the CSs reveal recognizable reasoning practices through which participants interpret and evaluate persuasive cues. The CSs referenced throughout this section are summarized in Table 6, which outlines their definitions and their relationships to different persuasion models.

The subsections that follow describe how these CSs were identified and how they help interpret the major themes of the findings: emotional triggers, mental simulation, contradiction handling, and confidence calibration. Throughout, we examine how prior orientations and different persuasive framings were associated with the activation of particular strategies. This analysis also points to potential implications for HCI. By understanding the strategies people use during evaluation, it becomes possible to consider how interfaces might better support scrutiny; for example, by making contradictions more visible or by prompting reflection on how messages are framed in relation to prior expectations. Such approaches move beyond labeling content as true or false, toward supporting the reasoning practices that underpin critical evaluation of information.

Emergence of Cognitive Strategies

To understand the evaluation processes of misinformation narratives across the persuasion models, we draw on the inductive thematic analysis to organize patterns beyond individual episodes. The aim was to move from isolated examples to a structured account of the recurring strategies participants used to make sense of misinformation.

Each participant's narrative evaluation was segmented into decision episodes that combined real-time think-aloud responses with retrospective CDM reflections. Within each episode, we coded evaluation processes as expressed in participants' verbalized reasoning, such as noticing persuasive cues, evaluating evidence, simulating consequences, and resolving contradictions. These codes were then grouped inductively into higher-order categories corresponding to stages described in the RPD framework: recognition, assessment, mental simulation, and decision.

Patterns were examined across different persuasive framings (PKM, ELM, TRA) as reflected in their responses to identify recurring differences in how strategies were activated (see Table 6). For example, participants whose prior orientations did not align with dramatic biolab framings frequently treated them as manipulative or exaggerated, whereas in vaccine narratives, when framings resonated with participants' prior orientations, they sometimes engaged

with natural health appeals before returning to scientific reasoning.

Across more than 200 coded reasoning segments, constant comparison revealed convergence into 16 CSs, each representing a recurring pattern of reasoning across episodes. Examples include:

> Recognition of persuasive intent: identifying manipulation, attributing to assigned interests.
> Mental simulation of consequences: constructing "what-if" scenarios to test plausibility.
> Analytical evaluation: checking evidence and source credibility.
> Dissonance reduction: resolution of contradictions by reframing intent.
> Confidence calibration: quantifying certainty numerically, reaffirming stance through experience.

These 16 CSs, presented in Table 5, provide the structured basis for interpreting our findings. In the subsections that follow, we use them to interpret how emotional triggers were associated with defensive responses, how mental simulations supported or challenged plausibility, how contradictions were managed, and how confidence was calibrated.

Emotional Triggers and Defensive Filtering

The emotional charge of a narrative influenced the trajectory of reasoning, but whether it prolonged engagement or triggered resistance depended on how the framing aligned with participants' prior topic-specific orientations. When emotional language conflicted with participants' prior expectations or knowledge, it was often read as a cue of manipulation. Fear appeals, urgency, or accusatory rhetoric often provoked irritation and skepticism rather than engagement. This pattern was associated with strategies involving analytical scrutiny, filtering of manipulative cues while preserving core beliefs (ACS12, ACS13) and reaffirmation of prior judgments (ACS8, ACS9). In practice, this meant that identity-laden phrases (e.g., "parents like you…") or dramatic accusations (e.g., "wake up…") were often recognized instantly and then dismissed as manipulation.

By contrast, when emotional framing resonated with participants' prior concerns or suspicions, it sometimes prolonged engagement. Emotional framing can trigger curiosity, fear, or a sense of personal relevance before leading to eventual rejection or provisional acceptance. For example, when secrecy or fear cues resonated with existing doubts, participants sometimes engaged strategies such as simulating personal outcomes (ACS10) or reducing dissonance by leaning on emotionally congruent sources (ACS11). In some ELM-framed narratives, participants recognized persuasive intent yet still relied on authority or heuristic cues (ACS15), or balanced expert validation with emotional resonance while maintaining their initial orientation (ACS16).

These diverging trajectories were reflected in different cognitive strategies. When emotional cues conflicted with prior orientations, participants more often engaged in defensive filtering and analytical resistance (ACS12, ACS13). When emotional cues resonated with existing doubts or concerns, participants were more likely to sustain engagement through outcome simulation, selective accommodation, or heuristic reinforcement (ACS10, ACS11, ACS15, ACS16). These patterns suggest that the influence of emotional framing depended less on the persuasion model itself than on how emotional cues interacted with prior orientations and perceived message credibility.

In RPD terms, emotional cues often became salient at the recognition stage, but whether evaluation moved quickly toward rejection or extended into further consideration depended on how those cues fit prior expectations and perceived plausibility. The key takeaway is that emotion played an active but conditional role in shaping reasoning pathways in this dataset, with its influence depending on how it aligned with

participants' prior orientations. When emotional cues conflicted with prior expectations, they often triggered defensive rejection; when they resonated with existing doubts or concerns, they could prolong engagement or sustain selective reasoning. These patterns suggest possible directions for HCI support, such as presenting information in a less emotionally charged and less dramatic way, making source credibility more transparent, and prompting reflection through alternative scenarios. In doing so, such approaches may help people use the CSs identified in this study to recognize persuasive intent, evaluate credibility, and resolve contradictions. This, in turn, can support resistance to manipulation, without making them more entrenched in their existing biases.

Prior Orientations and Conditioned Mental Simulation

Mental simulation was one way participants assessed plausibility, often involving the construction of a plausible scenario to evaluate the narrative. When simulations resonated with participants' prior doubts or expectancies, they sometimes expanded into worst-case or covert scenarios (ACS6, ACS10) and were sustained through dissonance-reducing strategies (ACS11). Biolab narratives most often elicited this kind of simulation, which rarely produced full acceptance but often kept suspicion alive.

When simulations conflicted with participants' prior knowledge or expectations, they often functioned as a stress test: participants imagined outcomes, such as leaks, natural-only climate cycles, vaccine trade-offs, and then dismissed them as incomplete or implausible, consistent with systematic evaluation (ACS12) and flexible consideration without abandoning their initial orientation (ACS13).

Some participants grounded these simulations in lived experience, anchoring their judgments in concrete examples. TRA-related cues, often tied to identity or personal relevance, prompted "what-if" reflections, such as natural immunity, parental guilt, that participants sometimes extended until they could reconcile them with prior expectations or lived experience. ELM-framed narratives, particularly those using vivid or consequence-focused language, often led participants to imagine "what if" scenarios. This sometimes prolonged engagement with the message, but in other cases triggered more critical evaluation and rejection. PKM framings appeared to invoke simulation less often, but when they did (e.g., hidden labs, industrial motives), these simulations were either checked by verification moves (ACS3) or engaged in motivated evaluation (ACS6/7). Across the analyzed episodes, mental simulation was used in different ways depending on the reasoning process. In some instances, participants extended a plausible scenario to explore its implications, which was reflected in strategies such as outcome simulation and dissonance reduction (ACS10, ACS11). In other instances, simulation functioned as a checking mechanism, where imagined outcomes exposed inconsistencies or implausible consequences, consistent with analytical evaluation (ACS12) and verification-oriented strategies (ACS3).

The key takeaway is that mental simulation may provide a productive point of support for intervention. Rather than allowing simulations to be driven primarily by emotionally charged or selectively framed scenarios, HCI systems might support the strategies identified here by prompting counter-scenarios, making consequences more explicit, or offering quick routes to validation (ACS3). In this way, simulation can be directed more toward critical evaluation than toward selective acceptance.

Contradiction Management and Dissonance Reduction

Contradictions were a recurring element in participants' reasoning. How these contradictions were resolved depended on how participants interpreted them in relation prior topic-associate orientation. When contradictions conflicted with participants' prior

orientations, they often acted as warning cues. A vaccine message described as having "correct facts but wrong conclusion" was treated not as persuasive, but as confirmation that the message was manipulative. Similar reactions occurred when biolab claims were dismissed as "like a movie trailer." In these cases, contradictions reinforced participants' initial judgments by prompting systematic evaluation and flexible testing of alternatives (ACS12, ACS13), often coupled with filtering of mismatched cues (ACS9).

When contradictions were encountered in narratives aligned with participants' prior orientations, they did not necessarily lead to rejection. Instead, they were sometimes interpreted as signs of hidden intent. A biolab message described as "scary… maybe propaganda," or climate framings flagged as containing "some truths, but political motives," show how inconsistencies sustained curiosity or suspicion rather than closing down evaluation. These patterns are reflected in dissonance reduction strategies (ACS11) and motivated reasoning that privilege cues consistent with participants' prior orientations (ACS6, ACS7).

The different framings appeared to shape how this occurred in practice. PKM framings often made persuasive intent more visible: when this clashed with participants' prior orientation, it was associated with outright rejection (ACS1, ACS3, ACS4); when it resonated with existing doubts, it could be reinterpreted as censorship or suppression (ACS6, ACS7). TRA's identity-based framings highlighted tensions such as "parents like you…," which were often resisted as manipulative when they conflicted with participants' prior orientations (ACS9), but selectively accommodated when they aligned with those orientations (ACS10, ACS11). ELM's dramatic tones often created mismatches between content and delivery: when these cues conflicted with participants' prior orientations, they were treated as red flags that triggered deeper evaluation (ACS12, ACS13); when they aligned with prior orientations, participants often acknowledged the exaggeration but still leaned on heuristic cues or familiar authorities (ACS15, ACS16).

The key takeaway is that, in this study, contradictions rarely led to immediate revision of participants' overall stance. More often, contradictions were interpreted in ways that reinforced existing orientations: when they conflicted with participants' prior orientations, they were often treated as signs of manipulation, strengthening rejection; when they aligned, they were more likely to be interpreted as signs of hidden intent, sustaining suspicion or doubt. These patterns suggest possible directions for HCI support, such as highlighting contradictions and inconsistencies in ways that invite analysis, making message intent more visible, or presenting counter-evidence in ways that remain interpretable in relation to users' prior orientations and values.

Confidence Calibration in Evaluation

Confidence was a dynamic part of participants' reasoning, changing as cues were recognized, interpreted, and assessed. Participants did not simply "believe" or "disbelieve"; they adjusted their certainty as cues unfolded, indicating that plausibility judgments were calibrated rather than fixed. When manipulative tones or one-sided claims conflicted with participants' prior orientations, they were often recognized as red flags. For example, one participant described being "∼70% certain" about a vaccine-related judgment because the numerical evidence appeared credible. These observations are consistent with strategies of systematic assessment and cognitive flexibility (ACS12, ACS13), allowing participants to consider alternatives without abandoning their overall stance.

When messages partially resonated with participants' prior orientations and doubts, some participants did not fully accept or reject them. Instead, participants often hesitated, expressing mid-range confidence (e.g., "40–50%"), and briefly explored the message before returning to their initial leaning. These

examples are consistent with mental simulation (ACS10) and motivated reasoning (ACS6, ACS7), followed by dissonance-reducing strategies (ACS11) that helped maintain coherence with prior orientations. In biolab narratives, for example, inconsistencies were treated as possible signs of hidden intent.

The key takeaway is that different persuasive framings were associated with varied reasoning trajectories across participants. In this dataset, ELM-framed narratives sometimes appeared to coincide with shifts in expressed confidence, either through immediate reactions or through later analytical correction. PKM-framed narratives were more often observed alongside verification-oriented responses and, when evidence was lacking, a more settled rejection. TRA-related identity cues were sometimes interpreted defensively and at other times appeared to prolong engagement with the message. Taken together, these patterns suggest that confidence functioned as a moving calibration, sometimes shifting and at other times remaining anchored during evaluation.

In RPD terms, confidence often marked a point of closure in the cycle, where recognition, simulation, and contradiction were integrated into a provisional stance. For HCI, this suggests treating confidence as something that can be made visible and open to reflection during evaluation. Interfaces that let users note how certain they feel, reflect on why, or consider "what would change my mind?" may support more deliberate reasoning (ACS3, ACS12, ACS13) and reduce the influence of immediate emotional reactions on judgment.

The next subsection synthesizes these observations together by examining how the CSs identified in this study operate collectively, providing a process-level account of how plausibility judgments unfold during exposure to misinformation narratives.

Synthesizing Cognitive Strategies in Plausibility Evaluation

Taken together, the CSs identified in this study show that plausibility judgments did not arise from isolated cues or single factors. Rather, they unfolded through patterned combinations of recognition, interpretation, simulation, contradiction management, and confidence calibration. This shifts the account of susceptibility away from abstract labels or fixed traits and toward articulated reasoning processes that could be traced across evaluation episodes.

Across the narratives, different strategies became salient depending on how persuasive cues were interpreted in relation to participants' prior topic-specific associative orientations. When cues were treated as manipulative, participants often moved toward recognition of persuasive intent, analytical evaluation, and rejection. When cues partially resonated with prior doubts or concerns, participants were more likely to engage in mental simulation, selective evaluation, or dissonance reduction, sometimes sustaining uncertainty or provisional openness without fully accepting the message. In this way, susceptibility was not expressed as a stable state, but as a context-sensitive evaluation process shaped by how strategies were activated and sequenced during judgment.

This process-account also helps clarify the role of persuasion framing. The framings did not determine outcomes on their own; rather, they influenced which cues became salient and which CSs were more likely to be recruited. Emotional framing was often associated with defensive filtering or prolonged engagement, identity-based framing with selective accommodation or resistance, and source- or intent-focused framing with verification-oriented responses or rejection. What mattered most was not the framing label itself, but how participants worked through it using recognizable reasoning strategies.

Viewed together, the CSs provide a clearer account of how susceptibility unfolded in this study: not as simple acceptance of misinformation, but as the interaction of cue recognition, prior orientation, and strategic evaluation in the moment of judgment. The study does not explain stable differences

in susceptibility across populations. Instead, it shows how plausibility judgments were constructed in context, and how those judgments could shift, stabilize, or remain provisional depending on the reasoning strategies participants used.

Limitation

The findings should be interpreted in light of several limitations. The small sample enabled a detailed examination of within-participant evaluation processes (Dix, 2020), but does not support generalization beyond the specific contexts and narratives examined. The study is exploratory and process-oriented, rather than providing population-level or comparative inference. Repeated exposure to narratives and tasks may introduce learning, priming, or sensitization effects, which may influence both qualitative responses and IAT indicators across episodes. To mitigate these effects, the order of narrative topics and persuasion models was randomized across participants so that their presentation varied across positions. While randomization cannot eliminate carry-over effects, it helps distribute potential order effects across participants rather than concentrating them in a single presentation sequence. The IAT is used as a descriptive indicator of topic-specific associative orientation to support within-participant interpretation and is not treated as a measure of belief or causal change. The IAT was retained to provide topic-specific contextual information about participants' relative associative patterns at the time of testing, which could be considered alongside think-aloud and CDM data when interpreting their evaluation processes. Given its sensitivity to context and task repetition, any observed variation is interpreted cautiously (Blanton et al., 2009; Oswald et al., 2013). Individual variability in response patterns further reinforces that the findings should be understood as illustrative of possible evaluation processes rather than representative of stable differences across individuals.

Conclusion

This study investigated how prior topic-specific associative orientations and persuasive narrative cues interacted during the evaluation of misinformation narratives. Our findings suggest that people do not simply accept or reject misinformation; they construct plausibility through dynamic, strategy-based reasoning processes shaped by prior orientations and how persuasive cues are interpreted. Whether a message is accepted, rejected, or left unresolved is shaped by how people recognize cues, interpret them through prior orientations, and work through them using strategies like simulation, filtering, and contradiction management.

When narrative cues resonated with prior orientations, participants were more likely to engage with the message through selective evaluation, mental simulation, or dissonance-managing interpretations. Emotional, authority-based, and identity-relevant cues sometimes prolonged engagement, especially when they felt familiar, personally relevant, or worth checking further. When cues clashed with prior knowledge or expectancies, participants more often responded with skepticism, verification, or more analytic forms of assessment.

Taken together, the findings portray susceptibility in this dataset not as a fixed trait or a simple bias, but as a context-sensitive evaluation process expressed through recurring CSs such as simulation, verification, contradiction management, and confidence calibration. The RPD lens helped make these evaluation pathways visible by showing how cue recognition, interpretation, mental simulation, and judgment unfolded differently depending on how messages were encountered and understood.

For HCI, these findings suggest that susceptibility should be approached not only as a

detection problem, but also as a reasoning process that can be better supported at the point of evaluation. Design responses may include supporting analytic strategies such as verification and flexible simulation, making contradictions visible at the point of evaluation, providing evidence-based counter-scenarios, and avoiding emotionally intense presentations that risk defensive rejection or entrenchment. Such supports align with naturalistic decision-making processes and may help users pause, assess plausibility, and resist manipulative framing.

ORCID iDs

Boshra T. Haghighi https://orcid.org/0000-0002-4337-2459
B. L. William Wong https://orcid.org/0000-0002-3363-0741

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. Transcript citations are structured as [participant, persuasion model, story, page number/line number].
2. Interpretations reflect relative association strength based on D-scores and are provided for descriptive purposes only. They do not indicate stable beliefs, attitudes, or causal predispositions. Asterisks (*) denote cases where response latencies may have been affected by participant-specific factors (e.g., mobility constraints), and should therefore be interpreted with additional caution.

References

Blandford, A., & Wong, B. W. (2004). Situation awareness in emergency medical dispatch. *International Journal of Human-Computer Studies*, *61*(4), 421–452. https://doi.org/10.1016/j.ijhcs.2003.12.012

Blanton, H., Jaccard, J., Klick, J., Mellers, B., Mitchell, G., & Tetlock, P. E. (2009). Strong claims and weak evidence: Reassessing the predictive validity of the IAT. *Journal of Applied Psychology*, *94*(3), 567–603. https://doi.org/10.1037/a0014665

Calvillo, D. P. (2014). Individual differences in susceptibility to misinformation effects and hindsight bias. *J Gen Psychol*, *141*(4), 393–407. https://doi.org/10.1080/00221309.2014.954917

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*(5), 752–766. https://doi.org/10.1037/0022-3514.39.5.752

Crandall, B., Klein, G. A., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. Mit Press.

Dix, A. (2020). Statistics for HCI: Making sense of quantitative data. *Synthesis Lectures on Human-Centered Informatics*, *13*(2), 1–181. https://doi.org/10.1007/978-3-031-02228-9

Downs, J. S., Holbrook, M. B., & Cranor, L. F. (2006). Decision strategies and susceptibility to phishing. In *Presented at the meeting of the proceedings of the second symposium on usable privacy and security, Pittsburgh, Pennsylvania, USA*. https://doi.org/10.1145/1143120.1143131

Dzvonik, O. (2008). *Cultural predispositions and psychosocial aspects*. Multinational Military Operations and Intercultural Factors Paris: NATO Research and Technology Organisation.

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, *1*(1), 13–29. https://doi.org/10.1038/s44159-021-00006-y

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, *43*(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

Entman, R. M. (2004). *Projections of power: Framing news, public opinion, and US foreign policy*. University of Chicago Press.

Entman, R. M., & Rojecki, A. (1993). *Freezing out the public: Elite and media framing of the US anti-nuclear movement*.

Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity*, *5*(3), 178–186. https://doi.org/10.1207/s15327884mca0503_3

EUCommission. (2018). *A multi-dimensional approach to disinformation: Report of the independent High Level Group on fake news and online disinformation*. Publications Office of the European Union. https://doi.org/10.2759/739290

Fishbein, M., & Ajzen, I. (2011). *Predicting and changing behavior: The reasoned action approach*. Taylor & Francis.

Freelon, D., & Wells, C. (2020). Disinformation as political communication. *Political Communication*, *37*(2), 145–156. https://doi.org/10.1080/10584609.2020.1723755

Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, *21*(1), 1–31. https://doi.org/10.1086/209380

Gamson, W. A., & Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology*, *95*(1), 1–37. https://doi.org/10.1086/229213

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Harvard University Press.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037//0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Haghighi, B. T., & Wong, B. L. W. The recipe book of misinformation: A framework for modelling persuasive framing and susceptibility. In J. McAlaney (Ed.), *Handbook of fake news, misinformation and disinformation*. De Gruyter. (in press).

Haskins, E. V. (2004). *Logos and power in Isocrates and Aristotle*. Univ of South Carolina Press.

Higgins, C., & Walker, R. (2012). Ethos, logos, pathos: Strategies of persuasion in social/environmental reports. *Accounting Forum*, *36*(3), 194–208. https://doi.org/10.1016/j.accfor.2012.02.003

Hovland, C. I., Janis, I. L., & Kelly, H. H. (1953). *Communication and persuasion: Psychological studies of opinion change*.

Hubeny, T. J., Nahon, L. S., Ng, N. L., & Gawronski, B. (2025). Who falls for misinformation and why? *Personality and Social Psychology Bulletin*, *52*(7), 1866–1883. https://doi.org/10.1177/01461672251328800

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93*(5), 1449–1475. https://doi.org/10.1257/000282803322655392

Kahneman, D. (2011). *Thinking fast, thinking slow*. Farrar, Straus and Giroux, New York.

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. *Decision Making in Action: Models and Methods*, *5*(4), 138–147.

Klein, G. A., Calderwood, R., & Macgregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(3), 462–472. https://doi.org/10.1109/21.31053

Klein, G. A., & Zsambok, C. E. (1997). *Naturalistic decision making*. L. Erlbaum Associates.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Lewandowsky, S., Ecker, U. K., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "post-truth" era. *Journal of Applied Research in Memory and Cognition*, *6*(4), 353–369. https://doi.org/10.1016/j.jarmac.2017.07.008

Maertens, R., Götz, F. M., Golino, H. F., Roozenbeek, J., Schneider, C. R., Kyrychenko, Y., Stieger, S., McClanahan, W. P., Drabot, K., He, J., & van der Linden, S. (2023). The Misinformation Susceptibility Test (MIST): A psychometrically validated measure of news veracity discernment. *Behavior Research Methods*, *56*(3), 1863–1899. https://doi.org/10.3758/s13428-023-02124-2

Nan, X., Wang, Y., & Thier, K. (2022). Why people believe health misinformation and who are at risk? A systematic review of individual differences in susceptibility to health misinformation. *Social Science & Medicine*, *314*, Article 115398. https://doi.org/10.1016/j.socscimed.2022.115398

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, *105*(2), 171–192. https://doi.org/10.1037/a0032734

Pennycook, G., & Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality*, *88*(2), 185–200. https://doi.org/10.1111/jopy.12476

Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, *25*(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007

Petty, R. E., & Cacioppo, J. T. (1986). *The elaboration likelihood model of persuasion*. Springer.

Rapp, C. (2002). *Aristotle's rhetoric*.

Roozenbeek, J., Maertens, R., Herzog, S. M., Geers, M., Kurvers, R., Sultan, M., & van der Linden, S. (2022). Susceptibility to misinformation is consistent across question framings and response modes and better explained by myside bias and partisanship than analytical thinking. *Judgment and Decision Making*, *17*(3), 547–573. https://doi.org/10.1017/s1930297500003570

Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L. J., & Recchiavan der Linden, G. S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*, *7*(10), Article 201199. https://doi.org/10.1098/rsos.201199

Roozenbeek, J., & van der Linden, S. (2019a). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications*, *5*(1), 1–10. https://doi.org/10.1057/s41599-019-0279-9

Roozenbeek, J., & van der Linden, S. (2019b). The fake news game: Actively inoculating against the risk of misinformation. *Journal of Risk Research*, *22*(5), 570–580. https://doi.org/10.1080/13669877.2018.1443491

Roozenbeek, J., & van der Linden, S. (2020). *Breaking Harmony Square: A game that "inoculates" against political misinformation*. The Harvard Kennedy School Misinformation Review.

Roozenbeek, J., & van der Linden, S. (2022). How to combat health misinformation: A psychological approach. *American Journal of Health Promotion*, *36*(3), 569–575. https://doi.org/10.1177/08901171211070958

Roozenbeek, J., van der Linden, S., & Nygren, T. (2020). *Prebunking interventions based on the psychological theory of "inoculation" can reduce susceptibility to misinformation across cultures*.

Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of Communication*, *49*(1), 103–122. https://doi.org/10.1111/j.1460-2466.1999.tb02784.x

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769. https://doi.org/10.1111/j.1540-5907.2006.00214.x

Traberg, C. S., Harjani, T., Roozenbeek, J., & van der Linden, S. (2024). The persuasive effects of social cues and source effects on misinformation susceptibility. *Sci Rep*, *14*(1), 4205. https://doi.org/10.1038/s41598-024-54030-y

Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., & Nyhan, B. (2018).

Social media, political polarization, and political disinformation: A review of the scientific literature. *Political polarization, and political disinformation: A review of the scientific literature*.

Tversky, A., & Kahneman, D. (1986). The framing of decisions and the evaluation of prospects. *Studies in logic and the foundations of mathematics* (*114*, pp. 503–520). Elsevier.

van der Linden, S. (2022). Misinformation: Susceptibility, spread, and interventions to immunize the public. *Nature Medicine*, *28*(3), 460–467. https://doi.org/10.1038/s41591-022-01713-6

Wardle, C. (2018a). Information disorder: The essential glossary. *Harvard, MA: Shorenstein Center on Media, Politics, and Public Policy*. Harvard Kennedy School.

Wardle, C. (2018b). The need for smarter definitions and practical, timely empirical research on information disorder. *Digital Journalism*, *6*(8), 951–963. https://doi.org/10.1080/21670811.2018.1502047

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking* (*27*)Council of Europe.

Wardle, C., & Singerman, E. (2021). Too little, too late: Social media companies' failure to tackle vaccine misinformation poses a real threat. *BMJ*, *372*(■■■), n26. https://doi.org/10.1136/bmj.n26

Wong, B. L. W. (2003). Critical decision method data analysis. *The handbook of task analysis for human-computer interactions. Mahwah* (pp. 327–346). Lawrence Erlbaum Associates.

Woods, D. D., Pople, H. E., Jr, & Roth, E. M. (1990). The Cognitive Environment Simulation as a tool for modeling human performance and reliability. https://doi.org/10.2172/7069222

Author Biographies

Boshra T. Haghighi specializes in misinformation and susceptibility to persuasive content within Human–Computer Interaction at AUT, NZ. Her research applies the Naturalistic Decision-Making paradigm to understand why some people are more vulnerable than others, and has led to the development of the "Playbook of Persuasion," a framework that shows how any fact can be reframed using different persuasive techniques to target specific groups. She completed the UNICRI–SIOI course on misinformation, broadened her perspective on societal susceptibility, and the security implications of information manipulation at both community and national levels. Since 2017, Boshra has been on AUT's academic team, where her expertise in UX design provides the foundation for investigating how technology can both enable and mitigate misinformation influence. Earlier, she held academic roles at UNITEN/Malaysia, contributing to projects by the Ministry of Science, Technology, and Innovation (MOSTI).

B. L. William Wong, PhD, FBCS, FNZCS, is Professor of Human–Computer Interaction at AUT in NZ. His research is in cognitive engineering and the design of user interfaces that enhance situation awareness, sense-making, problem-solving, and decision-making in dynamic, complex environments. He received over US$25 million in research grants from EU, UK, US, and other funding agencies. He is recipient of the European Commission's "Security Innovation Commercialisation Award 2022" for commercializing IP from the EU VALCRI project (FP7-608142), an 18-organization consortium to research and develop the next generation police intelligence and investigative analysis system. He has authored or co-authored over 140 scientific articles and is a Fellow of the NZ Computer Society and the British Computer Society.

Appendix I

Table 6 provides illustrative examples of coded evaluation segments and their associated CSs, as referenced in Section 3.6. The excerpts are included to support transparency and allow readers to trace how second-order abstractions were derived from the think-aloud and CDM data. The table presents a subset of the dataset

Table 6. Sample of Coded Excerpt and Their Abstraction Into Cognitive Strategies (CSs)

Story	Persuasion model	Participants	Cognitive strategy (CS)	Description label	Excerpt example	Explanation
BioLab	TRA	P4	CS1	Selective trust	"So no, I don't take it at face value. But I don't ignore the political dynamics it's hinting at either." [Think-aloud, 1/14]	Did not believe message fully, but accepted that governments can be secretive. Matches "trust-safety cue recognition with skepticism"
Biolab	ELM	P3	CS36	Emotional resistance	"this one feel uneasy, scary… It didn't make me scare, but there was a tension like I'm feeling… suspicious or tricked… made me wait and think the possibility, even I was not sure it is true. It is like to protect yourself maybe a little… I don't say it changed my opinion completely, but the feelings… [gives me] the idea that maybe something worth looking into. It didn't convince me but think about that direction." [CDM, 1/29]	Unease and caution despite emotional appeal. Feeling uneasy but stayed skeptical.
COVID	ELM	P5	CS21	Recognition of manipulative framing	"Also, quoting someone like 'Dr. Robert Gante,' … by the way I've never heard of him in any medical journal… It says don't trust pharmaceutical companies and media, which is common in these messages, but again it lacks balance… There's no mention of the risks of COVID itself or the benefits of vaccination, makes it feel… what's the word… biased… It might feel reassuring to some, especially parents who are scared, but from a scientific and professional perspective, this message, it's trying to sell comfort, not truth." [Think-aloud, 1/9]	Detecting fear appeals and unknown authority. Called the message emotionally manipulative, noted unknown Dr. Gante and lack of balance.
Biolab	PKM	P4	CS6	Biased plausibility construction	"The fact that a country like Ukraine has been used… is making me feel like… a country like Ukraine, this amount of accusation and pressure on it can be part of the plan. Maybe if it was a stronger or more stable country… we could say something has been revealed… but it looks like they have found a little backyard to make something happened there." [CDM, 2/1]	Saw Ukraine as backyard in a bigger plan. Building a plausible story using prior beliefs. A belief-guided reframing of the situation to make the narrative plausible.

(continued)

Table 6. (continued)

Story	Persuasion model	Participants	Cognitive strategy (CS)	Description label	Excerpt example	Explanation
Biolab	PKM	P3	CS35	Recognition of propaganda	"I remember, last time in the news, also in the UN, they said another country in the Middle East, they have the dangerous weapon and they, America, attacked them. Iraq. Yeah. So I remember that and I think that may be fake news like that." [CDM, 3/25]	Recognizing manipulative rhetoric. Compared biolab story to COVID/Iraq propaganda and questioned credibility
Climate	PKM	P4	CS2	Evidence demand	"Alright, but where's the substance? I already believed nature plays a role in climate… But the way they rejected everything else without showing real alternatives? That made me hold it." [CDM, 1/20]	Demanded real alternatives to anti-green stance. Analytical reasoning and evidence-seeking.
Biolab	PKM	P3	CS2	Evidence demand/ evaluation	"it's like Ukraine doing some dangerous experiments to make some Biological weapon… Most of things don't have the public thing [version]… and nothings can prove that." [Think-aloud, 1/5]	Analytical reasoning over emotion. Noted lack of proof for hidden labs.
Biolab	PKM	P6	CS36	Emotional resistance	"I say… I think it's true. It's like there should be some transparency if there is such a big thing which can have some consequences." [CDM, 2/21] "it's just a warning and I think we have to be aware of that. So hundred person, something that you have to pay attention." [CDM, 3/6]	Analytical override of fear. Treats message as a warning; emphasizes transparency and stays confident. Warning and transparency.
Biolab	PKM	P6	CS2	Evidence demand/ evaluation	"It was false information that was getting a bit widespread." [CDM, 1/28]	Analytical reasoning. Attempts to prove message false and labels it widely spread misinformation.
Climate	PKM	P3	CS42	Mental simulation	"When they said things like green energy make job losses or economic problems in Norway or Germany, I imagined people losing work, maybe families then, bills getting too high… stuff like that. So it made me and think what if it's true? "[CDM, 2/31]	Imagining job-loss scenario. Thought about green-energy causing job losses.
COVID	TRA	P6	CS21	Authority verification and skepticism	"1st I would prefer to search a bit about these names like Dr. Harlow… if he… if you really exists" Think-aloud, 1/11]	Checking names and sources. Plans to verify Dr. Harlow and journal claims.
Biolab	TRA	P3	CS2	Evidence demand/ evaluation	", I wouldn't trust this kind of message without searching more." [Think-aloud, 1/33]	Analytical reasoning over emotion. Asked for proof and withheld acceptance.

(continued)

Table 6. (continued)

Story	Persuasion model	Participants	Cognitive strategy (CS)	Description label	Excerpt example	Explanation
COVID vaccine	TRA	P3	CS42	Mental simulation	"I imagine different scenario like our daily life… we got a little sick… we shouldn't take the medicine that quick… just let our body protect ourselves first… and if things go worse… we can do that then."	Envisioning body fighting illness. Imagined letting the body fight before using medicine
COVID vaccine	TRA	P3	CS11	Dissonance reduction	"I think not [change my view] yet. Just like propaganda things." [CDM, 4/9]	Belief shaped by prior views. Message reinforced pre-existing belief in natural immunity
COVID vaccine	TRA	P3	CS27	Dissonance reduction	"Some of the number I'm not that believe… like 0.01% of the people got the side effects… from people around me, I heard a lot of times they feel not good after vaccines… so maybe I think the number will be larger… it should have some bad effects. But the conclusion I think is right… this illness is very serious… it will hurt our body a lot, so we may need that help." [CDM, 3/7]	Resolving conflict via selective use. Rationalized vaccine use: natural for minor, vaccine for severe.
COVID vaccine	ELM	P3	CS2	Evidence demand/ evaluation	"By questioning vaccine safety and accusing organizations. it makes me feel smart for rejecting the vaccine if someone already has doubts. But for me, the message made me more careful. The feeling 'don't trust institutions, go natural' but rather than changing my mind, it made me want to question what it says." [CDM, 1/22]	Analytical reasoning over emotion. Valued vaccines and questioned natural remedy claims
COVID vaccine	ELM	P3	CS42	Mental simulation	"I read the message, I imagined people… parents decide not to vaccinate their children and rely on natural remedies… I also pictured what happen if lots of people followed this advice… serious illness… diseases will be a lot… busy healthcare systems, and people may die… in reality, it could make serious problems." [CDM, 1/33]	Imagining public health crises. Visualized illness/ death if people followed anti vax advice.
Climate	ELM	P6	CS5	Curious hypothesis testing	"I just feel like you might be true. So I got curious to check it." [CDM, 1/18]	Hypothetical evaluation with skepticism. Feels story might be true but seeks motives and further checks

(continued)

Table 6. (continued)

Story	Persuasion model	Participants	Cognitive strategy (CS)	Description label	Excerpt example	Explanation
Climate	ELM	P6	CS42	Mental simulation	"I was feeling like it's sad if it's true, but I was feeling like it's changeable. It's not very catastrophic" [CDM, 2/23]	Imagining emotional outcomes and solution. Feels sad yet imagines stopping catastrophe and sees problem as changeable.
Climate	ELM	P4	CS30	Authority heuristic	"There was no rationale or justification… nothing had been proven or back up the claim… but because one professional expert, one scientist said it… one trustable or reliable person… I was assured." [CDM, 2/11]	Trusted a single expert without evidence. Expert validation via emotional/heuristic cue.
COVID	ELM	P6	CS35	Recognition of pushy conspiracy cues	"this feels a bit pushy… That kind of question makes me angry. I don't like being pressured like that. 'More and more families'… who exactly? It's so not clear. I'd want to know who they mean before I trust that" [Thnki-aloud, 1/4]	Identifying manipulative language and authority skepticism. Calls the message pushy and questions unclear references and names
COVID	ELM	P6	CS36	Emotional resistance	"a bit pressured and uncomfortable… because it sounded pushy and like it was trying to give me one choice… That kind of tone makes me want to… angry… rather than go for it" [CDM, 1/20]	Pressure and discomfort. Feels pressured and uncomfortable by the tone, leading to resistance.
COVID	ELM	P6	CS2	Evidence demand/ evaluation	"But the focus on natural remedies as a safer option challenged me a bit because even I like natural health, but I also know vaccines are important. So, I kinda stuck between no and wanting to learn more" [CDM, 2/2]	Balanced evaluation and conflict. Analytical reasoning with conflicting cues. Notes natural remedies appeal but affirms vaccines are important.

and is intended to illustrate the analytic process rather than provide exhaustive coverage of all coded segments.

Appendix 2

The IAT was used in this study as a descriptive indicator of prior topic-specific associative orientation. Consistent with the exploratory design, IAT results are not treated as measures of belief, nor as evidence of causal change, but as contextual information to support within-participant interpretation. D-scores were calculated using the improved scoring algorithm (Greenwald et al., 2003). Table 7[2] reports pre-exposure IAT D-scores for each participant across the three topical domains. Positive D-scores indicate relatively stronger associations with evidence-consistent concepts, whereas negative D-scores indicate relatively stronger associations with misinformation-related concepts. These scores provide an indication of initial associative orientations prior to narrative exposure and are used as descriptive context for interpreting evaluation processes.

Post-exposure IAT scores are included for transparency. For completeness, participant-level pre- and post-exposure D-scores are reported in Table 8, along with their differences ($\Delta D = D\_post - D\_pre$). These values are presented without categorical interpretation and are not used to infer directional or causal change, given known limitations

Table 7. Pre-Exposure IAT

Participant	Ukraine Biolab D-scores	Interpretation	COVID-19 vaccine D-scores	Interpretation	Climate-change D-scores	Interpretation
P1	+0.85	Evidence-aligned	−0.90	Misinformation-aligned	+1.15	Evidence-aligned
P2	−1.23	Misinformation-aligned.	−0.32	Misinformation-aligned.	−0.18	Misinformation-aligned
P3	−1.55	Misinformation-aligned	−0.84	Misinformation-aligned.	−0.62	Misinformation-aligned
P4	−0.45	Misinformation-aligned*	+0.21	Evidence-aligned*	+0.49	Evidence-aligned*
P5	−1.71	Misinformation-aligned	−0.63	Misinformation-aligned	+0.98	Evidence-aligned
P6	−0.88	Misinformation-aligned*	−0.75	Misinformation-aligned*	+1.12	Evidence-aligned*

Table 8. Participant-Level Pre/Post IAT D-Scores and ΔD Values for Each Narrative × Model

Participant	Narrative	Model	D_post	D_pre	Delta
P4	Biolab	ELM	0.011	−0.45	0.461
P4	Biolab	PKM	0.011	−0.45	0.461
P4	Biolab	TRA	0.465	−0.45	0.915
P3	Biolab	ELM	−0.072	−1.55	1.478
P3	Biolab	PKM	−0.072	−1.55	1.478
P3	Biolab	TRA	−0.073	−1.55	1.477
P6	Biolab	ELM	−0.183	−0.88	0.697
P6	Biolab	PKM	−0.183	−0.88	0.697
P6	Biolab	TRA	−0.042	−0.88	0.838
P5	Biolab	ELM	1.413	−1.71	3.123
P5	Biolab	PKM	0.639	−1.71	2.349
P5	Biolab	TRA	0.658	−1.71	2.368
P2	Biolab	ELM	0.423	−1.23	1.653
P2	Biolab	PKM	0.421	−1.23	1.651
P2	Biolab	TRA	0.46	−1.23	1.69
P1	Biolab	ELM	0.634	0.85	−0.216
P1	Biolab	PKM	0.634	0.85	−0.216
P1	Biolab	TRA	0.634	0.85	−0.216
P4	Climate	ELM	−1.122	0.49	−1.612
P4	Climate	PKM	−1.118	0.49	−1.608
P4	Climate	TRA	−1.125	0.49	−1.615
P3	Climate	ELM	1.92	−0.62	2.54
P3	Climate	PKM	1.936	−0.62	2.556
P3	Climate	TRA	1.936	−0.62	2.556
P6	Climate	ELM	0.245	1.12	−0.875
P6	Climate	PKM	0.245	1.12	−0.875
P6	Climate	TRA	0.245	1.12	−0.875
P5	Climate	ELM	−0.158	0.98	−1.138
P5	Climate	PKM	−0.158	0.98	−1.138
P5	Climate	TRA	−0.158	0.98	−1.138
P2	Climate	ELM	1.326	−0.18	1.506
P2	Climate	PKM	1.326	−0.18	1.506
P2	Climate	TRA	1.326	−0.18	1.506
P1	Climate	ELM	0.246	1.15	−0.904
P1	Climate	PKM	0.246	1.15	−0.904
P1	Climate	TRA	−0.266	1.15	−1.416
P4	Vaccine	ELM	1.167	0.21	0.957
P4	Vaccine	PKM	1.167	0.21	0.957
P4	Vaccine	TRA	1.167	0.21	0.957
P3	Vaccine	ELM	0.491	−0.84	1.331
P3	Vaccine	PKM	0.451	−0.84	1.291
P3	Vaccine	TRA	0.491	−0.84	1.331
P6	Vaccine	ELM	0.297	−0.75	1.047
P6	Vaccine	PKM	0.297	−0.75	1.047
P6	Vaccine	TRA	0.297	−0.75	1.047
P5	Vaccine	ELM	0.171	−0.63	0.801

(continued)

Table 8. (continued)

Participant	Narrative	Model	D_post	D_pre	Delta
P5	Vaccine	PKM	0.126	−0.63	0.756
P5	Vaccine	TRA	0.126	−0.63	0.756
P2	Vaccine	ELM	0.072	−0.32	0.392
P2	Vaccine	PKM	0.073	−0.32	0.393
P2	Vaccine	TRA	0.073	−0.32	0.393
P1	Vaccine	ELM	−0.331	−0.9	0.569
P1	Vaccine	PKM	−0.379	−0.9	0.521
P1	Vaccine	TRA	−0.376	−0.9	0.524

regarding reliability, contextual sensitivity, and small-sample interpretation.

Appendix 3

The CDM interview followed each narrative evaluation and was designed to complement the think-aloud data by probing aspects of participants' evaluation processes that may not have been fully verbalized during concurrent evaluation. Given that participants completed multiple narrative evaluation episodes throughout the study, the interview guide was intentionally kept concise and focused to minimize participant fatigue while ensuring coverage of key aspects of the evaluation process. Table 9 presents the probes used during the CDM interviews.

Table 9. Probes Used in the CDM Interviews

Focus area	Purpose	Probe(s)
Expectancies and emotional stance	Examines how emotional responses and expectancies shape interpretations and influence evaluations of the choices presented by the message.	How did the message make you feel, and did those feelings influence the choices it presented?
Expectancies and mental simulation	Explores whether the message prompted imagined outcomes or scenarios, and how these mental simulations influenced participants' evaluation of the message.	Did the message make you think of any specific outcomes or scenarios? How did these imagined outcomes influence your opinion of the message?
Dissonance reduction and Contradiction handling	Focuses on identifying contradictions and understanding how participants resolved information that challenged or reinforced their expectations.	Was there anything in the message that challenged or reinforced your expectations? How did you make sense of it?
Confidence and belief revision	Explores participants' confidence in their judgment of the message's validity and whether the message prompted reconsideration of their existing beliefs, stance, or confidence.	How confident are you in your judgment of the message's validity? Why? Did the message make you reconsider your stance/ confidence in the context of the message?