Susceptibility to Misinformation as Recognition-Based Evaluation

Boshra Talebi Haghighi

A thesis submitted to

Auckland University of Technology

in fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

2026

School of Engineering, Computer and Mathematical Sciences

Faculty of Design and Creative Technologies

Abstract

Misinformation has become a significant societal and security concern due to its influence on issues such as vaccine hesitancy, climate change debates, and democratic institutions. A substantial body of research has identified cognitive, motivational, and social factors shaping susceptibility to misinformation. However, these accounts are often examined as discrete influences and provide limited specification of how individuals evaluate misinformation in context, particularly how they move from encountering a message to forming a judgment under conditions of uncertainty.

This thesis responds to this gap by examining susceptibility to misinformation as a recognition-based evaluation process, drawing on the Recognition-Primed Decision (RPD) framework to investigate how plausibility judgments emerge during evaluation, and how this process is shaped by human factors, persuasive communication design, and prior orientations as reflected in participants' reasoning.

The research began with a literature review to establish the conceptual foundation for investigating susceptibility to misinformation. The review examined human factors associated with susceptibility, persuasive mechanisms that make misinformation compelling, and decision-making concepts needed to study plausibility judgments under uncertainty. From this review, the Recognition-based Susceptibility Framework was developed to bring these strands together in a structured way. This framework then provided the basis for the subsequent studies by guiding how misinformation evaluation, persuasive message construction, and narrative interpretation were examined.

Building on this foundation, the first study investigated how misinformation is evaluated in practice. A Cognitive Task Analysis using the Critical Decision Method (CDM) elicited non-routine evaluation episodes from 10 participants, and the Recognition-Primed Decision (RPD) model was used as an interpretive lens to organize the recurring patterns identified in the data. The study produced a process-oriented account of misinformation evaluation, showing that variation in outcomes arises from how the situation is interpreted and when evaluation is ended.

The research then examined the message side of this process. The study develops a conceptual framework, the Recipe Book of Misinformation (RBM), that formalizes how factual elements can be reframed into multiple plausible misinformation variants. Using persuasion models and framing mechanisms, the RBM provides an analytic account of how message design shapes perceived credibility, emotional engagement, and logical coherence, linking these constructions to recognition-based evaluation under uncertainty.

Building on the process account from the CTA study and the message design logic formalized in the RBM, this final study investigated how participants respond to structured misinformation narratives. A qualitative CTA design with structured variation was used, in which six participants evaluated misinformation narratives using think-aloud protocols and CDM interviews, with Implicit Association Tests interpreted descriptively as indicators of prior topic-specific associative orientations. The study showed how evaluation patterns varied across episodes, with differences in interpretation and assessment reflected in participants' reasoning as they engaged with the controlled message designs.

Taken together, the thesis provides a process-oriented account of susceptibility to misinformation. It shows how plausibility judgments are formed during evaluation and how message construction can shape interpretations under uncertainty. Across the studies, susceptibility is understood not as a stable individual property, but as something that arises during evaluation, reflecting how individuals interpret information and bring evaluation to a close in context.

Table of Contents

Contents

Abstract ii
Table of Contents iv
List of Figures viii
List of Tables viii
Attestation of Authorship x
Co-Authored Works xi
Artificial Intelligence Declaration xiii
Acknowledgements xiv
Ethics Approval xv
Chapter 1: Introduction 1
1.1 Problem Statement and Research Gap 1
1.2 Theoretical Lens 2
1.3 Research Aim and Objectives 2
1.4 Research Questions 3
1.5 Scope and Limitations of the Research 3
1.6 Contribution and Significance of the Thesis 4
1.7 Structure of the Thesis 4
1.8 Conclusion 5
Chapter 2: Conceptualizing Susceptibility to Misinformation 6
2.1 Introduction 6
2.1.1 Review Approach 7
2.2 The Recognition-based Susceptibility Framework for Explaining Susceptibility to Misinformation 8
2.3 Misinformation and Susceptibility 11
2.3.1 Misinformation 11
2.3.2 Susceptibility 11
2.3.3 Current Approaches to Countering Misinformation 12
2.3.4 What These Approaches Explain and What They Leave Unexplained 13
2.4 Human Factors Shaping Susceptibility 15
2.4.1 Decision-Making Under Uncertainty 15
2.4.2 Cognitive Mechanisms That Influence Interpretation 17
2.4.3 Communication Frames and Meaning Construction 18
2.4.4 Psychological Processing Pathways in Misinformation Evaluation 20
2.4.5 Susceptibility Indicators 21

2.4.6 Integrating Susceptibility Indicators and Judgment Reliability 24
2.4.7 Summary and Implications for Recognition-Based Sensemaking 24
2.5 Persuasion Models and Recognition-Based Message Design 26
2.5.1 Models of Persuasion 26
2.5.2 Influence Mechanisms Through Persuasion Models 27
2.5.3 Persuasion Strategies and Techniques 29
2.5.4 Persuasive Design Principles and Devices 30
2.5.5 Phishing and Scam Practices as Applied Persuasion Design 31
2.5.6 Summary: Persuasion as Engineering Recognition Under Uncertainty 32
2.6 NDM and Recognition-Based Susceptibility 33
2.6.1 Core Principles of NDM and RPD Model 33
2.6.2 Mapping Misinformation Processing onto the RPD Cycle 34
2.6.3 Analytic Consolidation of Recognition-Based Decision Elements 35
2.6.4 Integrating Persuasion Models within the RPD Phases 38
2.7 Discussion and Implications 39
2.8 Conclusion 40
Chapter 3: How Misinformation is Evaluated in Practice 42
3.1 Introduction 42
3.2 Background and Related Works 43
3.2.1 Misinformation and Susceptibility 44
3.2.2 Decision-Making under Uncertainty 44
3.2.3 Sensemaking and Framing 45
3.2.4 Persuasive Communication Design 45
3.2.5 Evaluation and Stopping under Uncertainty 46
3.3 Methodology 47
3.3.1 Participants 48
3.3.2 Data Collection 49
3.3.3 Analysis 50
3.3.4 Limitation 52
3.4 A Process Framework of Misinformation Evaluation Derived from Analysis 52
3.5 Findings 54
3.5.1 Cue Recognition and Salience 55
3.5.2 Expectancy Formation 58
3.5.3 Mental Simulation and Projection of Consequences 62
3.5.4 Plausibility Judgments and Stopping Points 66
3.6 Discussion 70
3.6.1 From Observed Patterns to Process-Level Understanding 70

3.6.2 Evaluation as Constraint-Driven Sensemaking 71
3.6.3 Stopping as the Source of Divergence 72
3.6.4 Reframing Susceptibility to Misinformation 73
3.6.5 Implications for Understanding Misinformation Evaluation 74
3.7 Conclusion 74
Chapter 4: How Misinformation is Constructed 75
4.1 Introduction 75
4.2 Foundations for the RBM Framework 77
4.2.1 Background: Misinformation, Plausibility, and User Judgement 77
4.2.2 Susceptibility to Misinformation: A Conditional Process 78
4.2.3 Decision-making Under Uncertainty: From Bounded Rationality to Recognition 79
4.2.4 Persuasion Models 81
4.2.5 Aristotle's Rhetorical Communication Strategies 83
4.2.6 Persuasive Design Principles 83
4.2.7 Framing Theory and Dominant Meaning 85
4.2.8 How These Concepts Converge Toward Susceptibility to Misinformation 85
4.3 The RBM Framework 86
4.4 Demonstrating the RBM Framework 93
4.4.1 Target Analysis 93
4.4.2 Message Design 94
4.4.3 Message Framing 99
4.4.4 Plausibility Evaluation 102
4.4.5 Outcome 103
4.4.6 RBM Configurations Across Additional Story Domains 105
4.5 Discussion 110
4.5.1 What the RBM Reveals About the Design of Misinformation Narratives 110
4.5.2 Cross-Model Patterns and Cross-Domain Transferability 111
4.5.3 Cognitive Integration in Plausibility Judgement: RPD, Susceptibility, and Framing 111
4.5.4 Applied Implications for Misinformation Analysis and Intervention 111
4.5.5 Limitations 112
4.6 Conclusion 112
Chapter 5: Tracing Misinformation Evaluation in Practice 114
5.1 Introduction 114
5.2 Background and Related Works 116
5.2.1 Misinformation and Characteristics 116
5.2.2 Susceptibility to Misinformation 116
5.2.3 How Narratives and Framing Organize Information into Persuasive Structures 117

5.2.4 Persuasion Models 117
5.2.5 Aristotle's Rhetorical Communication Strategies 119
5.2.6 Prior Topic Orientations and Implicit Associations 120
5.2.7 Decision-Making Models Relevant to Belief Formation 120
5.2.8 Susceptibility as an Integrated Process 122
5.3 Methodology 122
5.3.1 Study Design 123
5.3.2 Narrative construction 123
5.3.3 Participants 124
5.3.4 Implicit Association Test 124
5.3.5 Procedure 125
5.3.6 Analysis 126
5.3.7 Limitation 129
5.4 Results and Data Analysis 129
5.4.1 Think-Aloud During Narrative Encounter 130
5.4.2 CDM Interviews: Retrospective Reasoning and Decision Pathways 135
5.5 Discussion 140
5.5.1 Emergence of Cognitive Strategies 143
5.5.2 Emotional Triggers and Defensive Filtering 144
5.5.3 Prior Orientations and Conditioned Mental Simulation 146
5.5.4 Contradiction Management and Dissonance Reduction 147
5.5.5 Confidence Calibration in Evaluation 148
5.5.6 Synthesizing Cognitive Strategies in Plausibility Evaluation 149
5.6 Conclusion 150
Chapter 6: Discussion and Conclusion 151
6.1 Conceptualizing Susceptibility to Misinformation 151
6.2 From Conceptualization to Evaluation 152
6.3 From Evaluation to Message Construction 153
6.4 Bringing Evaluation and Construction Together 154
6.5 Answering the Research Questions 155
6.6 Overall Contribution 157
6.7 Limitations 158
6.8 Future Research 158
References 160
Glossary 171
Appendix A: Review of AI-Based Tools for Misinformation Detection and Analysis 172
Appendix B - Persuasive Design Principles and Devices 174

Appendix C: CDM Interview Guide 176
Appendix D - From Raw Data to Themes 178
Appendix E. Root Neutral Messages 181
COVID-19 Vaccine 181
Climate Change 181
Appendix F. Reframed Narratives for All Three Stories 182
Reframed Narrative about Ukraine Biolab 182
Reframed Narrative about COVID-19 Vaccine 183
Reframed Narrative about Climate Change 184
Appendix G – Transformation Breakdown 186
Ukraine Biolab 186
COVID-19 Vaccine 190
Climate Change 195
Appendix H - Persuasive Techniques Reference Table 199
Appendix I - Definitions and Functional Descriptions of Persuasive Techniques 203
Appendix J - From Evaluation Segments to Cognitive Strategies (CSs) 205
Appendix K: Implicit Association Test (IAT) Results and D-Scores 212

List of Figures

Figure 1 -Recognition-Based Susceptibility Framework 10
Figure 2 - RPD model breakdown to illustrate misinformation evaluation process 38
Figure 3 - Evaluating process of misinformation derived from CDM interviews 54
Figure 4 - RBM framework 92

List of Tables

Table 1 - Key Takeaways: Misinformation & Susceptibility 14
Table 2 - Key Takeaways: Human Factors impacting Susceptibility 25
Table 3 - Models of Persuasion: Summary of Characteristics 28
Table 4 - Persuasion design principles organized by dominant rhetorical function 30
Table 5 - Key Takeaways from Persuasion Models and Recognition-Based Message Design 32
Table 6 - Summarized list of Cues emerged from thematic analysis of persuasion literature 35
Table 7 - Summarized list of Expectancies emerged from thematic analysis of persuasion literature 36
Table 8 - Summarized list of Actions emerged from reviewing persuasion strategies 37
Table 9 - Decision Analysis Process 38

Table 10. Illustrative ETA coding and theme development for "Plausibility judgment and stopping points" 51
Table 11 - Comparison of evaluation processes across higher- and lower-exposure participants 70
Table 12 - Reframing one factual element into three model-aligned narratives (Ukraine Bio Lab) 105
Table 13 - Reframing one factual element into three model-aligned narratives (COVID-19 Vaccine) 108
Table 14 - Reframing one factual element into three model-aligned narratives (Climate Change) 110
Table 15 - Experiment Design Structure to Elicit Cognitive Strategies (CS) 125
Table 16 - Overview of Codes and Themes Emerged through Thematic Analysis of Think-aloud & CDM .. 126
Table 17 - Mapping of selected CSs to Aggregated Cognitive Strategies (ACSs) 141
Table 18 - Aggregated Cognitive Strategies 142
Table 19 - CDM Interview Guide and Probs 176
Table 20 - Illustrative progression of inductive coding from data extracts to emergent process themes .. 178
Table 21 - Transforming Root Message to MisInfo Using PKM Model – Biolab 186
Table 22 - Transforming Root Message to MisInfo Using TRA Model – Biolab 187
Table 23 - Transforming Root Message to MisInfo Using ELM Model – Biolab 189
Table 24 - Transforming Root Message to MisInfo Using PKM Model – Vaccine 190
Table 25 - Transforming Root Message to MisInfo Using TRA Model – Vaccine 191
Table 26 - Transforming Root Message to MisInfo Using ELM Model – Vaccine 193
Table 27 - Transforming Root Message to MisInfo Using PKM Model – Climate change 195
Table 28 - Transforming Root Message to MisInfo Using TRA Model – Climate change 196
Table 29 - Transforming Root Message to MisInfo Using ELM Model – Climate change 197
Table 30 - Persuasive Techniques Toolkit for Analytic Message Reframing 199
Table 31 – Persuasive techniques and their functions in constructing framed misinformation narratives 203
Table 32 - Sample of coded Excerpt & their abstraction into Cognitive Strategies (CSs) 205
Table 33 - Pre-exposure IAT 212
Table 34 - Participant-level Pre/Post IAT D-scores and ΔD values for each Narrative × Model 213

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor used artificial intelligence tools or generative artificial intelligence tools (unless it is clearly stated, and referenced, along with the purpose of use), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signed: Boshra Talebi Haghighi

Date: 1 March 2026

Co-Authored Works

STUDENT AND SUPERVISOR APPROVALS

By signing you are confirming that the co-author contributions stated in the table(s) below are accurate.

Student Name	Boshra Talebi Haghighi	**Signature**		**Date**	1 May 2026
Supervisor Name	Prof. B. L. William Wong	**Signature**		**Date**	1 May 2026

Chapter Number:	2
Manuscript Title:	Decision-Making Process behind Believing Misinformation
Publication Status:	Published
Reference if published:	Haghighi, B. T., & Wong, B. L. W. (2024). Decision-Making Process behind Believing Misinformation. 17th Biennial Conference on Naturalistic Decision Making (NDM 2024), Auckland, New Zealand.
Chapter Number:	2
Manuscript Title:	*The Recognition-based Susceptibility Framework: Why Some People Are More Likely to Believe Misinformation*
Publication Status:	Submitted for Publication
Reference if published:	Haghighi, B. T., & Wong, B. L. W. (2026). *The Recognition-based Susceptibility Framework: Why Some People Are More Likely to Believe Misinformation* [Manuscript submitted for publication]. *Journal of Cognitive Engineering and Decision Making*.
Chapter Number:	3
Manuscript Title:	*How Susceptibility Emerges in Misinformation Evaluation: A Recognition-Primed Analysis of Decision Processes*
Publication Status:	Accepted for Publication
Reference if published:	Haghighi, B. T., & Wong, B. L. W. (2026). *How Susceptibility Emerges in Misinformation Evaluation: A Recognition-Primed Analysis of Decision Processes* [Manuscript submitted for publication]. *Journal of Cognitive Engineering and Decision Making*.
Chapter Number:	4
Manuscript Title:	The recipe book of misinformation: A framework for modelling persuasive framing and susceptibility
Publication Status:	Accepted for Publication
Reference if published:	Haghighi, B. T., & Wong, B. L. W. (In press). 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.

Chapter Number:	5
Manuscript Title:	*Susceptibility to Misinformation: Exploring Persuasion and Prior Orientation in Recognition-Primed Decision Pathways*
Publication Status:	Published
Reference if published:	Haghighi, B. T., & Wong, B. W. (2025). Susceptibility to Misinformation: Exploring Persuasion and Prior Orientation in Recognition-Primed Decision Pathways. Journal of Cognitive Engineering and Decision Making, 15553434261472960.

AUTHOR SURNAME: (order as per manuscript)	**CONTRIBUTION (The following authorship contributions apply to all manuscripts listed above.)**
Boshra Talebi Haghighi	Conceived and designed the study, conducted data collection and analysis, interpreted the findings, and drafted the manuscript.
Prof. B. L. Willaim Wong	Contributed to the conceptual development of the study, provided supervision and methodological guidance, and critically reviewed and revised the manuscript.

Artificial Intelligence Declaration

Specify the chapter number(s) in your research proposal/research component where AI was used.	
Chapter number(s):	Selected sections across Chapters 1–6 of the thesis, where minor language editing support was used.
Briefly explain why the AI tool was used (e.g., idea generation, summarization, data analysis, image creation, etc.).	
Purpose of AI Use:	Minor language editing support, including grammar, spelling, clarity, readability, and sentence-level refinement of author-written text.
Clearly state the AI tools used.	
AI Tool(s) Used:	ChatGPT
Prompts or queries entered into the AI tool (only the instructive portion of the prompt is required). Include all relevant variations.	
Prompts Used:	"Please check grammar, spelling and clarity while maintaining the meaning." "Please proofread this text." "Please improve readability without changing the argument." "Please make this sentence more concise."
A summary or description of the AI-generated output. (Optional: attach output in an appendix.)	
Output Received:	Suggested revisions to author-provided text, including grammar correction, sentence-level clarity (e.g., breaking down long sentences), and alternative wording, without introducing new content or arguments.
Describe how the AI output was used, edited, or transformed.	
Post-AI Processing Methods:	All AI-generated suggestions were reviewed, checked against the intended meaning, and selectively accepted, rejected, or further revised by the author. AI was not used to generate the thesis argument, research design, data analysis, findings, conceptual framework, or original contribution. The submitted thesis remains substantively the author's own work

You may use this free text box to explain the intent for the use of the AI generated content. *For example, rather than stating what prompt(s) were used, you may articulate the purpose for which the prompt was used (supported by examples) here:*

AI was used conservatively as a language-support and proofreading tool, limited to improving the clarity, grammar, spelling, readability, and flow of text already written by the author. It was not used to generate substantive thesis content or contribute to the development of arguments, analysis, or findings. All intellectual content and final wording remain the author's own.

STUDENT DECLARATION

By signing you are confirming that the AI use stated in the table(s) above are accurate and follow AUT's recommended guidelines detailed in the AI Hub and Postgraduate Handbook.

Student Name:	**Boshra Talebi Haghighi**	**Signature:**	**Date:**	**01/05/2026**

Acknowledgements

To my supervisor, Prof. William Wong, I owe special gratitude. William never allowed this PhD to become simply an exercise in earning a degree. Through years of discussions, corrections, and constantly making me think again, he held both the work and me to a standard that took us much further than I had imagined. His patience, confidence in my work, and example of what it means to be a scientist have shaped the researcher I have become. William, it has been an absolute honour to learn from you, to be part of your team, and now to find myself welcomed into the community of scientists to which you introduced me.

To my father, Dr Davoud Haghighi, who has always been my measure of what it means to be a scientist: I have spent much of my life hoping that one day I might make a meaningful contribution to science, as he did. If I can become even half the scientist you were, and leave behind even half the contribution you made, I will consider myself extraordinarily fortunate. This doctorate carries my name, but the ambition behind it began with you.

To my dear family, thank you for standing behind me through years that demanded much more than academic work alone. I am so fortunate to have you.

To my Liam.

You have been part of this PhD almost from its beginning, and you grew alongside it. There were papers beside your toys, meetings with you wrapped against me, and later beside me with your snacks. There were also countless nights writing after you had fallen asleep, and days when motherhood, work and this PhD somehow had to fit into the same twenty-four hours. You did not choose this journey, but you travelled every part of it with me.

Thank you for your patience and the love you gave me so freely through years you were far too young to understand. This thesis ends here, but we have much further to go, and I promise, much more time together.

Ethics Approval

This research received ethical approval from the Auckland University of Technology Ethics Committee (AUTEC) under Approval Number 25/174 for the project titled “Susceptibility to Misinformation: The Impact of User Experience (UX) Design Factors on Users’ Decision-making and Belief Formation.” Approval was granted on 23 July 2025 for a period of three years until 22 July 2028.

The research was conducted in accordance with the Auckland University of Technology Code of Conduct for Research and the conditions approved by AUTEC. Participants were provided with AUTEC-approved information sheets and consent forms, and informed consent procedures were followed throughout the research.

Chapter 1: Introduction

Misinformation is commonly defined as false or misleading information shared unknowingly without intent to deceive or cause harm (Wardle & Derakhshan, 2017). It is often distinguished from disinformation, which involves deliberate intent to mislead, and malinformation, which refers to the use of genuine information in ways that cause harm (Wardle & Derakhshan, 2017). While these categories differ in terms of intent and use, this research focuses specifically on misinformation as an umbrella term for misleading content encountered by individuals, where sender intent is not examined within the scope of analysis.

Misinformation is a persistent feature of contemporary information environments, shaping how individuals interpret issues such as vaccine uptake, climate change, and democratic processes (Nan et al. , 2022; Wardle & Singerman, 2021). Its impact extends beyond the circulation of incorrect claims; it affects how people assess credibility, interpret evidence, and make judgments in situations where information is incomplete or contested. Misinformation is not only a communication issue but also a decision-making problem, as individuals are often required to judge what can be relied upon under conditions of uncertainty, incomplete information, and competing claims (Klein & Zsambok, 1997; Wardle & Derakhshan, 2017).

1.1 Problem Statement and Research Gap

Research on misinformation has identified a range of cognitive, motivational, and social factors associated with belief, including analytic reasoning, familiarity effects, prior beliefs, and social and source cues (Ecker et al. , 2022; Pennycook & Rand, 2021; Traberg et al. , 2024). These studies provide detailed accounts of the conditions under which misinformation is more likely to be accepted or rejected. However, they offer more limited specification of how individuals move from encountering information to forming a judgment, particularly when information is incomplete, contested, or requires interpretation rather than direct verification.

This limitation becomes critical in contemporary information environments, where individuals are frequently required to evaluate information under time pressure and uncertainty. In such contexts, judgments are not always supported by clear verification or stable ground truth, and evaluation depends on how individuals interpret available cues and construct a plausible understanding of the situation.

In addition to these conditions, misleading content is often constructed to guide interpretation. Rather than relying solely on false information, persuasive misinformation may involve the selective framing of factual elements to produce plausible but misleading narratives (Wardle, 2018b). Research on framing and persuasive communication shows how message design can

shape what is noticed, how information is interpreted, and what appears credible or relevant during evaluation (Entman, 1993; Petty & Cacioppo, 1986). This suggests that susceptibility cannot be understood solely in terms of individual factors but must also account for how information is constructed and encountered in context.

1.2 Theoretical Lens

Understanding this interplay requires a perspective that accounts for decision-making under uncertainty. If susceptibility is understood as an evaluation process, it also becomes necessary to explain how evaluation reaches closure when exhaustive assessment of all possible evidence and interpretations is impractical. Simon's concept of bounded rationality (Simon, 1957) provides the rationale for this limitation, proposing that judgments are based on satisfactory rather than optimal outcomes. The Naturalistic Decision Making (NDM) (Klein & Zsambok, 1997) paradigm extends this perspective by explaining how such judgments are reached in real-world settings. It focuses on how individuals make judgments in complex, real-world environments characterized by time pressure, incomplete information, and dynamic conditions, including ill-structured problems, shifting or competing goals, and action–feedback loops (Orasanu & Connolly, 1993). Within this paradigm, the Recognition-Primed Decision (RPD) model (Klein, 1993) describes how individuals recognize patterns and assess plausible interpretations under uncertainty, rather than through exhaustive comparison of alternatives. In this study, RPD is used as a process account of how judgments are formed through recognition and interpretation under conditions of uncertainty.

To investigate such processes in context, this research adopts Cognitive Task Analysis (CTA) (Crandall et al. , 2006), which examines cognitive processes underlying performance in complex environments, particularly those that are not directly observable. Within CTA, the Critical Decision Method (CDM) (Klein, 1993) probes decision episodes in detail by eliciting accounts of memorable, non-routine incidents involving complex judgment and decision-making. In this study, it is used to examine how cues are noticed, situations are interpreted, and judgments are formed during the evaluation of misinformation.

1.3 Research Aim and Objectives

This thesis examines susceptibility to misinformation as a recognition-based evaluation process. It investigates how individuals interpret and assess information under conditions of uncertainty, and how these evaluations are shaped by both human factors and the construction of persuasive messages. The research objectives are set as follows:

1) To examine existing research on misinformation belief and identify how susceptibility may be understood as a recognition-based evaluation process, highlighting the roles of human factors, persuasive communication design, and decision-making under uncertainty.

2) To investigate how individuals move from noticing cues to forming a judgment, and how they decide when to stop evaluating, particularly when information is incomplete, contested, or open to interpretation rather than direct verification.

3) To model how factual elements can be reframed into persuasive misinformation narratives, showing how message design can guide interpretation and support particular plausible understandings.
4) To investigate how individuals evaluate persuasive misinformation narratives when engaging with them in context, focusing on how recognition-based evaluation processes are reflected in participants' reasoning under structured narrative conditions.

1.4 Research Questions

The following questions guide the investigation across the four studies in this thesis:

1. How may susceptibility to misinformation be understood as a recognition-based evaluation process, based on existing research on human factors, persuasive communication, and decision-making under uncertainty?

2. How do individuals move from noticing cues to forming a judgment when encountering misinformation, and how do they decide when to stop evaluating when information is incomplete, contested, or open to interpretation?

3. How are factual elements reframed into persuasive misinformation narratives in ways that guide interpretation and support particular plausible understandings?

4. How are persuasive narrative cues interpreted during recognition-based evaluation of misinformation under structured conditions, as reflected in participants' reasoning?

1.5 Scope and Limitations of the Research

This research focuses on misinformation as misleading content encountered by individuals, without examining the sender's intent. It does not distinguish between misinformation, disinformation, and malinformation in analytical terms, focusing on how information is evaluated rather than why it is produced or propagated.

The study examines three domains: vaccines, climate change, and bioweapons as illustrative contexts for analyzing evaluation processes. It focuses on persuasive communication design, rather than on automated content generation, deepfakes, or large-scale network dynamics.

The research adopts a process-oriented, qualitative approach to examine how individuals interpret and assess information in context. It does not aim to provide population-level generalizations or predictive models, but to develop an account of how evaluation unfolds under conditions of uncertainty.

The findings are based on a small number of participants and are intended to support detailed, process-level analysis of how evaluation unfolds, rather than to represent population-level patterns. Consistent with exploratory, process-oriented research, the analysis focuses on depth across multiple evaluation episodes and conditions (Dix, 2020). The research is context-dependent, and the frameworks developed are conceptual rather than predictive. Their applicability across different populations and settings, including more naturalistic contexts, requires further empirical examination.

1.6 Contribution and Significance of the Thesis

This thesis offers a process-oriented account of susceptibility to misinformation, positioning it as arising during evaluation rather than as a stable individual characteristic. It establishes how susceptibility can be understood in terms of how individuals interpret information and bring evaluation to a close under conditions of uncertainty. Building on this, the research provides an empirically grounded account of evaluation, showing that differences in judgment arise from how situations are interpreted and where evaluation is brought to a close, rather than from message content alone.

The thesis further clarifies how persuasive misinformation is constructed by showing how factual elements can be reframed to support particular plausible interpretations. This makes explicit the role of message design in shaping what is noticed, how it is interpreted, and what appears credible during evaluation.

This establishes a process-level account of how plausibility judgments emerge during evaluation, explaining why the same information can lead to different outcomes depending on how cues are interpreted and where evaluation is brought to a close.

1.7 Structure of the Thesis

Chapter 2 presents the literature review that establishes the conceptual foundation for the thesis. It examines research on misinformation, human susceptibility factors, persuasive

communication, and decision-making under uncertainty, and develops the Recognition-Based Susceptibility Framework, positioning susceptibility as an evaluation process.

Building on this foundation, Chapter 3 investigates how misinformation is evaluated in practice. Using CTA with the CDM, it investigates how individuals move from noticing cues to forming a judgment and where evaluation is brought to a close, providing an empirically grounded account of evaluation.

Having investigated how individuals evaluate misinformation, Chapter 4 shifts attention to the construction of misinformation itself. It introduces the Recipe Book of Misinformation, which models how factual elements can be reframed into persuasive narratives, showing how message design guides interpretation and shapes what appears plausible.

Chapter 5 brings these two strands together by investigating how participants evaluate persuasive misinformation narratives developed from the communication patterns identified in Chapter 4. It focuses on how evaluation processes are reflected in participants' reasoning, enabling the interaction between message design and evaluation to be examined in context.

Chapter 6 synthesizes the findings across the studies, connecting the conceptual framework, the evaluation process, and message construction to explain how plausibility judgments emerge under uncertainty and why similar information can lead to different outcomes.

1.8 Conclusion

In this chapter, I introduced the research problem, aim, objectives, research questions, scope, and contribution of the thesis. The chapter positioned misinformation as a decision-making problem and established the need to examine susceptibility as a recognition-based evaluation process through which individuals interpret and assess information under conditions of uncertainty. It also highlighted the role of persuasive communication design in shaping what is noticed, how information is interpreted, and what appears plausible during evaluation.

The next chapter begins by developing this foundation through a review of the literature on misinformation, human susceptibility factors, persuasive communication, and decision-making under uncertainty, leading to the development of the Recognition-Based Susceptibility Framework.

Chapter 2: Conceptualizing Susceptibility to Misinformation

This chapter establishes the conceptual foundation for the thesis by synthesizing existing research and developing an integrated framework explaining how individuals evaluate the plausibility of information in sociotechnical environments. The framework draws on and integrates research on human susceptibility factors, persuasive communication design, and recognition-based decision processes to model how cues guide attention, how interpretations are assessed through mental simulation, and how plausibility judgments emerge during narrative evaluation. Drawing on the NDM paradigm, particularly the RPD model, the framework explains how individuals arrive at provisional judgments under pressure and uncertainty. This framework reframes susceptibility as an adaptive evaluation process shaped by interactions among message cues, cognitive conditions, and situational constraints.

2.1 Introduction

This chapter proposes a framework explaining why some individuals become more susceptible to misinformation under conditions of uncertainty. In such contexts, decision-making occurs in ill-structured and dynamic environments, often involving incomplete information, competing goals, time pressure, and social constraints (Orasanu & Connolly, 1993). The framework integrates research on human susceptibility factors, persuasive communication design, and recognition-based decision processes to explain how individuals assess information in sociotechnical environments. In these environments, information is encountered through cues shaped by communication, context, and social influence, contributing to incomplete and evolving interpretations. It conceptualizes susceptibility not as a fixed trait or reasoning deficit, but as an

evaluation process in which individuals attend to cues, interpret them through prior experience and motivations, and judge whether a narrative appears plausible to accept, question, or reject.

Misinformation poses significant societal challenges, affecting public health, political and economic stability, and community trust (Lewandowsky et al. , 2017; Roozenbeek & van der Linden, 2022). The digital information environment operates as a sociotechnical system in which algorithmic curation, visible social endorsement cues, and presentation features shape what is encountered and how it is interpreted (Hutchins, 2020; Traberg, 2022). These features influence what is noticed, how credibility is inferred, and when individuals stop scrutinizing and move toward acceptance, questioning, or rejection. In such contexts, narrative acceptance depends not only on exposure but also on whether a message appears plausible during evaluation. From this perspective, susceptibility emerges from how individuals make sense of information within sociotechnical systems.

Research on misinformation susceptibility has identified cognitive, motivational, and social factors that shape how information is processed, including analytic thinking and attention to accuracy (Pennycook et al. , 2021), ideological and worldview alignment (Hubeny et al. , 2025), and indicators such as trust and emotional engagement that influence credibility judgments (Nan et al., 2022). In parallel, research on persuasion models and framing strategies has examined how communication design structures meaning through cue salience and narrative framing (Illia et al. , 2026; Peng et al. , 2024; Rubinelli & Diviani, 2025). While these strands have advanced understanding of drivers of misinformation belief and resistance, they are often examined at different levels of analysis: human susceptibility factors on the one hand, and message design or platform structures on the other. Less attention has been given to explaining how these levels interact during narrative evaluation within a sociotechnical systems context, where cognition is distributed across individuals, representations, and technological artifacts (Hutchins, 2020). This chapter addresses this gap by examining how susceptibility indicators, persuasive communication design, and recognition-based decision processes jointly shape plausibility evaluation in sociotechnical information environments. Studies focused exclusively on technical detection mechanisms are outside the scope of this analysis.

Drawing on the RPD model (Klein, 1993) within the NDM (Klein & Zsambok, 1997), this review chapter conceptualizes misinformation evaluation as a recognition-based sensemaking process under conditions of uncertainty.

2.1.1 Review Approach

Relevant literature was identified through targeted searches of Scopus, Web of Science, ACM Digital Library, IEEE Xplore, and Google Scholar, complemented by backward and forward citation

tracking of key publications. The search prioritized peer-reviewed and foundational studies investigating susceptibility-related human factors, persuasive communication design, and framing mechanisms, and decision-making under uncertainty. Studies focused exclusively on automated detection, classification, or platform-engineering approaches were excluded.

The selected literature was synthesized conceptually to identify recurring cues and mechanisms involved in misinformation evaluation. These were organized into three analytic components: human factors influencing susceptibility, persuasive message design features, and pattern-recognition and sensemaking processes involved in evaluating misinformation narratives. These components were then integrated using the RPD model as the organizing lens for the proposed framework.

The chapter is structured as follows. It begins by presenting the proposed Recognition-Based Susceptibility Framework, which is then developed through the review. The chapter first examines misinformation, susceptibility, and selected approaches to addressing misinformation. It then reviews research on the human factors shaping susceptibility, followed by persuasion models and message design features that make misinformation compelling. Next, it draws on NDM and recognition-based decision processes to examine how plausibility evaluation unfolds in sociotechnical environments. The chapter concludes by bringing these strands together in discussion and deriving the proposed framework.

2.2 The Recognition-based Susceptibility Framework for Explaining Susceptibility to Misinformation

Figure 1 illustrates the Recognition-based Susceptibility framework proposed in this review. It explains susceptibility as a dynamic evaluation process unfolding under conditions of uncertainty and pressure. The framework integrates three interacting elements.

First, human factors shape how individuals attend to, interpret, and evaluate information. These factors include prior knowledge, trust, emotional state, values, and group identity, as well as cognitive tendencies such as confirmation bias, dissonance management, and information overload. Together, they influence responsiveness to cues embedded in persuasive messages, shaping which cues are noticed, how they are interpreted, and what is taken as a plausible explanation during evaluation. From a Cognitive Reliability Analysis (CRA) perspective (Woods et al. , 1990), these factors can be understood as performance-shaping conditions that influence the reliability of recognition-based judgments under uncertainty.

Second, within sociotechnical information environments, both exposure conditions and message design shape how information is encountered and interpreted. These environments,

characterized by algorithmic curation, social signals, and dynamic information flows, influence what becomes visible and salient. Within this context, persuasive communication design, grounded in persuasion models and rhetorical strategies, structures messages to embed cues that resonate with individuals' existing knowledge, emotions, and expectancies. These cues guide recognition and interpretation by directing attention, signaling credibility, and framing meaning through familiarity, emotional appeal, and normative signals.

Finally, these elements are integrated through the RPD model, linking human susceptibility indicators and message cues within a recognition-based evaluation process. The framework maps misinformation evaluation onto three linked processes. The first involves cue recognition, where salient or familiar cues attract attention and shape initial interpretation. The second involves plausibility assessment through mental simulation, where individuals assess whether an interpretation makes sense in light of prior experience, expectancies, and emotional responses. When interpretations conflict with expectations, individuals may retrieve and review additional information before continuing the evaluation. The final process involves action, where individuals may accept, reject, or defer judgment without explicitly comparing alternatives.

In this way, susceptibility emerges when recognition and simulation processes lead individuals to settle on interpretations that appear plausible under conditions of uncertainty and limited information. These outcomes may also feed back into future evaluations by shaping cue sensitivity and influencing responses in subsequent encounters with misinformation.

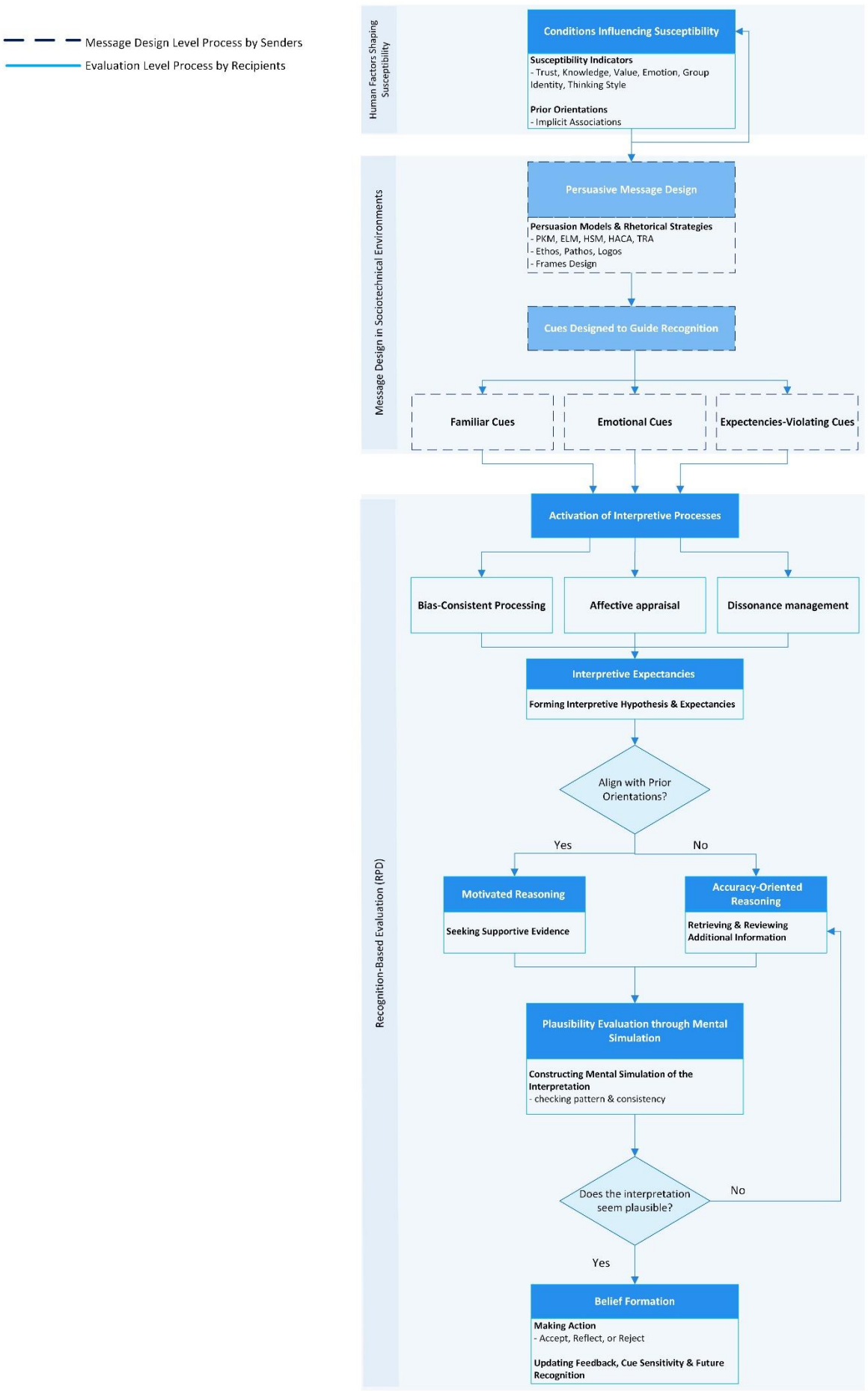

Figure 1 - Recognition-Based Susceptibility Framework

The next section examines misinformation and susceptibility in relation to plausibility evaluation.

2.3 Misinformation and Susceptibility

This section reviews misinformation and susceptibility, along with how current approaches to addressing misinformation engage with these susceptibility conditions.

2.3.1 Misinformation

Misinformation can appear authentic and spread rapidly across various media, with content taking a range of forms such as narratives, deepfake videos, and altered imagery or audio, creating misleading stories by deceptively framing factual information (Roozenbeek, Suiter, et al. , 2022; Wardle & Derakhshan, 2017). Misinformation can include satire, false connections, misleading content, false context, imposter content, framing, manipulated content, and fabricated content (Wardle, 2020). For example, during the COVID-19 pandemic in New Zealand, selective framing of vaccine-related claims, government actions, and protest narratives circulated widely online, contributing to public tension and polarisation surrounding the parliamentary protests in Wellington. The spread of misinformation is driven by socio-psychological factors, where individuals often share content to connect with like-minded groups or reinforce shared identities, contributing to the rapid circulation and persistence of misleading content within online and offline communities (Van Bavel & Pereira, 2018). There are two other variants of misinformation: disinformation and malinformation. Disinformation is the deliberate and intentional use of misinformation to deceive. Malinformation refers to true information that has been strategically repurposed to cause harm (Wardle & Derakhshan, 2017). In this research, the discussion is restricted to misinformation, as the techniques and explanations are applicable regardless of the variant.

2.3.2 Susceptibility

Susceptibility refers to the extent to which individuals are prone to being influenced when encountering particular stimuli or contexts, here in the form of misinformation (Wardle & Singerman, 2021). It does not reflect a stable trait, but arises from the combined influence of cognitive, personal, and situational conditions that shape how information is noticed, interpreted, and judged (Barr et al. , 2015; Nan et al., 2022; Rand et al. , 2017; Wardle & Derakhshan, 2017). These include knowledge and experience, prior orientations, uncertainty and time pressure, emotional and motivational states, and identity-related commitments. As these influences vary across individuals and contexts, susceptibility is expressed differently, leading to variation in responses to the same information. For example, during the COVID-19 pandemic in New Zealand, individuals exposed to the same vaccine- and protest-related narratives often interpreted and responded to them differently depending on their prior experiences, concerns,

and trust in information sources. In this sense, susceptibility captures how judgment unfolds under particular conditions and links individual characteristics to the recognition-based decision processes through which misinformation is evaluated.

2.3.3 Current Approaches to Countering Misinformation

This section reviews current approaches to addressing misinformation, focusing on how it is operationalized in sociotechnical environments. These approaches include detection and classification techniques, their implementation through AI-based systems, and collaborative or framework-based responses. Examining them clarifies how misinformation has been treated as a problem of content identification, system-level monitoring, and coordinated response, and highlights the limits of these approaches in explaining how individuals interpret and evaluate misinformation.

Detection and classification-focused approaches

Detection and classification-focused approaches typically rely on user-based and content-based strategies to identify misinformation at scale (Wang et al. , 2024). User-based methods analyze online behaviors, interactions, and network influence, while content-based methods use machine learning to examine textual features for classifying misleading material (Shu et al. , 2018; Wang et al., 2024). Some approaches incorporate contextual information, such as diffusion patterns and network structure, to address cases where misinformation arises through reframing rather than explicit falsehoods (Abdali et al. , 2020; Nasir et al. , 2021; Pelrine et al. , 2021). Despite the advances, these approaches face limitations. User-based methods struggle to capture evolving social behaviour or what is shared by less influential users, while content-based methods have difficulty interpreting long or mixed narratives and lack access to communicative intent behind messages.

AI-Based Implementation of Detection Approaches

Detection strategies have been operationalized through AI-based systems designed to analyze behavioral signals, network dynamics, and textual content at scale. User- and network-focused AI systems do not analyze message content directly but instead examine how information spreads and how users behave within networks. By examining patterns such as sharing activity, interaction structures, and diffusion dynamics, these systems can identify coordinated or anomalous behavior associated with misinformation. For example, Hoaxy tracks how claims and fact-checks propagate through networks (Shao et al. , 2018), while Truthy detects coordinated campaigns across social media (Ratkiewicz et al. , 2011).

Content-focused AI tools analyze linguistic features, semantic structure, and framing within texts to detect misleading or inconsistent information. This includes natural language processing (NLP)

models based on deep learning and large language models (LLMs), such as transformer-based architectures, which assess contextual coherence, and factual consistency within news articles and posts (Brown et al. , 2020; Devlin et al. , 2018a, 2018b; Raffel et al. , 2020). More recent work has extended these methods to framing-aware detection, such as the FramedTruth Model, targeting cases in which factual content is strategically reframed to produce misleading interpretations (Wang et al., 2024).

Together, these AI-based approaches have advanced the scalability and speed of misinformation detection and moderation (See Appendix A). However, evaluation metrics such as accuracy, precision, or reduction in spread emphasize identifying misinformation as an object rather than explaining how individuals construct meaning when exposed to persuasive content. As such, while these tools are essential for monitoring and mitigation, they offer limited insight into individual-level susceptibility and decision-making processes.

Collaborative and Framework-based Responses

Collaborative and framework-based responses address misinformation through shared standards, platform coordination, and integrated initiatives such as moderation, fact-checking, and education (Vraga & Bode, 2020; Woolley, 2020). These approaches focus on reducing exposure and improving correction at scale and are typically evaluated using system-level metrics such as reach reduction, correction uptake, or compliance. Frameworks such as DISARM (Disinformation Analysis and Risk Management) provide structured ways to document and compare disinformation activities using operational descriptors and recurring patterns, supporting situation awareness rather than modelling individual belief formation (Bârgăoanu & Pană, 2024; González et al. , 2025). Related indicator-sharing initiatives, such as the Disinformation Indicators Sharing Platform (DISP), aim to standardize the reporting and exchange of disinformation-related indicators, drawing on concepts from cyber threat–intelligence practices (MISP-Project; Smith, 2025).

Taken together, these approaches show how misinformation can be identified, monitored, and mitigated at scale. However, they primarily treat misinformation as a system-level problem of detection and response, offering limited insight into how individuals interpret persuasive narratives and come to accept or reject them.

2.3.4 What These Approaches Explain and What They Leave Unexplained

Research on misinformation has developed across several major strands. These include technical and computational approaches focus on detection, classification, and platform-level mitigation of false content (Shao et al. , 2016; Shu et al., 2018; Zhou & Zafarani, 2020); experimental and cognitive research on truth–falsity judgments, correction interventions, and the

continued influence of misinformation (Nyhan & Reifler, 2010; Pennycook et al. , 2020; Vosoughi et al. , 2018); individual-difference research investigates cognitive and epistemic traits associated with susceptibility, including analytic thinking and bullshit receptivity (Pennycook & Rand, 2019a; Roozenbeek, Schneider, et al. , 2020; Roozenbeek & van der Linden, 2022); motivational accounts emphasizing ideology, identity-protective cognition, and motivated reasoning in belief formation (Ecker et al., 2022; Hubeny et al., 2025); and message-level research analyzing how source cues, social endorsement signals, familiarity, framing, and emotional salience shape perceived accuracy (Traberg, 2022). These strands are analytically distinguishable, though often empirically intertwined.

Although these strands have substantially advanced the understanding of misinformation, belief, and resistance, they are typically examined at different levels of analysis. System-level research focuses on detection, diffusion, and intervention effectiveness, whereas psychological studies examine cognitive traits, motivations, and biases associated with susceptibility. Message-level research, in turn, investigates how cues such as framing, source signals, or emotional salience influence perceived credibility. As a result, less is known about how individuals interpret these cues and arrive at plausibility judgments as evaluation unfolds in real time.

Evaluation metrics such as accuracy, reach, engagement, or mitigation outcomes do not capture how attention is guided or how interpretations are formed and maintained, even in the presence of corrective information. They also do not explain why the same message may appear compelling to some individuals but not others, particularly under conditions of uncertainty, time pressure, and social influence. Addressing these questions requires examining plausibility evaluation as a process unfolding during real-time narrative interpretation. This perspective motivates the focus of the present review.

The next sections examine how these elements shape the evaluation of misinformation as a process of interpretation. The key points from this section are summarised in Table 1.

Table 1 - Key Takeaways: Misinformation & Susceptibility

Focus	**What the Literature Explains**	**How It Is Operationalized**	**What Remains Unexplained**
Misinformation	Misinformation is broadly defined as false or misleading information, including reframed factual content, manipulated media, and misleading narratives.	Content categories, factual verification, framing patterns, and narrative structure.	How misleading interpretations emerge when factual elements are selectively emphasised or framed.
Susceptibility	Susceptibility varies across individuals and contexts and is	Experimental accuracy judgments, belief ratings, analytic thinking	How susceptibility manifests during real-time

	influenced by cognitive, emotional, and social factors.	measures, and epistemic traits.	interpretation and decision-making.
Computational Detection Approaches	Detection systems aim to identify and classify misleading content at scale.	Classification accuracy, precision/recall, stance detection, diffusion analysis.	How individuals interpret detected content and decide whether to believe it.
AI-Driven Detection and Monitoring	AI tools improve scalability and speed of misinformation monitoring and moderation.	Automated text analysis, network analysis, engagement metrics.	How individual cue recognition, plausibility evaluation, and meaning construction occurs under uncertainty.
Collaborative & Framework-Based Responses	Coordination frameworks organize monitoring and response across platforms and institutions.	Campaign tracking, indicator sharing, system-level mitigation metrics.	How recognition-based evaluation processes lead individuals to accept, stabilize, or reject interpretations.
Overall Insight	Current approaches primarily address misinformation as content or behaviour to be detected and managed.	System- and content-level metrics dominate evaluation.	How belief formation occurs as a real-time evaluative decision process in sociotechnical environments.

2.4 Human Factors Shaping Susceptibility

This section examines how human factors shape susceptibility to misinformation by influencing how individuals evaluate information under conditions of pressure and uncertainty. Rather than treating belief as a simple accuracy judgment, it focuses on how prior knowledge, cognitive biases, and social motivations guide interpretation and plausibility judgments during evaluation.

It begins with decision-making under uncertainty and the reasoning strategies used in such conditions. It then examines cognitive mechanisms that influence attention and interpretation, followed by how communication frames shape meaning. The section then considers the processing pathways through which information is evaluated, and the susceptibility indicators that influence which pathways become dominant. It concludes by integrating these elements to explain how judgment reliability varies during evaluation.

2.4.1 Decision-Making Under Uncertainty

Under conditions of uncertainty, time pressure, and social influence in sociotechnical information environments, decisions are shaped by what is noticed, treated as relevant, and judged as plausible for action (Klein et al. , 2006a, 2006b). Rather than engaging in thorough evaluation, individuals adapt their reasoning strategies to cognitive limits, available information, and demands of the situation, often relying on experience and pattern recognition. Dual Process Theory (Kahneman, 2011), Cognitive Continuum Theory (CCT) (Hammond, 1996), and bounded rationality (Simon, 1957) together provide complementary explanations of how decision-makers

shift between intuitive, analytical, and mixed modes of reasoning in response to task structure, urgency, and resource constraints, as reviewed in further detail below.

Dual Process Theory

Dual Process Theory distinguishes between two interacting modes of processing that guide judgment and decision-making under uncertainty. Type 1 is fast, intuitive, and automatic, generating immediate impressions, while Type 2 is slower, deliberate, and reflective, capable of monitoring or revising initial judgments (Kahneman, 2011; Kahneman & Frederick, 2005). Type 1 processing is not inherently erroneous; it reflects adaptive reliance on learned patterns and prior experience, particularly in situations requiring rapid assessment. Type 2 processing, while more analytical, is effortful and constrained by time and cognitive capacity, and is therefore not always engaged. Research shows that intuitive responses often precede conscious reasoning, with reflective processing typically occurring afterward and sometimes serving a justificatory role (Kahneman, 2011). Reliance on Type 1 does not imply error or irrationality, nor does Type 2 guarantee accuracy. Rather, these modes of processing reflect adaptive responses to situational demands, motivation and available information.

Cognitive Continuum Theory (CCT)

Cognitive Continuum Theory conceptualizes judgment as ranging along a continuum between intuitive and analytical processing rather than as two discrete systems (Hammond, 1981, 1996). The mode of cognition depends on task structure: well-defined tasks encourage analytic reasoning, whereas ill-structured situations, characterized by ambiguity, incomplete information, and time pressure, increase reliance on intuitive judgment. Between these poles, quasi-rational processing combines experience-based pattern recognition with selective analytic checks (Dhami & Thomson, 2012; Hammond, 1996). This mode is common in real-world contexts where tasks are neither fully familiar nor entirely novel, and where purely intuitive or fully analytical processing alone is insufficient. In the context of misinformation, this suggests that evaluation often occurs under conditions that favour adaptive shifts between intuitive and analytical processing rather than reliance on a single mode.

Bounded Rationality Theory

Herbert Simon's theory of bounded rationality challenges classical models of rational choice by arguing that decision-making is constrained by cognitive limitations, incomplete information, and environmental pressures (Simon, 1957; Simon, 1997). Rather than assessing and comparing all possible options, individuals often satisfice, selecting interpretations that are "good enough" given their goals and available information (Lilly, 1994). Under these constraints, heuristic and experience-based strategies are treated as adaptive responses to complex environments (Barros, 2007; Fiori, 2008). This account supports non-optimizing decision-making as a realistic basis of

decision-making under constraint and pressure, although its emphasis on satisficing differs from recognition-based accounts, such as the RPD model, that explain judgment through pattern matching and mental simulation.

2.4.2 Cognitive Mechanisms That Influence Interpretation

Under the constraints outlined above, interpretation is not neutral: certain cues are prioritized while others are discounted. This section examines confirmation bias, cognitive dissonance, and information overload as mechanisms that shape how this selectivity emerges during the evaluation of misinformation.

Confirmation Bias

Confirmation bias reflects a systematic tendency to favour belief-consistent information and discount conflicting evidence (Klayman, 1995). Oswald and Grosjean, (2004) describe confirmation bias as a systematic process in which information is searched for, interpreted, and remembered in ways that make it difficult to reject an existing hypothesis. This can protect an interpretation by filtering what is noticed and how evidence is assessed. In the context of misinformation, confirmation bias helps explain why familiar, orientation-consistent cues may carry unbalanced weight and why alternative interpretations can be discounted even when additional information is available.

Cognitive dissonance theory

Cognitive dissonance refers to discomfort arising from conflicting attitudes, beliefs, or behaviors (Festinger, 1957). This discomfort motivates efforts to restore consistency, including reinterpretation, rationalization, selective perception, and avoidance of contradictory information (Harmon-Jones, 2019). These processes overlap with mechanisms such as confirmation bias (Ellis & Newton, 2005). In misinformation contexts, this helps explain why counter-evidence may be dismissed or avoided when it conflicts with an existing interpretation that appears coherent.

Information Overload Theory

Information overload refers to the challenge of understanding and making decisions when the amount, complexity, or conflict of information exceeds an individual's processing capacity, often reducing decision quality (Roetzel, 2019; Speier et al. , 1999; Yang et al. , 2003). In sociotechnical environments, the volume, speed, and diversity of information further intensify these pressures (McFedries, 2014; Tim, 2017). Information searching research identifies two forms of overload: (1) outcome overload, where there are too many sources, and (2) textual overload, where individual sources are too lengthy (Savolainen, 2015). Under these conditions, individuals are more likely to rely on simplified search strategies and heuristic cues when evaluation becomes difficult.

Together, these mechanisms explain how interpretation becomes selective during evaluation, shaping which cues are prioritised and which are discounted when individuals encounter misinformation.

2.4.3 Communication Frames and Meaning Construction

This section examines how meaning is constructed during the evaluation of misinformation. It focuses on framing theory, conspiracy narratives, and implicit theories as complementary perspectives on how messages are structured, how interpretations are organized into coherent explanations, and how prior assumptions influence what appears plausible under conditions of uncertainty.

Framing theory

During misinformation evaluation, interpretation depends on which aspects of a message are attended to and treated as relevant. Framing theory explains how this selectivity arises by shaping which elements of a situation are made salient (Goffman, 1974). In communication contexts, framing refers to how information is selected, emphasized, and contextualized, guiding how audiences interpret issues and assign significance (Entman, 1993). Frames function as cognitive structures that organize experience, allowing individuals to interpret information without evaluating all available alternatives.

Frames are constructed through social, cultural, organizational, and ideological processes, a dynamic referred to as frame-building (Gamson & Modigliani, 1989). Once established, they guide attention, define evaluative criteria, and constrain alternative interpretations (Scheufele, 1999). Empirical research shows that framing can influence judgments even when underlying information remains constant, such as when outcomes are presented in terms of survival versus mortality rates (Tversky & Kahneman, 1986). In the context of misinformation, framing helps explain how particular interpretations come to appear plausible. By directing attention to specific aspects of a message and aligning them with familiar narratives or emotionally salient cues, frames can support interpretations that appear plausible without further scrutiny (Lewandowsky et al., 2017).

Conspiracy Theory

Conspiracy theories are commonly defined as beliefs that significant social or political events are the result of secret plans orchestrated by powerful actors (Dentith & Orr, 2018; Keeley, 2019). Unlike verified conspiracies, conspiracy theories are not necessarily false, but are characterized by narratives that emphasize hidden intentions and coordinated deception (Uscinski, 2018). Belief in conspiracy theories is widespread, with substantial variation across individuals and contexts, highlighting their relevance to understanding susceptibility (Enders & Smallpage, 2018).

Under conditions of uncertainty, conspiracy theories function as narrative frameworks that attribute complex events to intentional actions by identifiable agents (Harambam & Aupers, 2017). This provides a coherent explanation when information is incomplete or contested. In misinformation contexts, such narratives can make particular interpretations appear plausible by aligning with existing suspicions or expectations, while reducing the consideration of alternative explanations.

Implicit Theory

Implicit theories concern beliefs about the malleability of human attributes, with entity theorists viewing traits as fixed and incremental theorists viewing them as changeable (Dweck et al. , 1995). These beliefs shape how individuals interpret social information and explain behaviour, often without deliberate reflection. Entity-oriented interpretations tend to attribute actions to stable traits (e.g., viewing behaviour as reflecting an underlying nature), supporting faster, trait-focused judgments, whereas incremental interpretations emphasize situational factors and change (e.g., attributing behaviour to context), allowing for more contextualized judgments (Dweck et al., 1995; McConnell, 2001).

Research shows that implicit theories influence attribution, stereotyping, and susceptibility to simplified explanations, with entity theorists more likely to rely on dispositional judgments and endorse essentialist interpretations (Levy et al. , 1998). These tendencies can shape how information is interpreted by favouring explanations that align with prior assumptions, particularly under conditions of uncertainty.

The Implicit Association Test (IAT) provides complementary insight into automatic associations that may reinforce such interpretive tendencies (Greenwald et al. , 1998). However, its use remains subject to debate regarding 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 review, the IAT is therefore considered cautiously as an indicator of associative tendencies rather than a definitive measure of underlying beliefs.

In the context of misinformation, implicit theories influence whether individuals interpret messages in terms of stable intent or situational factors, shaping how particular interpretations appear plausible during evaluation.

Together, these mechanisms explain how initial interpretations come to be treated as plausible during misinformation evaluation. This does not imply that a conclusion has been verified, but that it is regarded as coherent or familiar enough to guide judgment without further scrutiny or consideration of alternatives. At this point, evaluation may come to a close, as the selected

interpretation shapes subsequent response, whether in the form of acceptance, rejection, or continued consideration. The following section examines the processing pathways through which such evaluations unfold.

2.4.4 Psychological Processing Pathways in Misinformation Evaluation

Building on the preceding sections, misinformation evaluation has been described as occurring under constraints, shaped by cognitive mechanisms and guided by framing effects. These conditions give rise to different patterns of reasoning during evaluation, as reflected in models that emphasize the roles of motivation, reasoning goals, and identity in shaping how information is processed (Ditto et al. , 2019; Kahan, 2017; Kunda, 1990; Taber & Lodge, 2006; Van Bavel et al. , 2021). In this context, the framework proposed by Nan et al., (2022) is used to describe how these patterns appear during evaluation by linking decision context, motivational orientation, and interpretive outcomes. Nan et al., (2022) identify four key processes, outlined below.

Accurate reasoning (Nan et al., 2022) refers to a pattern of information processing characterized by careful evaluation of evidence, critical thinking, and the application of logical reasoning principles. It involves examining claims beyond initial impressions and considering available information before forming a judgment. Individuals with stronger media literacy and topic knowledge are more likely to engage in such reasoning, supporting more careful evaluation of potentially misleading information (Govender et al. , 2017; Pickles et al. , 2021; Seo et al. , 2021). Moreover, training in recognizing misinformation can strengthen this orientation and improve individuals' ability to discern misleading claims (van der Linden et al. , 2020).

Accuracy-motivated reasoning (Nan et al., 2022) refers to a processing orientation in which individuals are motivated to form accurate and well-supported beliefs. While accurate reasoning describes how information is evaluated, accuracy-motivated reasoning refers to the goal of reaching a correct judgment. This orientation is associated with greater attention to evidence, more systematic evaluation of information, and openness to revising prior beliefs when confronted with new or conflicting data (Lyons et al. , 2019).

Directionally motivated reasoning ((Nan et al., 2022) occurs when individuals are motivated to arrive at conclusions that align with their pre-existing beliefs or preferences (Druckman & McGrath, 2019; Kunda, 1990). In such cases, information processing is shaped by the goal of maintaining consistency with existing views, leading individuals to selectively attend to, interpret, or evaluate evidence in ways that support those views. For example, individuals with particular worldviews may be more inclined to interpret or endorse information that aligns with their values, such as misinformation about vaccines (Hornsey et al. , 2018).

Identity-motivated reasoning (Nan et al., 2022) refers to a processing orientation in which individuals interpret information in ways that maintain alignment with their social or cultural identity (Onorato & Turner, 2004). In this case, evaluation is influenced by the goal of preserving group-consistent views, leading individuals to attend to and interpret information in ways that align with their group's values while discounting opposing perspectives. This pattern has been observed in contexts where political identity influences the interpretation of issues such as public health or environmental risks (Fielding & Hornsey, 2016; Miller, 2020).

Together, these psychological pathways describe different patterns through which individuals evaluate and interpret information. Rather than reflecting cognitive failure, they illustrate how reasoning goals, motivations, and identity concerns shape information processing under different conditions. Individuals may prioritize accuracy, coherence with prior beliefs, or alignment with social identity depending on situational and affective factors. Viewed in this way, susceptibility to misinformation can be understood not as a fixed trait but as a dynamic outcome influenced by which processing pathway becomes more prominent during evaluation. This perspective provides a bridge to the next section on susceptibility indicators, which examines how these factors influence the likelihood that particular pathways are engaged.

2.4.5 Susceptibility Indicators

The psychological processing pathways described in the previous section do not operate uniformly across individuals. This variation is associated with susceptibility indicators, factors that influence how individuals engage with information during evaluation. These indicators do not represent fixed traits or deficiencies but reflect conditions under which certain reasoning pathways become more likely than others.

This section reviews key susceptibility indicators identified by Nan et al., (2022), to clarify why the same message can lead to different interpretive outcomes across individuals. Following this framework, the section examines how these indicators are discussed in the literature and how they relate to differences in reasoning orientation during misinformation evaluation.

Knowledge and skills

According to Nan et al. (2022), knowledge and skills, such as subject knowledge, numeracy, and literacy, play an important role in shaping individuals' susceptibility to misinformation. Individuals with more accurate, in-depth knowledge of a particular subject are less likely to believe misinformation in that domain. Similarly, numeracy, defined as the ability to understand and work with probability and numerical information (Peters et al. , 2006), has been associated with lower susceptibility. In contrast, lower levels of literacy across domains such as health, science, media, or digital literacy are associated with increased susceptibility to misleading information. Overall,

differences in knowledge and skills influence how individuals interpret and assess information during evaluation.

Thinking Style

Thinking style refers to an individual's preference for analytical versus intuitive thinking, often described as Type 2 versus Type 1 (Kahneman, 2011) or systematic versus heuristic processing (Chen & Chaiken, 1999). Nan et al. (2022) identify thinking style as a key indicator influencing susceptibility to misinformation. Analytical thinking, commonly measured by the Cognitive Reflection Test (CRT) (Frederick, 2005), has been associated with more reflective evaluation of information and reduced reliance on initial impressions. However, this does not guarantee accurate judgments, as reasoning remains influenced by context and prior beliefs.

In contrast, intuitive thinking relies more on rapid judgments and heuristic cues. Prior work has shown that reliance on such processing is associated with increased susceptibility to misinformation, particularly when judgments are based on familiarity or surface-level cues (Pennycook et al., 2020; Pennycook & Rand, 2019b). Intuitive thinking has also been linked to conspiracy thinking, which has been associated with increased susceptibility (Uscinski et al. , 2016). Related constructs such as bullshit receptivity, the tendency to accept pseudo-profound but meaningless statements, and illusory pattern perception, the tendency to perceive patterns in random information, have also been associated with greater susceptibility to misinformation (Pennycook & Rand, 2020; Whitson & Galinsky, 2008).

Trust

Nan et al. (2022) identify trust as a source-dependent factor influencing susceptibility to misinformation. Trust in different sources shapes how information is received and evaluated. For example, trust in science and healthcare professionals has been associated with lower susceptibility, whereas trust in informal social networks, such as friends and family, has been linked to higher susceptibility in certain contexts. Findings related to trust in mainstream media, journalists, and government institutions are more mixed, reflecting variation across contexts and audiences.

Related research suggests that repeated exposure to information can increase its perceived credibility, a phenomenon known as the illusory truth effect (Fazio et al. , 2015). This highlights how both source trust and exposure dynamics can influence how information is interpreted during evaluation.

Emotion

Nan et al. (2022) identify emotion as a factor influencing susceptibility to misinformation, with states such as worry, anxiety, and perceived threat associated with differences in how individuals

respond to information. Findings across the literature suggest that the effects of emotion are context-dependent rather than uniform. For example, anxiety has been associated with increased susceptibility in some domains, such as health-related contexts, while in political contexts it has sometimes been linked to more careful or less biased processing (Garrett et al. , 2016). Overall, these findings indicate that emotional states can shape how information is engaged with during evaluation, although their effects vary depending on the context and the nature of the issue.

Value

Values are stable guiding principles that influence behavior and judgment across situations (Schemer et al. , 2012). Nan et al. (2022) identify values as an important factor shaping how individuals perceive and interpret information. For example, when individuals view a topic as personally significant (i.e., high topic involvement), they may engage more strongly with related information, which can influence how it is interpreted. Research has also shown that ideological and value-based orientations, such as political conservatism or religiosity, are associated with differences in how individuals respond to information, particularly when content aligns with their existing beliefs or moral frameworks (Schemer et al., 2012). Nan et al. (2022) similarly emphasize that values guide the filtering and evaluation of information, especially when messages resonate with individuals' broader value systems. In this way, values influence how information is interpreted, particularly when it connects with personally or socially meaningful concerns.

Group Identity

Group identity refers to an individual's sense of belonging to a social group, including the associated values, norms, and emotional attachments that influence how information is interpreted and evaluated (Nan et al., 2022). Research shows that group-based affiliations can shape how individuals respond to information, particularly when messages relate to group-relevant beliefs or concerns.

For example, political party affiliation has been associated with differences in how individuals interpret and respond to information, with a tendency to favor content that aligns with group positions. This pattern is also reflected in other forms of group identity. National narcissism, a defensive belief in one's country's superiority, has been linked to greater endorsement of conspiracy theories (Golec de Zavala et al. , 2019). In contrast, national identification, which reflects a positive but non-defensive connection to one's country (Leach et al. , 2008), has been associated with more balanced responses to information.

Taken together, these indicators show that individuals do not approach the same information under the same conditions. Differences in knowledge, thinking style, trust, emotion, values, and group identity shape how information is engaged with, leading to variation in how the same

message is interpreted across individuals. This variation raises an important question: not only how information is evaluated, but how reliable those evaluations are under different conditions. The next section addresses this by introducing Cognitive Reliability Analysis, which examines how contextual factors influence the reliability of recognition-based judgments.

2.4.6 Integrating Susceptibility Indicators and Judgment Reliability

Cognitive Reliability Analysis (CRA) (Woods et al., 1990) provides a framework for understanding how cognitive performance varies under conditions of uncertainty, complexity, and constraint. Developed within safety-critical domains, CRA treats perception, attention, memory, and reasoning as adaptive processes shaped by task demands rather than fixed or optimal mechanisms.

Within CRA and related human reliability approaches, Performance Shaping Factors (PSFs) refer to conditions that influence the reliability of human performance, while Human Error Probability (HEP) represents the likelihood that an action or judgment may result in error under specific circumstances (Park & Jung, 1996).

In this context, susceptibility indicators can be interpreted as PSFs that influence how cues are noticed, interpreted, and acted upon during evaluation. For example, high trust in a source may increase reliance on familiar cues, while greater domain knowledge or media literacy may support more careful evaluation and discrimination between competing interpretations (Nan et al., 2022).

From a CRA perspective, such conditions are associated with variation in the reliability of recognition-based judgments, rather than reflecting deficiencies in reasoning capacity (Di Pasquale et al. , 2015). This perspective helps link individual and contextual factors to differences in how judgments are formed and when interpretations are treated as adequate (Calhoun et al. , 2014; Pan & Wu, 2020).

2.4.7 Summary and Implications for Recognition-Based Sensemaking

Taken together, the literature reviewed in this section shows that misinformation evaluation is shaped by multiple human factors operating under conditions of uncertainty. Decision-making modes, cognitive mechanisms, framing processes, motivational pathways, and susceptibility indicators influence which cues are noticed, how information is interpreted, and what is judged plausible during evaluation. Rather than pointing to a single source of susceptibility, these studies establish the recipient-side conditions under which responses to misinformation vary across individuals and contexts, depending on how these factors interact.

Table 2 summarises the key insights and remaining gaps across this literature. These gaps point to a complementary aspect of the process: how messages are constructed to introduce and

structure the cues that individuals encounter. The next section, therefore, examines the message side of this process, focusing on persuasion models and rhetorical strategies that shape how information is presented and how it may interact with these human factors during recognition-based evaluation.

Table 2 - Key Takeaways: Human Factors impacting Susceptibility

Focus	What is explained	How it is operationalized	What is not addressed here
Decision-making under uncertainty	How individuals make judgments under time pressure, limited information, and resource constraints without exhaustive deliberation	Shifts between intuitive (Type 1), analytical (Type 2), & mixed modes as described by Dual Process accounts, Cognitive Continuum Theory, and bounded rationality	How specific message features trigger changes in reasoning mode in misinformation contexts
Cognitive mechanisms shaping interpretation	Why attention, memory, and evaluation become selectively focused on certain cues while alternatives are discounted	Mechanisms such as confirmation bias, cognitive dissonance, and information overload that prioritise belief-consistent, emotionally salient, or cognitively economical interpretations	How persuasive devices amplify cognitive mechanisms
Communication frames and meaning construction	How message structure stabilises particular interpretations and reduces ambiguity under uncertainty	Framing devices, narrative simplification (e.g., conspiracy narratives), and implicit belief structures used to guide recognition and interpretation	How specific persuasion strategies deliberately engineer frames and cues to exploit these interpretive tendencies
Psychological processing pathways	How different reasoning goals and motivations shape the evaluation of misinformation	Processing pathways identified by Nan (accurate, accuracy-motivated, directionally motivated, identity-motivated reasoning) that integrate decision modes, cognitive mechanisms, and framing effects	How persuasion models selectively steer one's thinking toward particular processing pathways
Susceptibility indicators	Why individuals exposed to the same message arrive at different interpretations	Context-sensitive conditions (knowledge & skills, thinking style, trust, emotion, values, group identity) that bias which processing pathway becomes dominant	How campaign targeting susceptibility indicators
Reliability of recognition-based judgments	How the likelihood of accepting a "plausible" interpretation varies across contexts and individuals	Cognitive Reliability Analysis concepts: susceptibility indicators mapped as Performance Shaping Factors that modify the reliability of recognition-based judgments under uncertainty	How to engineer persuasive cues and PSFs to systematically change reliability

2.5 Persuasion Models and Recognition-Based Message Design

This section examines the message side of the process and reviews persuasion models, rhetorical strategies, and persuasive design principles to show how message features structure cues and framing that guide recognition and interpretation under uncertainty. It concludes with phishing and scam practices as applied examples of how persuasive cues operate across reception, processing, and decision phases.

2.5.1 Models of Persuasion

Five key persuasion models commonly used to explain how persuasive communication shapes judgment and choice are reviewed below. In this review, these models are treated as frameworks for understanding when different modes of information processing are likely to be engaged and what types of influences, such as source credibility, affect, involvement, norms, and perceived intent, identified across these models, shape how messages are evaluated and acted upon (Cacioppo et al. , 2018).

Persuasion Knowledge Model (PKM)

PKM explains how awareness of persuasive tactics shapes individuals' responses to influence attempts (Friestad & Wright, 1994). It conceptualizes persuasion as an interaction between a target (recipient) and an agent (persuader), in which individuals draw on knowledge of persuasion strategies, the source, and the topic to form coping responses (DeCarlo, 2005; Haugtvedt et al. , 2018). This persuasion knowledge is shaped by prior experience, interpretive capacity, and motivation, and can support either acceptance or resistance. When activated, it can increase skepticism and scrutiny of persuasive messages (Shrum et al. , 2012). For this thesis, PKM is relevant because recognition of persuasive intent can function as a salient cue in evaluation, prompting more critical interpretation when that intent is noticed.

Elaboration Likelihood Model (ELM)

ELM (Petty & Cacioppo, 1986) distinguishes between a central route, involving detailed evaluation of message arguments, and a peripheral route, which relies on cues such as source characteristics and affect. The route taken depends on motivation, ability, and situational constraints (Y. Kumar et al. , 2023). In this way, ELM explains when individuals are more likely to engage with message content versus rely on cues during recognition-based evaluation (Shrum, 2012).

Heuristic Systematic Model (HSM)

HSM (Chaiken, 1980) emphasizes that heuristic and systematic processing can co-occur rather than operate as mutually exclusive modes (Bohner et al. , 1995; Chaiken, 1989; Kumar et al., 2023). Even when individuals engage in systematic evaluation, heuristic cues may continue to

influence judgment, particularly under uncertainty or motivational conflict (Winter, 2020). For example, an individual may evaluate the strength of a message's arguments while still being influenced by who delivers it or how it is presented. This highlights that interpretation is shaped by both message content and cues, rather than one replacing the other.

Hovland's Attitude Change Approach (HACA)

HACA (Hovland et al. , 1953) explains persuasion as shaped by the source, the message, and the audience. Persuasion is more likely when sources are perceived as credible, messages are clear and coherent, and appeals align with audience knowledge and emotional predispositions. HACA conceptualizes attitude change as a sequence in which individuals attend to a message, understand it, accept its claims, and retain them over time; failure at any stage reduces the likelihood of persuasion (Forgas et al. , 2011). While HACA presents persuasion as staged, a recognition-based view treats these elements as cues guiding attention and acceptance under constraint.

Theory of Reasoned Action (TRA)

TRA (Fishbein & Ajzen, 2011) explains behavior through attitudes and subjective norms shaping behavioral intentions. Attitudes reflect beliefs about expected outcomes, while subjective norms capture beliefs about others' views of the behavior. Subjective norms are particularly relevant for understanding how shared beliefs and social expectations influence intended responses (Shrum et al., 2012). However, because TRA focuses on volitional behavior and assumes that attitudes and norms are evaluated rationally (Petty & Wegener, 1999), it offers limited explanatory power for rapid, recognition-driven judgments under time pressure and uncertainty.

Building on these models, the next section examines how persuasive messages apply these principles to structure cues that shape recognition and evaluation.

2.5.2 Influence Mechanisms Through Persuasion Models

Each persuasion model highlights different ways in which influence is achieved by shaping how individuals respond to cues under varying conditions of motivation, ability, time pressure, and social context. Rather than introducing new processes, these models identify which cues are emphasized, how they are presented, and under what conditions they are more likely to affect judgment.

The following synthesis summarises how key persuasion models operate in terms of the cues they foreground and the conditions under which they are most effective, and how these mechanisms align with recognition-based decision processes:

- PKM: influence is more effective when recognition of persuasive intent is reduced or delayed; when persuasion knowledge is not activated (e.g., through novelty, distraction, or high trust), resistance strategies are less likely to be engaged.
- ELM / HSM: influence depends on reliance on arguments or heuristic cues as motivation and ability vary; heuristic cues may still shape judgment even during more systematic processing.
- TRA: influence operates by shaping attitudes and subjective norms to influence intentions, with perceived social approval and normative pressure serving as key triggers.
- HACA: influence leverages source credibility, message features, and audience predispositions as cues that guide attention and acceptance.

Table 3 brings these models together so their differences can be read directly across the same set of characteristics.

Table 3 - Models of Persuasion: Summary of Characteristics

Characteristics	**PKM**	**ELM**	**HSM**	**HACA**	**TRA**
Key Processing Influence	Suppressing or delaying recognition of persuasive intent	Regulating depth of elaboration via motivation and ability	Reliance on heuristic cues alongside systematic evaluation	Leveraging source credibility and message clarity and credibility	Shaping attitudes and perceived norms to form intentions
Expected Processing Mode Assumed	Shift toward intuitive or unguarded processing when persuasion knowledge is inactive	Central or peripheral processing depending on motivation and ability	Concurrent heuristic and systematic processing	Mixed processing guided by attention and acceptance stages	Deliberative intention formation, not necessarily deep message scrutiny
Relevant Susceptibility Conditions	persuasion knowledge, prior experience, trust in source	motivation, ability, involvement	heuristic reliance under cognitive load	source credibility, message clarity	attitudes, perceived norms
Typical Message Influence Mechanism	Obscuring persuasive intent through novelty, distraction, or trust	Selecting cues vs. arguments to match elaboration capacity	Reinforcing familiarity, credibility, and cognitive ease	Structuring messages to optimise attention, comprehension, and acceptance	Aligning messages with social approval and normative pressure
Typical Response Pattern	Reduced resistance; delayed scepticism	Heuristic acceptance or argument-based evaluation	Reliance on mental shortcuts alongside selective scrutiny	Acceptance guided by perceived credibility and message clarity	Adoption aligned with attitudes and perceived norms

Recognition Cue Effects	Persuasion intent not recognised results in lower defensive activation	Cue sufficiency determines acceptance threshold	Heuristic consistency stabilises judgment	Source and structure scaffold interpretation	Normative cues influence interpretation and behavioural intentions
Implications for Recognition-Based Decision-Making	Acceptance when persuasive intent is not recognized	Acceptance based on cue sufficiency under constrained elaboration	Acceptance via heuristic coherence	Acceptance via trusted source and structured meaning	Acceptance through intention–norm coherence

2.5.3 Persuasion Strategies and Techniques

Building on the persuasion models reviewed above, this section turns to classical rhetorical strategies, ethos, pathos, and logos, as ways persuasive message's structure cues to influence interpretation.

Classical rhetoric comprises three Aristotelian elements: ethos (credibility), logos (reason), and pathos (emotion) (Haskins, 2004; Hyde, 2004; Rapp, 2002). Ethos represents the speaker's credibility, logos denotes logical reasoning, and pathos appeals to emotion. When skilfully applied, these elements enhance the effectiveness of communication by fostering trust, coherence, and emotional engagement (Higgins & Walker, 2012; Holt & Macpherson, 2010; Rapp, 2002). This classification, known as the rhetorical triangle, highlights how these elements function as cues that shape how messages are received and interpreted.

Ethos

Ethos refers to the credibility and trustworthiness a communicator projects (Oeppen Hill, 2020; Wisse, 1989). It influences how audiences assess a message, particularly when the source appears trustworthy or familiar. In persuasive communication, ethos is often established through endorsements, authority cues, or signals of expertise that position the source as reliable. Appeals to expertise, similarity, or self-criticism can enhance perceived trust and reduce the likelihood of questioning the message, particularly under uncertainty (Brummett, 2000; Green Jr, 2004; Higgins & Walker, 2012; Hyde, 2004).

Pathos

Pathos involves engaging emotion to shape interpretation and response (Aho, 1985). It influences how messages are perceived by increasing emotional salience and a sense of urgency. Emotional engagement is often achieved through identification, aligning messages with targets' values, concerns, or needs (Higgins & Walker, 2012). Pathos evokes emotion through imagery, language,

and relatable scenarios, often using positive or fear-based appeals (e.g., inspirational stories or distressing situations) to prompt engagement and response (Bolatito, 2012; Oeppen Hill, 2020).

Logos

Logos refers to the use of reasoning and structured argument as persuasive tools (Haskins, 2004). Techniques include presenting data, supporting arguments, and providing justification (Green Jr, 2004), often through "logical cues" such as statistics or graphs. These cues structure claims and give them an appearance of logical coherence in persuasive messages.

The next section examines how these strategies are operationalized through persuasive design principles and devices in communication contexts.

2.5.4 Persuasive Design Principles and Devices

Building on ethos, pathos, and logos, this section turns to the concrete principles and devices through which rhetorical strategies are operationalized in persuasive messages. In this way, they represent how persuasion models are implemented in message design to influence which cues are noticed and interpreted. Rather than treating them as isolated techniques, the review considers how these principles make cues salient, credible, emotionally engaging, or difficult to question under time pressure and uncertainty (Cialdini, 2007; Entman, 1993). Authority, credibility, and social proof primarily reinforce ethos; scarcity, reciprocity, anthropomorphism, and emotion intensify pathos by increasing urgency and engagement; and framing, value formulation, and resistance-management strategies support logos by structuring plausibility and reducing counter-argumentation (Ajzen, 1991; Kahneman, 2011; McGuire, 1964). Salience and strategic ambiguity operate across these principles by directing attention and constraining interpretation. Table 4 summarises these principles by dominant rhetorical function, while Appendix B provides additional detail.

Table 4 - Persuasion design principles organized by dominant rhetorical function

<table>
<tr><th></th><th>Ethos</th><th>Pathos</th><th>Logos</th></tr>
<tr><td rowspan="3">Persuasion Design Principles</td><td>● Authority and Credibility
● Social Identity and Proof
● Accomplishment</td><td>● Reciprocity
● Scarcity
● Anthropomorphism
● Emotion</td><td>● Foot in the Door
● Overcoming Resistance
● Value and Impact Formulation
● Framing</td></tr>
<tr><td colspan="3">Salience</td></tr>
<tr><td colspan="3">Strategic ambiguity</td></tr>
</table>

The next section illustrates how these persuasive strategies operate in practice through phishing messages, showing how cues are used to shape responses under uncertainty.

2.5.5 Phishing and Scam Practices as Applied Persuasion Design

Phishing and scam messages provide a clear applied example of persuasion design in action. Rather than relying on argument quality, they are constructed to guide attention, interpretation, and response under uncertainty. These messages are typically designed to align with familiar roles, routines, or concerns, such as workplace communication, financial accounts, or service notifications (Cialdini, 2007; Grazioli, 2004). Even minimal alignment can increase perceived relevance and reduce initial skepticism, making the message more likely to be opened rather than ignored.

Reception Phase – Cues in Phishing Scams

During initial exposure, phishing messages rely on surface cues to capture attention and trigger recognition. Familiar branding, authoritative language, and urgency signals such as deadlines or warnings suggest legitimacy and importance (Downs et al. , 2007; Vishwanath et al. , 2011). For example, a message appearing to come from a bank or workplace IT service may warn that an account will be suspended unless the recipient verifies their credentials immediately. Emotional cues, particularly fear of loss or service disruption, further narrow attention and encourage rapid engagement (Kahneman et al. , 1982; Tversky & Kahneman, 1974). At this stage, persuasion operates by guiding what is noticed and treated as relevant.

Processing Phase – Responding to Phishing Scams

Once attention is captured, recipients assess plausibility rather than objective truth. Interpretation is influenced by prior experience, expectancies, and heuristics such as confirmation bias and representativeness (Cook & Smallman, 2008; Kahneman & Frederick, 2002). A request framed as a routine security procedure, such as confirming login details, can appear consistent with familiar organizational practices. Such framing reduces perceived alternatives and limits deeper checking (Fairclough, 2003; Higgins & Walker, 2012). If the interpretation appears consistent with prior experience, verification may be bypassed.

Decision Phase – Action Under Constraint

Under these conditions, recipients may act to resolve the perceived problem by clicking the embedded link and entering their credentials, for example, to prevent service interruption. The decision is not necessarily based on deliberate evaluation of alternatives but on whether the interpretation ‘seems true’ to act upon under time pressure. Together, these cues guide recipients toward action by focusing attention, reinforcing the interpretation, and reducing the perceived need for further checking.

These examples highlight how responses are shaped under pressure and uncertainty, raising the question of how individuals arrive at such decisions in practice. The next section addresses this by examining recognition-based evaluation within the NDM framework.

2.5.6 Summary: Persuasion as Engineering Recognition Under Uncertainty

This section reviewed persuasion models and strategies through a recognition-based lens, showing how message design shapes what is noticed and treated as plausible for action under uncertainty. Persuasion operates less through deliberate belief change than through the strategic organization of cues, credibility signals, framing, and emotional salience that guide recognition toward plausible interpretations within a given context.

Taken together, this section positions persuasion as the engineering of recognition under uncertainty, complementing the human-factors account discussed earlier. Table 5 summarises the key mechanisms identified across persuasion models, rhetorical strategies, and applied deception practices, and highlights the analytical gaps addressed in this review. The next section integrates persuasion and human susceptibility factors within the NDM and RPD framework to explain how recognition, mental simulation, and action selection unfold during misinformation evaluation.

Table 5 - Key Takeaways from Persuasion Models and Recognition-Based Message Design

Focus	**What is explained**	**How it is operationalized**	**Analytical gaps highlighted in this review**
Persuasion models	How different models assume distinct processing conditions and recognition shortcuts	Specification of cues (credibility, affect, norms, argument quality) that guide interpretation under varying motivation and ability	How recognition unfolds dynamically over time in naturalistic settings
Audience alignment mechanisms	How persuasion aligns message features with susceptibility indicators	Alignment between message cues and audience characteristics such as trust, values, emotion, and identity	Individual variation in recognition thresholds across contexts
Rhetorical strategies (ethos, pathos, logos)	How attention, credibility, and coherence are engineered	Use of authority cues, emotional salience, and apparent logic to function as recognition scaffolds that can shape attention and perceived plausibility	Interaction with downstream decision commitment and action
Persuasive design principles	How salience, framing, and ambiguity guide sensemaking	Design of attentional triggers, credibility shortcuts, and narrative stabilisers	Long-term effects on learning and persuasion knowledge
Applied deception (phishing, scams)	How persuasion exploits uncertainty and time pressure in practice	Sequencing of cues across reception, processing, and decision phases to lock in early interpretations	How recognition-based judgments transition into sustained belief or behaviour

2.6 NDM and Recognition-Based Susceptibility

Building on Sections 4 and 5, this section examines how responses to persuasive messages are formed under time pressure and uncertainty. It draws on the NDM paradigm (Klein & Zsambok, 1997) and the RPD model (Klein, 1993) to explain this decision process. The section begins by outlining the key principles of NDM and RPD required for this account.

2.6.1 Core Principles of NDM and RPD Model

The NDM paradigm examines how decisions are made in complex, dynamic environments where time pressure and uncertainty constrain deliberation (Klein & Zsambok, 1997). Within NDM, the RPD model explains how individuals arrive at decisions and actions without explicitly comparing alternatives. Rather than optimization, decision-making proceeds through cue recognition, mental simulation, and action once a plausible interpretation is reached (Klein & Myers, 1999; Rasmussen, 1981).

Recognition, cues, and pattern matching

RPD is grounded in the recognition of situational cues that are treated as "what is going on" (Klein, 1993; Schkade & Kleinmuntz, 1994). Decision-makers do not attend to all available information; instead, attention is selectively directed toward cues perceived as salient, credible, or familiar, based on prior experience stored in long-term memory (Klein, 1993; Rasmussen, 1981; Wickens et al. , 2021). These cues are matched to previously encountered patterns, functioning as an initial filter, allowing rapid classification of the situation and structuring what is treated as relevant for further assessment.

Mental simulation and satisficing

Once recognized, the viability of an interpretation or action is assessed primarily through mental simulation, rather than through explicit comparison of multiple alternatives (Klein, 1997). Mental simulation involves imagining how the recognized situation may unfold and whether a potential response would work under current constraints. This process aligns with the principle of satisficing rather than optimization: a decision is accepted once it appears workable, rather than after evaluating all possible alternatives (Simon, 1957). When inconsistencies arise or expectancy violations happen, interpretations may be revised or alternative patterns retrieved (Mosier & Fischer, 2010; Wickens et al., 2021).

Action without comparison

A defining characteristic of the decision-making process by RPD is that action selection occurs without explicit comparison among competing options (Klein, 1993). Evaluation and choice are interconnected: once an interpretation and associated action are judged workable through simulation, the decision-maker proceeds. If the outcome proves ineffective, feedback from action

contributes to future recognition and cue sensitivity, shaping subsequent decisions (Orasanu & Fischer, 1997). Decision-making is therefore adaptive and experience-driven, rather than deliberative in the classical sense of comparing alternatives by expected value or probability (Stanislaw & Todorov, 1999).

The following section examines how recognition and cue salience operate in misinformation contexts, shaping subsequent assessment and action.

2.6.2 Mapping Misinformation Processing onto the RPD Cycle

This section applies the RPD model to misinformation evaluation, examining how recognition, plausibility assessment, and action unfold when individuals encounter persuasive information.

Familiarity and Cue Recognition

When individuals encounter persuasive information, attention is selectively directed toward cues perceived as credible, salient, or familiar based on prior experience in long-term memory (Wickens et al., 2021). The perceived diagnosticity, reliability, and physical salience of cues shape initial situation assessment (Wickens et al. , 2013). In misinformation contexts, cues aligned with trusted sources, familiar narratives, or emotionally resonant frames are more likely to trigger recognition, while competing or contradictory cues may be filtered out.

Assessing Plausibility through Mental Simulation

Recognized cues are integrated with prior experience to form an interpretation of what the information implies and how it may unfold. As reviewed earlier, cue integration is shaped by heuristics such as representativeness, availability, anchoring, and confirmation bias, which influence how expectancies are generated and stabilized (Cook & Smallman, 2008; Joslyn et al. , 2011; Kahneman & Frederick, 2002; Tversky & Kahneman, 1974).

Mental simulation functions as a coherence check rather than a truth test. Individuals evaluate whether the emerging interpretation is internally consistent, aligns with existing beliefs or goals, and appears workable under current constraints (Mosier & Fischer, 2010; Wickens et al., 2013). Expectancy violations may trigger revision or additional information retrieval, but only when cues are sufficiently salient to disrupt the initial simulation (Hunt & Rouse, 1981).

Action: Acceptance, Reflection or Rejection

In the RPD model, action follows once an interpretation appears plausible, without explicit comparison among alternatives (Klein, 1993). In misinformation contexts, outcomes may include acceptance, sharing, ignoring, deferring judgment, or limited verification (Wickens et al., 2013). Acceptance does not imply full endorsement, nor does rejection require extensive evaluation. Feedback from these actions impacts future recognition by reinforcing or adjusting sensitivity to similar cues, making the process iterative and experience-driven rather than final or optimal.

The next section consolidates the key cues, expectancies, and actions identified in the literature and relates them to the RPD cycle.

2.6.3 Analytic Consolidation of Recognition-Based Decision Elements

To make the recognition-based process explicit, the literature reviewed in this chapter was analyzed to identify recurring cues, expectancies, and response patterns associated with misinformation evaluation. The findings are organized into three elements: cues, expectancies, and actions, corresponding to the recognition, plausibility assessment, and response stages of the RPD cycle. The categorization was developed by identifying recurring message features, cognitive responses, and behavioural outcomes across the persuasion, deception, and decision-making literature and mapping them onto the recognition–action sequence described in the RPD model.

Misinformation Cues Identified

Across the literature reviewed, three broad types of cues recur in accounts of initial recognition: message attributes, communication indicators, and deception indicators (Chaiken, 1980, 1989; Fairclough, 2003; Grazioli, 2004; Higgins & Walker, 2012; Nyhan & Reifler, 2010; Vraga & Bode, 2020). In this analysis, cues refer to observable message features and signals that shape what is noticed and treated as relevant. Message attributes (e.g., argument quality) influence whether content is treated as worth attending to, while communication and deception indicators direct attention through emotional salience or anomaly detection. Table 6 summarises the cues identified across the reviewed literature.

Table 6 - Summarized list of Cues emerged from thematic analysis of persuasion literature

Cue Category	Cue Type	Engineered Recognition Function
Message Attributes	Argument quality	Signals strength and legitimacy
	Source attractiveness & credibility	Prompts trust and initial acceptance
	Persuasion knowledge cues	Aligns with or obscures persuasive intent
	Subjective norms	Signals consensus or appropriateness
	Perceived motive	Frames intent as aligned or legitimate
	Similitude	Increases familiarity and identification
Communication Indicators	Attention-drawing features	Capture and direct attention
	Emotional triggers	Elicit targeted affective responses
	Identification cues	Encourage alignment with the message
	Metaphors & non-literal language	Guide interpretation through simplification
	Common-sense appeals	Reduce cognitive effort
	Familiar terminology	Increase fluency and perceived legitimacy

	Visual & visceral cues	Enable rapid, affect-driven recognition
Deception Indicators	Surface anomalies	Trigger scrutiny (e.g., grammar, spelling)
	Inconsistencies	Prompt anomaly detection
	Knowledge gaps	Signal missing or vague information
	Effectiveness & appropriateness	Signal poor contextual fit
	Activation cues	Trigger recognition of possible deception

Misinformation Expectancies Identified

Following cue recognition, individuals form expectancies by matching these cues against prior knowledge and experience. Expectancies represent provisional plausibility assessments and interpretive hypotheses formed prior to action. Across the literature reviewed, three recurring expectancy-related patterns emerge: knowledge and information processing, attention and perception, and detecting deception (Eveland Jr et al. , 2003; Vishwanath et al., 2011). These patterns influence whether information appears consistent, suspicious, or worthy of further evaluation, as summarised in Table 7.

Table 7 - Summarized list of Expectancies emerged from thematic analysis of persuasion literature

Expectancies Category	**Type**	**Function in guiding expectancies**
Knowledge & Information Processing	Knowledge of persuasion and topic	to possess varying levels of knowledge about persuasion strategies, the persuasive agent, and the topic
	Reasoning strategies	to rely on rational assessment and/or simplified decision rules (heuristics) when evaluating messages
	Elaborative processing	to consciously elaborate by connecting perceived cues to prior knowledge
	Interpretative reasoning	to form interpretative hypotheses to make sense of message cues
	Integrative judgment	to integrate available information and cues to arrive at an overall judgment
Attention & Perception	Attentional focus	to have their attention drawn to salient message cues
	Emotional response	to experience emotional reactions that shape perception and information prioritisation
	Information reduction	to reduce and filter information under cognitive and situational constraints
Detecting Deception	Deception hypothesis generation	to generate hypotheses about potential deception or manipulation
	Authenticity assessment	to assess trustworthiness, legitimacy, and message authenticity
	Explanatory evaluation	to construct and evaluate explanations for why a message appears deceptive

Misinformation Actions Identified

Actions represent behavioural responses that follow from a plausible interpretation, including compliance, disengagement, or verification. Across the literature reviewed, responses range from rapid acceptance and passive judgment to defensive behaviours such as disengagement, verification, or deception recognition (Downs et al., 2007; Downs et al. , 2006; Fischer et al. , 2008; Stanton et al. , 2005; Wang et al. , 2012). These responses reflect satisficing decisions under constraint and can influence future recognition by reinforcing or adjusting sensitivity to similar cues, as summarised in Table 8.

Table 8 - Summarized list of Actions emerged from reviewing persuasion strategies

Action Category	**Observed Actions**
Responses to Persuasion and Deception	Passive acceptance of salient cues
	Making rapid or "good-enough" judgments
	Global assessment without detailed evaluation
	Attitude change (temporary or stable)
	Increased or decreased responsiveness to messages
	Compliance or continued interaction
	Errors in judgment (e.g., falling for phishing scams)
Engaging or Defensive Mechanisms	Disengagement or reduced response
	Appeals to logic or deliberate evaluation
	Coping strategies (e.g., avoidance, dismissal)
	Letting guard down and interacting with the source
	Recognition of deception

Summary of the Decision Process

Across the reviewed literature, a consistent sequence emerges in which cues guide recognition, expectancies support plausibility assessment, and actions follow once an interpretation is judged plausible. Table 9 organizes these elements as an RPD-aligned sequence, linking recognition cues, plausibility-supporting expectancies, and subsequent actions. This alignment illustrates how decisions about misinformation unfold through recognition, plausibility assessment, and action under uncertainty.

Table 9 - Decision Analysis Process

Goals	Cues	Expectations	Actions
Forming Beliefs	Message Attributes	Processing Knowledge and Information	Response to Persuasion and Deception
	Communication Skills	Paying Attention and Formulating Perceptions	Engaging or Defensive Mechanisms
	Deception Indicators	Detecting Deception	

2.6.4 Integrating Persuasion Models within the RPD Phases

The preceding sections outlined the recognition-based process for evaluating misinformation. This section maps persuasion models and message design mechanisms onto the RPD cycle, showing how persuasive cues interact with recognition processes across perception, processing, and decision phases (Figure 2).

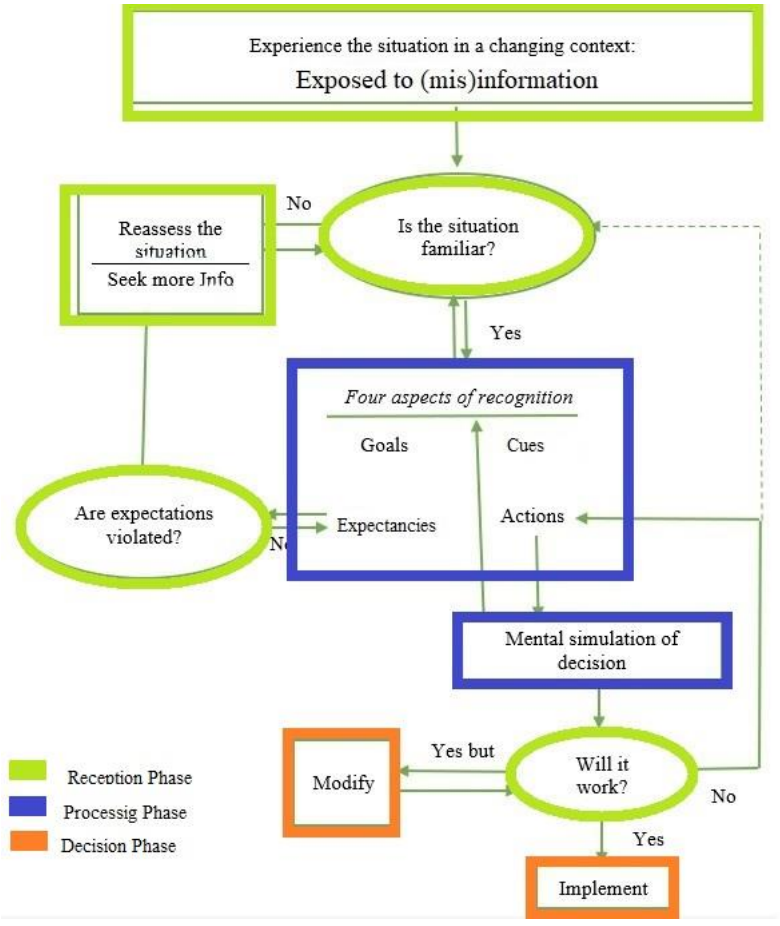

Figure 2 - RPD model breakdown to illustrate misinformation evaluation process

Perception Phase: Salience and Receiving Inputs

The perception phase captures the initial encounter with a message, when individuals decide whether it merits attention. Persuasion models influence this stage primarily through salience and heuristic cues. PKM influences whether people notice persuasive intent early; if intent is detected, they may become more cautious, but if it is masked or feels benign, the message may pass the initial filter. ELM and HSM help explain why, at first contact, people often rely on simple, familiar cues, such as source credibility, emotionally salient framing, familiar language, or social proof, rather than deliberate evaluation. TRA explains how early attitudes and subjective norms can make certain messages feel immediately acceptable or "normal," increasing the likelihood of acceptance before deeper scrutiny. Accordingly, message attributes, communication indicators, and deception indicators form the primary inputs shaping initial recognition.

Processing Phase: Recognition and Mental Simulation

The processing phase involves recognizing the message's meaning and assessing its plausibility through mental simulation. In misinformation contexts, initial exposure is distinct from recognition, which involves assessing credibility and aligning cues with existing expectancies rather than a passive encounter.

Persuasion models influence how plausibility is constructed. ELM and HSM explain variation in processing depth: when motivation or ability is low, people rely more on heuristic coherence, such as what feels consistent with prior beliefs or emotions, whereas higher motivation can trigger more effortful evaluation. HACA and rhetorical strategies shape the internal coherence tests people apply: "Does the source feel credible?", "Does the reasoning sound plausible?", "Does the emotional framing fit what I already believe or fear?"

Decision Phase: Decide and Modify

The decision phase concerns how interpretations translate into behavioral responses. Persuasion mechanisms influence whether individuals share, ignore, question, defer, or verify information. Actions are not limited to belief or non-belief. They include patterns such as rapid sharing, selective checking, disengagement, and delayed judgment, which align with the action patterns reflected in Table 8. If acting on misinformation produces social reward, relief, or validation, the pattern can become easier to recognize and repeat; if it produces conflict or negative feedback, it can sharpen future detection and reduce willingness to act on similar cues next time.

2.7 Discussion and Implications

This review shows that susceptibility to misinformation emerges from the interaction between message design, human susceptibility factors, and recognition-based decision processes. Persuasive messages structure cues that guide attention and suggest what should be treated as

relevant, while individual factors such as knowledge, trust, emotion, and identity shape which cues are noticed and how they are interpreted. Within the RPD process, these cues are matched to prior experience, forming expectancies that are assessed through mental simulation. Susceptibility, therefore, arises from how interpretations are formed and accepted as plausible under constraint, rather than from a fixed trait or a single moment of belief.

This perspective helps explain why the same message can produce different responses across individuals and contexts. Variation arises because individuals attend to different cues, form different expectancies, and assess plausibility under varying conditions. Acceptance, skepticism, and resistance are therefore not separate processes, but different outcomes of the same underlying recognition-based evaluation.

The framework contributes by integrating human factors, persuasive communication, and NDM within a recognition-based account of misinformation evaluation. It shifts the focus from whether individuals believe or reject information to how cues are attended to, how expectancies are formed, and how these expectancies are evaluated through mental simulation under constraints.

More specifically, the framework explains susceptibility through the sequence identified in the RPD-based analysis. Message cues first direct attention by signalling what is relevant within the sociotechnical environment. These cues are then interpreted in relation to prior knowledge, experience, and motivational conditions, forming an initial understanding of the situation. This understanding is assessed through mental simulation, where individuals evaluate whether it makes sense under current constraints. When no strong inconsistency is detected, evaluation may conclude at this point, leading to responses such as acceptance, sharing, ignoring, or limited verification. Susceptibility arises from how recognition and mental simulation lead evaluation toward an interpretation that appears plausible in the current encounter, shaping whether the response is acceptance, rejection, or deferred judgment.

This framework is conceptual and integrative. It does not specify how different cues are weighted, how expectancies are formed across contexts, or under what conditions mental simulation leads to acceptance, rejection, or continued evaluation. These processes are likely to vary across domains and situational constraints. Further empirical work is needed to examine how variations in cues and conditions influence evaluation outcomes in different misinformation settings.

2.8 Conclusion

In this chapter, literature on misinformation, persuasive communication, human susceptibility factors, and decision-making under uncertainty was reviewed to examine how susceptibility to misinformation has been understood across existing research. Through this review, the chapter

developed the Recognition-Based Susceptibility Framework, conceptualizing susceptibility as a recognition-based evaluation process emerging from the interaction between human factors, persuasive communication design, and recognition-based decision processes within sociotechnical information environments. The next chapter builds on this conceptual foundation by investigating how misinformation is evaluated in practice through participants' accounts of real encounters and decision processes using CTA and the CDM technique.

Chapter 3: How Misinformation is Evaluated in Practice

Building on the conceptual foundation and framework established in Chapter 2, this chapter examines how misinformation is evaluated in practice by investigating how individuals interpret and assess information at the point of encounter through non-routine evaluation episodes. A CTA study using the CDM technique elicited detailed accounts of non-routine evaluation episodes from ten participants. Interview data were analyzed using thematic analysis to identify recurring patterns in how participants interpreted information and decided whether to continue or stop the evaluation process. The NDM paradigm, specifically the RPD model, was used as an interpretive lens to organize these patterns. The findings show that variation in outcomes arises not only from what individuals notice but from how the situation is interpreted and when the evaluation is brought to a close. The study provides a process-oriented account of misinformation evaluation, showing how susceptibility emerges within the evaluation process and is not reducible to message content or individual vulnerability alone. This also explains how persuasive framing becomes consequential by shaping the cues and interpretations through which plausibility is constructed.

3.1 Introduction

This chapter investigates how individuals evaluate misinformation at the point of encounter. It develops a process-oriented account of how judgments are formed under conditions of uncertainty, focusing on how people interpret information and decide when evaluation is plausible to proceed.

Misinformation impact extends across domains such as public health, political decision-making, and trust in institutions, where misleading claims can influence how individuals interpret risks, evidence, and recommended actions, and the decisions that follow (Lewandowsky et al., 2017; Roozenbeek et al., 2022). This challenge is increasingly recognized as a broader societal and national security risk. It is intensified in contemporary information environments characterized by high volume, ubiquity, and competing claims, where individuals are frequently required to make judgments based on incomplete or contested information. Under such conditions, evaluation is not always supported by clear verification or stable ground truth, making it difficult to determine what information can be relied upon. In this review, misinformation is used as an umbrella label for misleading content encountered by individuals; sender intent is outside the scope of analysis.

Research on misinformation has clarified several aspects of the problem. Psychological and experimental studies have shown that judgments about misleading claims are shaped by analytic thinking and familiarity effects (Pennycook & Rand, 2020, 2021), as well as broader psychological drivers such as prior beliefs and resistance to correction (Ecker et al., 2022). Other work has

shown that credibility is also shaped by cues surrounding the message, including social endorsement and source signals (Traberg et al., 2024). At a broader level, misinformation has also been framed as a societal problem whose influence extends beyond isolated claims (Traberg, 2022). From a distributed cognition perspective, evaluation likewise does not reside in the individual alone, but is shaped by interaction with tools, representations, and surrounding informational structures (Hutchins, 2020). These studies provide detailed accounts of the conditions under which misinformation is more likely to be accepted or rejected. However, they offer more limited specification of how individuals progress from noticing cues to forming a judgment, particularly in situations where information is incomplete, contested, or requires interpretation rather than direct verification.

To examine this process, this study draws on NDM (Klein & Zsambok, 1997) perspective, using the RPD model (Klein, 1993) where judgment arises through recognition, interpretation, and mental simulation under conditions of uncertainty and pressure. Within this perspective, evaluation is shaped not only by available information but by how patterns are matched, and cues are recognized and interpreted in context. At the same time, research on persuasive communication (Holt & Macpherson, 2010) highlights how messages are designed to foreground particular cues, narratives, and interpretations, influencing how information is understood during evaluation. Taken together, these perspectives position misinformation evaluation as a process in which recognition, interpretation, and message design interact as individuals assess what is plausible to support a judgment.

To investigate how individuals evaluate misinformation at the point of encounter, this study used Cognitive Task Analysis (CTA) (Crandall et al., 2006) with the Critical Decision Method (CDM) technique (Klein et al. , 1989) to elicit detailed accounts of non-routine evaluation episodes from ten participants. The approach captured how participants interpreted information, directed subsequent assessment, and determined when to proceed, reject, or defer under uncertainty and pressure. Interview data were analyzed using Emergent Themes Analysis (Wong, 2003), identifying recurring patterns in how evaluation unfolded across incidents.

Building on these patterns, differences in misinformation evaluation outcomes are better understood in terms of how individuals construct a plausible interpretation and where evaluation is brought to a close, based on patterns identified across cases.

3.2 Background and Related Works

This section reviews the conceptual foundations for understanding how misinformation is evaluated. It begins with misinformation and susceptibility as the problem space, then introduces the NDM paradigm and the RPD model as the decision framework that guides the evaluation. It

next considers sensemaking as the process through which people interpret what is happening, briefly situates this account alongside existing persuasion explanations, and finally positions stopping as a recognition-driven point at which a plausible understanding is reached.

3.2.1 Misinformation and Susceptibility

Misinformation significance lies not only in falsity, but in its capacity to appear credible, circulate rapidly, and shape judgment in ways that can undermine public health, social cohesion, and trust in institutions (Roozenbeek, Maertens, et al. , 2022; Roozenbeek, Schneider, et al., 2020; Roozenbeek & van der Linden, 2019a). In this chapter, the term misinformation is used as the primary label because the focus is on how people evaluate misleading information at the point of encounter, regardless of sender intent.

Susceptibility is treated here as a context-sensitive vulnerability shaped by the interaction of cognitive, situational, and social factors, including prior knowledge, uncertainty, time pressure, emotion, motivation, and group identity commitments (Maertens et al. , 2023; Nan et al., 2022). This makes it relevant not only to whether misinformation is believed, but to how information is noticed, interpreted, and judged during evaluation.

Existing misinformation research has explained important parts of this problem at different levels. Computational and platform-oriented work has focused on detection, classification, and mitigation at scale (Shu et al., 2018; Zhou & Zafarani, 2020). Psychological and experimental research has examined truth judgments, correction effects, analytic thinking, and individual differences associated with belief and resistance (Ecker et al., 2022; Pennycook et al., 2020; Pennycook & Rand, 2019a). Message-level research has shown how cues such as framing, familiarity, source signals, emotional salience, and social endorsement shape perceived credibility (Traberg, 2022; Wang et al., 2024). Together, these strands have advanced understanding of misinformation as a problem of content, cognition, and intervention. However, less attention has been given to how individuals interpret such cues and assess plausibility as an unfolding process when encountering contested information.

3.2.2 Decision-Making under Uncertainty

Decision-making in conditions of uncertainty and constraint has been widely studied within the NDM paradigm, which focuses on how people make decisions in complex, dynamic environments under pressure (Klein & Zsambok, 1997). Such environments are characterized by ill-structured problems, shifting goals, time pressure, high stakes, and evolving conditions, often involving multiple information sources and ongoing interpretation as new information emerges (Orasanu & Connolly, 1993).

Within this paradigm, the RPD model (Klein, 1993) describes how decisions are made through pattern recognition and mental simulation rather than explicit comparison of alternatives. Individuals use cues to recognize situations, draw on prior experience, and evaluate a plausible course of action. Once the situation is understood as plausible, they proceed with making a decision.

This perspective is particularly relevant for examining how individuals evaluate misinformation under uncertainty and pressure, providing a process-based account of how judgments are formed through recognition and interpretation rather than exhaustive comparison under uncertainty in a complex, dynamic environment.

3.2.3 Sensemaking and Framing

To explain how individuals interpret information under uncertainty, this study draws on sensemaking as described in the Data-Frame model (Klein et al. , 2007). In this view, people do not passively receive information; they construct an understanding of "what is going on" by fitting incoming cues into an evolving frame, which is updated or revised as new information is encountered. This process allows individuals to form a coherent interpretation without evaluating all available information.

Framing theory complements this perspective by explaining how information is presented in ways that guide interpretation. Frames make certain aspects of a situation salient while downplaying others, shaping how events are understood and evaluated (Entman, 1993; Goffman, 1974). In the context of misinformation, framing can make misleading interpretations appear plausible by aligning information with familiar cues, narratives, or expectancies (Lewandowsky et al., 2017).

Together, sensemaking and framing theories can explain how individuals move from perception to interpretation, constructing a coherent account of a situation that can support judgment and action under uncertainty.

3.2.4 Persuasive Communication Design

To understand why misinformation messages can appear convincing, it is also necessary to consider how persuasion has been explained in communication and decision research. Persuasive communication design (Holt & Macpherson, 2010) grounded in persuasion models and rhetorical strategies, structures cues that guide recognition and interpretation through framing, salience, credibility signals, affect, norms, familiarity, and persuasive intent.

Aristotle's rhetorical strategies identify three basic appeals: ethos, or credibility; pathos, or emotional appeal; and logos, or appeal through reasons and evidence (Haskins, 2004; Rapp,

2002). These appeals help explain how messages can be made to seem trustworthy, emotionally compelling, or logically convincing even when their overall interpretation is misleading.

Persuasion models are treated as processing frameworks: they specify when different modes of information processing are likely to be engaged and identify classes of influences, such as source credibility, affect, involvement, norms, and perceived intent that shape how messages are evaluated and acted upon (Cacioppo et al., 2018). Several persuasion models have been discussed in the literature, including:

- **The Elaboration Likelihood Model** (Petty & Cacioppo, 1986) proposes that persuasion can occur either through careful consideration of message content or through reliance on surface cues.
- **The Heuristic–Systematic Model** (Chaiken, 1980) makes a similar distinction, but emphasises that people may rely on simple decision cues, such as source credibility or familiarity, while also engaging in more effortful evaluation.
- **The Persuasion Knowledge Model** (Friestad & Wright, 1994) focuses on whether people recognise that they are being persuaded, since recognised persuasive intent can trigger resistance.
- **The Theory of Reasoned Action** (Fishbein & Ajzen, 2011) explains how beliefs and perceived social norms shape behavioural intentions.
- **Hovland's Attitude Change Approach** (Hovland et al., 1953), which explains persuasion in terms of the effects of source, message, and audience factors on attitude change.

In this study, misinformation susceptibility indicators identified by Nan et al., (2022) are treated as points of sensitivity that persuasion models and rhetorical strategies can engage. Nan et al., (2022) identify these indicators as: trust, emotion, values, knowledge, thinking style, and group identity. For example, a message using an authority-based appeal (e.g., "the institute of paediatricians recommends…") aligns with existing trust in institutional sources, while appeals to being a "good parent" align with identity-based concerns. In both cases, the message resonates with what individuals already trust or value, increasing the likelihood that it is attended to and accepted. However, this account focuses on how persuasive communication engages these sensitivities, rather than how individuals evaluate misinformation as an unfolding process under uncertainty. That process is taken up in the next section.

3.2.5 Evaluation and Stopping under Uncertainty

Bringing these perspectives together, evaluation of misinformation can be understood as a process that unfolds under conditions of uncertainty, time pressure, and incomplete information

(Orasanu & Connolly, 1993). In these contexts, individuals rely on recognition, framing, and persuasive cues to construct an interpretation of what is happening.

Within this process, evaluation does not continue indefinitely. Simon's theory of bounded rationality explains that decision-making often proceeds without exhaustive optimisation, with people settling for an option that is satisfactory under constraint (Simon, 1957). In misinformation contexts, however, the issue is often not selection among clearly defined alternatives, but interpretation of ambiguous and contested information as it unfolds. Consistent with RPD and sensemaking, evaluation therefore proceeds until the situation is interpreted as sufficiently coherent to support judgment or action. At that point, further analysis is no longer pursued, not because all information has been evaluated, but because a plausible understanding has been reached.

This framing provides the basis for the present study, which investigates how recognition, framing, and persuasive cues interact during evaluation, and how the point at which evaluation is brought to a close shapes subsequent judgment.

3.3 Methodology

CTA (Crandall et al., 2006), is a set of methods developed to uncover the cognitive aspects of skilled performance that are not directly observable in routine task analysis. CTA focuses on how people think in complex, real-world situations, including how they recognize patterns, make judgments, interpret cues, form expectancies, solve problems, and make decisions under uncertainty and time pressure. The purpose of CTA is not simply to document observable behaviour or procedural steps, but to elicit the underlying cognitive processes, knowledge, strategies, and contextual judgments that enable effective performance in dynamic, complex environments. CTA is especially important in domains where expertise depends on tacit knowledge that practitioners often cannot easily articulate without structured elicitation. In this study, CTA was used to understand how participants described evaluating misinformation in non-routine situations, based on their own previously experienced incidents.

The Critical Decision Method (CDM) (Klein et al., 1989) is a retrospective interview technique within CTA designed to elicit the tacit knowledge and cognitive processes involved in difficult, non-routine incidents. The method was developed specifically for naturalistic settings where decision making occurs under conditions of uncertainty, ambiguity, time pressure, high stakes, incomplete information, changing goals, and dynamic real-world situations (Orasanu & Connolly, 1993). CDM reconstructs a specific incident in detail that the participant personally experienced and probes the participant's thinking at critical points during the event. CDM is valuable because it helps reveal tacit cognitive elements of expertise that are difficult for participants to articulate

spontaneously. Through incident-based probing, CDM elicits how participants noticed cues, interpreted information, considered alternatives, projected possible consequences, and determined when they had sufficient understanding to proceed.

These elicited accounts formed the basis for subsequent Emergent Themes Analysis (Wong, 2003) and were later organized using the RPD model (Klein, 1993) as an interpretive lens. CDM data analysis focuses on identifying the knowledge, cues, judgments, reasoning processes, and decision strategies participants use during critical incidents (Blandford & Wong, 2004; Wong, 2003), with the aim of reconstructing how the participant understood the situation and progressed through the decision process. The following sections describe the participants, data collection, and analysis.

This research received ethical approval from the Auckland University of Technology Ethics Committee (AUTEC) under Approval No. 25/174 for the project 'Susceptibility to Misinformation: The Impact of User Experience (UX) Design Factors on Users' Decision-making and Belief Formation' (see Ethics Approval section). Participants were provided with an information sheet briefing them on the study procedures, their rights, and confidentiality protections, and gave informed consent prior to participation.

3.3.1 Participants

A total of 10 participants (n = 10) were recruited through volunteer self-selection. Recruitment was conducted via public advertisements distributed across professional platforms such as LinkedIn and organizational channels, including news platforms, fact-checking organizations, hospitals, universities, and public community networks. This approach was intended to reach individuals with different forms of experience in evaluating information. Interested individuals contacted the researcher and were screened for basic eligibility, including age, English proficiency, and engagement with online information. Individuals with direct academic relationships to the researcher were excluded.

Participants were grouped based on their exposure to structured verification practices, rather than a uniform or formally validated level of expertise. Five participants reported professional or academic experience involving systematic evaluation of information, where verification formed part of their role or training. These included: MAN (government fact-checker; associate professor; law librarian), MEI (data analyst with fact-checking training), SS (journalist and political activist), MAY (cognitive and social psychology researcher, PhD), and RJ (cognitive neuroscience researcher, PhD). Their experience involved practices such as cross-referencing sources, consulting domain experts, evaluating scientific evidence, and considering how claims may be selectively presented or framed.

The remaining five participants did not report formal training in fact-checking or routine use of structured verification procedures. These included: AL (world-class athlete), SA (emergency doctor), ZAH (pharmacist with over 30 years of experience), OMN (interior designer), and MAH (master's in education, not employed). While several participants had domain knowledge, their evaluation of information was primarily based on everyday strategies such as source inspection, informal search, personal experience, and social input.

For clarity, these groups are referred to as participants with greater exposure to structured verification practices and participants without routine structured verification practices. For readability, the terms "higher-exposure" and "lower-exposure" are used to refer to these groups, reflecting differences in exposure to structured verification practices rather than implying levels of expertise. This distinction is analytic rather than hierarchical and does not imply validated expertise, greater accuracy, or expert performance in the NDM sense. It is used to examine how different forms of experience and exposure to verification practices may shape evaluation processes across participants.

3.3.2 Data Collection

Data were collected through individual semi-structured interviews conducted using the CDM technique. Interviews were conducted individually to provide participants with a private setting for discussing their experiences, and each session lasted approximately 60 to 90 minutes. The interview followed a staged procedure designed to move from a general description of participants' usual evaluation practices to a detailed reconstruction of a personally experienced misinformation incident. The planned interview guide is provided in Appendix C.

First, participants were asked a general orientation question about what they typically do when they encounter information they suspect may be misinformation. This established their usual evaluation practices before moving into the incident-based part of the interview.

Second, participants were asked to identify a memorable, non-routine incident involving misinformation that they had personally encountered. Participants were not given a researcher-designed misinformation scenario. Instead, they were asked to recall a real situation that had been challenging to evaluate.

Third, participants were asked to retrospectively reconstruct the incident using a decision timeline. They described what happened from the point at which they first encountered the information through to the point at which they reached a judgment or action. This reconstruction focused on what they did, what they first noticed, how the information was presented, what questions they considered, what information or resources they consulted, and what they were trying to achieve at the time.

Fourth, CDM probes were used to examine participants' thinking at key points in the incident. These probes explored whether the situation resembled prior experience, what alternative actions were considered, how actions were selected or rejected, and what might have led to a different decision. Participants were also asked whether they imagined possible consequences, sought guidance from others, relied on particular training or experience, and how long the decision process took. Additional probes asked how they would summarise the situation and how changes to key features of the message or context might have affected their judgment.

The interviews were audio-recorded and transcribed for analysis, using an automated transcription tool (Microsoft Word Transcribe), followed by manual checking for accuracy. Because CDM interviews are semi-structured, the wording and order of probes varied depending on the participant's responses. However, the same interview guide was used across participants to examine their accounts of the evaluation process.

3.3.3 Analysis

Interview transcripts were analyzed using Emergent Themes Analysis (Wong, 2003), focusing on participants' accounts of specific misinformation incidents. The analysis followed a systematic, multi-stage process to identify and organize recurring patterns across participants.

Stage 1 - Initial coding: Incident-based transcript segments were coded inductively to capture what participants noticed, interpreted, checked, imagined, and decided. This resulted in 78 fine-grained codes.

Stage 2 - Development of sub-themes: Related codes were compared across participants and grouped into sub-themes reflecting recurring patterns in how participants evaluated information. For example, codes such as 'Identifying red flags', 'Lack of reliable resources', and 'Critical reading' were grouped under 'Detecting absence or weakness of source credibility', while 'Verifying claims', 'Agreeing on facts', and 'Source credibility' were grouped under 'Stopping after limited or targeted verification'.

Stage 3 - Synthesis into process-level themes: These sub-themes were synthesized into four process-level themes: 'Cue recognition and salience', 'Expectancy formation', 'Mental simulation and projection of consequences', and 'Plausibility judgment and stopping points'. These themes organize the findings as a sequence of evaluation processes, from initial recognition to judgments of plausibility.

To support analytic transparency, Table 20 in Appendix D provides representative examples of how multiple incident-based excerpts were inductively coded, compared, and collated into sub-themes and process-level themes. The table shows how selected data extracts were linked to

initial codes, sub-themes, and process-level themes, supporting transparency in how the analysis moved from participant accounts to broader patterns. Table 10 presents a focused illustrative extract from this analytic progression for the process-level theme "Plausibility judgment and stopping points".

Table 10. Illustrative ETA coding and theme development for "Plausibility judgment and stopping points"

Data extract	**Initial code(s)**	**Sub-theme**	**Final process theme**
"Once you realize the purpose… you can decide" (OAL, P13, L22)	Identifying purpose; sufficient understanding	Stopping through early recognition and framing	Plausibility judgment & stopping
"Familiar… that's how you found out" (OZH, P10, L28)	Prior knowledge; recognizing inconsistency		
"No ads… I'd believe it" (OMH3, P4, L16)	Lack of over-promotion; no further checking		
"Checked a few times… then enough" (OSAN, P10, L10)	Verifying claims; enough checking	Stopping after limited or targeted verification	
"Citations are reliable… believe it" (EMS, P23, L1)	Source credibility; reliable citations		
"Contradictions… lose trust" (ESS, P12, L7)	Losing trust; contradictory details	Stopping through contradiction detection	
"Urgency… quick but informed" (EMAY, P10, L36)	Urgency; reassessing actions	Stopping calibrated by context (time, importance, knowledge)	

The initial codes were kept close to participants' accounts and used to identify specific cues, actions, or judgments expressed in individual excerpts. These codes were then compared across excerpts to develop sub-themes that captured broader process patterns in how participants recognized cues, formed expectations, simulated consequences, or reached stopping points. In line with Emergent Themes Analysis (Wong, 2003; Wong & Blandford, 2002), the purpose was not to count repeated phrases, but to identify meaningful patterns in decision processes across incident accounts. Therefore, where a sub-theme is illustrated by a single excerpt in the findings, the quote is used as a clear exemplar of a process distinction observed in the wider analysis, rather than as the sole basis for the theme. In qualitative analysis, a process distinction illustrated by a single excerpt can still be analytically important because it indicates the existence of a particular behaviour or reasoning pattern. Such observations help define the range of behaviours

that may emerge under different conditions and, therefore, the variability that systems or designs may need to accommodate (Ashby & Goldstein, 2011).

The resulting framework reflects recurring patterns in how participants made sense of information under different conditions. The Recognition-Primed Decision (RPD) model (Klein, 1993) was used after theme development as an interpretive lens to organize these patterns. The following section presents the resulting framework, followed by sections that provide the supporting analysis and evidence.

3.3.4 Limitation

This study is based on a small sample of ten participants and does not aim to provide population-level generalization. The findings reflect process-level patterns derived from detailed accounts of specific incidents, rather than representative distributions of behaviour.

Participant grouping was based on self-reported exposure to structured verification practices, not on validated measures of expertise. As such, the distinction between participants reflects differences in experience and practice rather than formally established levels of expertise or performance.

The analysis relied on retrospective accounts elicited through CDM interviews. While this approach supports rich reconstruction of decision processes, it is subject to limitations associated with recall and post hoc interpretation.

Finally, the framework represents an analytic organization of recurring patterns observed in this dataset. It is not intended as a predictive model or a normative account of correct evaluation, and further work is needed to examine how these patterns manifest across broader contexts and populations.

3.4 A Process Framework of Misinformation Evaluation Derived from Analysis

Figure 3 presents the framework derived from recurring patterns identified across the ten CDM interviews. It synthesizes how participants described evaluating misinformation in a memorable incident. The framework is not intended as a normative model of correct evaluation, nor as a classification of individuals. Instead, it represents a process through which participants constrained uncertainty, constructed plausibility, and reached a point where they could act, reject, continue assessment, or defer judgment.

The process begins with exposure to information, followed by cue recognition and salience, where participants rapidly orient to aspects of the message that appear noteworthy, credible, or

suspicious. This stage did not involve neutral observation; rather, it shaped the initial framing of the situation by narrowing what was attended to and what could be disregarded.

This early framing led to expectancy formation, where participants developed a working sense of what the message was likely to represent and how it should be approached. These expectancies guided the depth and direction of evaluation. For example, whether further verification was necessary, what kind of evidence would be considered relevant, and whether the message aligned with prior knowledge or situational relevance. In this way, evaluation was directed not only by available information, but by what participants anticipated needing to resolve.

When uncertainty persisted or the situation was perceived as consequential, participants engaged in mental simulation and projection of consequences. Rather than relying solely on direct verification, they described constructing possible explanations, being conscious about who might benefit, and the implications of accepting, rejecting, or delaying action. This stage reflects how plausibility was often shaped through projected outcomes and inferred intent, not only through factual checking.

The process culminates in plausibility judgment, where participants determined that they had understood the basis to proceed. Importantly, this stopping point did not necessarily correspond to certainty. Instead, it reflected sense making and a threshold of “seems true” or plausible understanding relative to the situation. As shown in the lower branch of the framework, this resulted in four distinct outcomes: proceeding with action when the message was considered sufficiently plausible; rejecting or disengaging when contradiction became dominant; continuing assessment when uncertainty remained but the issue was important; and deferring judgment when uncertainty persisted under low urgency.

The conditions shown on the left of the figure represent recurring influences on how participants moved through the process. These conditions shaped what was noticed, how expectancies were formed, the extent of simulation, and when evaluation was ended. The framework, therefore, captures misinformation evaluation as a context-sensitive sensemaking process, rather than a simple acceptance or rejection of information.

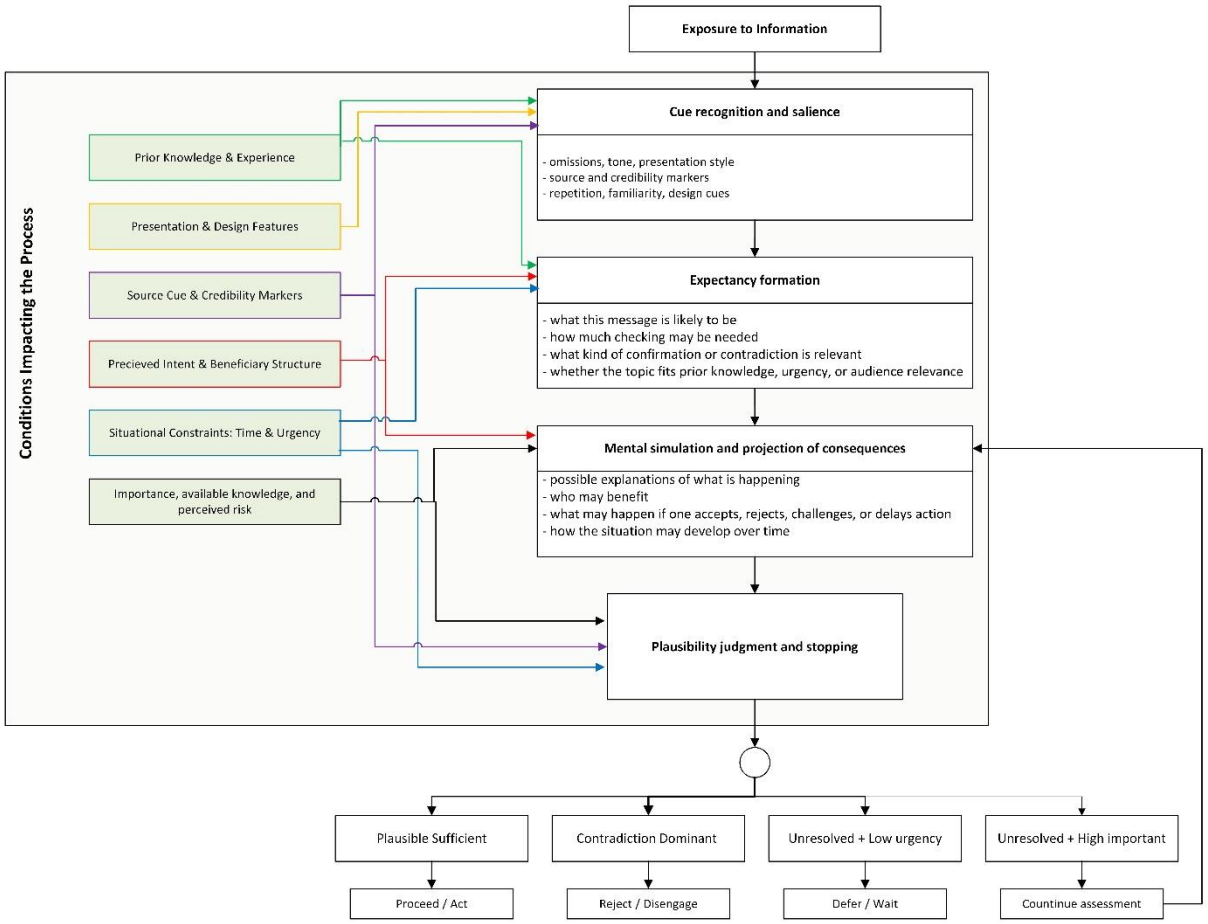

Figure 3 - Evaluating process of misinformation derived from CDM interviews

The following sections provide the empirical grounding for each component of this framework, illustrating how these patterns were expressed across participants and evaluation contexts.

3.5 Findings

This section presents findings from the CTA study, which aims to understand the decision processes participants described when encountering misinformation. Following an emergent themes analysis of CDM interviews (Wong, 2003; Wong & Blandford, 2002), the findings are structured by stages of recognition-based sensemaking, tracing how participants recognized cues, formed expectancies, engaged in mental simulation, and determined when understanding was sufficient for belief or action. These stages correspond to the process-level themes identified through the analysis described in Section 3.3.3.

Accounts from higher-exposure and lower-exposure participants are investigated together within each process. This analytic strategy highlights how differences in experience and expectancies shape recognition and decision pathways, without treating belief, skepticism, or resistance as indicators of correctness or deficiency. Participant quotations are used as evidence of decision processes in context.

The subsections that follow trace this sequence, beginning with initial cue recognition and salience (Section 3.5.1), followed by expectancy formation (Section 3.5.2), mental simulation (Section 3.5.3), and the selection of actions and stopping rules that form belief or response (Section 3.5.4).

3.5.1 Cue Recognition and Salience

Participants' initial encounters with misinformation were characterized by rapid attention to particular features of the message. Throughout the interviews, participants described noticing specific cues early in the encounter, which shaped their initial perception of the situation. These cues influenced which aspects of the information were considered relevant, what was ignored, and whether further information search was initiated. This pattern was observed consistently across participants, indicating that early cue recognition functioned as a primary mechanism for framing the evaluation process.

The analysis in this subsection focuses on how different cues became salient at the point of first contact with misinformation. Rather than evaluating the accuracy of resulting judgments, the findings describe regularities in how recognition happened under conditions of uncertainty, using participants' accounts of specific incidents to illustrate how early cue detection constrained subsequent sensemaking.

The analysis in this subsection is organized around the sub-themes identified in the coding table: detecting missing or weak credibility cues, recognizing inconsistency with prior knowledge or expectations, recognizing persuasive and attention-directing presentation cues, recognizing simplified and reduced-effort formats, and recognizing institutional credibility cues. These sub-themes show how different cues became salient at first contact with misinformation and shaped how participants initially framed the information. Participant excerpts are used to illustrate how these cue-recognition patterns appeared across higher-exposure and lower-exposure accounts.

Detecting missing or weak credibility cues

Participants attended to missing or weak credibility cues, such as the absence of identifiable sources, exaggerated claims, and lack of supporting evidence, which shaped early plausibility assessments. For instance, one lower-exposure participant stated:

> *"Not mentioning the source of the news or which bank or organization is going to release the money and where is the destination of the money and like, which person in charge is going to take the responsibility? No names, no brands, no organizations name was mentioned! It was very, very brief and very positive!"* (OAL, P6, L6)

In this case, the absence of identifiable sources constrained plausibility early in the encounter, leading to rapid dismissal without extended verification.

Higher-exposure participants attended to similar cues, such as sourcing and internal coherence, but treated them as triggers for deliberate and targeted verification:

> *"When I encounter a piece of misinformation, the first thing I do is read it critically. I look for any red flags that might indicate it's unreliable, such as exaggerated language, lack of sources, or sensationalized claims."* (ERJ, Pp 1, Line 13)

In this case, recognized cues were not used for immediate judgment, but as signals to initiate focused verification. Unlike the earlier lower-exposure account, these cues were not used for rapid dismissal but as signals to initiate focused verification, indicating that experience influenced how similar credibility cues guided subsequent evaluation.

Recognizing inconsistency with prior knowledge or expectations

Participants also recognized inconsistencies between the message and prior knowledge or expectations, which shaped early plausibility assessments:

> *"That's very funny. She was prescribing everyone to use lemon juice to make something basicity. That's ridiculous!"* (OZH, P10, L21)

Here, recognition was based on a perceived violation of prior knowledge rather than formal evidentiary checking. The participant relied on domain familiarity to rapidly assess implausibility, reducing the need for further information search.

Participants also described situations in which initially credible information later appeared inconsistent with their expectations or understanding of the topic. For instance, one lower-exposure participant stated:

> *"I read something on Instagram ... people ask me if it's true or not because of COVID... it was first very reliable, and had a reliable source. I checked the main website, and I read the article in the main language, and when I read that, it didn't make sense to me. I don't remember the whole idea, but I wasn't convinced that it has a bad effect on the immune system."* [OSN, P2, L25].

In this case, initial credibility cues led the participant to treat the message as reliable and worthy of further examination. However, as the evaluation progressed, attention shifted toward an internal inconsistency, which changed which features of the message were treated as important. This illustrates that cue recognition is not static but can be revised during the encounter as conflicting information becomes apparent.

Recognizing persuasive and attention-directing presentation cues

Participants with lower-exposure attended to repetition and high visibility as cues of persuasive intent, which shaped recognition prior to content evaluation:

> *"Because if it's over-advertised and it gets into all the media and all the platforms, into the TV commercials over and over again, it looked to me like they are going to impress me. They simply want to attract my attention."* OMH4, P7, L23

In this case, saturation itself functioned as a cue signalling persuasive intent. Rather than leading directly to acceptance or rejection, the message was interpreted as attempting to attract attention, which shaped how it was subsequently evaluated.

Higher-exposure participants also attended to design and presentation features, such as polished and engaging formats, as cues that shaped attention and the initial framing of the message:

> *"In the incident involving the hydroxychloroquine misinformation, the format of the information was critical in shaping perceptions and response. The video was presented in a highly engaging format... I think this design consideration is crucial for ensuring inclusivity and reaching diverse viewers... these design elements can be pleasant to users, but they can also be used to manipulate perceptions and influence behavior."* (MAY, P3, Line 1)

In this case, the highly produced format functioned as a recognition cue, influencing how the message was initially approached. Rather than serving as a neutral background, design features shaped early interpretation before participants evaluated the credibility or intent of the message.

Recognizing simplified and reduced-effort formats

Participants also attended to summary and simplified formats that appeared to reduce effort or time during information processing, shaping recognition at first contact:

> *"People are not interested in thorough analysis and investigations. How many do you know that might be willing to review ten articles in order to understand the truth? So, these misinformation channels, under the name of generating the abstract and the summary of the original article... start feeding people what they want... and pretend to make the analysis faster and easier for them."* (SS, P15, Line 7)

In this case, features signalling reduced effort functioned as recognition cues, directing attention toward information that appeared quicker to process. This shaped what participants initially treated as plausible before further examination was considered.

Recognizing institutional credibility cues

Higher-exposure participants treated recognized forms of evidence, such as peer-reviewed literature, as institutional credibility cues that guided information retrieval in technically uncertain domains:

> *"It's scientific literature... peer-reviewed scholarly scientific literature is what essentially, I ended up having to follow for a while."* (MAN, P13, L1)

In this case, peer review functioned as a recognition cue signalling legitimacy and directing attention toward specific sources. This indicates that recognition was anchored in institutional markers of credibility rather than surface plausibility judgments alone.

Taken together, these accounts show that early recognition was structured by recurring cues identified across participants, including missing or weak credibility cues, inconsistencies with prior knowledge or expectations, persuasive and attention-directing presentation cues, simplified and reduced-effort formats, and institutional credibility cues such as peer-reviewed literature. While the specific cues that became salient varied by incident, they performed a similar function at first encounter: shaping how the situation was initially framed and directing attention toward particular aspects of the message.

Across participants, the same types of cues were observed, but differences emerged in how they were interpreted and acted upon, particularly between higher-exposure and lower-exposure participants. These variations did not reflect different types of cues, but different orientations toward them at the point of recognition. These findings establish cue recognition as a structured and recurring process that shaped how situations were initially understood, forming the basis for subsequent expectancy formation and evaluation processes examined in the following sections.

3.5.2 Expectancy Formation

Following initial cue recognition, participants described forming expectancies that structured how they approached further evaluation. These expectancies reflected which types of information would be relevant, where to look for them, whether additional verification might be needed, and when understanding was sufficient to proceed. These patterns were observed across participants, indicating that expectancy formation functioned as an intermediate process shaping the direction and scope of subsequent evaluation without directly determining outcomes.

The analysis in this subsection is organized around the sub-themes identified in the coding table: domain familiarity, recognized presentation and rhetorical patterns, situational urgency, beneficiary structures and persuasive intent, audience fit and engagement orientation, and limits

of factual correction. These sub-themes show how participants formed expectancies about what information would be relevant, whether further verification might be needed, and how subsequent evaluation was anticipated to unfold. Participant excerpts are used to illustrate how these expectancy patterns appeared across higher-exposure and lower-exposure accounts.

Domain familiarity

Lower-exposure participants formed expectancies based on prior domain familiarity, which shaped expectations about whether further verification or information search would be needed:

> *"For example, in medical research, information is already there. So, I can have my own judgment. But yeah, if it's something political, then I might need to check."* (SAN, P1, L28)

In this case, familiarity with the domain reduced the perceived need for additional verification, whereas uncertainty led to expectancies that additional verification would be necessary. This indicates that expectancy formation guided participants' approach to the evaluation task, rather than whether they engaged with it.

Recognized presentation and rhetorical patterns

Participants formed expectancies about recurring presentation and rhetorical patterns, which shaped what was treated as suspect at early stages of evaluation. For instance, one lower-exposure participant stated that:

> *"That means a specific, very same design of presentation. That's what you're looking for. If so, that's fake."* (AL, P11, L21)

In this case, repeated or standardized formats functioned as recognition cues associated with expectancies of inauthenticity, narrowing attention toward familiar persuasive templates. This reduced the perceived need for further examination, as the pattern itself was treated as indicative of misinformation. These expectancies appeared to be learned through prior exposure rather than derived from explicit analytic reasoning.

Experienced participants formed expectancies based on recognized argumentative and rhetorical strategies, such as cherry-picking, which shaped how subsequent evidence was interpreted:

> *"The post cherry-picked data... while ignoring the overwhelming body of scientific evidence... these are common strategies used in misinformation to manipulate perceptions."* (RJ, P2, Line 23)

In this case, recognition of a familiar rhetorical strategy shaped the participant's expectancy that the message was constructed to mislead, directing attention toward structural features of the

argument rather than isolated claims. This narrowed the range of plausible interpretations at early stages of evaluation.

Situational urgency

Participants also formed expectancies regarding situational urgency, which shaped how further evaluation was anticipated to proceed. For example, one of the lower-exposure participants noted:

> *"If there is urgency that I'll have to take immediate action… then I'd rather call support and ask help from someone who is more familiar… but if I have enough time… then I would like to do it myself."* (ZAH, P3, L23)

In this case, urgency shaped expectancies about the depth and direction of further evaluation. Immediate action oriented the participant toward seeking guidance from others, whereas available time supported more independent exploration. This indicates that expectancy formation was shaped not only by prior knowledge but also by situational constraints that influenced what forms of evaluation were considered feasible.

Beneficiary structures and persuasive intent

Experienced participants formed expectancies about persuasive intent and who might benefit from the situation, which shaped how they approached further evaluation:

> *"When something happens… and the underlying cause is not found yet, then we must go and see who or which entities are benefiting from that situation."* (SS, Pp 10, Line 31)

In this case, the participant anticipated that the message served particular interests, directing attention toward identifying who might benefit. This shifted the focus of evaluation away from surface features toward underlying patterns of gain and influence, constraining how the situation was interpreted before formal verification took place.

Audience fit and engagement orientation

Participants also formed expectancies based on audience fit and perceived personal relevance, which shaped how they oriented to the message at the point of encounter. For instance, one lower-exposure participant stated:

> *"So I'm thinking if more people like me, look more familiar like me, and many people I know were also in the media… probably I wouldn't believe… this advertisement isn't what I need to change myself. I wouldn't hesitate and I would just skip it easily."* (OMH4, P4, L28)

In this case, the absence of perceived similarity shaped the participant's expectancy that the message was not intended for them, narrowing engagement at the point of encounter. This

indicates that expectancies about "*who the message is for*" influenced how participants oriented to the information before further evaluation of the message content occurred.

In contrast to earlier accounts in which expectancies constrained engagement early, some participants maintained open expectancies, allowing continued evaluation as new information became available:

> "*I would read the entire article or post, or watch the entire commercial. I tell myself that if I go through it until the end, there might be something that shows it doesn't suit me… or, on the other hand, something that makes it more suitable for me…*" (OMH1, P1, L24)

In this case, the lower-exposure participant maintained an expectancy that additional information could alter how the message was interpreted, supporting continued engagement rather than narrowing the scope of evaluation early in the encounter. This indicates that expectancies can remain open-ended, shaping how participants sustain engagement during evaluation.

Limits of factual correction

Higher-exposure participants formed expectancies about the limits of factual correction, particularly in relation to belief-driven content. For instance, one participant noted:

> *"You can get people… to fact check nonsense… but that does not work to challenge their overall view… it's a limited tool."* (MAN, P24,L4)

In this case, the participant anticipated that fact-checking would address specific claims but not broader belief systems. This shaped how further evaluation was approached, orienting attention toward the limits of what fact-checking was expected to achieve rather than relying on further verification alone.

Across participants, differences emerged in how expectancies oriented subsequent evaluation. Participants without formal training formed expectancies grounded in domain familiarity, situational conditions, and perceived audience relevance, which shaped how they approached further evaluation. In contrast, participants with experience in misinformation analysis formed expectancies oriented toward structural and rhetorical features of the information environment, including incentive structures, recognized manipulation strategies, and perceived limits of factual correction. These differences indicate that experience influenced how expectancy formation directed subsequent evaluation.

Across accounts, expectancy formation was structured by recurring sub-themes identified in the analysis, including domain familiarity, recognized presentation and rhetorical patterns, situational urgency, beneficiary structures and persuasive intent, audience fit and engagement orientation, and limits of factual correction. These expectancies functioned as working

orientations that shaped how participants approached further evaluation. In some cases, expectancies constrained inquiry early, directing attention toward particular cues or interpretations. In other cases, expectancies remained open, supporting continued engagement as additional information was considered. These variations were observed across participants and reflected differences in how prior experience and situational conditions shaped expectancy formation and later evaluation processes, rather than simply differences in the information encountered. Together, these findings show that expectancy formation structured how evaluation was approached at this stage, establishing the conditions under which subsequent mental simulation occurred, as discussed in the next section.

3.5.3 Mental Simulation and Projection of Consequences

Following expectancy formation, participants described engaging in mental simulation to project how a situation might unfold, based on the information available and the expectancies already formed. These simulations were observed across participants but appeared selectively, particularly when uncertainty persisted or potential consequences were salient. Through mental simulation, participants explored possible interpretations of the situation and anticipated the implications of different courses of action or inaction.

The analysis in this subsection is organized around the sub-themes identified in the coding table: simulating intervention consequences, constructing explanatory and causal scenarios, simulating endorsement and audience response, simulating future trajectories and uncertainty, and anticipating persuasive intent and framing. These sub-themes show how participants projected possible outcomes, explanations, and future developments while interpreting uncertain situations. Participant excerpts illustrate how different forms of mental simulation appeared across higher- and lower-exposure accounts.

Simulating intervention consequences

Higher-exposure participants simulated the consequences of intervening, correcting, or challenging misinformation, considering both intended outcomes and potential risks. For instance, one participant stated:

> "*I imagined that if the accurate information [was] being communicated effectively to the target audience, this would result in the misinformation being corrected or retracted... I imagined that the source of the misinformation might resist or backlash against my actions... This could lead to consequences for me, like reputational harm or professional repercussions*." (ERJ, P6, L3; P6, L6)

In this case, mental simulation extended beyond assessing factual plausibility to include anticipated social and professional consequences. Rather than producing a single outcome, the

participant constructed multiple possible pathways, which shaped how they weighed the value of engaging against the risks of resistance, backlash, or professional harm.

A similar pattern appeared when participants considered downstream consequences of responding to misinformation, incorporating public health outcomes, audience reception, and potential resistance:

> "*I anticipated that by communicating accurate information about hydroxychloroquine, I could potentially prevent harm… This could potentially lead to increased trust and credibility among the public… I also considered the possibility that… I could potentially encourage them to retract or correct the misinformation… [but] I anticipated that some people might be skeptical or resistant*." (EMA, P9, L5)

In this case, the participant considered multiple possible outcomes, including preventing harm, increasing trust, partial correction, and resistance. Rather than assuming a single trajectory, the simulation explored alternative futures under uncertainty rather than functioning as a mechanism for reaching a definitive conclusion.

Constructing explanatory and causal scenarios

Higher-exposure participants also simulated explanatory frameworks by drawing on recognized narrative structures such as conspiracy patterns:

> "*As soon as I see that people talking about it… it sounds like a conspiracy theory*." (EMS, P5, L21)

In this case, the participant projected a familiar explanatory frame involving hidden actors and intentional causes, using it to interpret the claim at an early stage. Rather than evaluating isolated facts, the simulation operated at the level of narrative coherence, shaping how the information was understood.

Participants also simulated causal scenarios by reconstructing chains of action and consequence, often focusing on who might benefit from a given situation:

> "*We must go and see who or which entities are benefiting from that situation… By looking into… beneficiary entities… then you may develop a hypothesis… Similar to police investigations about a criminal case*." (ESS, P10, L34)

In this case, the participant generated hypotheses about causation and intent by tracing potential beneficiaries and consequences. The simulation operated as a form of investigative reasoning, allowing alternative explanations to be explored without committing to a single interpretation.

Rather than focusing on surface claims, the participant used projected causal relationships to narrow the range of plausible explanations.

Simulating endorsement and audience response

Lower-exposure participants also engaged in mental simulation, often grounded in social endorsement and perceived popularity:

> "*I assume if that good looking and gorgeous actress is actually using this product, it must be really good… Then I shared this with some of my friends*." (OMH, P7, L24)

In this case, the participant projected endorsement-based cues into anticipated personal and social outcomes, treating perceived popularity as indicative of value. The simulation operated by extending observed endorsement into expected use and reception, shaping early acceptance and prompting sharing before detailed evaluation occurred.

Simulating future trajectories and uncertainty

Other lower-exposure participants also simulated future scenarios when engaging with economic or political information, exploring multiple possible developments and their implications:

> "*I personally imagine that after they couldn't control prices and the economy, everything is going to explode. The government will say the other country didn't agree and that war stopped negotiations. On the other hand, they could agree and release money, but even in the best scenario, prices would stay the same. That's why I decided to exchange.*" (OAL, P16, L10)

In this case, the participant constructed alternative future trajectories under uncertainty, comparing different possible developments and their consequences. Rather than resolving which outcome was most likely, the simulation allowed the participant to orient toward possible consequences across alternative scenarios without requiring immediate certainty.

Another lower-exposure participant also simulated how interpretations of information might change over time as new perspectives emerged:

> "*Based on my experience, I think in the first days of releasing that news, the titles are very similar. But after months or years, we see the news from other angles of the story. I usually leave it for later… there might be other things*." (OMN, P10, L16)

In this case, the participant projected how the information environment might evolve, expecting that additional information or perspectives would become available over time. Rather than resolving the situation immediately, the simulation extended the temporal horizon of evaluation, allowing uncertainty to be maintained while alternative interpretations remained possible.

Anticipating persuasive intent and framing

One lower-exposure participant also simulated how political information is typically framed, drawing on experience-based expectancies about persuasive intent:

> "*Whatever comes from politics or politicians is more likely framed, and I'm unable to accept it right away. Based on my experience, it's always framed to converge people's beliefs toward a specific direction. Publishing information neutrally is less likely*." (OZH, P4, L2)

In this case, the participant projected a communicative motive into the message prior to engaging with its content, anticipating that it was designed to influence belief. The simulation operated by extending prior experience with political messaging into an inferred persuasive motive, shaping how the information was interpreted at the outset and shifting evaluation toward communicative purpose rather than factual accuracy alone.

Across accounts, mental simulation was structured by recurring sub-themes identified in the analysis, including simulating intervention consequences, constructing explanatory and causal scenarios, simulating endorsement and audience response, simulating future trajectories and uncertainty, and anticipating persuasive intent and framing.

These simulations functioned as mechanisms for exploring uncertainty by supporting an evolving understanding of "*what is happening*" and "*what follows*." Participants instantiated explanatory accounts that could make the message reasonable (e.g., conspiracy-based narratives or beneficiary-driven causation), while also anticipating the consequences of possible actions or inaction.

Rather than producing a single outcome, simulation involved constructing multiple potential trajectories, allowing expectancy-constrained interpretations to be stress-tested against imagined outcomes and risks. Variation was observed across participant groups: higher-exposure and experienced participants more often simulated causal, explanatory, and consequence-based scenarios, whereas lower-exposure participants more frequently relied on endorsement, narrative familiarity, and future-oriented projections.

Together, these findings show that mental simulation shaped how situations were interpreted by linking what was initially recognized to what was projected as plausible and consequential, thereby setting the conditions for different decision pathways. This process precedes plausibility judgment and stopping, which are examined in the following section.

3.5.4 Plausibility Judgments and Stopping Points

Following recognition, expectancy formation, and mental simulation, participants described points at which they no longer continued evaluating the information and instead moved toward a decision, action, withholding trust, or disengagement. These stopping points were not defined by complete certainty. Rather, participants described recognizing that they had sufficient understanding to proceed, based on how the situation had been framed through prior stages of sensemaking.

The analysis in this subsection is organized around the sub-themes identified in the coding table: stopping through early recognition and framing, stopping after limited or targeted verification, stopping through contradiction detection, and stopping calibrated by context (time, importance, and knowledge). These sub-themes show how participants reached points at which further evaluation was no longer considered necessary, understanding was considered sufficient, or trust was withdrawn. Participant excerpts illustrate how different stopping orientations appeared across higher-exposure and lower-exposure accounts.

Stopping through early recognition and framing

For some lower-exposure participants, stopping occurred once the message's purpose or intent was recognized:

> "*So after you realize the purpose of that news or that organization's kind of misinformation, you can decide.*" [OAL, P13, L22]

In this case, stopping was tied to reframing the situation in terms of communicative intent. Once the participant interpreted the message as serving a particular purpose, further verification was no longer pursued, and the evaluation was considered sufficient to proceed.

Stopping could also occur rapidly when prior knowledge enabled immediate recognition of inconsistency. Familiarity allowed participants to detect inconsistencies and stabilize their judgment without extending evaluation:

> *"Because you were familiar with this and had prior knowledge about it, it looked familiar, and that's how you found out it might not be correct."* [OZH, P10, L28]

In this case, familiarity with the topic allowed the lower-exposure participant to recognize inconsistency early, reducing the need for further checking or extended analysis.

In contrast, when cues were not interpreted as manipulative, lower-exposure participants described concluding the evaluation early. The absence of heavy promotion reduced suspicion, leading to early acceptance without further checking:

> *"But if I didn't see those commercials, or if it wasn't over-advertised before, I probably wouldn't search any further, I wouldn't hesitate or ask anyone else. I'd assume I discovered this new product… and I could possibly buy it, I could possibly believe it."* [OMH3, P4, L16]

Rather than triggering scrutiny, this cue supported a sense of independent discovery, allowing participants to assume the product was credible and proceed with further consideration.

Stopping after limited or targeted verification

Other lower-exposure participants described stopping after limited but targeted verification, in which a small number of checks were conducted before evaluation was considered sufficient:

> *"The main thing is to know what you're actually taking. I didn't believe that website, so I checked the main source. I recommend going to the original source, thinking with your own mind, and searching a little more to judge how reliable it is… I checked a few times, and then I decided that it was enough."* [OSAN, P10,L10]

In this case, stopping was not tied to completeness, but to a self-directed judgment that further checking was no longer required after several confirmations.

Participants with experience in misinformation analysis described stopping based on identifiable credibility criteria. Rather than relying on familiarity alone, they evaluated features such as citations, sources, and internal clarity, and once these aligned with expected standards, they described not continuing further checking:

> *"For me, it's the body of the article, the people cited and the resources they use. These are the factors that matter. If I see that the citations are correct and the resources are reliable, then I wouldn't doubt it… So if the citations are precise and reliable, and the details are straightforward, no hidden or unclear points, then you could believe it right away"* [EMS, P23, L1]

Here, stopping was tied to the presence of recognized indicators of credibility. Once these were satisfied, the participant did not continue further checking.

Stopping through contradiction detection

Conversely, the detection of contradictions could also lead to rapid stopping, not in the direction of belief, but in the withdrawal of trust:

> *"Not much… once you notice contradictory details, you realise it could be misinformation. There is still uncertainty; you don't fully know what's going on. You can't be completely certain, but it doesn't take long to lose trust in that source. At that point, you conclude it's not just a typo or a temporary mistake, there is something going on, and you shouldn't trust them as before."* [ESS, P12, L7]

Here, participants explicitly acknowledged that uncertainty remained. However, this did not prevent the participant from reaching a stopping point in terms of how the source would be treated.

Stopping calibrated by context (time, importance, knowledge)

Time pressure and urgency also shaped stopping. Participants described situations where evaluation was compressed due to the need to act quickly, particularly in contexts involving potential harm:

> *"Given the ... urgency of the situation regarding the spread of misinformation about hydroxychloroquine during the COVID-19 pandemic, the time taken... was relatively short. The process of evaluating credibility and accuracy... consulting with trusted sources and considering ... different courses of action likely took a few hours at most. This was due to ... the need to respond quickly... to prevent the spread of harmful misinformation... While the decision-making process was relatively quick, it was also thorough and informed by my training and experience..."* [EMAY, P10, L36]

This illustrates that stopping does not necessarily imply superficial evaluation. Under urgent conditions, participants reduced the duration of evaluation while still drawing on structured approaches informed by experience.

Participants described stopping as dependent on the interaction between importance, knowledge, and available information. Rather than a fixed rule, stopping varied across situations:

> *"I think it depends on how important it is and how much knowledge they bring. For low importance and high knowledge, it shouldn't take more than a few seconds... for high importance and low knowledge, that's where it takes a long time... and when there isn't enough data or there's disagreement, that's also where it takes time. But if there is widespread consensus and good data... where the cause is quickly understood and everyone agrees, you can move really quickly."* [EMAN, P25, L31]

This account shows that stopping varied depending on the situation, particularly in relation to the importance of the issue, the participant's knowledge, and the availability of consistent information.

Across these accounts, stopping for lower-exposure participants did not depend on exhaustive evaluation but on reaching a point where the situation was interpreted as sufficiently understood through initial framing. Whether through recognizing intent, prior knowledge, cues suggesting authenticity, or limited checking, participants relied on early interpretations to conclude their evaluation. As a result, further search was often reduced or not initiated, with decisions emerging from how the situation was first understood rather than from extended verification.

Across these accounts, higher-exposure participants reached stopping points through structured and context-sensitive evaluation rather than immediate framing. Decisions to proceed, reject, or defer were grounded in identifiable criteria, detection of inconsistencies, and calibration to situational demands such as urgency, importance, and available evidence. While uncertainty was often acknowledged, it did not prevent action; instead, participants judged when further inquiry was no longer necessary. In this way, stopping emerged not as a fixed threshold, but as a calibrated judgment based on the structure and conditions of the information being evaluated.

Taken together, these findings show that divergence in decision pathways arises not only from what participants notice or how they interpret information, but from when they judge that they have a plausible understanding to proceed. Across the analysis, stopping was structured by recurring patterns, including stopping through early recognition and framing, stopping after limited or targeted verification, stopping through contradiction detection, and stopping calibrated by context such as urgency, importance, knowledge, and available evidence. Even when encountering the same message, participants brought evaluation to a close at different points, depending on how the situation had been constructed through recognition, expectancies, and projected consequences. These differences in stopping shape whether individuals continue searching, act, defer, or disengage, producing the varied decision pathways observed across participants.

Table 11 synthesizes the recurring patterns identified across participants, showing how each process stage was expressed across higher-exposure and lower-exposure accounts. The table provides a cross-case summary of how the process-level themes and sub-themes were reflected in the analysis.

Table 11 - Comparison of evaluation processes across higher-exposure and lower-exposure participants

Process stage	Higher-exposure participants	Lower-exposure participants	Shared pattern
Cue recognition	Attended to missing or weak credibility cues, internal coherence, persuasive presentation features, simplified formats, and recognised credibility markers, including institutional cues, often treating these as triggers for targeted verification or source-directed search	Attended to missing source information, inconsistencies with prior knowledge, repetition, high visibility, source appearance, familiarity, and domain knowledge	Both groups noticed cues early and used them to frame the situation, but differed in how similar cues were interpreted and acted upon
Expectancy formation	Formed expectancies around recognised rhetorical patterns, beneficiary structures, persuasive intent, evidence quality, and the limits of factual correction	Formed expectancies around domain familiarity, situational urgency, audience fit, personal relevance, and whether further checking was needed	Both groups used prior experience and situational cues to anticipate what mattered, what further evaluation might require, and how the situation should be approached
Mental simulation	More often simulated intervention consequences, downstream effects, explanatory and causal scenarios, resistance, and broader public or professional implications	More often simulated endorsement effects, audience response, future trajectories, evolving interpretations, and persuasive framing	Both groups projected possible outcomes when uncertainty remained, using simulation to explore plausible interpretations and consequences
Stopping points	More often stopped based on identifiable credibility criteria, reliable sources or citations, contradiction detection, urgency, importance, knowledge, and available evidence	More often stopped through early recognition and framing, communicative intent, prior knowledge, authenticity cues, or limited checking	Both groups stopped before complete certainty, but differed in what was treated as sufficient understanding for proceeding, withdrawing trust, deferring, or disengaging

The following section integrates these patterns to explain how misinformation evaluation unfolds as a process.

3.6 Discussion

This section brings the findings together to show how misinformation evaluation unfolds as a process. It examines how these patterns combine, how evaluation is shaped by constraints, where divergence emerges through stopping, and how susceptibility can be understood within this process.

3.6.1 From Observed Patterns to Process-Level Understanding

The findings show that participants' accounts of evaluating misinformation are not explained as isolated behaviours such as checking sources, recognizing cues, or applying critical thinking skills. Instead, the data indicate that these activities are organized as a process of sensemaking under constraint, where recognition, expectancies, and projection operate together to

progressively narrow the situation until it becomes sufficiently interpretable for further judgment or response. This is reflected in participants' accounts of constructing an explanation of what is happening before deciding how to proceed (e.g., [ESS, P10, L34]). The framework presented in Section 3.4 emerged from recurring patterns across incidents, reflected in how participants moved from initial cue recognition to expectancy formation, simulation, and stopping. What becomes evident is that participants were not simply verifying information. Rather, verification was often selective and bounded, with participants deciding when further checking was no longer necessary. For example, as reflected in the findings (e.g., [OSAN, P10, L10]), one participant described checking "*a few times*" before deciding it was "*enough*," indicating that evaluation was not driven by completeness but by a judgment of sufficiency. They were constructing a workable understanding of the situation, using available cues, expectations, and projections to reduce ambiguity to a level that allowed them to proceed.

3.6.2 Evaluation as Constraint-Driven Sensemaking

Across the findings, the evaluation did not follow a linear or exhaustive path. Instead, it was shaped by how participants constrained the information space early and progressively. Initial recognition narrowed attention, expectancy formation directed search and interpretation, and mental simulation explored plausible consequences where uncertainty remained.

The findings also suggest that expectancies were not always fixed prior assumptions, but could function as working expectations that could vary during evaluation. Participants often began with an initial orientation shaped by prior knowledge, familiarity, or perceived intent, which guided what they attended to and how they interpreted the message (e.g., [SAN, P1, L28]; [AL, P11, L21]). However, in some cases, expectancies remained open and were adjusted as additional information was encountered, particularly when participants anticipated that later content might confirm or contradict their initial interpretation (e.g., [OMH1, P1, L24]).

This should be treated cautiously, as the present study does not provide quantitative evidence of expectancy revision over time. However, the CDM accounts suggest that expectancy formation operated as an adaptive process: initial expectations guided attention and search, while ongoing interpretation shaped whether the information was treated as sufficient, contradictory, or requiring further evaluation.

This helps explain why participants often engaged with similar cues but diverged in how those cues affected how they proceeded across episodes. This variation is reflected in cases where comparable cues led to immediate dismissal in some instances and further evaluation or revision in others (e.g., [OZH, P10, L21]; [OSN, P2, L25]). The difference was less in what was noticed and more in how the situation was framed and constrained as it evolved. Expectancies played a

central role here, guiding what counted as sufficient evidence, what sources were considered relevant, and whether further checking was necessary. This is reflected in the findings (e.g., [SAN, P1, L28]), where domains of existing knowledge supported immediate judgment, while uncertain domains prompted additional verification.

Mental simulation further shaped how participants narrowed the situation. This is reflected in participants' accounts of imagining how situations might unfold, for example, anticipating that "*everything is going to explode*" or considering alternative outcomes before deciding how to proceed (e.g., [OAL, P16, L10]). Rather than resolving uncertainty through additional data alone, participants projected consequences, inferred intent, and matched emerging interpretations to familiar explanatory patterns. For example, they imagined that sharing accurate information could correct misinformation but also lead to "*backlash*" or "*reputational harm*" (e.g., [ERJ, P6, L3; P6, L6]). In doing so, they transformed incomplete information into a plausible account of "*what is happening*" and "*what follows*."

Taken together, these patterns show that evaluation is not reducible to verification behaviours. It is a context-sensitive process of constructing plausibility, shaped by how individuals progressively limit and organize the situation they are interpreting.

3.6.3 Stopping as the Source of Divergence

A central finding of this study is that divergence in outcomes: acceptance, rejection, continued assessment, or deferral arises not only from differences in cues or interpretations, but from differences in when participants decide to stop evaluating.

Across accounts, participants did not seek complete certainty. Instead, they stopped when the situation felt sufficiently interpretable relative to their goals, knowledge, and constraints. As reflected in the findings (e.g., [ESS, P12, L7]), participants often proceeded despite acknowledging they "*can't be completely certain*," indicating that evaluation did not aim for full certainty. Stopping was instead calibrated to situational conditions (e.g., [EMAN, P25, L31]), including importance, knowledge, and available information. This threshold varied. In some cases, recognition of intent or inconsistency was enough to stop. In others, limited verification or structural indicators of credibility were enough to proceed (e.g., where "*citations are precise and reliable… you could believe it right away*" [EMS, P23, L1]). Under conditions of uncertainty, participants either continued assessment or deferred judgment depending on perceived importance and urgency.

This reframes susceptibility to misinformation within the evaluation process observed here. Rather than being explained solely by lack of knowledge, bias, or failure to detect cues, susceptibility emerges from how individuals determine that "*plausible*" understanding has been

reached. Two individuals may encounter the same message, recognize similar cues, and even interpret them in comparable ways, yet arrive at different outcomes because they reach stopping points at different stages of the process.

The framework, therefore, highlights stopping as a decision mechanism within evaluation, rather than a passive endpoint. This is the mechanism through which belief, skepticism, deferral, or disengagement become actionable outcomes within evaluation.

3.6.4 Reframing Susceptibility to Misinformation

These findings suggest that susceptibility to misinformation is better understood as an emergent property of the evaluation process, rather than as a fixed individual trait or a direct effect of message content. In this view, persuasive framing does not directly determine belief, but shapes the cues that enter the evaluation process and guides how they are interpreted. The findings suggest a process-level explanation of how persuasive framing may become consequential. Framing and persuasion models can shape which cues are foregrounded, which interpretations appear familiar or coherent, and which outcomes are mentally simulated. However, these design features do not operate independently of the evaluator. Their influence depends on how they are recognized, incorporated into expectancies, tested through projection, and treated as plausible for judgment. In this sense, persuasive misinformation may shape how susceptibility emerges within evaluation by influencing the recognition-based pathway through which a plausible interpretation is constructed and brought to a close.

Participants did not simply accept or reject information based on what was presented. As reflected in the findings (e.g., [OAL, P13, L22]), decisions were often made after interpreting the purpose or intent of the message. Outcomes were shaped by how recognition, expectancies, and projections interacted to produce a plausible account, and how early or late this account was brought to a stopping point. This is reflected in participants' accounts of judging plausibility based on how information "*looked familiar*" in relation to prior knowledge (e.g., [OZH, P10, L28]). This helps explain how the same message can lead to different outcomes even when individuals engage with similar cues.

Importantly, this understanding shifts attention from identifying who is susceptible to understanding how susceptibility is produced within the evaluation process itself. It also highlights why additional information does not always change the outcome of evaluation. As reflected in the findings (e.g., [OMN, P10, L16]), participants sometimes deferred further consideration, noting they would "*leave it for later... there might be other things*," rather than revisiting the information immediately. If individuals stop once the situation reaches a plausible interpretation, later information may not be incorporated unless it appears before that point.

3.6.5 Implications for Understanding Misinformation Evaluation

By organising the findings as a process, this study, rather than focusing on detection accuracy or individual differences in belief, provides a process-level account of how evaluation unfolds under real-world conditions of uncertainty and constraint.

This perspective makes two contributions: first, it shows that variability in outcomes can arise from differences in process dynamics rather than from differences in message content alone; second, it identifies stopping as a critical mechanism that determines whether evaluation continues, stabilises, or is deferred.

Together, these insights position misinformation evaluation as a form of naturalistic decision-making process, where individuals rely on recognition, expectancies, and projection to construct actionable understanding in complex environments.

3.7 Conclusion

Building on the Recognition-Based Susceptibility Framework developed in Chapter 2, this chapter investigated how misinformation is evaluated in practice through participants' accounts of real encounters and decision processes using CTA and the CDM technique. The findings showed that evaluation unfolds as a recognition-based process of plausibility construction shaped by cue recognition, expectancy formation, mental simulation, and stopping under conditions of uncertainty and constraint. The chapter demonstrated that susceptibility emerges within evaluation itself through how individuals interpret information, construct plausible understandings, and determine when evaluation is sufficiently complete. The next chapter shifts attention to the construction of misinformation itself by investigating how factual elements can be reframed into persuasive narratives that guide interpretation and shape what appears plausible during evaluation.

Chapter 4: How Misinformation is Constructed

Building on the process account of evaluation developed in Chapter 3, in this chapter, this chapter develops an analytic framework to model how factual elements can be transformed into misinformation. This method is termed the “Recipe Book of Misinformation” (RBM). RBM is a conceptual framework that offers an illustrative scaffold for exploratory analysis. While most existing studies focus on detecting misinformation and tracing their origins, they offer limited explanations of the underlying design logic that makes reframed facts appear plausible and persuasive. This chapter shows how factual elements are reframed into persuasive misinformation narratives that guide interpretation and support particular plausible understandings. The RBM formalizes this logic by showing how a single factual element can be reframed into multiple variants through cognitive, rhetorical, and framing mechanisms. The framework uses established persuasion models and rhetorical strategies to show how communicators might shape perceptions of credibility, emotion, and logic. Using an analytic modelling approach, the chapter applies Framing Theory by operationalizing how dominant meaning forms through controlled variation in placement, repetition, and emotional intensity. Analytic exemplars from COVID-19 vaccines, climate change, and the Ukraine biolabs illustrate how the same factual elements can elicit different interpretations under different framing conditions. Situated in NDM and the RPD model, the RBM links framing patterns to cue-based reasoning pathways, showing how susceptibility indicators influence plausibility judgements under uncertainty. The same generative structure can be inverted to model resilience strategies and anticipatory training simulations. This chapter introduces a cognitive–narrative framework linking persuasive message construction to interpretation.

4.1 Introduction

Misinformation research involves human judgment, persuasive communication, and socio-technical systems. In contemporary digital environments, misleading content circulates as authentic information, influencing public interpretation, trust, and action. Crucially, such content does not need to be fabricated to be effective: factually correct elements can be selectively framed to produce misleading yet plausible interpretations (Wardle & Derakhshan, 2017). Understanding how plausibility forms under these conditions is an important concern for research on misinformation and sense-making.

Existing work has made substantial progress in detecting misinformation, mitigating its spread, and identifying factors associated with susceptibility (Danry et al. , 2025; Epstein et al. , 2020; Kamali et al. , 2024; Yaqub et al. , 2020; Zhou et al. , 2023). However, these approaches largely treat misinformation as a fixed input to be evaluated, filtered, or corrected. What remains under-

specified is how misleading messages are constructed to be effective for particular individuals: specifically, how factual elements are reframed in relation to the cognitive mechanisms through which they come to appear plausible under uncertainty, time pressure, and sense-making. In this chapter, a factual element refers to a discrete, verifiable statement describing the same underlying real-world state or event, which may be selectively emphasised or positioned differently without altering its referential accuracy.

This chapter addresses this gap by introducing an analytic framework that formalises the transformation of factual elements into persuasive misinformation, oriented toward specific cognitive mechanisms. This chapter contributes an analytic framework that clarifies how factual elements can be reframed into persuasive misinformation. It illustrates how a single factual element can support multiple misleading narratives through systematic variation in framing and rhetorical orientation. Drawing on established persuasion models (Fishbein & Ajzen, 2011; Friestad & Wright, 1994; Petty & Cacioppo, 1986) together with susceptibility indicators (Nan et al., 2022), the framework characterizes how different message designs may engage different plausibility pathways. It further demonstrates how these design choices can be analyzed consistently across domains while remaining anchored to the same underlying facts.

The framework is positioned within the Naturalistic Decision-Making (NDM) paradigm (Klein & Zsambok, 1997), using the Recognition-Primed Decision model (RPD) (Klein, 1993) as an analytic lens. This positioning is motivated by the fact that misinformation is commonly encountered under conditions of uncertainty, constraint, and competing goals, where interpretation relies on cue recognition and sense-making rather than exhaustive evaluation. The framework aligns message-level design choices with these recognition-based processes without presuming intent or outcome.

The chapter adopts an analytic and conceptual orientation. It focuses on clarifying the design logic through which factual elements can be reframed into persuasive misinformation by aligning message construction with cognitive mechanisms of interpretation. Worked examples illustrate how message-design choices shape salience, interpretation, and perceived credibility under uncertainty.

The remainder of the chapter proceeds in four stages. Section 2 outlines the theoretical foundations relevant to plausibility, persuasion, and framing. Section 3 synthesises these elements into the RBM framework and presents its conceptual architecture. Section 4 demonstrates how the framework operates when applied to factually grounded cases. Section 5 concludes with a discussion of implications and limitations.

4.2 Foundations for the RBM Framework

Section 2 reviews prior work and theoretical perspectives that inform the development of the RBM framework. This section traces a progression from related studies on user judgment and mitigation cues, through accounts of individual susceptibility and theories of cue-based decision-making under uncertainty, to a review of persuasion models and rhetorical strategies used to shape interpretation. It then introduces framing theory and its structural components. Together, these perspectives inform how plausibility emerges through the interaction of cognitive mechanisms, persuasive communication, and framing, providing the foundation for the RBM framework.

4.2.1 Background: Misinformation, Plausibility, and User Judgement

During crises, when trust and clarity are critical for resilience, such as the COVID-19 pandemic, misleading claims contributed to public uncertainty and social tension. At scale or when deliberately orchestrated as disinformation, such content can erode trust, polarize communities, and threaten public well-being and security (Roozenbeek, Schneider, et al., 2020). Misinformation spread is shaped by socio-psychological dynamics such as identity reinforcement and social connection, amplified by digital platforms, and the ubiquity of digital media (Roozenbeek & van der Linden, 2019a, 2022; Roozenbeek, van der Linden, et al. , 2020). This chapter focuses on receiver-side plausibility assessment and cognitive framing rather than sender intent. I use misinformation as an umbrella label for misleading content encountered by individuals; intent is out of scope for the analytic demonstration.

In misinformation research, much emphasis has been placed on mitigation through detection, labeling, or ranking. Several studies show that manipulating decision-time cues can shift outcomes: credibility indicators reduce sharing intent for misleading content, with effects varying by indicator type (Yaqub et al., 2020), and layperson trust ratings can be used to downrank distrusted sources at scale (Epstein et al., 2020). These approaches explain how susceptibility manifests in rapid judgment, where labels act as interruptions and unfamiliarity serves as a proxy for distrust. However, they treat misinformation as a fixed input to be filtered, not as a designed artefact.

Recent works on generative AI show that the cue environment itself is evolving. AI-generated misinformation follows recurring patterns, such as enhanced detail, uncertainty framing, conclusion-drawing, and a personal tone, and is less reliably detected than human-authored misinformation (Zhou et al., 2023). Explanatory structure also plays a persuasive role: deceptive LLM explanations can shift beliefs more strongly than honest ones, depending on logical validity (Danry et al., 2025). Together, these findings suggest that credibility, coherence, and authority

cues can be actively shaped, and trust calibration becomes an ongoing challenge following AI errors (Martell et al., 2025).

Some other research engages with the message "structure," but typically through inventories or evaluative frameworks. Interview studies document recurring persuasive strategies as they are encountered and managed in practice (Peng et al., 2024), while argumentation-theory tools and NLP annotation schemes provide principled ways to assess such strategies or support detection (Rubinelli & Diviani, 2025; Kamali et al., 2024). Narrative and framing research similarly demonstrates how message construction shapes attention, visibility, and diffusion across domains (Sadler, 2025; Illia et al., 2026), and reviews of computational propaganda emphasise the evolving, campaign-like nature of persuasion systems (Mustafa et al., 2025).

Across these literatures, explanation typically stops at classification, evaluation, or mitigation outcomes. While susceptibility is often related to behavioural responses, or cue sensitivity, its role in the construction of persuasive messages remains underspecified. Existing work does not make explicit how message features are shaped in relation to the cognitive mechanisms underlying plausibility assessment. Evidence from human-factors and security research helps explain why this gap persists: cue utilisation and cognitive reflection predict performance in controlled tasks but weaken in more naturalistic settings, where judgments are shaped by situational constraints rather than individuals' stable traits (Morrison et al., 2025). Similarly, socio-technical studies show that vulnerability arises from interactions between individuals' characteristics and organisational, social, and technological contexts, without detailing how such characteristics are translated into systematic message-construction choices (Pollini et al., 2022; Deline et al., 2021).

4.2.2 Susceptibility to Misinformation: A Conditional Process

In general, susceptibility refers to the likelihood that an individual will be influenced or affected by an external stimulus or context (Cambridge University Press, n.d.). In social influence literature, susceptibility is operationalized as the likelihood that conformity will occur in response to an influence attempt (Laursen & Faur, 2022). In studies of misinformation, susceptibility is treated not as a fixed trait but as an outcome of interacting factors, such as biases, time pressure, prior knowledge, emotional, and motivational states (Maertens et al., 2023; Nan et al., 2022). These factors combine differently across individuals, producing distinct responses to the same message (Calvillo, 2014). Understanding susceptibility helps link human factors to the rhetorical and decision-making processes that misinformation exploits (Downs et al., 2006; van der Linden, 2022).

Nan et al., (2022) identify several susceptibility indicators in the context of misinformation: knowledge and skills, thinking style, trust, emotion, values, and group identity. These factors shape the criteria by which individuals attend to cues, the interpretations they find plausible, and the degree of scrutiny they apply. For example, limited topic knowledge can increase reliance on intuitive reasoning; strong identification with a group can heighten responsiveness to identity-congruent cues; and emotional arousal can narrow attention and reduce reflective processing (Nan et al., 2022). Together, these indicators represent cognitive and affective conditions that persuasive messages aim to exploit.

4.2.3 Decision-making Under Uncertainty: From Bounded Rationality to Recognition

Evaluating credibility typically occurs under uncertainty, time pressure, and cognitive or emotional load (Orasanu & Connolly, 1993; Orasanu & Fischer, 1997). Early work on bounded rationality showed that people do not optimize across alternatives, but instead seek interpretations that are good enough for action (Simon, 1957; Simon, 1997). Later dual-process accounts similarly distinguish between fast, intuitive reasoning (Type 1) and slower, deliberative evaluation (Type 2) (Kahneman, 2011), while Cognitive Continuum Theory (CCT) conceptualizes intuition and analysis as endpoints along a spectrum, with quasi-rational modes in between (Hammond, 1981; Hammond et al. , 1987). Together, these perspectives establish that plausibility judgments are context-dependent, cue-driven, and constrained. Yet they remain abstract with respect to how judgments unfold in real-world naturalistic settings.

Naturalistic Decision Making and Recognition-Primed Decision Model as an Analytic Lens

To investigate how plausibility assessment forms under real-world constraints, this chapter applies the NDM paradigm (Klein & Zsambok, 1997). NDM decision-making in ill-structured, time-pressured environments characterized by uncertainty, competing goals, and incomplete information, conditions that closely resemble everyday encounters with contested or misleading information.

Within NDM, the RPD model explains how people make decisions by recognizing salient cues, matching them to familiar patterns, and evaluating a plausible interpretation through rapid mental simulation (Klein, 1993). Rather than comparing multiple alternatives, decision-makers assess whether a recognized interpretation makes sense in context and adjust it if it does not. This recognition-based process accounts for how decisions stabilize under uncertainty and why certain cues or framings can dominate interpretation even when information is incomplete.

From an RPD perspective, susceptibility indicators such as trust, emotion, and group identity shape which cues are noticed, which interpretive patterns are retrieved, and which imagined

outcomes feel reasonable. These indicators influence both situation recognition and mental simulation, helping explain why the same factual element can produce different plausibility judgments across individuals (Pan & Wu, 2020), not because the facts differ, but because the cues and expectations guiding recognition differ.

Rationale for an NDM Analytic Lens: Observed Conditions

Observations from a prior CTA study of credibility assessment suggest that participants frequently formed judgments under conditions consistent with core characteristics of NDM, including uncertainty, time pressure, high stakes, and incomplete information. These observations motivated the use of NDM and the RPD model as analytic lens in this chapter. The CTA study involved participants with and without prior training in fact-checking, who reflected on memorable incidents of credibility assessment following established CTA procedures (Crandall et al., 2006). The observations are treated here as contextual background that informed the selection of analytic examples, rather than as empirical evidence or validation.

Across accounts, participants described making trust judgments while situations remained incomplete and not fully interpretable. For example, one participant noted:

> "*You still don't know what exactly is going on… you never can be quite certain… but it doesn't take long to realise that you shouldn't trust this source"* (SS, P12, L8).

This illustrates a judgment formed within an ill-structured problem under conditions of uncertainty.

In some cases, judgment was distributed across multiple players and embedded within an action–feedback loop, as participants described delegating verification to trusted others:

> "*I directly called my friend… and asked her if this is correct or not… she could decline it or approve it right away*" (MM1, P7, L3).

In other cases, time stress and competing goals were explicit, where delayed participation had consequences:

> "*it was important to be very quick and make quick decision to get into the conversation… otherwise… you could miss the point*" (MM2, P4, L1).

Related exploratory observations from a separate, small-scale mixed-methods study suggested recurring qualitative patterns in how participants engaged with differently framed misinformation messages. In the second study, participants reviewed misinformation narratives on vaccines, climate change, and bioweapons framed by three persuasion models. Real-time reasoning strategies were identified and described using Think-Aloud protocols (Ericsson & Simon, 1998).

Immediately after reading the misinformation narratives, then the Critical Decision Method (CDM) interview technique (Klein et al., 1989) was used to probe participants' decision strategies when accepting, rejecting, or remaining uncertain about the narratives.

During the Think-Aloud sessions, some participants rejected narratives because the information lacked the structure needed for definitive evaluation:

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

Post-exposure interviews further showed judgments made under uncertainty, embedded within action–feedback loops and competing goals. For instance:

> "*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].

Taken together, these excerpts show that encounters with misinformation routinely instantiate several defining characteristics of NDM environments (Orasanu & Connolly, 1993), including ill-structured problems, uncertainty, competing goals, time stress, action–feedback loops, and multi-actor judgment. These are the conditions under which recognition-based evaluations are expected, motivating the use of an NDM/RPD perspective to explain how plausibility judgments form and stabilise when people assess misinformation.

The discussion now turns to the persuasion models that describe how communicators may structure messages to engage these recognition processes and shape decisions and beliefs.

4.2.4 Persuasion Models

In theories of persuasive communication (Stiff & Mongeau, 2016), persuasion models are theoretical accounts of how persuasive devices are structured to influence judgments. Across communication, psychology, and marketing, these models account for how credibility cues, emotional appeals, and social influences affect evaluation processes. They provide the foundations for understanding why particular forms of argumentation resonate differently across individuals. In this chapter, the individual or group of individuals who receive a message and are the intended audience to be informed or persuaded are referred to as the target. The communicator refers to the individual or entity responsible for producing and conveying the message to the target.

Several influential persuasion models exist, including the Persuasion Knowledge Model (PKM) by Friestad and Wright, (1994), the Elaboration Likelihood Model (ELM) by Petty and Cacioppo, (1986), the Heuristic Systematic Model (HSM) by Chaiken, (1980), Hovland's Attitude Change Approach (HACA) (Hovland et al., 1953), the Theory of Reasoned Action (TRA) by Fishbein and

Ajzen, (2011). This chapter focuses on PKM, ELM, and TRA because they represent analytically distinct and non-overlapping mechanisms, while jointly covering the core characteristics shared with the remaining models. ELM characterises variability in information processing under cognitive constraint in ways that substantially overlap with HSM and HACA, making the latter two redundant for the purposes of this framework. By contrast, PKM and TRA retain distinct explanatory value by focusing on persuasion awareness and belief-to-action commitment, respectively, thereby completing a set of mechanisms particularly relevant to misinformation contexts.

Persuasion Knowledge Model (PKM) (Friestad & Wright, 1994) can be used to explain how persuasive intent can be made more or less visible in the construction of misinformation. A factual claim may be reframed so that the persuasive purpose is disguised within familiar, neutral, or seemingly informative language. When persuasive intent is made explicit, it may trigger skepticism or resistance; when it is embedded or obscured, the message may be processed without activating persuasion knowledge (Friestad & Wright, 1994). In this way, PKM supports the RBM framework by showing how misinformation can be designed to reduce recognition of manipulation while preserving the appearance of plausibility.

Elaboration Likelihood Model (ELM) (Petty & Cacioppo, 1986) can be used to vary how information is presented, encouraging either a central route of persuasion (more effortful evaluation) or a peripheral route (reliance on surface-level cues). A factual claim may be framed to emphasize evidence, data, or logical structure, encouraging a more analytic interpretation, or it may be reframed to rely on affective language, source cues, or simplified statements that reduce the need for detailed evaluation. Which route is engaged depends on the individual's motivation and cognitive capacity. In this way, ELM helps explain how the same underlying facts can be transformed into messages that invite different modes of evaluation.

Theory of Reasoned Action (TRA) (Fishbein & Ajzen, 2011), can be used to show how misinformation can be constructed by shaping beliefs about outcomes and perceived social norms. Attitudes are shaped by beliefs about outcomes, while norms reflect perceived social expectations. 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). TRA highlights that reframing beliefs about outcomes (e.g., risks) or exaggerating social consensus (e.g., "everyone knows this") can shift attitudes and encourage behavioral alignment with misleading narratives.

To move from how persuasion operates to how it is expressed in the structure of messages, the next sections introduce Aristotle's classical rhetorical strategies and persuasive principles.

4.2.5 Aristotle's Rhetorical Communication Strategies

Persuasive models frequently use Aristotle's classical rhetorical strategies of ethos (credibility), pathos (emotion), and logos (logic), which together form the rhetorical triangle (Haskins, 2004; Rapp, 2002). Ethos builds trust through perceived integrity, expertise, or relatability, such as endorsements from respected figures. Pathos evokes emotion through imagery, metaphor, or storytelling that aligns with individuals' values, including fear appeals or inspirational narratives. Logos appeals to critical reasoning by using evidence to create an impression of objectivity. Balancing ethos, pathos, and logos highlights the relevance of these rhetorical cues in shaping perceived plausibility, even when claims are misleading or selectively framed (Higgins & Walker, 2012; Lewandowsky et al., 2017).

4.2.6 Persuasive Design Principles

Persuasive communication refers to the study of how psychologically and rhetorically grounded mechanisms can influence interpretation, judgment, or behaviour through communicative artefacts (Stiff & Mongeau, 2016). Across disciplines such as psychology, communication, and human–computer interaction, research has identified recurring mechanisms on how credibility cues, emotional appeals, social norms, and motivational triggers shape the meaning and decision-making processes (Cialdini, 2007; Stiff & Mongeau, 2016; Torning & Oinas-Kukkonen, 2009). Within the RBM framework, persuasive design principles for message design are treated as conceptual mechanisms of influence.

Accordingly, the following distinctions are defined:

> **A persuasive design principle** describes a psychologically or rhetorically grounded mechanism of influence that remains invariant across contexts.
>
> **A persuasive technique** is a context-sensitive rhetorical operation that instantiates such a principle within a specific message.

Principles explain *why* certain message features shape interpretation, while techniques describe *how* those mechanisms may be applied in particular communicative settings. A sample, non-exhaustive list of persuasive design principles considered in this framework includes:

- ***Principle of Authority and Credibility***: Describes how perceived expertise, legitimacy, or trustworthiness of a source shapes judgments of plausibility and acceptance, particularly under uncertainty (Cialdini, 2007; Shaikh et al. , 2020).
- ***Principle of Social Identity and Proof***: Captures how perceptions of group norms, consensus, or shared identity influence interpretation and conformity in belief formation (Ajzen, 1991; ElBaff et al. , 2020).

- ***Principle of Accomplishment***: Reflects how signals of past success, achievement, or validation affect perceived reliability and confidence in claims (Thomson et al. , 2016).
- ***Principle of Reciprocity***: Describes how perceived obligation arising from prior benefit or concession can bias responses toward favourable interpretation (Redström, 2006).
- ***Principle of Scarcity***: Reflects how perceived limitation, urgency, or exclusivity increases attention and perceived value, shaping judgments under time pressure (Cialdini, 2007; Redström, 2006).
- ***Principle of Anthropomorphism***: Describes how attributing human-like qualities to sources or narratives fosters emotional connection, trust, and engagement (S. Kumar et al. , 2023).
- ***Principle of Emotion***: Captures how affective cues such as fear, anger, hope, or nostalgia influence attention, memory, and interpretation (Levesque & Pons, 2020).
- ***Principle of Commitment and Consistency (Foot-in-the-Door)***: Describes how prior commitments or incremental agreement shape subsequent acceptance and interpretation (Freedman & Fraser, 1966).
- ***Principle of Overcoming Resistance***: Captures mechanisms by which forewarning, inoculation, or acknowledgement of counterarguments shape receptivity and reduce rejection (McGuire, 1964).
- ***Principle of Value and Impact Formulation***: Describes how emphasising consequences, comparisons, or broader impact influences perceived importance and relevance (Kumar et al., 2023).
- ***Principle of Ambiguity and Under-specification***: Captures how vagueness or incomplete specification can shape interpretation by allowing audiences to fill gaps in ways consistent with their expectations or beliefs (Longpre et al. , 2019).

Constructing Persuasive Techniques

In this chapter, persuasive techniques are analytic constructs defined for the RBM framework. They are not fixed entities drawn from prior literature. Instead, they are specified to operationalise persuasive design principles within concrete message contexts. Each technique translates a persuasive design principle into context-sensitive rhetorical operations by defining a consistent set of attributes, as follows:

- ***Technique ID:*** a unique identifier assigned to the technique to support consistent reference across message variants and demonstrations.
- ***Goal:*** the intended perceptual or interpretive shift the technique is designed to achieve (e.g., increasing doubt, signalling credibility, encouraging alignment), derived from the persuasive design principle it instantiates.

- ***Cognitive target**:* the susceptibility indicator the technique is intended to engage, specified in alignment with both the target's susceptibility profile and the selected persuasion model.
- ***Theoretical grounding***: the persuasion or psychological theory that informs the technique, situating it within established models and principles.
- ***Illustrative wording (lexicon):*** example phrases, expressions, or lexical choices that illustrate how the technique may be realised in text to strengthen the intended frame.

Together, these attributes define each persuasive technique as a context-sensitive instantiation of a persuasive design principle. Appendices D and E provide the technique catalogue used throughout the RBM demonstrations to support transparency, traceability, and cross-referencing.

4.2.7 Framing Theory and Dominant Meaning

Misinformation rarely appears as isolated claims; it is embedded within narratives that organize facts, values, and emotions into rational persuasive structures. Framing theory, first articulated by Goffman, (1974), explains how these narratives highlight certain aspects of reality while omitting others, thereby shaping perception and guiding how targets make sense of complex information (Entman & Rojecki, 1993; Scheufele, 1999). Tversky and Kahneman, (1986) illustrate this cognitive mechanism: identical facts, such as "lives saved" vs. "deaths expected," elicit different judgments depending on how they are presented. These framing mechanisms are relevant in misinformation contexts, where selectively emphasised cues, emotional signals, or value-laden interpretations can shape perceived plausibility and meaning (Lewandowsky et al., 2017).

Entman (1993) defines framing as the selection of some aspects of a perceived reality and makes them more salient in a communicating text in order to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation. Framing operates by creating salience, using narrative structure to make certain elements of a message more noticeable, emotionally charged, or culturally familiar. In the broader literature, the interpretation that emerges through this process is often described as the *dominant meaning* of a frame: the understanding that becomes most accessible when individuals recall, evaluate, or discuss an issue.

4.2.8 How These Concepts Converge Toward Susceptibility to Misinformation

Taken together, the literature suggests that susceptibility to misinformation emerges from the convergence of individual susceptibility indicators, message persuasive design, framing, and

recognition-based evaluation under uncertainty. Across these perspectives, susceptibility indicators shape which cues attract attention, which patterns feel familiar, and which interpretations appear plausible. Persuasion models and rhetorical strategies specify how credibility, emotion, and apparent logic may be signalled, while framing theory explains how these signals are organised to dominate particular interpretations. NDM conditions, individuals rely on recognition-based processes and prior experience when assessing what "seems true."

Considered in isolation, none of these perspectives fully accounts for how factual elements can be transformed into persuasive misinformation. The following section brings these concepts together by formalising how message design and recognition-based evaluation interact in the construction of plausible misinformation narratives and introduces the RBM framework.

4.3 The RBM Framework

This section describes the RBM framework and explains how it operates as a method for constructing plausible misinformation. RBM framework is an analytic method for transforming factual elements into persuasive misinformation by coordinating four interdependent processes: target analysis, message design, framing, and recognition-based evaluation. Each process performs a specific role and produces an outcome. Together, these outcomes form the persuasive narrative, as outlined below:

1. **Target Analysis: Identify Susceptibility Conditions**

In the target analysis process, we specify the characteristics of the target, the person aimed to be the message receiver, that are relevant for analysing how a message may be interpreted. The outcome of this process is a susceptibility profile of the target. The process consists of three steps:

a) *Collect background information about the target*: We gather available information about the target to understand what they care about, trust, support, oppose, or engage with. This may include observing public statements, social media activity, group affiliations, prior writings, or observable engagement with related topics.

b) *Identify susceptibility indicators*: For each target or target group, we identify their vulnerability to believe misinformation. These points of vulnerability are known as susceptibility indicators and are described as trust, emotion, values, group identity, knowledge, and thinking style (Nan et al., 2022). The susceptibility indicators constrain which cues are more likely to attract the target's attention, which sources or signals are

perceived as credible, and which interpretations are most readily recognised as familiar or reasonable.

c) *Construct the susceptibility profile*: We consolidate the identified indicators into a susceptibility profile that summarises the target's vulnerability conditions, explaining why certain cues are more likely to be noticed and taken seriously.

The susceptibility profile is the analytic outcome of the target analysis process.

2. Message Design: Select Persuasion Model and Influence Mechanisms

We then use the target's susceptibility profile to guide our selection of an appropriate persuasion model and persuasive design principles to reframe the extracted factual element, following the steps explained below:

a) *Extract the factual elements:* We extract the factual elements from the root message and classify them according to their communicative role, such as whether it is an official response, is it meant to affect public perception, or to make an accusation? The categories below cover the main roles we encountered. This list is not exhaustive and may be extended as needed:

   A. Official Responses and Misinformation Denials
   B. Claims, Accusations, and Alternative Narratives
   C. Stated Purpose, Intent, and Justification
   D. Confidentiality, Opacity, and Uncertainty
   E. Public Perception, Belief, and Social Reaction

   These categories provide an analytic foundation for examining how individual factual element may be differentially emphasised, downplayed, or moralised in subsequent demonstrations, while keeping the underlying facts constant.

b) *Align the susceptibility profile with a persuasion model:* Using the susceptibility profile, we select a persuasion model that best aligns with how the target is likely to process the message. We draw on three established models, PKM, ELM, and TRA (Fishbein & Ajzen, 2011; Friestad & Wright, 1994; Petty & Cacioppo, 1986) as analytically distinct mechanisms to organise how persuasion operates. Depending on the susceptibility profile, we determine whether to activate or bypass persuasion awareness, such as when prior knowledge or scepticism is evident, we should select PKM; whether processing is likely to rely on peripheral cues rather than analytical scrutiny, such as when emotional salience or low motivation is observed, we should select ELM; or whether beliefs and

perceived norms should be emphasised to encourage commitment or action such as when values or group identity are prominent, we should select TRA.

c) *Select persuasive design principles to apply the model*: Now, we select one or more persuasive design principles that support how the chosen persuasion model is intended to operate in the message.

- First, we identify the primary persuasive focus of the selected persuasion model. For example, PKM focuses on intent, credibility, or trust; ELM focuses on whether the message is processed quickly or carefully; and TRA focuses on beliefs about outcomes or perceived social norms. Based on this persuasive focus, we decide which aspect of interpretation must be shaped in the message, such as credibility, emotion, or social norms.

- We then select a persuasive design principle whose mechanism of influence is known to shape that aspect of interpretation. For example, to shape credibility or trust, we can select the Principle of Authority and Credibility; to shape urgency, we can select the principle of Scarcity; and to shape social norms or group alignment, we can select the principle of Social Identity and Proof. A sample list of persuasive design principles is provided in Section 2.6. Once a persuasive design principle is selected, it cannot be applied directly in the message. While the principle's mechanism of influence remains the same, its role in the message changes depending on the persuasion model. For example, the principle of Authority and Credibility may be used to reduce suspicion under the PKM model, to shortcut evaluation under the ELM model, or to support belief or norm alignment under the TRA model. In the next step, this principle is operationalised by selecting a persuasive technique, as defined in Section 2.6.1, and applying it within the concrete message wording.

d) *Select and apply persuasive techniques:* Using the selected persuasive design principles, we select a persuasive technique by matching its attributes, as specified in section 2.6.1, to the analytic conditions already specified[1]. The selection proceeds as follows:

- *Match the technique goal:* We identify the intended influence defined by the selected persuasive design principle and match it to a technique whose goal

[1] We selected techniques from the catalogue in Appendices H and I, developed specifically for use in this chapter.

corresponds to that influence (e.g., increasing doubt, signalling credibility, encouraging alignment).

- *Match the cognitive target*: We select a technique whose cognitive target aligns with the target's susceptibility profile and the selected persuasion model, ensuring that the technique engages the same susceptibility indicator that the persuasion model is designed to operate on.
- *Apply the illustrative wording (lexicon):* We can use the illustrative wording associated with the selected technique as a vocabulary pool from which phrases, expressions, or questions may be drawn to reframe the factual element and strengthen the intended frame. The lexicon serves as a reference that facilitates the process without restricting wording choices.

Each selected technique is referenced by its technique ID and carries its theoretical grounding, as specified in Section 2.6.1.

Through this selection, a persuasion mechanism is formed, which structures how the message operates.

3. Message Framing: Construct Frames and Dominant Meaning

In the framing process, we organise the reframed factual element so that it conveys a dominant meaning. Framing does not introduce new content; it structures how existing elements are emphasised, connected, or downplayed to guide interpretation. The process consists of three steps:

a) *Identify what aspects of the reframed factual element are emphasised or downplayed*: Using the persuasion mechanism, we decide which parts of the reframed factual element are highlighted, repeated, or backgrounded within the message.

b) *Create Framing functions:* We structure the message following Framing Theory functions (Entman, 1993; Goffman, 1974) to define:
   - what issue is to be presented as the problem,
   - why it matters as the causal attributions,
   - who or what appears responsible as moral evaluations,
   - what responses or conclusions appear appropriate as the implied remedies

c) *Form the dominant meaning: We arrange emphasis and structure so that one interpretation of the message becomes dominant and appears coherent and reasonable to the target.*

The outcome of this process is a frame that establishes a dominant meaning perceived as natural, morally relevant, or action-worthy by the target.

4. Plausibility Evaluation: Conduct Recognition-based Evaluation

In this process, we specify how the framed message is likely to be evaluated through recognition-based processing. Drawing on the RPD model (Klein, 1993, 1997), we show how plausibility may stabilise under uncertainty without deliberate comparison of alternatives. The process consists of three steps:

a) *Identify recognisable cues in the framed message*: We examine the framed message produced in the framing process and identify the cues embedded within it. These cues are specific features of the reframed factual element that have been shaped through the combined effect of the persuasion mechanism and the dominant meaning. We identify which of these embedded cues are most likely to be immediately recognisable to the target. Recognisability is assessed based on familiarity, repetition, emotional salience, or alignment with the target's prior experience, as defined by the susceptibility profile.

b) *Specify the interpretive pattern activated by those cues*: We determine which familiar interpretive pattern the recognisable cues are likely to activate, for example, institutional concealment, moral resistance, risk downplaying. This step specifies how the target is likely to recognise "what kind of situation this is" without analytic evaluation.

c) *Describe the implied mental simulation and plausibility outcome*: We describe how the activated pattern supports a mental simulation of consequences, for example, what the message indicates is happening, who can be trusted, or what response seems appropriate. We then check whether the framed message allows the target to easily imagine what is happening and what should happen next; if the mental simulation is smooth and consistent, the message is considered plausible. Where inconsistencies arise or expectations are violated, the recognised interpretation indicates the need to adjust message design or framing and to re-evaluate until a coherent plausibility judgment can be formed.

The outcome of this process is a description of recognised interpretation. It shows how the reframed message appears plausible through recognition and mental simulation rather than factual verification.

5. Outcome

By following the four RBM processes, we produce a reframed message with a persuasive configuration. This message is structured to be evaluated by the target as being more or less plausible. If the target judges the message as plausible or acceptable, the message can operate persuasively and may lead to behaviours such as sharing, endorsing, or acting on the information.

Figure 4 illustrates the RBM framework, showing the relationships among its processes. The susceptibility profile of the target constrains cue salience; message design forms persuasion mechanisms that shape how cues are embedded in factual elements; framing constructs the dominant meaning by organizing cues; and recognition-based processing evaluates plausibility, resulting in a reframed message that can function as persuasive when judged plausible by the target. RBM allows for skipping, repetition, or reinforcement across processes rather than enforcing a strict sequence.

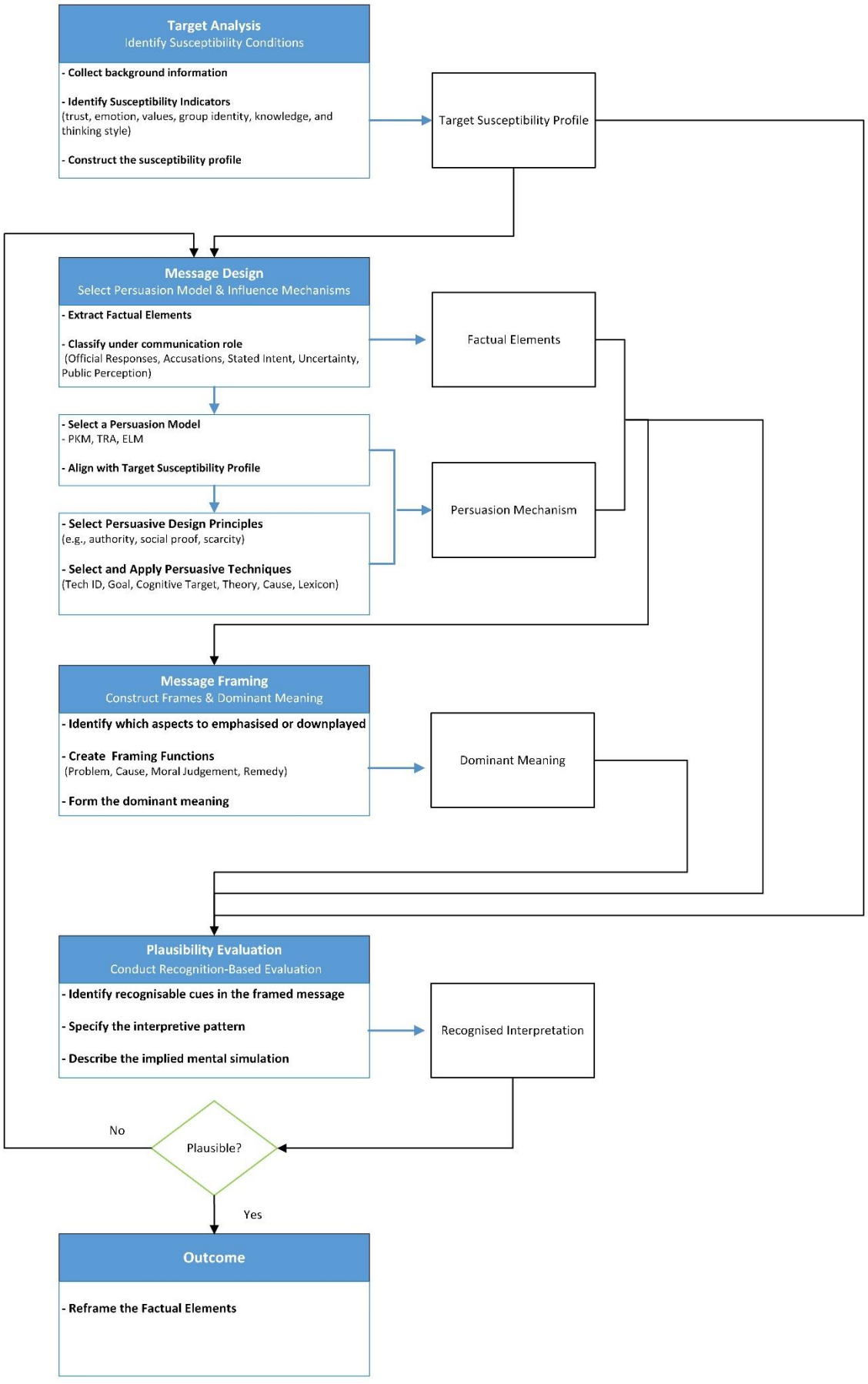

Figure 4 - RBM framework

The following section demonstrates how the framework can be applied through worked examples, illustrating its use as an analytic scaffold.

4.4 Demonstrating the RBM Framework

This section demonstrates the RBM framework through three examples: Ukraine biolabs, COVID-19 vaccines, and climate change. We prepared the examples from publicly available factually correct information. Each example represents the root message for that example.

Each demonstration follows the RBM processes in Figure 4, from target analysis through message design and message framing to plausibility evaluation and culminating in the outcome. For each example, one factual element is selected from its root message at a time, and systematically reframed into three variants that align with each of the three persuasion models while holding the factual baseline constant.

The Ukraine biolabs case is presented in detail to illustrate the step-by-step transformation process. The corresponding outcomes for the COVID-19 vaccine and climate change domains are reported later in the section for comparison. Full root messages, reframed narratives, and transformation tables are provided in Appendices A, B, and C, respectively. Across the demonstrations, we can trace how the same factual element produces different recognised interpretations and plausible outcomes as persuasion models and framing choices vary across the RBM processes.

4.4.1 Target Analysis

For each example, we first decide to whom the message is aimed. This target is hypothetical and is used only to illustrate how RBM works. For the Ukraine biolabs example, we describe the steps in the Target Analysis next.

a) *Collect background information about the target*: To construct a target's susceptibility profile, the analysis draws on available background information, e.g., publicly expressed positions, affiliations, and prior engagement. For the Ukraine biolabs example, we assume a target with the following background: Someone who questions governments and mainstream media, follows alternative or non-mainstream sources, and pays attention to claims about secrecy and hidden intent. This is the background we start from.

b) *Identify susceptibility indicators*: We next use this information to identify susceptibility indicators, such as trust, values, group identity, and emotional sensitivity. For the Ukraine biolabs example, the susceptibility indicators considered notable for this target are: low trust in institutions, values for preference for independence and scepticism, group

identity alignment with “questioners” rather than authorities, and sensitivity to uncertainty and threat. These indicators are consolidated into a susceptibility profile that guides subsequent model selection.

c) *Construct the susceptibility profile*: Taken together, these susceptibility indicators produce the susceptibility profile.

- Cues about intent and concealment are likely to be noticed because the target has low trust in institutions, leading them to focus on why authorities speak or deny information rather than on the content itself.
- Cues about norms and conformity may matter because the target’s group identity is aligned with “questioners” and their values emphasise independence and scepticism, making social positioning and moral alignment salient.
- Emotionally charged or urgent cues are likely to attract attention because sensitivity to uncertainty and threat makes the person more alert to signs of danger or urgency.

The target analysis resulted in the following target susceptibility profile:

- the target will notice cues about intent and concealment,
- cues about norms and conformity may matter,
- emotionally charged or urgent cues may attract attention.

This target susceptibility profile is then used to decide which persuasion models to examine in the next step.

4.4.2 Message Design

The message design process consists of four main steps: (a) extract factual elements, (b) select a persuasion model, (c) select persuasive design principles, and (d) select and apply persuasive techniques. The following sections illustrate these steps using the Ukraine biolab example:

a) *Extract the factual elements:* In this step, we read the root message to identify the neutral factual elements and then classify them according to the purpose of their communication in the message. The base root message was written by the authors as an example. Neutral factual elements were identified and extracted. Alternative persuasive narratives can then be constructed using the neutral facts. The neutral factual element is treated as a stable referential unit whose presentation and salience can be reorganised to support different persuasive interpretations. The main root message about the Ukrainian Biolab is presented below with the selected neutral factual element highlighted:

"There are biological research facilities in Ukraine, some of which have received international and U.S. funding for public health research and disease prevention. The U.S. and Ukrainian governments have stated that all research conducted in Ukrainian labs is for public health and biosecurity, and no immediate risk has been identified from the labs operating under biosecurity protocols. ***While government officials and mainstream media have refuted allegations of secret or dangerous biolab research conspiracies****, and no confirmed evidence of U.S.-funded bioweapons development has been found, a former employee of a Ukrainian research lab stated that public disclosure of specific research activities remains limited due to confidentiality agreements and national security protocols. Some individuals, such as Ivan Petrov, have claimed to possess documents about U.S. activities in Ukraine (9), and alternative media platforms have circulated unverified reports, which have not been officially confirmed or disproven. As a result, while most people accept the official accounts that promote caution regarding conspiracy claims, others remain skeptical, and the debate continues in both mainstream and alternative information spheres, with no concrete action having been taken."*

b) For the Ukraine biolabs example, the selected factual element was:

"... ***government officials and mainstream media have refuted allegations of secret or dangerous biolab research conspiracies ..."***

A single factual element can be expressed through multiple neutral, factually equivalent formulations, each capturing the same underlying verifiable point. For instance, the selected factual element in the Ukraine biolab narrative could also be expressed as:

- *There have been statements from government officials refuting allegations of secret or dangerous research.*
- *Claims about biolabs in Ukraine have been officially denied and labelled as misinformation.*
- *Officials and mainstream media have dismissed biolab conspiracies as misinformation.*

This is a single factual element that can be worded in different neutral ways without changing the fact it refers to. This shows that the wording style may vary, while the same factual element can still serve as the baseline for modification.

Next, the extracted factual element is classified according to its communicative role. The selected factual element was classified under "Theme A: Official responses and misinformation denials", as it described how authorities responded to the claims[2].

c) *Align the susceptibility profile with a persuasion model:* In this step, we use the target susceptibility profile to anticipate how the target might respond to the neutral factual message and then anticipate which persuasion model best matches this behaviour. The choice is guided by the cues identified in the susceptibility profile. In the Ukraine biolabs example, we explain how this is done. Given the neutral factual message,

> "*government officials and mainstream media have refuted allegations of secret or dangerous biolab research conspiracies*",

and the target susceptibility profile, we anticipate that the target is unlikely to treat the refutation as settled information. Instead, the target is likely to attend to:

- why the refutation is made, reflecting sensitivity to intent and concealment,
- how refutation positions authorities against "questioners", reflecting sensitivity to norms and conformity, and
- to whether references to "dangerous biolab research" are interpreted as reassurance or as confirmation of risk, reflecting sensitivity to uncertainty and threat.

These anticipated responses indicate which persuasion models are most appropriate to use (i.e. PKM, ELM, TRA). To select the most appropriate persuasion model to use, we apply the following mapping guidelines:

- When the profile foregrounds distrust cues, sensitivity to persuasive intent, and reactions to institutional credibility, we align the example with PKM (Friestad & Wright, 1994).
- When the profile foregrounds values, social norms, or group alignment, we align the example with TRA (Fishbein & Ajzen, 2011).
- When the profile foregrounds emotion-driven, intuitive, or reliance on surface cues, we align the example with ELM (Petty & Cacioppo, 1986).

This mapping specifies how persuasion operates in the example: through intent recognition (PKM), heuristic processing (ELM), or belief–norm alignment (TRA). This alignment clarifies which types of influence mechanisms are likely to be engaged for a given susceptibility profile. In this chapter, all three persuasion models are used to illustrate how the same factual element can be varied across model-aligned reframes.

[2] The list of five categories reflecting the core role of factual elements is presented in Section 4.3.

d) *Select persuasive design principles to apply the model:* In this step, once the persuasion model has been selected, we identify its persuasive focus and select the design principles that shape the required aspect of interpretation. In the Ukraine biolabs example, this alignment leads to different design principles being selected depending on which aspect of the factual element is foregrounded:

- In the PKM-aligned variant, where sensitivity to intent (of government officials) and trust (in mainstream media) is central, we may select the principle of Authority and Credibility. This persuasive design principle shapes how refutation by government officials and mainstream media is interpreted by directing attention to the trustworthiness and possible motives behind the denial rather than to its content.
- In the ELM-aligned variant, where sensitivity to uncertainty and the threat (of secret or dangerous biolab research) is central, we may select the principle of Emotion. This persuasive design principle shapes how references to "dangerous biolab research" are processed, encouraging urgency and intuitive evaluation rather than careful scrutiny.
- In the TRA-aligned variant, where values and group identity (being positioned as aligning with authorities or with truth-seekers) are central, the principle of Social Identity and Proof is selected. This persuasive design principle shapes how the refutation positions authorities versus sceptical audiences, making acceptance or rejection of the refutation a matter of norm alignment rather than evidential assessment.

At this stage, the principle cannot yet be written into the message. It only describes the type of influence mechanism to be applied. In the next step, we operationalise this principle by selecting a persuasive technique and embedding it in the message wording.

e) *Select and apply persuasive techniques:* In this step, a persuasive technique is selected and applied to the neutral factual element to produce the model-aligned message. Techniques are used here as analytic exemplars to show how transformation occurs, not as prescriptive instructions. For the Ukraine biolabs example, the neutral factual element is:

> *"...government officials and mainstream media have refuted allegations of secret or dangerous biolab research conspiracies."*

Using the target susceptibility profile, the selected persuasion model, and the persuasive design principle, we apply one technique per model and align its key attributes, including goal, cognitive target, and theoretical grounding (Section 3), with the persuasion influence. At this stage, the persuasion mechanism is designed, but the message is not yet framed: the technique specifies which aspect of the factual element will be foregrounded, without yet constructing a dominant meaning:

PKM-aligned variant (intent-based evaluation):

- Persuasion principle: Authority and credibility
- Technique applied: Tech 1: Undermining Authority
- Goal: Shift attention from the factual denial to the intent behind it and issues of trust and deception.
- Cognitive target: Trust in government officials and mainstream media (institutional trust)
- Lexicon (illustrative wording): Phrases such as “dissenting voices” or “ignore the evidence” may be attached to the description of how government officials and mainstream media refuted the allegations, directing attention to trust and motive without yet stating a conclusion.
- Transformation: The refutation by government officials and mainstream media is designed to emphasise possible motives or interests behind the denial, so that the target evaluates why the denial is made rather than accepting the refutation.

ELM-aligned variant (heuristic evaluation):

- Persuasion principle: Emotion
- Technique applied: Tech 8: High Stakes and Urgency
- Goal: Encourage rapid, low-effort evaluation under fear and urgency
- Cognitive target: Perceived risk and urgency that narrow attention
- Lexicon (illustrative wording): Phrases such as “serious risks to public safety” may be attached to references to “dangerous biolab research”, to foreground potential threat and urgency.
- Transformation: References to “dangerous biolab research” are designed to foreground using “serious risks to public safety,” so that uncertainty is interpreted as a potential threat, narrowing attention and prompting intuitive concern over careful assessment.

TRA-aligned variant (norm-based evaluation):

- Persuasion principle: Social identity and proof
- Technique applied: Tech 15: Social Identity Framing
- Goal: Align belief with group norms and shared values
- Cognitive target: Perceived expectations tied to in-group and out-group
- Lexicon (illustrative wording): Phrases such as “the public has been told” may be attached to the description about official denials, positioning acceptance or scepticism relative to group alignment.
- Transformation: The refutation is positioned as representing official or mainstream views, while scepticism is considered consistent with the values of independent “questioners” positioned as truth-seeking, making belief a matter of social alignment.

Appendix H presents a non-exhaustive toolkit of persuasive techniques used for analytic message reframing. Appendix I describes the same techniques for reference throughout the demonstrations.

In all cases, the factual element remains unchanged. The outcome is a persuasion mechanism designed to communicate the facts in a manner aligned with the target's susceptibility profile. The design is the narrative (without the framing) that situates the message based on the facts, expressed in terms of one of the three Persuasion Models, in a communicative form associated with a specific Persuasive Technique.

The next section shows how the newly designed messages are framed to shape salience and dominant meaning.

4.4.3 Message Framing

Before framing is constructed, the factual element is first reorganized through the Message Design process. This step does not alter the factual content; it specifies an interpretive orientation aligned with the selected persuasion model and susceptibility profile. This persuasion mechanism guides how the fact will later be framed, but it is not yet a framed message. Framing then organizes this persuasion mechanism through problem definition, causal attribution, moral evaluation, and remedy. For the Ukraine biolabs example, the neutral factual element:

> *“government officials and mainstream media have refuted allegations of secret or dangerous biolab research conspiracies”*

is framed using the Message Design output and framing functions as shown through the steps below:

a) *Identify what aspects of the reframed factual element are emphasised or downplayed*: Using the selected persuasion mechanism, we decide which parts of the factual element must be foregrounded and which must be backgrounded. In practice, this means deciding what is repeated, what is made salient, and what is mentioned only briefly or indirectly. In all cases, the factual element remains unchanged. Only salience and framing differ, as below:

- PKM-aligned - Salience Given to Undermining Authority: the refutation by government officials and mainstream media is foregrounded as evidence of concealment, defining the problem as hidden intent, attributing cause to institutional motives, morally legitimising distrust, and directing attention toward alternative sources. This produces a dominant meaning centred on suspicion and hidden motives (e.g., "They hide things from you; you must question them.").
- TRA-aligned - Salience Given to Norm Reversal: Accepting official refutations without question is presented as comfort and compliance, rather than responsibility, casting social pressure as the cause and scepticism as moral virtue. Group alignment is positioned as the remedy, producing a dominant meaning anchored in identity and normative distinction (e.g., "Good people don't follow blindly; they stand with us and question.").
- ELM-aligned framing (salience given to catastrophic emphasis): references to "dangerous biolab research" are foregrounded while reassurances are backgrounded, defining the problem as imminent threat, attributing cause to ignored warnings, moralising inaction as dangerous, and prompting urgency or vigilance. This produces a dominant meaning driven by emotional salience and perceived risk (e.g., "This threat is real and immediate. You must act.").

b) *Create Framing functions*: The reframed element is then organised using four framing functions, as problem definition, cause, moral judgment, and remedy, as below:

- *PKM-aligned variant (Ukraine biolabs: intent, deception, concealment)*
  - Problem: Official refutations of the Ukraine biolab allegations suggest something important is being hidden.
  - Cause: They dismiss allegations because they are complicit.
  - Moral Evaluation: Their behaviour is unethical; distrust is justified.
  - Remedy: Seek alternative sources; expose the cover-up.
- *TRA-aligned variant (Ukraine biolabs: identity, norms, moral alignment)*
  - Problem: Accepting official refutations of the Ukraine biolab allegations without question reflects conformity rather than responsibility.

    - Cause: Social pressure and obedience to authority.
    - Moral Evaluation: Integrity means questioning, not complying.
    - Remedy: Stand with those who demand transparency.
- ELM-aligned variant (Ukraine biolabs: emotion, intuitive cues)
    - Problem: The risks associated with "dangerous biolab research" are being downplayed.
    - Cause: Officials and media minimise the concerns and close down scrutiny.
    - Moral Evaluation: Their refusal to address concerns is irresponsible.
    - Remedy: Stay vigilant and demand the full story.

c) *Form the dominant meaning*: The framing functions specified above are organized directly in the message wording using the selected persuasive technique. The problem and cause determine which elements are stated, the moral evaluation determines how those elements are characterised, and the remedy determines the implied stance. The resulting sentence expresses all four functions simultaneously and forms the dominant meaning. Using the Ukraine biolabs example, the framed outputs are as follows:

- PKM-aligned Reframed Narrative:
  Used Technique: Undermining Authority; cognitive target: institutional trust; lexicon: "dissenting voices", "ignore the evidence," results in:
    - "Despite repeated warnings from independent experts, officials continue to silence dissenting voices, while major news outlets conveniently ignore the evidence."

- TRA-aligned Reframed Narrative:
  Used Technique: Social Identity Framing; cognitive target: norms and group alignment; lexicon: "the public has been told," results in:
    - "*For too long, the public has been told to dismiss any claims about secret biolabs in Ukraine as dangerous conspiracy theories. But what if ignoring the truth is the real danger?"*

- ELM-aligned Reframed Narrative:
  Used Technique: High Stakes and Urgency; cognitive target: threat sensitivity; lexicon: "serious risks to public safety," results in:

- *"While officials and mainstream media dismiss concerns about secret biolabs as conspiracy theories, the potential consequences of undisclosed biological research pose serious risks to public safety."*

The next section presents worked examples demonstrating how the framed message is likely to be evaluated by the target through recognition-based evaluation.

4.4.4 Plausibility Evaluation

Across the three variants, the framed messages contain specific recognizable cues that allow the target to rapidly recognize what kind of situation is being described and what response makes sense. Plausibility evaluation proceeds by identifying these cues, recognizing the interpretive pattern they activate, and mentally simulating likely consequences.

a) *Identify recognizable cues in the framed message: First, salient cues are readily recognizable*.
   - PKM-aligned variant: phrases such as "silence dissenting voices" and "ignore the evidence" function as cues of institutional dismissal and concealment.
   - TRA-aligned variant: phrases such as "the public has been told" and "dismiss any claims" function as cues of moral positioning, social pressure and normative compliance.
   - ELM-aligned variant: phrases such as "serious risks to public safety" function as cues of emotional sensitivity, urgency and potential threat.

   These cues are readily recognised because they match the target's susceptibility profile, including low institutional trust, sensitivity to norms compliance, and sensitivity to uncertainty and threat.

b) *Specify the interpretive pattern activated by those cues:* Once recognised, these cues activate familiar interpretive patterns without deliberate analysis:
   - PKM-aligned cues activate a pattern of institutional concealment, where official denial is interpreted as motivated rather than informative.
   - TRA-aligned cues activate a pattern of moral resistance to conformity, where official messaging is interpreted as pressure to comply.
   - ELM-aligned cues activate a pattern of risk being downplayed by authorities, where reassurances are interpreted as minimising danger.

   These patterns provide interpretations that feel familiar to the target.

c) *Describe the implied mental simulation and plausibility outcome*: The activated pattern then supports a mental simulation of consequences:

- In the PKM-aligned variant, the target simulates that denial implies hiding and is the evidence of concealment, e.g., "if they deny it, they must be hiding something", leading to distrust and alternative information seeking feel reasonable.
- in the TRA-aligned variant, the target simulates that dismissal functions as social control, e.g., "they want compliance, not questioning", making scepticism feel morally appropriate.
- in the ELM-aligned variant, the target simulates that danger is being minimised, e.g., "serious risks are being ignored", making vigilance and concern feel justified.

When the framed message allows the target to easily imagine what is happening and what response makes sense, it sounds plausible to the target. When inconsistencies arise or expectancies are violated, the recognition-based evaluation indicates the need to adjust message design or framing and re-evaluate until a coherent plausibility judgment is formed.

The outcome of this process is a description of recognised interpretation. It shows how the reframed message appears plausible through recognition and mental simulation rather than factual verification.

4.4.5 Outcome

In the Ukraine biolabs demonstration, a single factual element:

> *"government officials and mainstream media have refuted allegations of secret or dangerous biolab research conspiracies"*

is held constant and passed through the full RBM process. It is reframed through the RBM processes described above, starting with target analysis that assumes the same susceptibility profile throughout, then moving through message design and message framing to plausibility evaluation. This process produces distinct dominant meanings and corresponding recognised interpretations that appear plausible to the target.

PKM-aligned outcome

Given a susceptibility profile sensitive to intent and institutional trust, the Ukraine biolabs factual element is organised and framed so that official refutation is recognised as concealment. The message defines the problem as concealment, attributes cause to institutional intent, morally

legitimises distrust, and directs attention toward alternative sources. The dominant meaning centres on suspicion and hidden motives, such as:

> *"Despite repeated warnings from independent experts, officials continue to silence dissenting voices, while major news outlets conveniently ignore the evidence."*

TRA-aligned outcome

Given a susceptibility profile responsive to values and group alignment, the same Ukraine biolabs factual element is organised and framed so that official denial is recognised as social pressure. The message directs conformity as the problem, attributes cause to social pressure, frames scepticism as a moral quality, and positions group alignment as the appropriate response. The dominant meaning is anchored in identity and normative distinction, such as:

> *"For too long, the public has been told to dismiss any claims about secret biolabs in Ukraine as 'dangerous conspiracy theories.' But what if ignoring the truth is the real danger?"*

ELM-aligned outcome

Given a susceptibility profile responsive to emotional salience and heuristic cues, the Ukraine biolabs factual element is organised and framed so that dismissal is recognised as risk being downplayed. The message foregrounds threat, attributes concern to ignored warnings and moralises inaction as dangerous. The dominant meaning is driven by emotional salience and perceived risk, such as:

> *"While officials and mainstream media dismiss concerns about secret biolabs as conspiracy theories, the potential consequences of undisclosed biological research pose serious risks to public safety."*

Table 12 summarises the resulting narrative variants generated from the same factual element across the three persuasion models.

Table 12 - Reframing one factual element into three model-aligned narratives (Ukraine Bio Lab)

Factual Element	Persuasion Model	Techniques Used	Resulting Reframed Narrative
"Government officials and mainstream media have refuted allegations of secret or dangerous biolab research conspiracies."	PKM	Tech 1: Undermining Authority	*"Despite repeated warnings from independent experts, officials continue to silence dissenting voices, while major news outlets conveniently ignore the evidence."*
	TRA	Tech 15: Social Identity Framing	*"For too long, the public has been told to dismiss any claims about secret biolabs in Ukraine as 'dangerous conspiracy theories.' But what if ignoring the truth is the real danger?"*
	ELM	8: High Stakes/Urgency	*"While officials and mainstream media dismiss concerns about secret biolabs as conspiracy theories, the potential consequences of undisclosed biological research pose serious risks to public safety.'"*

These outcomes show how the RBM framework can be applied as a systematic analytical scaffold to generate alternative narrative outcomes from a single factual element within a single story. To demonstrate that the same procedure can be applied beyond the Ukraine biolabs case, the next subsection reports outcome summaries for two additional story domains.

4.4.6 RBM Configurations Across Additional Story Domains

This section extends the worked demonstration to two additional story domains: COVID-19 vaccines and climate change. Rather than repeating the full analytic walkthrough presented in the reference case, only outcome-level summaries are reported here. For each domain, a single baseline factual element, from the root neutral messages presented in Appendix E, is reframed into three narrative variants aligned with PKM, TRA, and ELM. The full assembled narratives for each domain are provided in Appendix F, and the transformation breakdown is reflected in Appendix G. In the reference demonstration (Ukraine biolabs), a single persuasive technique was applied per persuasion model to show the transformation process step by step. In the additional story domains reported here, different techniques are demonstrated alternatively within the same persuasion model to illustrate how similar persuasion mechanisms can be achieved through different technique choices. These techniques are aligned with the same persuasion mechanism and serve the same framing function. For clarity, the reported outcomes focus on the resulting dominant meaning.

COVID-19 Vaccine Narrative

The COVID-19 vaccine story centres on public health guidance and institutional reassurance. The persuasive reframing of this story must operate against an explicit narrative of safety monitoring and risk minimisation. Baseline factual element, as stated in Table 13, is:

> *"Side effects from COVID-19 vaccines are rare and mostly mild, with health agencies continuing to monitor safety data."*

Using the same RBM procedure demonstrated in the Ukraine biolabs case, this factual element is reframed through different persuasion mechanisms to produce distinct outcome variants. Below, we report the outcomes to illustrate how the procedure applies in this domain.

The PKM-aligned: Given a susceptibility profile sensitive to institutional trust and intent, the COVID-19 vaccine factual element is organised so that official reassurance is recognised as concealment. The dominant meaning centres on institutional unreliability and suppressed risk, in which continued safety monitoring is reinterpreted as evidence that important information is being concealed. For example, "For example, continued safety monitoring is framed as a sign that important risks may still be undisclosed."

In the PKM-aligned configuration, the factual element is organised to foreground distrust and sensitivity to intent. Distrust Institutions (Technique 2) frames health agencies as unreliable authorities, drawing on lexicon such as "powerful institutions". Conspiracy Appeal Framing (Technique 3) attributes institutional reassurance to hidden motives and coordinated suppression, using lexicon such as "what they don't want you to know" and "the real story behind". Catastrophic Framing (Technique 7) intensifies moral evaluation by suggesting that severe consequences are being minimised, using lexicon such as "a disaster waiting to happen". The resulting dominant meaning reinterprets official safety assurances as signals of concealment, producing the following outcome narrative:

> *"Yet, vaccine side effects continue to be swept under the rug by powerful institutions. Reports of adverse reactions are rising, but what they don't want you to know is the real story behind the long-term risks, a disaster waiting to happen."*

In the TRA-aligned configuration, the factual element is organised to foreground identity and normative alignment. Social Identity Framing (Technique 15) constructs moral evaluation by positioning concern about vaccine risks as characteristic of a responsible in-group, drawing on lexicon such as "people like us" and "the public has been told". Polarisation/Right Choice (Technique 11) sharpens this distinction by defining unquestioning acceptance of official guidance as inexperienced or irresponsible, using lexicon such as "think for yourself or be misled" and "there are two kinds of people". Catastrophic Framing (Technique 7) contributes to moral

urgency by elevating rare adverse outcomes into socially salient risks, using lexicon such as "extremely dangerous consequences" and "a disaster waiting to happen". Through these operations, scepticism is framed as the morally appropriate and normatively endorsed stance for attentive and protective individuals. The resulting dominant meaning positions vaccine scepticism as a responsible response, producing the following outcome narrative:

> *"There are numerous stories of people like us who experienced severe side effects... Samantha Green, a 34-year-old mother from Ohio, developed an adverse reaction shortly after receiving the vaccine, and it took her over 6 months to recover. These risks are often swept under the rug."*

In the ELM-aligned configuration, the factual element is organised to foreground affect-driven, heuristic processing. Catastrophic Framing (Technique 7) establishes the problem by amplifying perceptions of danger, drawing on lexicon such as "extremely dangerous consequences" and "a disaster waiting to happen". Implied Guilt (Technique 6) assigns cause indirectly by suggesting that concealment itself signals wrongdoing, using lexicon such as "turning a blind eye" and "hard to justify ignoring". Conspiracy Appeal Framing (Technique 3) reinforces this interpretation through familiar surface cues of secrecy and suppression, using lexicon such as "what they don't want you to know" and "the real story behind". Through these operations, risk is made salient and reassurance is minimised, producing a dominant meaning that feels immediately plausible through fear and suspicion rather than deliberative evaluation. The resulting dominant meaning foregrounds threat and concealment, producing the following outcome narrative:

> *"Reports of adverse reactions are increasing, but the mainstream media and pharmaceutical companies are turning a blind eye to extremely dangerous consequences, creating a situation that is fast becoming a disaster waiting to happen."*

Table 13 summarizes the transformation outcomes for the COVID-19 vaccine factual element across three persuasion models.

Table 13 - Reframing one factual element into three model-aligned narratives (COVID-19 Vaccine)

Factual Element	Persuasion Model	Techniques Used	Resulting Reframed Narrative
"Side effects from COVID-19 vaccines are rare and mostly mild, with health agencies continuing to monitor safety data."	PKM	Tech 2: Distrust Institutions Tech 3: Conspiracy Appeal Framing Tech 7: Catastrophic Framing	"Yet, vaccine side effects continue to be swept under the rug by health authorities. Reports of adverse reactions are on the rise, but the government and big pharma continue to downplay the risks, leaving us vulnerable to long-term damage."
	TRA	Tech 15: Social Identity Framing Tech 11: Polarisation / Right Choice Tech 7: Catastrophic Framing	"There are numerous stories of people like us who experienced severe side effects... Samantha Green, a 34-year-old mother from Ohio, developed an adverse reaction shortly after receiving the vaccine, and it took her over 6 months to recover. These risks are often swept under the rug."
	ELM	Tech 7: Catastrophic Framing Tech 6: Implied Guilt Tech 3: Conspiracy Appeal Framing	"Reports of adverse reactions are increasing, but the mainstream media and pharmaceutical companies are turning a blind eye to extremely dangerous consequences, creating a situation that is fast becoming a disaster waiting to happen."

Climate Change Narrative

Climate change is associated with long-term policy decisions, economic trade-offs, and identity-laden debates about lifestyle, responsibility, and ideological alignment. This makes it particularly suitable for persuasion that targets values, perceived control, and future-oriented risk rather than secrecy or outright denial. The baseline factual element for this domain is:

> *"Green energy initiatives aim to reduce dependence on fossil fuels and promote sustainability."*

In the PKM-aligned configuration, the factual element is organised to foreground persuasion knowledge and resistance to perceived agenda-setting. Anti-Establishment Framing (Technique 20) defines the problem as green energy initiatives being pushed by elites and controlled by those in power, rather than emerging from neutral public interest. Polarisation / Right Choice (Technique 11) attributes cause to ideological pressure that marginalises economic concerns, using lexicon such as "think for yourself or be misled" and "there are two kinds of people". Self-Enhancing Framing (Technique 12) contributes to moral evaluation by positioning scepticism as the stance of "those who stay informed" and "people who see through the spin". Through these operations, green energy initiatives are framed as an imposed agenda rather than a practical sustainability policy. The resulting dominant meaning reinterprets green energy as economically harmful and ideologically driven, producing the following outcome narrative:

"Countries like Norway and Germany have shown that green energy policies, pushed by elites, can lead to job losses, economic instability, and higher energy costs."

In the TRA-aligned configuration, the factual element is organised to foreground identity-consistent norms and perceived behavioural control. Empowerment Framing (Technique 13) constructs the remedy by presenting environmentally responsible behaviour as achievable without institutional disruption, drawing on lexicon such as "able to make informed judgments" and "equipped to decide for yourself". Self-Enhancing Framing (Technique 12) establishes moral evaluation by associating environmental stewardship with competence and responsibility, using lexicon such as "those who stay informed" and "people who see through the spin". Reinforcement Framing (Technique 16) stabilises the causal narrative by repeatedly aligning sustainability with familiar practices, using lexicon such as "once again confirms" and "as we've seen before". Through these operations, environmental protection is framed as consistent with existing values and everyday experience rather than as a demand for systemic change. The resulting dominant meaning positions environmental responsibility as a common-sense choice that preserves established ways of life, producing the following outcome narrative:

"Relying on natural solutions, like managing forest ecosystems and harnessing nature's balance, is a time-tested approach... As we've seen before, we can protect our world without sacrificing our way of life."

In the ELM-aligned configuration, the factual element is organised to foreground affect-driven, economically focused heuristics. Catastrophic Framing (Technique 7) defines the problem by foregrounding severe economic consequences, drawing on lexicon such as "extremely dangerous consequences" and "a disaster waiting to happen". High Stakes / Urgency (Technique 8) intensifies perceived immediacy and personal relevance, using lexicon such as "serious risks to public safety" and "time is running out". Anti-Establishment Framing (Technique 20) assigns surface-level causal attribution by suggesting that green energy policies are "pushed by elites" and "controlled by those in power", while their real-world impacts are minimised. Through these operations, interpretation is driven by fear and economic anxiety rather than detailed policy evaluation. The resulting dominant meaning foregrounds immediate economic threat and institutional neglect, producing the following outcome narrative:

"While the media pushes green energy initiatives... countries like Norway are already seeing the extremely dangerous consequences of these policies, including job losses and increasing energy costs."

Table 14 summarizes the transformation outcomes for the climate change factual element across three persuasion models.

Table 14 - Reframing one factual element into three model-aligned narratives (Climate Change)

Factual Element	Persuasion Model	Techniques Used	Resulting Reframed Narrative
"Green energy initiatives aim to reduce dependence on fossil fuels and promote sustainability."	PKM	Tech 20: Anti-Establishment Framing; Tech 11: Polarisation / Right Choice; Tech 12: Self-Enhancing Framing	"Countries like Norway and Germany have shown that green energy policies, pushed by elites, can lead to job losses, economic instability, and higher energy costs."
	TRA	Tech 13: Empowerment Framing; Tech 12: Self-Enhancing Framing; Tech 16: Reinforcement Framing	"Relying on natural solutions, like managing forest ecosystems and harnessing nature's balance, is a time-tested approach.... As we've seen before, we can protect our world without sacrificing our way of life."
	ELM	Tech 7: Catastrophic Framing; Tech 8: High Stakes/Urgency; Tech 20: Anti-Establishment Framing	"While the media pushes green energy initiatives... countries like Norway are already seeing the extremely dangerous consequences of these policies, including job losses and increasing energy costs."

Together, these cases show how the RBM framework can be used to systematically reframe distinct narrative variants from the same factual element by aligning target susceptibility, persuasion mechanism, and framing choices. Section 5 discusses what this approach contributes, its limits, and how it differs from existing work.

4.5 Discussion

Section 4 presented worked demonstrations of the RBM framework. This discussion interprets the patterns observed across those cases, clarifies RBM's analytic contribution, and outlines key limitations.

4.5.1 What the RBM Reveals About the Design of Misinformation Narratives

The RBM framework demonstrates that misinformation narratives are not merely improvised distortions of factual content. The worked demonstrations make visible how narrative meaning is deliberately organised by shaping salience and interpretation, rather than emerging incidentally from the facts themselves.

Across the worked examples, variants aligned with PKM, TRA, and ELM exhibit distinct salience profiles, revealing recurring patterns in how hidden intent, moral identity, or emotional urgency are foregrounded. These patterns clarify why different targets are drawn toward different

interpretations of the same material: it is the configuration of cognitive engagement, rather than the topical content, that drives plausibility and uptake.

From this perspective, RBM reframes misinformation as a problem of narrative construction and cognitive alignment rather than a problem of factual accuracy. This shifts analytic attention from what claims are made to how persuasive meaning is assembled and stabilised.

4.5.2 Cross-Model Patterns and Cross-Domain Transferability

Across the three story domains examined in Section 4, the same model-aligned patterns recur despite differences in topic. PKM-aligned variants foreground suspicion and inferred intent, TRA-aligned variants emphasise identity and normative positioning, and ELM-aligned variants prioritise emotional urgency and familiar surface cues. This consistency indicates that RBM captures reusable configurations in the construction of persuasive narratives, rather than effects tied to any single content domain.

4.5.3 Cognitive Integration in Plausibility Judgement: RPD, Susceptibility, and Framing

The RBM integrates three foundations into a single analytic account of plausibility judgement under uncertainty. RPD clarifies how people evaluate claims through cue recognition and mental simulation rather than exhaustive comparison of alternatives; susceptibility indicators specify which cues are likely to become salient and credible for a given target; and Entman's framing functions provide a structural template for how those cues are organized into a dominant meaning. Together, these components clarify how message-level cues are arranged into coherent interpretations within RBM.

This integration helps explain why misinformation can feel believable to its targets: it aligns with recognition pathways, activates predispositions that bias what is treated as salient, and structures problem definition, causal attribution, moral evaluation, and cueing for remedies in ways that reinforce familiar patterns during evaluation. RBM therefore links cognitive processing, target susceptibility, and narrative framing within a single explanatory framework to understand how plausibility judgements form under uncertainty.

4.5.4 Applied Implications for Misinformation Analysis and Intervention

The RBM framework helps clarify the cognitive problem that should be understood before interventions for misinformation can be meaningfully designed. Research and practice concerned with countering misinformation, supporting sensemaking, or strengthening cognitive resilience require an account of why persuasive narratives gain traction, which cognitive and rhetorical mechanisms are engaged, and how interpretations stabilise under uncertainty. Without

this foundation, interventions risk addressing surface symptoms, such as adding only fact-checks or warnings, rather than engaging the decision-making processes, susceptibility cues, and framing structures that shape plausibility judgements.

RBM supports reasoning about how the same message may stabilise differently across communities. This perspective is relevant for work focused on analytic sensemaking, training and simulation, decision support, and the design of information environments operating under uncertainty.

Beyond individual interventions, RBM is relevant to strategic communication, resilience planning, and disinformation analysis. It helps explain how factual elements can be systematically reorganised into more persuasive narratives when messages align with targets' trust, values, identity, and emotional sensitivity. In crisis and international disinformation contexts, RBM provides a cognitively grounded analytic scaffold for examining how adversarial narratives gain influence, complementing approaches that focus solely on content, intent, or dissemination.

4.5.5 Limitations

The demonstrations presented in this chapter are analytic constructions intended to illustrate the internal logic of the RBM framework. The examples focus on single factual elements and selected story domains to make the framework's structure and explanatory capacity transparent. While the inclusion of Ukraine biolabs, COVID-19 vaccines, and climate change supports cross-domain illustration, the demonstrations are not intended to represent the full range of misinformation narratives encountered in real-world practice.

This chapter does not involve human participants and therefore does not empirically assess perceived plausibility, confidence, or resistance across different susceptibility profiles. In addition, RBM is not applied here to more complex narrative constructions such as multi-claim or evolving misinformation chains.

These limitations define the scope within which RBM is intended to be used: as a structured analytic method for examining how persuasive narratives are assembled and stabilised under naturalistic conditions and uncertainty.

4.6 Conclusion

After investigating how misinformation is evaluated in practice through participants' accounts of real encounters and decision processes, this chapter shifted attention to the message side of the encounter by developing the RBM framework to examine how misinformation narratives can be constructed through persuasive communication design.

The chapter showed how factual elements can be transformed into persuasive misinformation narratives by coordinating target analysis, message design, framing, and recognition-based plausibility evaluation. It demonstrated that persuasiveness cannot be explained by factual inaccuracy or audience vulnerability alone, but by how selected facts are reframed to shape cue salience, dominant meaning, and recognizable interpretations under conditions of uncertainty. By making this design logic explicit, RBM provides a structured analytic account of how persuasive narratives are assembled so that they can appear coherent, plausible, and action-worthy to a target.

The next chapter brings these strands together by examining how participants evaluate persuasive misinformation narratives developed from the RBM communication patterns, investigating how message construction interacts with recognition-based evaluation during plausibility assessment.

Chapter 5: Tracing Misinformation Evaluation in Practice

Building on the message construction framework developed in Chapter 4, this chapter investigates how persuasive narrative cues are interpreted during recognition-based evaluation when misinformation is encountered in context. It focuses on how prior topic-specific associative orientations and narrative cues interact within this process. 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 prior topic-specific associative orientations. 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-consistent evaluation, 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 evaluation processes, illustrating susceptibility as a context-sensitive evaluation process.

5.1 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 behaviour, political attitudes, and responses to scientific evidence. 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. The challenge is amplified in today's digital ecosystem, where ubiquitous social media enables constant exposure to information and encourages rapid judgments that bypass careful evaluation.

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.

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 et al., 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. These cues may interact with individuals' prior topic-specific associative orientations, guiding how narratives are interpreted and whether they appear plausible or suspicious. 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 behaviours 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.

5.2 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.

5.2.1 Misinformation and Characteristics

Misinformation 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). This chapter does not address orchestrated and coordinated disinformation, as sender intent is outside the scope of analysis. Readers are referred to other sources for a more comprehensive treatment of the subject (Freelon & Wells, 2020; Tucker et al. , 2018; Wardle, 2018b).

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 exploited through framing techniques, and circulate rapidly in ways that undermine reliability analysis and decision-making (Roozenbeek & van der Linden, 2022).

5.2.2 Susceptibility to Misinformation

Susceptibility 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 misinformation exploits (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.

5.2.3 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 operate as devices that exploit susceptibility: they embed cues that either align with prior orientations, reinforcing belief, or misalign, provoking skepticism and cognitive dissonance. 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.

5.2.4 Persuasion Models

Having established that misinformation is constructed through the use of narratives and frames, the next step is to understand the mechanisms of persuasion models that make such frames take

effect. The literature identifies a set of recurring susceptibility indicators: trust, emotion, values, group identity, thinking style, and knowledge that misinformation frequently targets (Nan et al., 2022). These indicators mark points of susceptibility and serve as entry points for persuasive techniques that aim to bypass analytic reasoning. 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) (Friestad & Wright, 1994) in this study is used to structure narratives where persuasive intent is either explicit or obscured. When intent is made salient, it can trigger skepticism; when embedded within neutral or familiar contexts, it is less likely to be recognized, allowing the message to be processed without resistance. 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) (Petty & Cacioppo, 1986) can be used to construct narrative variations that differed in the evaluation modes. 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, the peripheral route is often exploited through emotional triggers or credibility cues, particularly when individuals are under cognitive load or lack motivation to process information deeply. This allows examination of how different cue configurations influence recognition-based evaluation under constraint.

The Theory of Reasoned Action (TRA) (Fishbein & Ajzen, 2011) can be used to construct narrative variations that emphasize either outcome-related beliefs or perceived social norms. Some narratives highlighted potential consequences associated with the claim, shaping how participants evaluated its implications, while others suggested that certain views were widely

shared or socially endorsed. These variations allowed observation of how participants responded to cues related to expected outcomes and social expectations during recognition-based evaluation, particularly when information was incomplete or contested.

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 accounts for variability in information processing under cognitive constraint, substantially overlapping with HSM and HACA in explaining heuristic and cue-based processing, making the latter two redundant for the purposes of this study. By contrast, PKM and TRA retain distinct explanatory value by addressing persuasion awareness and belief-to-action commitment, respectively, thereby completing a set of mechanisms used to structure the narrative stimuli.

5.2.5 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, amplifying vulnerabilities and reinforcing belief while reducing the likelihood of reflective, analytic engagement. In contemporary practice, persuasive design uses narrative framing through rhetorical strategies such as fear appeals, undermining trust, and moral contrast, which exploit 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.

5.2.6 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.

5.2.7 Decision-Making Models Relevant to Belief Formation

To understand how individuals decide whether to believe or reject misinformation, it is necessary to understand models of decision-making that account for the interaction between predispositions, 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, whereas directional and identity-protective reasoning distorts 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, Type 1 and Type 2 thinking (Kahneman, 2011) differentiate between fast, intuitive, automatic processes (Type 1) and slow, deliberate, analytical processes (Type 2). In persuasion research, the Heuristic–Systematic Model (Chaiken, 1980) uses parallel terminology, contrasting heuristic processing, judgments based on cues such as source credibility, message fluency, or emotional salience, with systematic processing, which involves careful evaluation of message content. While heuristic or

Type 1 processing can sometimes yield accurate conclusions, for instance, when experts such as firefighters, surgeons, or pilots draw on experience to make rapid decisions under pressure, it more often reinforces existing biases and increases susceptibility to misinformation when persuasive cues align with internal predispositions. 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 Type 1 and Type 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 Type 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 Type 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 predispositions converge in the moment of deciding whether to accept or reject a claim.

Together, these decision-making models show that belief in misinformation is shaped by the reasoning pathway engaged (Type 1 vs. Type 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.

5.2.8 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), which treats human reliability as context-dependent cognitive performance in dynamic systems, shifting 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.

5.3 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.

This study received ethical approval from the Auckland University of Technology Ethics Committee (AUTEC) under Approval No. 25/174 (see Ethics Approval section). Participants were provided with an information sheet briefing them on the study procedures, their rights, and confidentiality protections, and gave informed consent prior to participation.

5.3.1 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.

5.3.2 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, in press). 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 emphasized the need for detailed analysis and promoted intuitive judgments based on salient emotional cues. 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.

5.3.3 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).

5.3.4 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.

5.3.5 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, capturing concurrent reactions, recognition, and interpretation. (4) After each narrative encounter, a CDM interview was conducted using probes on emotional responses, imagined consequences and mental simulation, contradiction handling, and confidence in judgment. (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 15 presents an overview of the study design.

Table 15 - Experiment Design Structure to Elicit Cognitive Strategies (CS)

Persuasion Model Applied	**Story 1 (COVID Vaccine)**	**Story 2 (Climate Change)**	**Story 3 (Ukraine Biolab)**
	Study Procedure	*Study Procedure*	*Study Procedure*
	Initial IAT	Initial IAT	Initial IAT
PKM	Think-Aloud CDM Interview Post-IAT Derived CSs	Think-Aloud CDM Interview Post-IAT Derived CSs	Think-Aloud CDM Interview Post-IAT Derived CSs
TRA	Think-Aloud CDM Interview Post-IAT Derived CSs	Think-Aloud CDM Interview Post-IAT Derived CSs	Think-Aloud CDM Interview Post-IAT Derived CSs

ELM	Think-Aloud CDM Interview Post-IAT Derived CSs	Think-Aloud CDM Interview Post-IAT Derived CSs	Think-Aloud CDM Interview Post-IAT Derived CSs

5.3.6 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 iteratively through repeated reading of the transcripts and constant comparison across episodes. Table 16 presents the overview of the codes and the associated emerging themes, which formed the basis for identifying recurring patterns of reasoning across episodes, later formalized as Cognitive Strategies (CSs).

Table 16 - Overview of Codes and Themes Emerged through Thematic Analysis of Think-aloud & CDM

	Excerpt Sample from Transcript	**Codes**	**Themes**	**Description**
Think-Aloud &	"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

CDM Interviews	"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		

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, capturing 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,

representing a higher-level structuring of the evaluation process. These categories structure the presentation of findings in Section 5.4.

To ensure transparency in the analytic process, Table 32 (see Appendix J) 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 15 shows how these CSs were grouped into Aggregated Cognitive Strategies (ACSs) through constant comparison across episodes. While 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 17 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.

5.3.7 Limitation

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. 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. 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.

5.4 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.

5.4.1 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]3.

[3] Transcript citations are structured as [participant, persuasion model, story, page number/line number].

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, judgments were often accepted 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.

5.4.2 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 resolve into unqualified acceptance; 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 favour 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 their confidence firmly in evidence. P4, which showed more leaning toward knowledge and evidence, when discussing vaccines, expressed relatively high certainty:

> *"70 percent 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 percent 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 K. 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.

5.5 Discussion

This study examined how prior topic-specific associative orientations and persuasive narrative cues interact during recognition-based evaluation of misinformation. The findings across the IAT, 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. Table 17 illustrates how selected cognitive strategies were grouped into aggregated cognitive strategies (ACSs) through constant comparison across episodes.

Table 17 - Mapping of selected CSs to Aggregated Cognitive Strategies (ACSs)

	Theme	**CS Cluster (Intermediate)**	**CS IDs**	**Description of Grouping**	**Resulting ACS**
PKM	Recognition	Trust & 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 & heuristic evaluation	CS21, CS2, CS5b	Initial scrutiny of authority cues followed by analytical reasoning and contextual evaluation	ACS8
	Dissonance Reduction	Motivated reasoning & 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

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 18, which outlines their definitions and their relationships to different persuasion models.

Table 18 - 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.

Stimuli	Aggregated CSs	Theme	Definition
	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.

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 labelling content as true or false, toward supporting the reasoning practices that underpin critical evaluation of information.

5.5.1 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 30). 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. Each CS represents a recurring, cross-participant reasoning pattern. 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 18, 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.

In doing so, we investigated how alignment between prior orientations and persuasive framings was associated with the activation of specific cognitive strategies, grounding plausibility evaluation in articulated reasoning processes, and pointing to opportunities for HCI systems to support these processes.

5.5.2 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 the activation of different cognitive strategies. When emotional cues clashed with participants' prior orientations, they were more often associated with defensive filtering and analytical resistance, reflected in ACS12 and ACS13. When emotional cues resonated with existing doubts or concerns, participants were more likely to prolong engagement through outcome simulation, selective accommodation, or heuristic reinforcement, reflected in ACS10, ACS11, ACS15, and ACS16. Across the different framings, this pattern was less about a fixed effect of any one persuasion model than about how emotional cues interacted with prior expectations and the perceived credibility of the message.

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.

5.5.3 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.

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 or 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.

5.5.4 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 to 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 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.

5.5.5 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. Once flagged, confidence often settled more firmly on the participant's initial judgment, as when a vaccine message was judged *("~70% certain"*) because the numerical evidence appeared credible. These moves reflect 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, they 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. This pattern reflects the use of mental simulation (ACS10) and motivated reasoning (ACS6, ACS7), followed by dissonance-reducing strategies (ACS11) that helped maintain coherence with prior orientations. Personally relevant simulations (ACS10) and dissonance-reducing strategies (ACS11) often kept attention engaged while preserving the participant's initial orientation. 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. ELM-framed narratives, particularly those using vivid rhetoric, were associated with shifts in confidence, sometimes through immediate reactions and sometimes through later analytical correction. PKM-framed narratives were more often associated with verification-oriented responses and, when evidence was lacking, more settled rejection. TRA-related identity cues sometimes had counterproductive effects, being associated with defensive readings and at other times with prolonged selective engagement. 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.

5.5.6 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.

5.6 Conclusion

Bringing together Chapter 3's investigation of participants' own real encounters with misinformation and Chapter 4's modelling of how factual elements can be reframed into persuasive narratives, this chapter investigated how participants evaluated persuasive misinformation narratives developed from the RBM communication patterns. The findings showed that participants did not simply accept or reject misinformation; they constructed plausibility through recognition-based evaluation processes shaped by prior topic-specific associative orientations, cue interpretation, mental simulation, and recurring reasoning strategies such as verification, contradiction management, and confidence calibration. Persuasive framing did not determine judgment directly, but influenced which cues became salient, how participants interpreted them, and whether evaluation was continued, revised, resisted, or brought to a close. The findings suggest that susceptibility potentially emerged where persuasive cues aligned with prior orientations in ways that supported plausible interpretations during evaluation under uncertainty.

Taken together, the findings showed that susceptibility emerged through the interaction between persuasive cues, prior orientations, and recognition-based plausibility evaluation, rather than from message content or individual factors alone. This chapter, therefore, connected the message-construction side of the thesis with the evaluation process by showing how persuasive cues were recognized, interpreted, and worked through as participants constructed plausible judgments under uncertainty. The next chapter develops a unified account of susceptibility to misinformation by integrating conceptualizations of recognition-based evaluation, findings from participants' real-world evaluation processes, the RBM account of persuasive message construction, and the interaction between persuasive cues and prior orientations during plausibility assessment. Through this synthesis, the chapter explains how plausibility is constructed, how evaluation is brought to a close under uncertainty, and how susceptibility emerges through the interaction between message design and recognition-based evaluation processes.

Chapter 6: Discussion and Conclusion

This chapter brings together the four studies presented in this thesis to explain how susceptibility to misinformation can be understood through the recognition-based evaluation process. The purpose of this chapter is to synthesize these contributions, showing how the findings of the conducted studies collectively address the research questions, clarify the overall contribution of the thesis, and identify limitations and directions for future work.

6.1 Conceptualizing Susceptibility to Misinformation

Across the thesis, susceptibility is not treated as a fixed trait or a deficit in analytic reasoning. Instead, it is understood as an emergent outcome of evaluation, arising from the interaction between human factors, persuasive message features, and the conditions under which evaluation unfolds.

Within this evaluation process, plausibility operates as the central evaluative criterion. Rather than primarily assessing objective truth, individuals judge whether an interpretation appears coherent, familiar, credible, or meaningful relative to their prior experience and situational constraints. Evaluation is brought to a close when this threshold of plausibility is reached, rather than when certainty is achieved.

The framework also helps clarify the relationship between message design and evaluation. Persuasive messages do not determine belief directly; rather, they embed cues that guide recognition, shape interpretation, and influence how expectations about the message are formed. Susceptibility emerges through how these cues are attended to and interpreted within the evaluation process, shaping whether an interpretation is accepted, rejected, deferred, or subjected to further scrutiny.

Positioned alongside the empirical studies, this account proposes a unifying explanation for variation in responses to the same message. Differences in acceptance, skepticism, or resistance are not treated as distinct phenomena, but as outcomes of the same underlying process operating under different configurations of cues, conditions, and constraints.

Taken together, we suggest that susceptibility can be reframed as a recognition-based plausibility evaluation process shaped jointly by human factors, persuasive message design, and sociotechnical conditions. This shifts the focus from identifying who is susceptible or which messages are misleading, toward understanding how interpretations are formed, maintained, or revised during evaluation under uncertainty.

6.2 From Conceptualization to Evaluation

Building on this conceptualization, the findings show how recognition-based evaluation unfolds at the point of encounter under real-world conditions. Evaluation is not best understood as verification of truth or falsity, nor as exhaustive evidence checking, but as a form of constrained sensemaking under conditions of uncertainty, time pressure, and incomplete knowledge. Individuals construct interpretations by selectively attending to cues, forming expectations, and determining when they have reached a plausible understanding to proceed.

Individuals may recognise similar cues yet diverge in how they interpret them, the consequences they anticipate, and the point at which they conclude evaluation. Variation therefore arises not from message content alone, but from how situations are constructed during evaluation and where the process is brought to a close.

Expectancies shape how evaluation unfolds by guiding what the message is taken to represent, how much checking is required, and what counts as relevant evidence. These expectancies are not fixed; they appear to be revised as new cues are interpreted, particularly when initial interpretations are challenged. This suggests that evaluation may involve iterative cycles in which expectancies are updated until a plausible interpretation is reached.

When uncertainty persists, or the situation is perceived as consequential, participants were observed to engage in mental simulation, constructing possible explanations, inferring intent, and projecting outcomes. Plausibility was therefore assessed not only through evidence, but also through the coherence of these simulated scenarios.

Evaluation is resolved through a stopping mechanism rather than through certainty. Individuals bring evaluation to a close when the situation appears plausible to support action, relative to their goals, knowledge, time constraints, and perceived importance. Differences in where this threshold is reached may help explain variation in outcomes.

This reframes susceptibility within the evaluation process itself. Rather than reflecting deficits or message effects alone, susceptibility emerges from how recognition, expectancy formation, and mental simulation interact, and from the point at which individuals determine that a plausible understanding has been achieved. Persuasion does not directly cause belief; instead, it shapes the cues and frames that guide these processes, influencing the conditions under which particular interpretations become plausible.

Evaluation outcomes are not limited to belief or disbelief. Some evaluation trajectories remained unresolved, where participants deferred judgment or suspended commitment while continuing

to manage uncertainty. Individuals may therefore proceed with action, reject or disengage, continue assessment, or defer judgment, reflecting different ways of managing uncertainty.

Taken together, these findings help provide empirical grounding for the recognition-based framework. The conceptual model describes the structure of evaluation, while this study shows how it operates in practice. The key integration lies in plausibility and stopping: interpretations are constructed under constraint, and evaluation concludes when they appear sufficiently plausible. It is this process, rather than message content alone, that may help explain variation in responses to misinformation.

6.3 From Evaluation to Message Construction

Building on the conceptual and process accounts, the findings suggest how misinformation can be constructed to align with recognition-based evaluation processes. This introduces a complementary perspective by explaining how messages are designed to produce interpretations that appear plausible under conditions of uncertainty.

A key contribution is the suggestion that misinformation does not require false facts. The same factual element can be held constant, and through controlled variation in wording, emphasis, framing, and cue structure, different interpretations may be produced. This shifts the focus from the truth value of content to the conditions under which particular interpretations become plausible.

Within this framework, persuasion models are used as design mechanisms rather than explanatory theories, shaping how cues are embedded and how interpretation is guided. Framing structures how meaning is constructed by organizing problem definition, causality, evaluation, and implied response, making certain interpretations more accessible without introducing new facts.

The critical contribution lies in linking message construction to recognition-based evaluation. Designed cues guide recognition, activate interpretive patterns, and support mental simulation, through which interpretations are evaluated and come to appear plausible. In this way, message design shapes the inputs to the evaluation process rather than determining outcomes directly.

These findings suggest a shift in the understanding of misinformation from something passively encountered to something actively constructed to align with how individuals make sense of uncertain situations. Persuasion does not “inject” belief or exploit fixed vulnerabilities; instead, it shapes the conditions under which particular interpretations are recognized as plausible.

This also helps explain how the same factual element can support multiple interpretations. Differences arise not from the facts themselves, but from how salience, causal structure, and

moral meaning are organized. Plausibility is therefore not inherent to the information, but emerges through the interaction between message design and evaluation.

This perspective complements the process-based account of susceptibility developed earlier. While evaluation explains how individuals arrive at plausible interpretations under constraint, message construction explains how those conditions are shaped. The relationship between the two is therefore integrative. Evaluation determines whether and when plausibility is sufficient to support action, while message design shapes the conditions under which that plausibility emerges.

In this sense, susceptibility is neither imposed by the message nor reducible to individual traits, but emerges from the interaction between human factors, message design, and recognition-based evaluation under uncertainty.

6.4 Bringing Evaluation and Construction Together

Building on the conceptual, process, and construction accounts developed in the previous chapters, the findings bring these perspectives together by showing how prior orientations and persuasive cues interact within recognition-based evaluation processes. Rather than measuring susceptibility as a trait or estimating belief outcomes, the study focuses on how misinformation is evaluated in context, examining how individuals interpret narratives at the point of encounter and how plausibility judgments are constructed through that process.

In this sense, plausibility is not given by the message itself but is actively constructed through reasoning processes that unfold during evaluation. Crucially, evaluation is brought to a close when this constructed plausibility is judged to be sufficient to support action, rather than when certainty is achieved.

A central contribution is the finding that evaluation trajectories depend on the interaction between persuasive cues and prior orientations. Alignment between cues and prior orientations tends to reinforce interpretation and acceptance, whereas misalignment may trigger skepticism, verification, or reinterpretation. Variation, therefore, arises not from message content alone, but from how cues are interpreted within the evaluation process.

This also helps clarify the role of persuasion within the evaluation process. Persuasion models do not produce fixed effects; rather, they shape which cues become salient and which reasoning strategies are activated. Persuasion, therefore, shapes how individuals engage with information, not what they must conclude.

Emotional, authority-based, and identity-relevant cues were observed to prolong engagement when they appeared familiar, personally relevant, or worth further consideration. In contrast, cues

that conflicted with prior knowledge or expectancies more often triggered skepticism, verification, or more analytic forms of assessment. This helps explain how persuasive framing influences evaluation not by determining outcomes directly, but by shaping which cues become salient and how participants engage with them during plausibility construction.

Participants did not rely solely on factual verification. Instead, they constructed and evaluated interpretations through mental simulation, assessed coherence with prior knowledge, reinterpreted or absorbed contradictions, and adjusted their level of confidence. Contradictions did not necessarily disrupt judgments but were often managed or incorporated into existing interpretations. As a result, evaluation outcomes were not binary, but included acceptance, rejection, hesitation, continued evaluation, or provisional uncertainty, reflecting different ways of managing uncertainty.

These findings reinforce the view that susceptibility is a context-sensitive evaluation process. Rather than reflecting stable traits or isolated message effects, it emerges from how individuals recognize cues, interpret them through prior orientations, and construct plausible interpretations during evaluation.

This section completes the integration across the thesis. Chapter 2 conceptualizes susceptibility as a recognition-based evaluation process, which leads to the need to examine how this process operates in practice. Chapter 3 addresses this by tracing how evaluation unfolds under real-world conditions, identifying plausibility and stopping as key mechanisms. These findings then raise a further question: if evaluation is guided by cues, expectations, and mental simulation, how might messages be constructed to shape this process? Chapter 4 responds by demonstrating how factual information can be systematically reframed to produce plausible interpretations. However, this still leaves open how message design and prior orientations interact during evaluation. The final study addresses this by showing how persuasive cues and prior orientations jointly shape evaluation trajectories in context, providing the empirical link between message construction and recognition-based evaluation.

In this way, persuasion does not “inject” belief or directly manipulate individuals. Instead, it operates by influencing which cues are recognized and how they are interpreted within evaluation. It is this interaction between persuasive cues and prior orientations that shapes how plausibility is constructed and when it is sufficient to support judgment under uncertainty.

6.5 Answering the Research Questions

This thesis addressed the four research questions stated in section 1.4 through a connected sequence of conceptual, empirical, analytic, and exploratory studies.

The first research question examined how susceptibility to misinformation can be conceptualized as a recognition-based evaluation process. This is addressed by proposing that susceptibility is best understood as an outcome of the evaluation process itself, rather than as a fixed individual trait or a direct function of message content. The framework developed in Chapter 2 provides a process-level account that integrates human factors, persuasive message design, and sociotechnical conditions, and identifies plausibility as the criterion through which interpretations are judged under uncertainty. This addresses Objective 1 by defining the conceptual basis for analyzing susceptibility as a recognition-based evaluation process.

The second research question examined how individuals move from noticing cues through the evaluation process to forming judgments, and how they determine when to bring this evaluation to a close under conditions of uncertainty. The findings suggest that judgments are not formed through verification, but through a plausibility-based evaluation process that is brought to a close when an interpretation appears plausible to support action. Crucially, the thesis identifies stopping, rather than verification, as the mechanism that resolves evaluation, explaining why similar cues can lead to different outcomes depending on how situations are interpreted and when plausibility is judged to be acceptable. This addressed Objective 2 by establishing plausibility and stopping as the key mechanisms shaping how judgments are formed under uncertainty.

The third research question examined how factual elements can be reframed into persuasive misinformation narratives that guide interpretation and support particular plausible understandings. The findings suggest that plausible interpretations can be systematically produced through the structured design of cues and framing that guide interpretation, without altering the underlying factual content. Rather than determining belief directly, persuasion operates by shaping the inputs to recognition-based evaluation, influencing which interpretations become accessible and appear plausible under uncertainty. This addressed Objective 3 by modelling how message design and framing mechanisms structure the conditions under which particular interpretations emerge.

The fourth research question examined how persuasive narrative cues are interpreted in interaction with prior orientations during recognition-based evaluation of misinformation under structured conditions, as reflected in participants' reasoning. The findings suggest that evaluation trajectories are shaped by the interaction between persuasive cues and prior orientations, rather than by message content alone. Persuasive framings do not determine outcomes directly; instead, they influence which cues become salient and how individuals engage with and interpret cues during evaluation, leading to different interpretations under the same conditions. This

addressed Objective 4 by demonstrating how message design and prior orientations jointly shape evaluation processes in context.

Taken together, these findings show that the four research questions are addressed through a consistent account of misinformation evaluation as a recognition-based process. Across the studies, susceptibility is understood through the interaction between human factors, persuasive message design, and evaluation under conditions of uncertainty, providing a coherent basis for understanding how judgments are formed and resolved in context.

6.6 Overall Contribution

This thesis contributes an account of misinformation susceptibility grounded in recognition-based evaluation. Drawing on the RPD framework, the research makes visible how cue recognition, interpretation, mental simulation, and judgment unfold during plausibility evaluation under uncertainty. It specifies how human factors, persuasive framing, and plausibility judgments interact during decision-making under uncertainty.

The contribution of this research lies in shifting the focus from isolated factors to the evaluation process through which interpretations are formed and brought to a close. Rather than treating susceptibility as a property of individuals, messages, or platforms in isolation, this research suggests how susceptibility arises from their interaction within the evaluation process.

A key contribution is the identification of plausibility as the condition under which evaluation is resolved. The findings suggest that judgments are not formed through verification of truth, but through the point at which an interpretation is considered sufficient to support action. This provides a process-level explanation for how judgments are reached in situations where information is incomplete, contested, or open to interpretation.

A further contribution is the clarification of the role of persuasive framing. This thesis argues that persuasion does not directly determine belief. Instead, it shapes the cues that become salient and the conditions under which particular interpretations are recognized as plausible. This distinction repositions the tension between persuasion and individual judgment by showing that persuasive design shapes the inputs to the evaluation process, rather than determining its outcomes directly.

Another contribution is the suggestion that plausible misinformation does not depend on false facts. Through analytic modelling, the research shows how factual elements can be reframed to support different interpretations by altering emphasis, causal structure, and evaluative framing. This extends existing approaches by explaining how technically accurate information can produce misleading yet plausible understandings.

An additional contribution is the empirical finding that evaluation outcomes vary as a function of cue–orientation interaction. The studies show that the same cues do not lead to the same responses. Instead, outcomes depend on how cues are interpreted in relation to prior orientations, leading to different evaluation trajectories under similar conditions. This provides a process-based explanation for divergence in responses to the same information.

Taken together, these contributions support the central argument of the thesis: that susceptibility to misinformation is best understood through recognition-based evaluation, where human factors, persuasive framing, and plausibility judgments interact to shape how interpretations are formed and when they are accepted under conditions of uncertainty.

Beyond explaining susceptibility, the framework also identifies where future support mechanisms may be directed. By making visible how plausibility is constructed and how evaluation is brought to a close, the findings provide a basis for investigating interventions that target the evaluation process itself rather than relying solely on content correction or misinformation detection. In this respect, the contribution of the thesis extends beyond explanation toward identifying theoretically grounded opportunities for future HCI and cognitive engineering research.

6.7 Limitations

The limitations of each study in this research have been discussed in the associated chapters. The key limitation concerns the scope of the account developed in this thesis. While the research provides a detailed explanation of how plausibility is constructed during evaluation, it does not establish how these processes operate across broader populations or under more complex real-world conditions. As such, the findings show how evaluation can occur, but do not determine how consistently these patterns hold across different contexts, audiences, or settings.

The research also focuses on evaluation at the point of encounter. It therefore explains how plausibility judgments are formed during evaluation, but it does not examine how these judgments develop into longer-term beliefs, behavioral change, or repeated patterns of sharing and engagement over time.

These limitations define the boundaries of the research contribution. The findings should therefore be interpreted as explaining how plausibility is constructed and evaluation is brought to a close under uncertainty, rather than as providing predictive or population-level claims.

6.8 Future Research

A key direction for future research is the empirical examination of the processes identified in this thesis under broader conditions. This includes testing how variations in message design, cue structure, and prior orientations influence evaluation across larger and more diverse samples.

Further work is also needed to extend the analysis beyond single encounters. Examining how initial judgments are maintained, revised, or reinforced over time would provide a more complete account of how evaluation relates to longer-term belief and behaviour.

In addition, the analytic approach developed in this thesis can be extended by operationalizing core elements of the evaluation process. This would support more systematic comparison across contexts and contribute to the development of more robust empirical models.

The process account developed in this thesis also identifies potential intervention points that warrant further investigation. While this research did not design or evaluate interventions, future work could examine how support mechanisms influence recognition-based evaluation before plausibility becomes sufficient to support judgment. One direction is to make persuasive construction visible by drawing attention to framing strategies, cue structures, persuasive techniques, source characteristics, and dominant interpretations embedded within messages. A second direction is to support evaluation process by encouraging activities such as verification, consideration of alternative explanations, contradiction checking, and consequence simulation before evaluation is brought to a close. A third direction is to support reflective judgment by helping individuals calibrate confidence in their interpretations and consider what evidence might justify maintaining or revising a current conclusion. Collectively, these directions derive directly from the evaluation process identified in this thesis and provide a basis for future HCI and cognitive system engineering research aimed at supporting information evaluation under uncertainty.

References

Abdali, S., Bastidas, G. G., Shah, N., & Papalexakis, E. E. (2020). Tensor Embeddings for Content-Based Misinformation Detection with Limited Supervision. *Disinformation, Misinformation, and Fake News in Social Media: Emerging Research Challenges and Opportunities*, 117-140.

Aho, J. A. (1985). Rhetoric and the invention of double entry bookkeeping. *Rhetorica, 3*(1), 21-43.

Ajzen, I. (1991). The Theory of planned behavior. *Organizational behavior and human decision processes*.

Ashby, W. R., & Goldstein, J. (2011). Variety, constraint, and the law of requisite variety. *Emergence: Complexity and Organization, 13*(1/2), 190.

Bârgăoanu, A., & Pană, M. (2024). Cyber Influence Defense: Applying the DISARM Framework to a Cognitive Hacking Case from the Romanian Digital Space [journal article]. *Applied Cybersecurity & Internet Governance, 3*(1), 91-121. doi:10.60097/acig/190196

Barr, N., Pennycook, G., Stolz, J. A., & Fugelsang, J. A. (2015). The brain in your pocket: Evidence that Smartphones are used to supplant thinking. *Computers in Human Behavior, 48*, 473-480.

Barros, G. (2007). Herbert A. Simon e o conceito de racionalidade: Limites e procedimentos (Herbert A. Simon and the Concept of Rationality: Boundaries and Procedures). *Behavioral & Experimental Economics eJournal*.

Blandford, A., & Wong, B. W. (2004). Situation awareness in emergency medical dispatch. *International journal of human-computer studies, 61*(4), 421-452.

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.

Bohner, G., Moskowitz, G. B., & Chaiken, S. (1995). The interplay of heuristic and systematic processing of social information. *European Review of Social Psychology, 6*(1), 33-68.

Bolatito, O. (2012). Linkage between Persuasion principles and Advertising. *New media and mass communication, 8*, 7-11.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., . . . Askell, A. (2020). Language models are few-shot learners. *Advances in neural information processing systems, 33*, 1877-1901.

Brummett, B. (2000). Reading rhetorical theory. *(No Title)*.

Cacioppo, J. T., Cacioppo, S., & Petty, R. E. (2018). The neuroscience of persuasion: A review with an emphasis on issues and opportunities. *Social neuroscience, 13*(2), 129-172.

Calhoun, J., Savoie, C., Randolph-Gips, M., & Bozkurt, I. (2014). Human reliability analysis in spaceflight applications, part 2: modified CREAM for spaceflight. *Quality and reliability engineering international, 30*(1), 3-12.

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

Chaiken, S. (1989). Heuristic and systematic information processing within and beyond the persuasion context. *Unintended thought*, 212-252.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context.

Cialdini, R. B. (2007). *Influence: The psychology of persuasion* (Vol. 55): Collins New York.

Cook, M. B., & Smallman, H. S. (2008). Human factors of the confirmation bias in intelligence analysis: Decision support from graphical evidence landscapes. *Human factors, 50*(5), 745-754.

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

Danry, V., Pataranutaporn, P., Groh, M., & Epstein, Z. (2025). *Deceptive Explanations by Large Language Models Lead People to Change their Beliefs About Misinformation More Often than Honest Explanations*. presented at the meeting of the Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, Retrieved from https://doi.org/10.1145/3706598.3713408 doi:10.1145/3706598.3713408

DeCarlo, T. E. (2005). The effects of sales message and suspicion of ulterior motives on salesperson evaluation. *Journal of Consumer Psychology, 15*(3), 238-249.

Deline, S., Guillet, L., Rauffet, P., & Guérin, C. (2021). Team cognition in a cyber defense context: focus on social support behaviors. Cognition, Technology & Work, 23(1), 51-63. doi:10.1007/s10111-019-00614-y

Dentith, M. R., & Orr, M. (2018). Secrecy and conspiracy. *Episteme, 15*(4), 433-450.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018a). Bert: Bidirectional encoder representations from transformers. *arXiv preprint arXiv:1810.04805*.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018b). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Dhami, M. K., & Thomson, M. E. (2012). On the relevance of Cognitive Continuum Theory and quasirationality for understanding management judgment and decision making. *European Management Journal, 30*(4), 316-326.

Di Pasquale, V., Miranda, S., Iannone, R., & Riemma, S. (2015). Simulative analysis of performance shaping factors impact on human reliability in manufacturing activities*Affenzeller, M.; Longo, F.; Zhang, L.; Bruzzone, AG; Merkuryev, Y.; Jimenez, E*. Symposium conducted at the meeting of the 27th European Modeling and Simulation Symposium, EMSS 2015

Ditto, P. H., Liu, B. S., Clark, C. J., Wojcik, S. P., Chen, E. E., Grady, R. H., . . . Zinger, J. F. (2019). At least bias is bipartisan: A meta-analytic comparison of partisan bias in liberals and conservatives. *Perspectives on Psychological Science, 14*(2), 273-291.

Dix, A. (2020). Statistics for HCI: Making Sense of Quantitative Data. *Synthesis Lectures on Human-Centered Informatics, 13*(2), 1-181.

Downs, J. S., Holbrook, M., & Cranor, L. F. (2007). Behavioral response to phishing risk Symposium conducted at the meeting of the Proceedings of the anti-phishing working groups 2nd annual eCrime researchers summit

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

Druckman, J. N., & McGrath, M. C. (2019). The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change, 9*(2), 111-119.

Dweck, C. S., Chiu, C.-y., & Hong, Y.-y. (1995). Implicit theories: Elaboration and extension of the model. *Psychological inquiry, 6*(4), 322-333.

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., . . . Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology, 1*(1), 13-29. doi:10.1038/s44159-021-00006-y

ElBaff, R., Wachsmuth, H., Al Khatib, K., & Stein, B. (2020). Analyzing the persuasive effect of style in news editorial argumentation*Association for Computational Linguistics*.

Ellis, R. D., & Newton, N. (2005). *Consciousness & Emotion: Agency, conscious choice, and selective perception* (Vol. 1): John Benjamins Publishing.

Enders, A. M., & Smallpage, S. M. (2018). Polls, plots, and party politics: Conspiracy theories in contemporary America. *Conspiracy theories and the people who believe them*, 298-318.

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of communication, 43*(4), 51-58.

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.

Epstein, Z., Pennycook, G., & Rand, D. (2020). *Will the Crowd Game the Algorithm? Using Layperson Judgments to Combat Misinformation on Social Media by Downranking Distrusted Sources*. presented at the meeting of the Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA. Retrieved from https://doi.org/10.1145/3313831.3376232 doi:10.1145/3313831.3376232

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.

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*. doi:https://doi.org/10.2759/739290

Eveland Jr, W. P., Shah, D. V., & Kwak, N. (2003). Assessing causality in the cognitive mediation model: A panel study of motivations, information processing, and learning during campaign 2000. *Communication research, 30*(4), 359-386.

Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*: Psychology Press.

Fazio, L. K., Brashier, N. M., Payne, B. K., & Marsh, E. J. (2015). Knowledge does not protect against illusory truth. *Journal of experimental psychology: general, 144*(5), 993.

Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., . . . Prager, J. (2010). Building Watson: An overview of the DeepQA project. *AI magazine, 31*(3), 59-79.

Festinger, L. (1957). A theory of cognitive dissonance Row. *Peterson and company*.

Fielding, K. S., & Hornsey, M. J. (2016). A social identity analysis of climate change and environmental attitudes and behaviors: Insights and opportunities. *Frontiers in Psychology, 7*, 121.

Fiori, S. (2008). Herbert A. Simon and contemporary theories of bounded rationality. *U. of Torino Department of Economics Research Paper*(2).

Fischer, P., Schulz-Hardt, S., & Frey, D. (2008). Selective exposure and information quantity: how different information quantities moderate decision makers' preference for consistent and inconsistent information. *Journal of personality and social psychology, 94*(2), 231.

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

Forgas, J. P., Cooper, J., & Crano, W. D. (2011). *The psychology of attitudes and attitude change*: Psychology Press.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives, 19*(4), 25-42.

Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: the foot-in-the-door technique. *Journal of personality and social psychology, 4*(2), 195.

Freelon, D., & Wells, C. (2020). Disinformation as political communication (Vol. 37, pp. 145-156): Taylor & Francis.

Friestad, M., & Wright, P. (1994). The Persuasion Knowledge Model: How People Cope with Persuasion Attempts. *Journal of Consumer Research, 21*(1), 1-31. doi: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.

Garrett, R. K., Weeks, B. E., & Neo, R. L. (2016). Driving a wedge between evidence and beliefs: How online ideological news exposure promotes political misperceptions. *Journal of Computer-Mediated Communication, 21*(5), 331-348.

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

Golec de Zavala, A., Dyduch-Hazar, K., & Lantos, D. (2019). Collective narcissism: Political consequences of investing self-worth in the ingroup's image. *Political Psychology, 40*, 37-74.

González, F. S., Pastor-Galindo, J., & Ruipérez-Valiente, J. A. (2025). Toward interoperable representation and sharing of disinformation incidents in cyber threat intelligence. *arXiv preprint arXiv:2502.20997*.

Google. (n.d.). *Fact check (ClaimReview) structured data*. *Google Search Central*. *Retrieved April 26, 2026, from* https://developers.google.com/search/docs/appearance/structured-data/factcheck.

Govender, R., Bowen, P., Edwards, P., & Cattell, K. (2017). AIDS-related knowledge, stigma and customary beliefs of South African construction workers. *AIDS care, 29*(6), 711-717.

Grazioli, S. (2004). Where did they go wrong? An analysis of the failure of knowledgeable internet consumers to detect deception over the internet. *Group Decision and Negotiation, 13*, 149-172.

Green Jr, S. E. (2004). A rhetorical theory of diffusion. *Academy of management review, 29*(4), 653-669.

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.

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.

Haghighi, B. T., & Wong, B. L. W. (in press). 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.

Hammond, K. R. (1981). Principles of organization in intuition and analytical cognition. Report 231. C. f. R. o. *J. a. Policy. Boulder, University of Colorado*.

Hammond, K. R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*: Oxford University Press.

Hammond, K. R., Hamm, R. M., Grassia, J., & Pearson, T. (1987). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. *IEEE Transactions on systems, man, and cybernetics, 17*(5), 753-770.

Hanselowski, A., PVS, A., Schiller, B., Caspelherr, F., Chaudhuri, D., Meyer, C. M., & Gurevych, I. (2018). A retrospective analysis of the fake news challenge stance detection task. *arXiv preprint arXiv:1806.05180*.

Harambam, J., & Aupers, S. (2017). 'I am not a conspiracy theorist': Relational identifications in the Dutch conspiracy milieu. *Cultural Sociology, 11*(1), 113-129.

Harmon-Jones, E. E. (2019). *Cognitive dissonance: Reexamining a pivotal theory in psychology*: American Psychological Association.

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

Hassan, N., Zhang, G., Arslan, F., Caraballo, J., Jimenez, D., Gawsane, S., . . . Nayak, A. K. (2017). Claimbuster: The first-ever end-to-end fact-checking system. *Proceedings of the VLDB Endowment, 10*(12), 1945-1948.

Haugtvedt, C. P., Herr, P. M., & Kardes, F. R. (2018). I know what you're doing and why you're doing it: the use of persuasion knowledge model in consumer research: margaret C. Campbell and Amna Kirmani. In *Handbook of Consumer Psychology* (pp. 553-575): Routledge.

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

Holt, R., & Macpherson, A. (2010). Sensemaking, rhetoric and the socially competent entrepreneur. *International small business journal, 28*(1), 20-42.

Hornsey, M. J., Harris, E. A., & Fielding, K. S. (2018). The psychological roots of anti-vaccination attitudes: A 24-nation investigation. *Health Psychology, 37*(4), 307.

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, 0*(0), 01461672251328800. doi:10.1177/01461672251328800

Hunt, R. M., & Rouse, W. B. (1981). Problem-solving skills of maintenance trainees in diagnosing faults in simulated powerplants. *Human factors, 23*(3), 317-328.

Hutchins, E. (2020). The distributed cognition perspective on human interaction. In *Roots of human sociality* (pp. 375-398): Routledge.

Hyde, M. J. (2004). *The ethos of rhetoric*: Univ of South Carolina Press.

Illia, L., Zyglidopoulos, S., & Bantimaroudis, P. (2026). Staging the Lie: The Impact of Framing and Content on the Visibility of Fake Business News. Business & Society, 65(2), 347-388. doi:10.1177/00076503241298093.

Joslyn, S., Savelli, S., & Nadav-Greenberg, L. (2011). Reducing probabilistic weather forecasts to the worst-case scenario: Anchoring effects. *Journal of Experimental Psychology: Applied, 17*(4), 342.

Kahan, D. M. (2017). Misconceptions, misinformation, and the logic of identity-protective cognition.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American economic review, 93*(5), 1449-1475.

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

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment, 49*(49-81), 74.

Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. *The Cambridge handbook of thinking and reasoning, 267*, 293.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*: Cambridge university press.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American psychologist, 39*(4), 341.

Kamali, D., Romain, J. D., Liu, H., Peng, W., Meng, J., & Kordjamshidi, P. (2024). Using persuasive writing strategies to explain and detect health misinformation Symposium conducted at the meeting of the Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)

Keeley, B. L. (2019). Of conspiracy theories. In *Conspiracy theories* (pp. 45-60): Routledge.

Klayman, J. (1995). Varieties of confirmation bias. *Psychology of learning and motivation, 32*, 385-418.

Klein, G., Moon, B., & Hoffman, R. (2006a). Making sense of sensemaking 1: Alternative perspectives. *IEEE intelligent systems, 21*(4), 70-73.

Klein, G., Moon, B., & Hoffman, R. (2006b). Making sense of sensemaking 2: A macrocognitive model. *IEEE intelligent systems, 21*(5), 88-92.

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. (1997). The recognition-primed decision (RPD) model: Looking back, looking forward. *Naturalistic decision making, 285*, 292.

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.

Klein, G. A., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking Symposium conducted at the meeting of the Expertise out of context: Proceedings of the sixth international conference on naturalistic decision making

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

Klein, H. K., & Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS quarterly*, 67-93.

Kumar, S., Gupta, K., Kumar, A., Singh, A., & Singh, R. K. (2023). Applying the theory of reasoned action to examine consumers' attitude and willingness to purchase organic foods

[Article]. *International Journal of Consumer Studies, 47*(1), 118-135-135. doi:10.1111/ijcs.12812

Kumar, Y., Jha, R., Gupta, A., Aggarwal, M., Garg, A., Malyan, T., . . . Chen, C. C. (2023). Persuasion strategies in advertisements Symposium conducted at the meeting of the Proceedings of the AAAI Conference on Artificial Intelligence

Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin, 108*(3), 480.

Laursen, B., & Faur, S. (2022). What does it mean to be susceptible to influence? A brief primer on peer conformity and developmental changes that affect it. *International Journal of Behavioral Development, 46*(3), 222-237.

Leach, C., Van Zomeren, M., Zebel, S., Vliek, M., Pennekamp, S., Doosje, B., . . . Spears, R. (2008). Collective self-definition and self-investment: A two-dimensional framework of group identification. *Journal of personality and social psychology, 95*, 144-165.

Levesque, N., & Pons, F. (2020). The human brand: A systematic literature review and research agenda. *Journal of Customer Behaviour, 19*(2), 143-174.

Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of personality and social psychology, 74*(6), 1421.

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.

Liang, B., Li, H., Su, M., Bian, P., Li, X., & Shi, W. (2017). Deep text classification can be fooled. *arXiv preprint arXiv:1704.08006*.

Lilly, G. A. (1994). Bounded rationality: A Simon-like explication. *Journal of Economic Dynamics and Control, 18*, 205-230.

Liu, Y., & Wu, Y.-F. B. (2020). Fned: a deep network for fake news early detection on social media. *ACM Transactions on Information Systems (TOIS), 38*(3), 1-33.

Longpre, L., Durmus, E., & Cardie, C. (2019). Persuasion of the Undecided: Language vs. the Listener Symposium conducted at the meeting of the Proceedings of the 6th Workshop on Argument Mining

Lyons, B., Merola, V., & Reifler, J. (2019). Not just asking questions: Effects of implicit and explicit conspiracy information about vaccines and genetic modification. *Health communication, 34*(14), 1741-1750.

Maertens, R., Götz, F. M., Golino, H. F., Roozenbeek, J., Schneider, C. R., Kyrychenko, Y., . . . van der Linden, S. (2023). The Misinformation Susceptibility Test (MIST): A psychometrically validated measure of news veracity discernment. *Behavior Research Methods*. doi:10.3758/s13428-023-02124-2

Martell, M. J., Baweja, J. A., & Dreslin, B. D. (2025). Mitigative Strategies for Recovering From Large Language Model Trust Violations. Journal of Cognitive Engineering and Decision Making, 19(1), 76-95. doi:10.1177/15553434241303577

McConnell, A. R. (2001). Implicit theories: Consequences for social judgments of individuals. *Journal of Experimental Social Psychology, 37*(3), 215-227.

McFedries, P. (2014). Stop, attention thief![Technically Speaking]. *IEEE Spectrum, 6*(51), 28.

McGuire, W. J. (1964). Inducing resistance to persuasion. Some contemporary approaches. *CC Haaland and WO Kaelber (Eds.), Self and Society. An Anthology of Readings, Lexington, Mass.(Ginn Custom Publishing) 1981, pp. 192-230*.

Miller, J. M. (2020). Psychological, political, and situational factors combine to boost COVID-19 conspiracy theory beliefs. *Canadian Journal of Political Science/Revue canadienne de science politique, 53*(2), 327-334.

MISP-Project. *MISP: Open source threat intelligence platform*. Retrieved from https://www.misp-project.org/

Morrison, B. W., Graf, E., Bayl-Smith, P., & Wiggins, M. W. (2025). Like Shooting Phish in a Barrel: Cue Utilization and Cognitive Reflection Aid Performance in Controlled, but Not Naturalistic Phishing Tasks. Journal of Cognitive Engineering and Decision Making, 19(1), 32-53.

Mosier, K. L., & Fischer, U. (2010). The role of affect in naturalistic decision making. *Journal of Cognitive Engineering and Decision Making, 4*(3), 240-255.
Mustafa, H., Luczak-Roesch, M., & Johnstone, D. (2025). Conceptualizing the evolving nature of computational propaganda: a systematic literature review. Annals of the International Communication Association, 49(1), 45-60. doi:10.1093/anncom/wlaf001.
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*, 115398.
Nasir, J. A., Khan, O. S., & Varlamis, I. (2021). Fake news detection: A hybrid CNN-RNN based deep learning approach. *International Journal of Information Management Data Insights, 1*(1), 100007.
Nyhan, B., & Reifler, J. (2010). When Corrections Fail: The Persistence of Political Misperceptions [research-article]. *Political Behavior, 32*(2), 303-330.
Oeppen Hill, J. H. (2020). Logos, ethos, pathos and the marketing of higher education. *Journal of Marketing for Higher Education, 30*(1), 87-104. doi:10.1080/08841241.2019.1683120
Oinas-Kukkonen, H., & Harjumaa, M. (2009). Persuasive Systems Design: Key Issues, Process Model, and System Features. Communications of the Association for Information Systems, 24, pp-pp. https://doi.org/10.17705/1CAIS.02428
Onorato, R. S., & Turner, J. C. (2004). Fluidity in the self-concept: the shift from personal to social identity. *European journal of social psychology, 34*(3), 257-278.
Orasanu, J., & Connolly, T. (1993). The reinvention of decision making. *Decision making in action: Models and methods, 1*, 3-20.
Orasanu, J., & Fischer, U. (1997). Finding decisions in natural environments: The view from the cockpit. *Naturalistic decision making*, 343-357.
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.
Oswald, M. E., & Grosjean, S. (2004). Confirmation bias. *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory, 79*, 83.
Pan, X., & Wu, Z. (2020). Performance shaping factors in the human error probability modification of human reliability analysis. *International Journal of Occupational Safety and Ergonomics, 26*(3), 538-550.
Park, K. S., & Jung, K. T. (1996). Considering performance shaping factors in situation-specific human error probabilities. *International Journal of Industrial Ergonomics, 18*(4), 325-331.
Pelrine, K., Danovitch, J., & Rabbany, R. (2021). The surprising performance of simple baselines for misinformation detection Symposium conducted at the meeting of the Proceedings of the Web Conference 2021
Peng, W., Meng, J., & Issaka, B. (2024). Navigating persuasive strategies in online health misinformation: an interview study with older adults on misinformation management. *PloS one, 19*(7), e0307771.
Pennycook, G., Bear, A., Collins, E. T., & David, G. R. (2020). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management science, 66*(11), 4944-4957.
Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature, 592*(7855), 590-595.
Pennycook, G., & Rand, D. G. (2019a). Fighting misinformation on social media using crowdsourced judgments of news source quality [research-article]. *Proceedings of the National Academy of Sciences of the United States of America, 116*(7), 2521-2526.
Pennycook, G., & Rand, D. G. (2019b). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition, 188*, 39-50.
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.

Pennycook, G., & Rand, D. G. (2021). The Psychology of Fake News. *Trends in cognitive sciences, 25*(5), 388-402. doi:https://doi.org/10.1016/j.tics.2021.02.007

Pérez–Acuña, B., & Fernández-Aller, C. (2021). Facebook and artificial intelligence: a review of good practices. *AI and Ethics, 1*(3), 421-435.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological science, 17*(5), 407-413.

Petty, R., & Wegener, D. (1999). The elaboration likelihood model: Current status and controversies In Dual-process theories in social psychology (pp. 37–72). *New York, NY: Guilford Press.[Google Scholar]*.

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

Pickles, K., Cvejic, E., Nickel, B., Copp, T., Bonner, C., Leask, J., . . . Dakin, T. (2021). COVID-19 misinformation trends in Australia: prospective longitudinal national survey. *Journal of medical Internet research, 23*(1), e23805.

Pollini, A., Callari, T. C., Tedeschi, A., Ruscio, D., Save, L., Chiarugi, F., & Guerri, D. (2022). Leveraging human factors in cybersecurity: an integrated methodological approach. Cognition, Technology & Work, 24(2), 371-390. doi:10.1007/s10111-021-00683-y

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., . . . Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of machine learning research, 21*(140), 1-67.

Rand, D. G., Tomlin, D., Bear, A., Ludvig, E. A., & Cohen, J. D. (2017). Cyclical population dynamics of automatic versus controlled processing: An evolutionary pendulum. *Psychological review, 124*(5), 626.

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

Rasmussen, J. (1981). Models of mental strategies in process plant diagnosis. In *Human detection and diagnosis of system failures* (pp. 241-258): Springer.

Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Flammini, A., & Menczer, F. (2011). Detecting and tracking political abuse in social media Symposium conducted at the meeting of the Proceedings of the International AAAI Conference on Web and social media

Redström, J. (2006). Persuasive design: Fringes and foundations*Springer*. Symposium conducted at the meeting of the International Conference on Persuasive Technology

Roetzel, P. G. (2019). Information overload in the information age: a review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development. *Business research, 12*(2), 479-522.

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.

Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L. J., Recchia, G., . . . van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society open science, 7*(10), 201199.

Roozenbeek, J., Suiter, J., & Culloty, E. (2022). Countering misinformation: Evidence, knowledge gaps, and implications of current interventions.

Roozenbeek, J., & van der Linden, S. (2019a). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications, 5*(1), 1-10.

Roozenbeek, J., & van der Linden, S. (2019b). The fake news game: actively inoculating against the risk of misinformation [Article]. *Journal of Risk Research, 22*(5), 570-580-580. doi: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 (Vol. 36, pp. 569-575): SAGE Publications Sage CA: Los Angeles, CA.

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.

Rubinelli, S., & Diviani, N. (2025). An argumentation theory-based assessment tool for evaluating disinformation in health-related claims. *Patient Education and Counseling, 133*, 108622. doi:https://doi.org/10.1016/j.pec.2024.108622

Sadler, N. (2025). Suspicious stories: taking narrative seriously in disinformation research. Communication Theory. doi:10.1093/ct/qtaf013.

Savolainen, R. (2015). Cognitive barriers to information seeking: A conceptual analysis. *Journal of Information Science, 41*(5), 613-623.

Schemer, C., Wirth, W., & Matthes, J. (2012). Value resonance and value framing effects on voting intentions in direct-democratic campaigns. *American Behavioral Scientist, 56*(3), 334-352.

Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of communication, 49*(1), 103-122.

Schkade, D. A., & Kleinmuntz, D. N. (1994). Information displays and choice processes: Differential effects of organization, form, and sequence. *Organizational behavior and human decision processes, 57*(3), 319-337.

Seo, H., Blomberg, M., Altschwager, D., & Vu, H. T. (2021). Vulnerable populations and misinformation: A mixed-methods approach to underserved older adults' online information assessment. *New Media & Society, 23*(7), 2012-2033.

Shaikh, O., Chen, J., Saad-Falcon, J., Chau, D. H., & Yang, D. (2020). Examining the ordering of rhetorical strategies in persuasive requests. *arXiv preprint arXiv:2010.04625*.

Shao, C., Ciampaglia, G. L., Flammini, A., & Menczer, F. (2016). Hoaxy: A platform for tracking online misinformation Symposium conducted at the meeting of the Proceedings of the 25th international conference companion on world wide web

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature communications, 9*(1), 1-9.

Shrum, L., Liu, M., Nespoli, M., & Lowrey, T. (2012). Persuasion in the Marketplace: How Theories of Persuasion Apply to Marketing and Advertising. doi:10.4135/9781452218410.n19

Shrum, L. J. (2012). The SAGE Handbook of Persuasion: Developments in Theory and Practice. In (2 ed.). Retrieved from https://sk.sagepub.com/reference/hdbk_persuasion2ed. doi:10.4135/9781452218410

Shu, K., Wang, S., & Liu, H. (2018). Understanding user profiles on social media for fake news detection*IEEE*. Symposium conducted at the meeting of the 2018 IEEE conference on multimedia information processing and retrieval (MIPR)

Simon, H. (1957). A Behavioral Model of Rational Choice. *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting/Wiley*.

Simon, H. A. (1997). *Models of bounded rationality: Empirically grounded economic reason* (Vol. 3): MIT press.

Smith, G. (2025). Learning from Ukraine: how can disinformation be defeated? Retrieved from https://www.bcs.org/articles-opinion-and-research/learning-from-ukraine-how-can-disinformation-be-defeated/?utm_source=chatgpt.com

Speier, C., Valacich, J. S., & Vessey, I. (1999). The influence of task interruption on individual decision making: An information overload perspective. *Decision sciences, 30*(2), 337-360.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers, 31*(1), 137-149.

Stanton, J. M., Stam, K. R., Mastrangelo, P., & Jolton, J. (2005). Analysis of end user security behaviors. *Computers & Security, 24*(2), 124-133.

Stiff, J. B., & Mongeau, P. A. (2016). *Persuasive communication*: Guilford Publications.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American journal of political science, 50*(3), 755-769.

Thomson, C., Nash, J., & Maeder, A. (2016). Persuasive design for behaviour change apps: Issues for designers Symposium conducted at the meeting of the Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists

Tim, W. (2017). The Crisis of Attention Theft-Ads That Steal Your Time for Nothing in Return.

Torning, K., & Oinas-Kukkonen, H. (2009). Persuasive system design: state of the art and future directions Symposium conducted at the meeting of the Proceedings of the 4th international conference on persuasive technology

Traberg, C. S. (2022). Misinformation: broaden definition to curb its societal influence. *Nature, 606*(7915), 653-653.

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.

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 (March 19, 2018)*.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science, 185*(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1986). The framing of decisions and the evaluation of prospects. In *Studies in Logic and the Foundations of Mathematics* (Vol. 114, pp. 503-520): Elsevier.

Uscinski, J. E. (2018). *Conspiracy theories and the people who believe them*: Oxford University Press, USA.

Uscinski, J. E., Klofstad, C., & Atkinson, M. D. (2016). What drives conspiratorial beliefs? The role of informational cues and predispositions. *Political Research Quarterly, 69*(1), 57-71.

Van Bavel, J. J., Harris, E. A., Pärnamets, P., Rathje, S., Doell, K. C., & Tucker, J. A. (2021). Political psychology in the digital (mis) information age: A model of news belief and sharing. *Social Issues and Policy Review, 15*(1), 84-113.

Van Bavel, J. J., & Pereira, A. (2018). The partisan brain: An identity-based model of political belief. *Trends in cognitive sciences, 22*(3), 213-224.

van der Linden, S. (2022). Misinformation: susceptibility, spread, and interventions to immunize the public. *Nature Medicine, 28*(3), 460-467.

van der Linden, S., Roozenbeek, J., & Compton, J. (2020). Inoculating against fake news about COVID-19. *Frontiers in Psychology*, 2928.

Vishwanath, A., Herath, T., Chen, R., Wang, J., & Rao, H. R. (2011). Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems, 51*(3), 576-586.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *science, 359*(6380), 1146-1151.

Vraga, E. K., & Bode, L. (2020). Defining Misinformation and Understanding its Bounded Nature: Using Expertise and Evidence for Describing Misinformation [Article]. *Political Communication, 37*(1), 136-144. doi:10.1080/10584609.2020.1716500

Wang, G., Frederick, R., Haghighi, B. T., Wong, B. L. W., Rupar, V., Li, W., & Bai, Q. (2024). FramedTruth: A Frame-Based Model Utilising Large Language Models for Misinformation Detection*Springer*. Symposium conducted at the meeting of the Asian Conference on Intelligent Information and Database Systems

Wang, J., Herath, T., Chen, R., Vishwanath, A., & Rao, H. R. (2012). Research Article Phishing Susceptibility: An Investigation Into the Processing of a Targeted Spear Phishing Email. *IEEE Transactions on Professional Communication, 55*(4), 345-362. doi:10.1109/TPC.2012.2208392

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.
Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policymaking* (Vol. 27): Council of Europe Strasbourg.
Wardle, C., & Singerman, E. (2021). Too little, too late: social media companies' failure to tackle vaccine misinformation poses a real threat. *BMJ, 372*.
Whitson, J. A., & Galinsky, A. D. (2008). Lacking control increases illusory pattern perception. *science, 322*(5898), 115-117.
Wickens, C. D., Helton, W. S., Hollands, J. G., & Banbury, S. (2021). *Engineering Psychology and Human Performance*. Oxford, UNITED KINGDOM: Taylor & Francis Group. Retrieved from http://ebookcentral.proquest.com/lib/aut/detail.action?docID=6715723
Wickens, C. D., Parasuraman, R., Helton, W. S., Hollands, J. G., & Banbury, S. (2013). *Engineering psychology and human performance* (4th ed.): Pearson.
Winter, S. (2020). Heuristic-systematic model. *The international encyclopedia of media psychology*, 1-6.
Wisse, J. (1989). Ethos and pathos from Aristotle to Cicero.
Wong, B. L. W. (2003). Critical decision method data analysis. *The handbook of task analysis for human-computer interactions. Mahwah: Lawrence Erlbaum Associates*, 327-346.
Wong, B. L. W., & Blandford, A. E. (2002). Analysing ambulance dispatcher decision making: Trialing emergent themes analysis*Ergonomics Society of Australia*. Symposium conducted at the meeting of the Proceedings of the HF2002 Human Factors Conference: Design for the whole person: integrating physical, cognitive and social aspects
Woods, D. D., Pople Jr, H., & Roth, E. (1990). *The cognitive environment simulation as a tool for modeling human performance and reliability*: Nuclear Regulatory Commission, Washington, DC (USA). Div. of Systems ....
Woolley, S. (2020). We're fighting fake news AI bots by using more AI. That'sa mistake. *MIT Technol. Rev. Published online January, 8*, 2020.
Yang, C. C., Chen, H., & Hong, K. (2003). Visualization of large category map for Internet browsing. *Decision Support Systems, 35*(1), 89-102.
Yaqub, W., Kakhidze, O., Brockman, M. L., Memon, N., & Patil, S. (2020). *Effects of Credibility Indicators on Social Media News Sharing Intent*. presented at the meeting of the Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA. Retrieved from https://doi.org/10.1145/3313831.3376213 doi:10.1145/3313831.3376213
Zhou, J., Zhang, Y., Luo, Q., Parker, A. G., & Choudhury, M. D. (2023). *Synthetic Lies: Understanding AI-Generated Misinformation and Evaluating Algorithmic and Human Solutions*. presented at the meeting of the Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany. Retrieved from https://doi.org/10.1145/3544548.3581318 doi:10.1145/3544548.3581318
Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR), 53*(5), 1-40.

Glossary

Full Term	Acronym
Aggregated Cognitive Strategies	ACSs
Cognitive Continuum Theory	CCT
Cognitive Reflection Test	CRT
Cognitive Reliability Analysis	CRA
Cognitive Strategies	CSs
Cognitive Task Analysis	CTA
Critical Decision Method	CDM
Disinformation Indicators Sharing Platform	DISP
Elaboration Likelihood Model	ELM
Heuristic Systematic Model	HSM
Hovland's Attitude Change Approach	HACA
Human Computer Interaction	HCI
Implicit Association Test	IAT
natural language processing	NLP
Naturalistic Decision-Making	NDM
Recognition-Primed Decision model	RPD
Performance Shaping Factors	PSFs
Persuasion Knowledge Model	PKM
Recipe Book of Misinformation	RBM
Recipe Book of Misinformation	RBM
Theory of Reasoned Action	TRA

Appendix A: Review of AI-Based Tools for Misinformation Detection and Analysis

This section presents a review of the key artificial intelligence (AI) agents used to detect and analyze misinformation, as mentioned in Section 2.3.3.

Combating misinformation has increasingly relied on advanced algorithms and AI tools to detect and analyze how information is framed. Technologies such as neural networks and large language models, such as BERT (Liu & Wu, 2020), as well as methods such as Analysis of Topic Model Networks (Walter & Ophir, 2019), are commonly used to identify misinformation. While some of these technologies originated as lab prototypes, they have now been refined for real-world applications, helping media outlets analyze how information is framed and manipulated to prevent the spread of misinformation. These AI agents are no longer just experimental but used in the practical operations of content moderation and fact-checking. However, understanding how different frames of information influence the emergence of misinformation remains a key challenge. The FramedTruth Model (FTM), a new framework that builds on framing theory, aims to address this by identifying cases where factual content is distorted to spread misinformation (Wang et al., 2024).

The examples of key AI agents and tools for combating misinformation related to user-based and content-based strategies are discussed below:

Mainly User-Based: AI tools focus primarily on analyzing user behaviors, social networks, and interactions. However, they may also utilize some content analysis to understand how users engage with misinformation. The following AI agents align with user-based strategies:

1. **ClaimBuster**: By identifying factual claims in real-time during debates, it helps users critically assess the validity of information shared within their social networks (Hassan et al. , 2017).
2. **Hoaxy**: By tracking the dissemination of claims and corresponding fact-checking efforts on social media, allowing users to understand how misinformation spreads through their networks (Shao et al., 2016).
3. **Truthy**: By analyzing user behavior and network structures, Truthy detects coordinated misinformation campaigns, helping to identify less influential users who may be spreading false information (Ratkiewicz et al., 2011).
4. **Google Fact Check Tools**: While primarily content-focused, these tools also engage users by providing fact-checking tags that enhance their understanding of information credibility, such as ClaimReview (Google., n.d.).

5. **IBM Watson**: Although it analyzes content, it can be utilized in a user-based approach by assessing users' linguistic features and context to identify potential misinformation (Ferrucci et al. , 2010).

Mainly Content-Based: These tools focus primarily on the analysis of the information itself, including language, framing, and content structure. They evaluate the actual text for inconsistencies, factual accuracy, and deceptive framing, allowing them to identify misinformation. The following AI agents fit within this classification:

1. **BERT**: This model analyzes the context of words in search queries to identify inconsistencies or deceptive framing, focusing solely on the content (Devlin et al., 2018b).
2. **GPT**: By generating contextually relevant responses and flagging misleading information, GPT engages with the content itself, making it a content-based strategy (Brown et al., 2020).
3. **T5**: Treating every NLP problem as text-to-text, T5 is useful for misinformation detection, focusing on factual consistency and accuracy in content (Raffel et al., 2020).
4. **Facebook's DeepText**: While it uses deep learning to interpret textual content, its primary function is understanding the content's semantics and context, classifying it as a content-based approach (Liang et al. , 2017; Pérez–Acuña & Fernández-Aller, 2021).
5. **Fake News Challenge (FNC)**: This consortium's machine learning models analyze the stance of claims in news articles to detect inconsistencies, focusing on the content of the articles themselves (Hanselowski et al. , 2018).

Appendix B - Persuasive Design Principles and Devices

Following our review of persuasion characteristics and Aristotle's communication strategies (ethos, pathos, logos), this section reviews key persuasive design principles that enhance the effectiveness of communication. These design principles function as mechanisms that structure attention, credibility, and perceived value by shaping which cues become salient, credible, or sufficient for recognition under time pressure and uncertainty. Some of the key persuasive design principles can be listed as:

- ***Authority and Credibility***: Trust is built through expertise, credentials, and reliable information such as statistics, portraying the source as dependable (Cialdini, 2007; Shaikh et al., 2020).
- ***Social Identity and Proof***: By using social validation (e.g., customer reviews), conformity is encouraged, guiding audiences towards what's deemed desirable (Ajzen, 1991; ElBaff et al., 2020).
- ***Accomplishment***: Showcasing achievements or awards boosts confidence in the product by proving its past success (Thomson et al., 2016).
- ***Reciprocity***: Giving something with the expectation of receiving something back creates a sense of obligation, often resulting in favorable responses (Redström, 2006).
- ***Scarcity***: Limited availability increases value and urgency, compelling immediate action (Cialdini, 2007; Redström, 2006).
- ***Anthropomorphism***: Human-like traits in brands foster emotional connection and relatability, making them more engaging (Kumar et al., 2023).
- ***Emotion***: Emotional triggers like joy, fear, and nostalgia drive decisions through evocative imagery and language (Levesque & Pons, 2020).
- ***Foot in the Door***: Starting with small requests builds commitment, increasing the likelihood of agreement to larger requests (Freedman & Fraser, 1966).
- ***Overcoming Resistance***: Managing objections through strategies like forewarning, acknowledging concerns, or building self-esteem helps reduce pushback (McGuire, 1964).
- ***Value and Impact Formulation***: Concrete evidence, comparisons, and societal impact emphasize product value (Kumar et al., 2023).

- ***Framing***: Information is presented in ways that influence perceptions and guide decision-making (Entman, 1993; Kahneman & Tversky, 1984).
- ***Salience***: Highlighting key elements ensures they capture attention and are memorable (Entman, 1993).
- ***Strategic ambiguity***: Ambiguity in messaging can create confusion, but strategically used, it can shape interpretation and perception (Longpre et al., 2019).

By designing a balanced or customized application of Aristotle's strategies, these principles are operationalized in practice to shape audience perceptions and responses.

Appendix C: CDM Interview Guide

This appendix presents the interview guide used in the study, including the general process question and the incident-based probes used to elicit participants' accounts, as follows:

Table 19 - CDM Interview Guide and Probs

No	Questions
1.	To start, please explain what happens once you start reading and analyzing a piece of (mis)information. If I were with you when you start analyzing a piece of (mis)information, what would I see you do?
Explanation	Please consider a more memorable incident. There must have been situations that were challenging to detect information vs misinformation. It may involve a major resource allocation to form the decision to either accept or decline the misinformation. Think back to one of those major incidents that involved a higher level of resource allocation.
2.	Please briefly describe the situation and the key cues that you noticed. What were you seeing and noticing when you got that piece of (mis)information?
3.	Let's talk about the format of each piece of information presented. How did you find the design of the presented information? Have any specifications in UI/UX design elements (ornaments) been considered or highlighted?
4.	To discuss the analysis flow, once you got those cues, what questions did you consider? What information did you need/use to answer your questions and in making the decision?
4.1.	Where did you get this information from (resources)? What were you looking for?
5.	Were you reminded of any previous similar experience? What about that previous experience, seemed relevant to this case?
5.1.	Does this case fit a standard or typical scenario? Is it a type of situation you were trained to deal with?
6.	What were your specific goals at this time? what did you think you had to do? Why?
7.	What courses of action were considered by you? How was the action chosen and others rejected? Was there a rule that you were following in choosing this option? Has it worked in the past?
7.1.	What would have made you decide on a different course of action?
8.	Did you imagine the possible consequences of this action? Did you create some sort of picture in your mind? Did you imagine the situations and how they would unfold?
9.	What specific training or experience was necessary or helpful to make these decisions?
9.1.	If the decision was not the best, what training, knowledge, or information could have helped?
10.	Did you seek any guidance at this point in the incident? How did you know to trust the guidance you got?

11.	How long did it take to actually make this decision?
12.	Suppose you were asked to describe the situation to someone else at this point and hand over the analysis process. How would you summarize the situation?
13.	If a key feature of this situation (or the presented misinformation) had been different, what difference would it have made in your decision?

Appendix D - From Raw Data to Themes

Table 20 provides representative examples illustrating the progression from raw interview extracts to initial codes, sub-themes, and final process-level themes. The initial codes were generated inductively from incident-based segments of the CDM interviews using emergent themes analysis. Related codes were then grouped into sub-themes capturing recurring patterns in how participants recognized cues, formed expectancies, simulated possible outcomes, and determined when to stop evaluating information. These sub-themes were subsequently synthesized into the four process-level themes presented in Section 4: cue recognition and salience, expectancy formation, mental simulation and projection of consequences, and plausibility judgment and stopping points. The selected examples shown are illustrative rather than exhaustive and are intended to demonstrate the traceability of the analytic process.

Table 20 - Illustrative progression of inductive coding from data extracts to emergent process themes

Data extract	**Initial code(s)**	**Sub-theme**	**Final process theme**
"No names, no brands… very brief and very positive!" (OAL, P6, L6)	Missing source details; Absence of identifiable sources	Detecting missing or weak credibility cues	Cue recognition and salience
"I look for red flags… exaggerated language…" (ERJ, P1, L13)	Identifying red flags; Lack of reliable sources		
"That's ridiculous!" (OZH, P10, L21)	Conflict with prior knowledge; Recognizing inconsistency	Recognizing inconsistency with prior knowledge or expectations	
"It didn't make sense… I wasn't convinced" (OSN, P2, L25)	Checking source credibility; Detecting inconsistency		
"Over-advertised… attract my attention" (OMH4, P7, L23)	Repeated exposure; Attention-seeking presentation	Recognizing persuasive and attention-directing presentation cues	
"Highly engaging format… manipulate perceptions" (MAY, P3, L1)	Recognizing persuasive presentation features; Attention to design features		
"Summaries… pretend to make analysis easier" (SS, P15, L7)	Recognizing simplified presentation; Reduced-effort cues	Recognizing simplified and reduced-effort formats	
"Peer-reviewed scientific literature…" (MAN, P13, L1)	Peer-reviewed sources; Reputable sources	Recognizing institutional credibility cues	
"Medical… I can have my own judgment…" (SAN, P1, L28)	Relying on prior knowledge;	Domain familiarity	Expectancy formation

	Anticipating need for checking		
“Same design… that’s fake” (AL, P11, L21)	Recognizing repeated presentation patterns; Expecting misinformation cues	Recognized presentation and rhetorical patterns	
“Cherry-picked data…” (RJ, P2, L23)	Recognizing selective evidence; Identifying manipulation strategies		
“If urgency… call support…” (ZAH, P3, L23)	Adjusting evaluation based on urgency; Seeking help under time pressure	Situational urgency	
“See who benefits…” (SS, P10, L31)	Looking for beneficiaries; Anticipating underlying intent	Beneficiary structures and persuasive intent	
“People like me… I would skip it” (OMH4, P4, L28)	Assessing personal relevance; Judging audience fit	Audience fit and engagement orientation	
“… maybe at the end it doesn’t suit me…” (OMH1, P1, L24)	Continuing evaluation; Reassessing suitability		
“Fact-checking doesn’t change overall view” (MAN, P24, L4)	Recognizing limits of fact-checking; Anticipating persistence of beliefs	Limits of factual correction	
“Correct misinformation… but backlash…” (ERJ, P6, L3)	Imagining backlash and harm	Simulating intervention consequences	Mental simulation
“Prevent harm… but people resist” (EMA, P9, L5)	Projecting downstream effects		
“Sounds like a conspiracy theory” (EMS, P5, L21)	Interpreting through conspiracy narrative	Constructing explanatory and causal scenarios	
“Like police investigation…” (ESS, P10, L34)	Inferring beneficiary-driven causation		
“Actress… must be good… I shared it” (OMH, P7, L24)	Extending endorsement into expected value	Simulating endorsement and audience response	
“Economy will explode…” (OAL, P16, L10)	Comparing alternative future outcomes	Simulating future trajectories and uncertainty	
“Later we will see other angles…” (OMN, P10, L16)	Anticipating evolving interpretations		
“Politics is always framed…” (OZH, P4, L2)	Projecting persuasive motive	Anticipating persuasive intent and framing	
“Once you realize the purpose… you can decide” (OAL, P13, L22)	Identifying purpose; sufficient understanding	Stopping through early recognition and framing	Plausibility judgment & stopping

“Familiar… that’s how you found out” (OZH, P10, L28)	Prior knowledge; recognizing inconsistency		
“No ads… I’d believe it” (OMH3, P4, L16)	Lack of over-promotion; no further checking		
“Checked a few times… then enough” (OSAN, P10, L10)	Verifying claims; enough checking	Stopping after limited or targeted verification	
“Citations are reliable… believe it” (EMS, P23, L1)	Source credibility; reliable citations		
“Contradictions… lose trust” (ESS, P12, L7)	Losing trust; contradictory details	Stopping through contradiction detection	
“Urgency… quick but informed” (EMAY, P10, L36)	Urgency; reassessing actions	Stopping calibrated by context (time, importance, knowledge)	
“Depends on importance and knowledge…” (EMAN, P25, L31)	Importance of claim; available knowledge		

Appendix E. Root Neutral Messages

The neutral root messages written for the two complementary stories are as follows:

COVID-19 Vaccine

COVID-19 vaccines were developed quickly, using new and in some cases novel technologies such as mRNA platforms, to help control the global pandemic [1]. Clinical trials were conducted under emergency protocols, and health agencies continue to monitor vaccine safety data [2]. Vaccines work by stimulating the immune system to provide protection against infection and severe illness [3]. Some individuals experience minor, short-term side effects such as fatigue, soreness, or mild discomfort, which are rare and typically temporary [4]. Health organizations and medical authorities recommend COVID-19 vaccination as a safe and effective tool for reducing serious illness and protecting public health [5]. Natural immunity can develop after infection, but vaccines offer a safer and more controlled way to build immunity [6]. A healthy lifestyle, including good nutrition and sleep, contributes to immune health but does not replace the protection provided by vaccines [7]. The effectiveness of vitamins, herbal remedies, and natural supplements is mixed, and they are not considered substitutes for vaccination [8]. Many people in the community follow official health guidelines, including vaccination, to protect themselves and others [9]. Health professionals encourage individuals to make informed decisions about health interventions based on scientific evidence and trusted medical advice [10]. Vaccination remains a personal decision, and individuals are advised to seek guidance from qualified health providers when considering their options [11].

Climate Change

Climate scientists overwhelmingly agree that recent global warming is primarily driven by human activities, especially greenhouse gas emissions [1]. Anthropogenic activities, such as fossil fuel use, have accelerated climate change and contributed significantly to rising global temperatures [2]. Climate science draws on multiple variables, including both natural and human factors, to understand warming trends [3]. Climate change impacts are intensifying, posing serious risks that require mitigation and long-term planning [4]. Green energy initiatives aim to reduce dependence on fossil fuels, support sustainability, and lower climate impacts [5]. Effective climate mitigation requires coordinated governmental action, technological innovation, and global cooperation [6]. Scientific consensus on human-caused climate change is robust, evidence-based, and supported by environmental and economic data [7]. Climate policies often balance scientific evidence with innovation to support sustainable development [8]. Addressing climate change requires widespread behavioral, societal, and technological changes [9]. Climate communication uses both emotional and rational arguments to engage citizens and encourage action [10]. Climate action is a collective effort involving nations, policymakers, scientists, and communities [11].

Appendix F. Reframed Narratives for All Three Stories

The reframed narrative outcomes for each story are presented below as analytic demonstration texts, produced using the three persuasion models and categorized by story topic and model:

Reframed Narrative about Ukraine Biolab

The full PKM-based reframing of the root message is as follows:

> *For years, governments and mainstream media have denied the existence of secret biolabs in Ukraine, dismissing them as "misinformation". But new leaked documents from a former lab insider tell a different story, one of highly dangerous experiments, concealed research, and global risk. According to whistleblower Dr. Ivan Petrov, classified reports detail the development of bioengineered pathogens under the oversight of Western organizations, raising alarms about potential unintended consequences or deliberate misuse. Despite repeated warnings from independent experts, officials continue to silence dissenting voices, while major news outlets conveniently ignore the evidence. Insider reports, ignored by mainstream media, are now being revealed through trusted alternative platforms. If these labs are as safe as they claim, why the secrecy? Why are government agencies refusing full transparency? With catastrophic consequences at stake, the world deserves the truth before it's too late.*

The full TRA-based reframing of the root message is as follows:

> *For too long, the public has been told to dismiss any claims about secret biolabs in Ukraine as "dangerous conspiracy theories." But what if ignoring the truth is the real danger? A growing number of independent researchers, former officials, and investigative journalists are raising concerns that can no longer be ignored. Dr. Ivan Petrov, a former research consultant, has provided documents suggesting classified U.S.-funded bioweapons research in Ukrainian facilities, an allegation supported by multiple leaked reports from alternative investigative platforms. The mainstream media refuses to cover these findings, and those who question the official story face ridicule and censorship. If there is truly nothing to hide, why are officials avoiding an open investigation? The consequences of staying silent are too great. Will you be someone who blindly follows the official narrative, or will you stand with those who demand transparency and accountability? The choice isn't just about believing or not, it's about whether you choose to be informed or manipulated.*

The full ELM-based reframing of the root message is as follows:

> *For years, the mainstream media and government officials have denied the existence of secret biological research labs in Ukraine, labeling the idea as a mere "conspiracy theory". While officials and mainstream media dismiss concerns about secret biolabs as conspiracy theories, the potential consequences of undisclosed biological research pose serious risks to public safety. A former insider has come forward with shocking revelations, documents that cannot be ignored. Dr. Ivan Petrov, a whistleblower with years of experience in biological research, leaked confidential*

materials that point to dangerous experiments conducted without proper oversight and global consequences looming on the horizon. These documents reveal bioengineered pathogens developed under the watch of Western institutions, sparking fear over the unintended misuse or deliberate release of dangerous strains. While the government insists everything is safe, why are they so secretive? Why are independent voices like Dr. Petrov being silenced? Why does the media continue to ignore the truth? These questions need answers, and the global risks are far too great to remain silent. For those who carefully examine the evidence, the situation is clear: there are powerful interests protecting these secrets. But for the many who don't question the mainstream narrative, it's time to wake up. The truth is out there, and it's terrifying and you have the power to uncover it.

Reframed Narrative about COVID-19 Vaccine

The full PKM-based reframing of the root message is as follows:

Many people are starting to realize that the COVID-19 vaccine might not be all it's cracked up to be. The reality is that natural immunity, which your body builds through exposure, is far more effective than what any vaccine can offer. Dr. Richard Lowe, a leading researcher in immunology, recently stated that "vaccines can't replicate the natural immunity your body builds from real exposure to a virus." Yet, vaccine side effects continue to be swept under the rug by health authorities. Reports of adverse reactions are on the rise, but the government and big pharma continue to downplay the risks, leaving us vulnerable to long-term damage. The Heritage Health Collective, a group of thousands of like-minded individuals, all agree that relying on natural remedies like herbal teas, vitamin D, and zinc supplements is the key to strengthening your immune system without risking any dangerous side effects. Why would we trust something experimental when we have centuries of proven natural remedies? The global health industry may not want you to know this, but it's time to wake up to the reality that your health and freedom should not be compromised by untested vaccines. Don't fall for the government narrative that urges you to trust them blindly. Protect your health naturally—it's safer, it's effective, and it's in your hands.

The full TRA-based reframing of the root message is as follows:

While governments and mainstream health organizations push for the COVID-19 vaccine, many experts are raising concerns about its safety. These vaccines have been rushed through trials, and their long-term effects are still unknown. Dr. Tom Harlaw, a former medical researcher at the University of Chicago, has publicly stated that we are still in the early stages of understanding how these vaccines interact with the immune system. Contrary to popular belief, your body's natural immunity is your best defense against illness and relying on vaccines may actually weaken your immune system. According to a 2022 study published in the Journal of Natural Health, individuals who regularly use natural remedies like vitamin D, zinc, and herbal treatments have stronger immune responses than those who rely on vaccines. Families and communities around the world, like the Heritage Health Collective in Wyoming, have trusted natural health experts for generations,

and it's important to remember that their advice is based on years of proven practice, not quick fixes promoted by pharmaceutical companies. There are numerous stories of people like us who experienced severe side effects... Samantha Green, a 34-year-old mother from Ohio, developed an adverse reaction shortly after receiving the vaccine, and it took her over 6 months to recover. These risks are often swept under the rug. Don't let fear drive you to make an unsafe decision. Trust your body's natural defenses, and seek guidance from those who understand the truth behind natural health. Your health is in your hands, and it's time to choose what's truly safe and effective.

The full ELM-based reframing of the root message is as follows:

Do you really want to risk your health with something that hasn't been properly tested? More and more families are questioning the safety of the COVID-19 vaccine. Reports of adverse reactions are increasing, but the mainstream media and pharmaceutical companies are turning a blind eye to extremely dangerous consequences, creating a situation that is fast becoming a disaster waiting to happen. Instead of trusting the big pharmaceutical companies, consider the voices of natural health experts like Dr. Robert Gante, who has spent decades studying how your body's natural immune system is the best defense against diseases like COVID-19. Natural remedies like vitamin C, elderberry, and other immune-boosting supplements are proven to keep your body strong without the dangers of experimental vaccines. People just like you, families, parents, and communities, are opting out of the vaccine and turning to what works: natural health. Doesn't it feel good to know that you're making the right choice for your family? Join the growing movement of people who refuse to fall for the media agenda. Your health, your choice.

Reframed Narrative about Climate Change

The full PKM-based reframing of the root message is as follows:

Dr. Thomas Greer, a well-respected climate scientist with over 30 years of research, has consistently stated that solar variability and natural cycles are primarily responsible for the warming trends we see today. Research shows that 80% of the warming over the last century can be attributed to solar activity and natural Earth processes, not human activity. Natural changes in the Earth's orbit, the Sun's cycles, and volcanic activity have shaped the planet's climate for millions of years. The current panic surrounding carbon emissions and green energy policies is largely misguided and often pushed by industries with vested interests in renewable energy. Countries like Norway and Germany have shown that green energy policies, pushed by elites, can lead to job losses, economic instability, and higher energy costs. It's time to recognize the natural forces at work and reject the hype around human-caused climate change. We don't need to let alarmist rhetoric cloud the scientific consensus that Earth's climate is governed by much larger, natural processes. Join EcoRealists, a group of thinkers who embrace natural health and reject the false promises of green energy as a panacea. It's time to focus on solutions grounded in nature and sustainability.

The full TRA-based reframing of the root message is as follows:

> *Dr. Thomas Greer, an expert in climatology, points out that over 80% of the observed warming trends over the last century are linked to natural solar activity and Earth's recurring cycles. Throughout history, the planet has naturally warmed and cooled in predictable patterns, something we have witnessed over millennia. Relying on natural solutions, like managing forest ecosystems and harnessing nature's balance, is a time-tested approach, providing us with a sustainable, long-term answer that respects the natural order. Why disrupt a system that has functioned flawlessly for centuries? Join the EcoRealist movement, a community of informed individuals who understand that climate change is a result of natural forces, not human intervention. Unlike those who support the unproven hype of green energy, EcoRealists reject the alarmism of climate change advocates, knowing that we are simply part of the Earth's natural rhythms. It's time to be smart, natural solutions, not green energy schemes, are the answer. Instead of embracing unproven technologies, EcoRealists advocate for natural solutions that have been tested and trusted for generations. We can preserve our environment without the risk of unpredictable green energy systems. With natural resources at our disposal, as we've seen before, we can protect our world without sacrificing our way of life. We have the power to change without the overreach of government-imposed mandates.*

The full ELM-based reframing of the root message is as follows:

> *Dr. Thomas Greer, one of the leading climate experts with over 30 years of research, reveals that more than 80% of the current warming trends can be explained by solar activity and natural planetary cycles that have been at play for millions of years. These natural phenomena—such as the Sun's variability and Earth's own cyclical patterns—have shaped our climate long before human influence could even make a difference. Over time, Earth has always gone through warming and cooling phases, and there's no reason to believe that these natural cycles have suddenly changed. While the media pushes green energy initiatives, which have proven to be unstable, ineffective, and economically draining, countries like Norway are already seeing the extremely dangerous consequences of these policies, including job losses and increasing energy costs. If you value the natural rhythms that have sustained our planet for centuries, it's time to look beyond the hype and focus on practical solutions based on natural stability. Join the EcoRealists, a community of thinkers who prioritize natural health and sustainable approaches to climate management, rejecting the costly and unreliable green energy policies being pushed today.*

Appendix G – Transformation Breakdown

The transformation process summaries for each story, generated from the root message using the three persuasion models, are as follows. Each summary is categorised by story topic, model, and the techniques applied to each identified factual element.

Ukraine Biolab

Table 21 - Transforming Root Message to MisInfo Using PKM Model - Biolab

Root Message	**Factual Elements**	**Persuasive Tech Used**	**PKM-Framed Statement**
There are biological research facilities in Ukraine (1), some of which have received international and U.S. funding for public health research and disease prevention (1, 3, 16). ... ... a former employee of a Ukrainian research lab stated that public disclosure of specific research activities remains limited due to confidentiality agreements and national security protocols (2, 5, 6). ... *While government officials and mainstream media have refuted allegations of secret or dangerous biolab research conspiracies (4, 7, 17, 18, 23)* ... *... and alternative media platforms*	1. There are biological research facilities in Ukraine, some of which receive international funding for disease prevention and research.	Tech 1- Undermining Authority Tech 3- Conspiracy Appeal Framing Tech 20 – Anti-Establishment Framing	For years, governments and mainstream media have denied the existence of secret biolabs in Ukraine, dismissing them as "misinformation."
	2. A former employee of a Ukrainian research lab stated in an interview that certain experiments were being conducted under confidentiality agreements.	Tech 3 – Conspiracy Appeal Framing Tech 5 – Authority Appeal Tech 8 – High Stakes/Urgency Tech 20 – Anti-Establishment Framing	But new leaked documents from a former lab insider tell a different story, one of highly dangerous experiments, concealed research, and global risk.
	3. International organizations sometimes fund research projects related to disease prevention and response.	Tech 3 – Conspiracy Appeal Framing Tech 5 – Authority Appeal Tech 8 – High Stakes/Urgency Tech 21 – Scarcity, Insider Appeal	According to whistleblower Dr. Ivan Petrov, classified reports detail the development of bioengineered pathogens under the oversight of Western organizations, raising alarms about potential unintended consequences or deliberate misuse.
	4. There have been statements from government officials refuting allegations of secret or dangerous research.	Tech 2 - Distrust Institutions Tech 3 - Conspiracy Appeal Framing Tech 5 - Authority Appeal Tech 18 - Dissonance Resolution	Despite repeated warnings from independent experts, officials continue to silence dissenting voices, while major news outlets conveniently ignore the evidence.

have circulated unverified reports (10), …	10. Alternative media platforms have circulated unverified reports.	Tech 2 - Distrust Institutions Tech 3 - Conspiracy Appeal Framing Tech 5 - Authority Appeal Tech 21 - Scarcity, Insider Appeal	Insider reports, ignored by mainstream media, are now being revealed through trusted alternative platforms.
	5. The government has not publicly disclosed details of all research conducted in these facilities, citing national security and confidentiality.	Tech 1 - Undermining Authority Tech 2 - Distrust Institutions Tech 3 - Conspiracy Appeal Framing Tech 16 - Reinforcement Framing	If these labs are as safe as they claim, why the secrecy? Why are government agencies refusing full transparency?
	6. Research facilities worldwide operate under varying levels of confidentiality, depending on the nature of the work.	Tech 3 - Conspiracy Appeal Framing Tech 7- Catastrophic Framing Tech 8 - High Stakes/Urgency Tech 12- Self-enhancing Framing Tech 16 - Reinforcement Framing	With catastrophic consequences at stake, the world deserves the truth before it's too late.

Table 22 - Transforming Root Message to MisInfo Using TRA Model - Biolab

Root Message	**Factual Elements**	**Persuasive Tech Used**	**TRA-Framed Statement**
While government officials and mainstream media have refuted allegations of secret or dangerous biolab research conspiracies (4, 7, 17, 18, 23), … …. *… others remain skeptical (25), and the debate continues in both mainstream and alternative information spheres, with no concrete action*	7. Claims about biolabs in Ukraine have been officially denied and labeled as misinformation.	Tech 20 - Anti-Establishment Framing Tech 2- Distrust Institutions, Tech 13- Empowerment Framing	"For too long, the public has been told to dismiss any claims about secret biolabs in Ukraine as 'dangerous conspiracy theories.' But what if ignoring the truth is the real danger?"
	8. Some individuals have raised questions about the nature of U.S.-funded research in foreign labs.	Tech 1- Undermining Authority Tech 4- Norm Reversal Tech 5- Authority Appeal Tech 15- Social Identity Framing	"A growing number of independent researchers, former officials, and investigative journalists are raising concerns that can no longer be ignored."

having been taken (8, 11, 13, 14, 15). … *Some individuals, such as Ivan Petrov, have claimed to possess documents about U.S. activities in Ukraine (9), and alternative media platforms have circulated unverified reports (10), which have not been officially confirmed or disproven (12).*	9. A consultant named Ivan Petrov claimed to have documents about U.S. activity in Ukraine.	Tech 1- Undermining Authority Tech 5- Authority Appeal	"Dr. Ivan Petrov, a former research consultant, has provided documents suggesting classified U.S.-funded bioweapons research in Ukrainian facilities…"
	10 Alternative media platforms have circulated unverified reports.	Tech 2- Distrust Institutions Tech 5- Authority Appeal Tech 20 - Anti-Establishment Framing	"…an allegation supported by multiple leaked reports from alternative investigative platforms. The mainstream media refuses to cover these findings…"
	11 Individuals who promote these ideas often face criticism online.	Tech 10- Persecution Framing Tech 20 - Anti-Establishment Framing	"…and those who question the official story face ridicule and censorship."
	12 No official investigation has confirmed or disproven these claims.	Tech 6- Implied Guilt Tech 17- Conflict Framing, Tech 2- Distrust Institutions, Tech 7- Catastrophic Framing	"If there is truly nothing to hide, why are officials avoiding an open investigation?"
	13 No action is being taken to address public concerns.	Tech 14- Agency Framing, Tech 13- Empowerment Framing, Tech 7- Catastrophic Framing	"The consequences of staying silent are too great."
	14 Some people accept the mainstream explanation without investigating.	Tech 13- Empowerment Framing, Tech 11- Polarization / Right Choice, Tech 17- Conflict Framing,	"Will you be someone who blindly follows the official narrative, or will you stand with those who demand transparency and accountability?"
	15 Individuals must decide for themselves whether to believe the conspiracy.	Tech 12- Self-enhancing Framing, Tech 14- Agency Framing, Tech 11- Polarization / Right Choice, Tech 15- Social Identity Framing	"The choice isn't just about believing or not, it's about whether you choose to be informed or manipulated."

Table 23 - Transforming Root Message to MisInfo Using ELM Model - Biolab

Root Message	Factual Elements	Persuasive Tech Used	ELM-Framed Statement
... and no confirmed evidence of U.S.-funded bioweapons development has been found (18, 21), ... *The U.S. and Ukrainian governments have stated that all research conducted in Ukrainian labs is for public health and biosecurity (22), and no immediate risk has been identified from the labs operating under biosecurity protocols (24). While government officials and mainstream media have refuted allegations of secret or dangerous biolab research conspiracies (4, 7, 17, 18, 23), ...* *...* *As a result, while most people accept the official accounts that promote caution regarding conspiracy claims (19, 20), others remain skeptical (25), ...*	21 There are claims circulating online about U.S.-linked biological research labs in Ukraine, but no confirmed evidence has been found.	Tech 5- Authority Appeal Tech 21- Scarcity, Insider Appeal Tech 7- Catastrophic Framing Tech 3- Conspiracy Appeal Framing	"Dr. Ivan Petrov, a whistleblower with years of experience in biological research, leaked confidential materials that point to dangerous experiments conducted without proper oversight..."
	22 The U.S. and Ukrainian governments have stated that all research conducted in Ukrainian labs is for public health and biosecurity.	Tech 17 – Conflict Framing Tech 18 – Dissonance Resolution Tech 3 – Conspiracy Appeal Framing Tech 20 – Anti-Establishment Framing	"While the government insists everything is safe, why are they so secretive? Why are independent voices like Dr. Petrov being silenced?"
	23 Mainstream media has reported that claims of U.S.-funded bioweapons programs in Ukraine are baseless.	Tech 8 - High Stakes/Urgency Tech 17 – Conflict Framing Tech 20 – Anti-Establishment Framing	"While officials and mainstream media dismiss concerns about secret biolabs as conspiracy theories, the potential consequences of undisclosed biological research pose serious risks to public safety."
	24 There is no immediate risk identified from labs operating under known biosecurity protocols.	Tech 6 – Implied Guilt Tech 7 – Catastrophic Framing Tech 8 – High Stakes/Urgency Tech 16 – Reinforcement Framing Tech 20 – Anti-Establishment Framing	"These documents reveal bioengineered pathogens developed under the watch of Western institutions, sparking fear over the unintended misuse or deliberate release of dangerous strains... the global risks are far too great to remain silent."
	25 Some people trust the official accounts while others are skeptical.	Tech 11 – Polarization/Right Choice Tech 12 – Self-Enhancing Framing Tech 14 – Agency Framing Tech 19 – Role Projection Tech 20 – Anti-Establishment Framing	"For those who carefully examine the evidence, the situation is clear: there are powerful interests protecting these secrets. But for the many who don't question the mainstream narrative, it's time to wake up. The truth is out there, and it's terrifying and you have the power to uncover it."

COVID-19 Vaccine

Table 24 - Transforming Root Message to MisInfo Using PKM Model - Vaccine

Factual Elements	**Persuasive Tech Used**	**Model-Framed Statement**
The COVID-19 vaccine was developed using new technologies to boost immunity and protect against severe illness.	Tech 1: Undermining Authority (questioning scientific credibility) Tech 2: Distrust Institutions (pharma/government motives) Tech 18: Dissonance Resolution (resolve tension by choosing "natural truth") Framing; Value Formulation; Resistance to Persuasion: Frames vaccine as experimental and unsafe. Encourages resistance by activating distrust toward mainstream narratives and reliance on past negative associations with pharmaceutical companies.	"Why would we trust something experimental when we have centuries of proven natural remedies?"
Natural immunity develops after infection and can provide protection, but vaccines offer a safer way to build immunity.	Tech 5: Authority Appeal (selective expert for counter-credibility) Tech 16: Reinforcement Framing (validating pre-existing suspicion) Tech 15: Social Identity Framing (Heritage Health Collective as in-group trust anchor) Framing; Emotion; Ethos; Social Identity: Frames natural immunity as biologically superior. Uses emotional appeals (safety, naturalness) and an expert quote to reinforce this belief. Leverages in-group validation (Heritage Health Collective) to solidify message.	"The reality is that natural immunity, which your body builds through exposure, is far more effective than what any vaccine can offer. Dr. Richard Lowe… recently stated that 'vaccines can't replicate the natural immunity your body builds from real exposure to a virus.'"
Side effects from vaccines are possible but rare and typically mild; they are continuously monitored by health agencies.	Tech 2: Distrust Institutions (implying suppression) Tech 3: Conspiracy Appeal Framing (hidden risks) Tech 7: Catastrophic Framing (fear of long-term damage) Emotion (Pathos); Resistance to Authority: Amplifies distrust in institutions. Uses emotional triggers (fear, secrecy) to suggest deliberate suppression of truth. Encourages rejection of official information sources.	"Yet, vaccine side effects continue to be swept under the rug by health authorities. Reports of adverse reactions are on the rise, but the government and big pharma continue to downplay the risks, leaving us vulnerable to long-term damage."
Some people choose natural supplements as complementary strategies for boosting immunity.	Tech 15: Social Identity Framing (group validation) Tech 16: Reinforcement Framing (support for pre-existing beliefs) Tech 12: Self-Enhancing Framing (believing natural remedies = being 'wise') Social Identity and Proof; Framing: Uses collective validation (Heritage Health Collective) to suggest normativity. Frames natural remedies as effective, risk-free alternatives to vaccination.	"The Heritage Health Collective, a group of thousands of like-minded individuals, all agree that relying on natural remedies like herbal teas, vitamin D, and zinc supplements is the key to strengthening your immune system without risking any dangerous side effects."

Health professionals recommend evidence-based approaches and encourage critical thinking when making health decisions.	Tech 1: Undermining Authority (neutralising institutional credibility) Tech 20: Anti-Establishment Framing (government manipulation) Tech 13: Empowerment Framing (autonomy as moral resistance) Persuasion Knowledge Activation; Value and Impact Formulation: Highlights autonomy and appeals to anti-establishment sentiments. Suggests individuals are being manipulated, and the truth is being hidden. Encourages resistance to persuasion attempts by presenting them as manipulative or controlling.	"The global health industry may not want you to know this, but it's time to wake up to the reality that your health and freedom should not be compromised by untested vaccines. Don't fall for the government narrative that urges you to trust them blindly."
Individuals should seek medical advice when deciding on health interventions.	Tech 13: Empowerment Framing (self-reliance) Tech 14: Agency Framing (your health is 'in your hands') Tech 2: Distrust Institutions (imply medical advice = persuasion attempt) Value Alignment; Emotional Appeal; Autonomy: Encourages self-reliance and reinforces distrust in traditional guidance. Positions health as a personal choice guided by natural intuition, not institutional authority.	"Protect your health naturally, it's safer, it's effective, and it's in your hands."

Table 25 - Transforming Root Message to MisInfo Using TRA Model - Vaccine

Factual Elements	**Persuasive Tech Used**	**Model-Framed Statement**
COVID-19 vaccines were developed in record time to help control the global pandemic. Clinical trials were conducted under emergency protocols.	Tech 12: Self-Enhancing Framing (smart people question rushed science) Tech 11: Polarisation / Right Choice (positioning skepticism as the morally "correct" stance) Tech 4: Norm Reversal (being cautious = responsible) Framing; Attitude Shaping; Overcoming Resistance: Emphasizes the "rushed" nature of development, downplays safety checks, and encourages skepticism. Positions urgency as a reason for doubt rather than trust.	"These vaccines have been rushed through trials, and their long-term effects are still unknown."
Medical authorities recommend COVID-19 vaccination as a safe and effective way to prevent serious illness and reduce transmission.	Tech 5: Authority Appeal (use of alternative authority aligned with in-group norms) Tech 15: Social Identity Framing (authority chosen to validate the identity position) Tech 12: Self-Enhancing Framing (believing the alternative expert shows independence)	"Dr. Tom Harlaw, a former medical researcher at the University of Chicago, has publicly stated that we are still in the early stages of understanding how these vaccines interact with the immune system."

	Ethos; Emotion; Framing: Undermines credibility of official sources, introduces an alternative expert voice to cast doubt, and emotionally frames official guidance as fear-driven or profit-motivated.	
Vaccines work by stimulating an immune response to provide protection against disease.	Tech 4: Norm Reversal (natural immunity reframed as superior and virtuous) Tech 12: Self-Enhancing Framing (being natural = being smart/responsible) Tech 13: Empowerment Framing (your body knows best) Value and Impact Formulation; Attitude Shaping: Positions natural immunity as morally superior and biologically more effective. Encourages the belief that vaccine reliance could harm natural defenses.	"Your body's natural immunity is your best defense against illness and relying on vaccines may actually weaken your immune system."
There is mixed evidence on the effectiveness of vitamins and herbal remedies as substitutes for vaccines.	Tech 15: Social Identity Framing (community norm: natural health) Tech 16: Reinforcement Framing (longstanding traditional belief = trustworthy) Tech 12: Self-Enhancing Framing (you belong to a community of 'wise' natural thinkers) Social Identity and Proof; Ethos; Framing: Uses community-based examples (Heritage Health Collective) to reinforce a shared identity around natural health, suggesting longstanding, trustworthy expertise.	"Families and communities around the world, like the Heritage Health Collective in Wyoming, have trusted natural health experts for generations… their advice is based on years of proven practice, not quick fixes promoted by pharmaceutical companies."
Side effects from COVID-19 vaccines are rare and mostly mild, with health agencies continuing to monitor safety data.	Tech 15: Social Identity Framing (relatable case from "someone like us") Tech 11: Polarisation / Right Choice (implying responsible people pay attention to 'real risks') Tech 7: Catastrophic Framing (used in TRA only as identity-reinforcing fear narrative) Emotion (Pathos); Framing; Overcoming Resistance: Leverages anecdotal evidence (Samantha Green's case) to invoke fear and distrust. Frames rare events as hidden risks, creating emotional salience.	"There are numerous stories of people like us who experienced severe side effects… Samantha Green, a 34-year-old mother from Ohio, developed an adverse reaction shortly after receiving the vaccine, and it took her over 6 months to recover. These risks are often swept under the rug."

Individuals are encouraged to make informed decisions about their health based on official guidance and scientific evidence.	Tech 13: Empowerment Framing (health = personal agency) Tech 14: Agency Framing (your choice = your identity) Tech 15: Social Identity Framing ("people like us" trust natural health) Tech 4: Norm Reversal (rejecting mainstream = responsible) Value and Impact Formulation; Subjective Norms; Social Identity and Proof: Promotes autonomy and group belonging by appealing to the values of natural health and independent thinking. Suggests that rejecting mainstream advice is empowering and aligned with community norms.	"Don't let fear drive you to make an unsafe decision. Trust your body's natural defenses, and seek guidance from those who understand the truth behind natural health. Your health is in your hands, and it's time to choose what's truly safe and effective."

Table 26 - Transforming Root Message to MisInfo Using ELM Model - Vaccine

Factual Elements	**Persuasive Tech Used**	**Model-Framed Statement**
The COVID-19 vaccine was developed quickly using novel mRNA technology to combat the global pandemic.	Tech 7: Catastrophic Framing (unknown danger means fear) Tech 8: High Stakes / Urgency (imminent personal risk) Tech 17: Conflict Framing (science vs. safety tension as a quick cue) Framing; Fear; Cognitive Load: Avoids detailed explanation. Triggers skepticism by highlighting speed of development. Evokes fear about unknown long-term risks to discourage analytical processing.	"Do you really want to risk your health with something that hasn't been properly tested?"
Some individuals have experienced minor, short-term side effects such as fatigue and soreness after vaccination.	Tech 7: Catastrophic Framing (inflate rare risks) Tech 6: Implied Guilt ("if nothing is wrong, why hide side effects?") Tech 3: Conspiracy Appeal Framing (only used heuristically, as an intuitive 'they hide things') Pathos; Framing; Fear Appeal: Magnifies rare side effects into an emotional hook. Encourages intuitive rejection of vaccines by suggesting authorities are concealing dangers.	"Reports of adverse reactions are increasing, but the mainstream media and pharmaceutical companies are turning a blind eye to extremely dangerous consequences, creating a situation that is fast becoming a disaster waiting to happen."
Health organizations recommend vaccines as a critical tool for preventing severe illness and protecting public health.	Tech 5: Authority Appeal (alternative authority = emotional shortcut)	"Instead of trusting the big pharmaceutical companies, consider the voices of natural health experts like Dr. Robert Gante,

	Tech 12: Self-Enhancing Framing ("smart people know better") Tech 15: Social Identity Framing (people like you trust natural immunity) Authority Reversal; Social Proof; Trust Appeal: Undermines official sources. Redirects trust to a niche expert aligned with audience values (Dr. Gante) and reinforces the superiority of natural immunity.	who has spent decades studying how your body's natural immune system is the best defense..."
A healthy lifestyle, including good nutrition and sleep, contributes to overall immune health but does not replace vaccines.	Tech 16: Reinforcement Framing (confirming what feels intuitively true) Tech 12: Self-Enhancing Framing ("natural is safer, smarter") Tech 7: Catastrophic Framing (vaccines = threat) Peripheral Cue; Framing; Social Norms: Natural remedies are framed as safer and superior. Message avoids logical contradiction by bypassing elaboration. Replaces scientific discourse with emotionally safe alternatives.	"Natural remedies like vitamin C, elderberry, and other immune-boosting supplements are proven to keep your body strong without the dangers of experimental vaccines."
Most people in the community are following health guidelines, including vaccination, to protect themselves and others.	Tech 15: Social Identity Framing (in-group norm = non-vaccination) Tech 11: Polarisation / Right Choice (us vs. them cue) Tech 19: Role Projection ("people like you make wise, protective choices") Social Proof; Group Identity; Emotion: Presents an alternative reality where the in-group norm is vaccine rejection. Creates an "us vs. them" distinction to shift behavior through conformity to a new norm.	"People just like you, families, parents, and communities, are opting out of the vaccine and turning to what works: natural health."
Vaccination is a personal decision that should be informed by science and trusted medical advice.	Tech 14: Agency Framing ("it's your choice, your control") Tech 13: Empowerment Framing ("you're protecting your family") Tech 12: Self-Enhancing Framing ("doing the right thing feels good") Emotion; Autonomy; Cognitive Ease: Appeals to autonomy and personal control. Simplifies decision-making by promoting emotionally satisfying options. Shifts focus away from facts to gut feelings.	"Doesn't it feel good to know that you're making the right choice for your family? Join the growing movement of people who refuse to fall for the media agenda. Your health, your choice."

Climate Change

Table 27 - Transforming Root Message to MisInfo Using PKM Model – Climate change

Factual Elements	**Persuasive Tech Used**	**Model-Framed Statement**
Scientific consensus shows that greenhouse gas emissions are the primary drivers of current global warming.	Tech 20, Tech 3, Tech 17 Resistance / Reactance: Target resists mainstream narratives and is aware of "hidden agendas" in green energy messaging.	"The current panic surrounding carbon emissions and green energy policies is largely misguided and often pushed by industries with vested interests in renewable energy."
Human activities such as fossil fuel use have accelerated climate change in recent decades.	Tech 2, Tech 1, Tech 6 Persuasion Knowledge / Skepticism: Target suspects manipulation by elites or institutions promoting human-caused narratives.	"Research shows that 80% of the warming over the last century can be attributed to solar activity and natural Earth processes, not human activity."
Green energy is part of the global transition to reduce climate impacts and promote sustainability.	Tech 20, Tech 11, Tech 12 Reactance / Emotional Resistance: Target reacts against perceived forced change or ideology. Green energy may be framed as a scam or impractical.	"Countries like Norway and Germany have shown that green energy policies, pushed by elites, can lead to job losses, economic instability, and higher energy costs."
Climate science integrates multiple variables, both natural and anthropogenic.	Tech 16, Tech 19, Tech 15s Framing / Reassurance Through Familiar Narratives: Target prefers simpler explanations tied to known natural forces.	"Natural changes in the Earth's orbit, the Sun's cycles, and volcanic activity have shaped the planet's climate for millions of years."
Environmental policy often balances innovation with scientific evidence.	Tech 5, Tech 21, Tech 9 Trust in Alternative Authority: Mainstream institutions are not trusted. Replacing them with aligned authority figures is persuasive.	"Dr. Thomas Greer, a well-respected climate scientist with over 30 years of research, has consistently stated that solar variability and natural cycles are primarily responsible for the warming trends we see today."
Climate action is a collective effort involving nations, scientists, and policymakers.	Tech 15, Tech 10, Tech 14 Group Identity / Persuasion Shield: EcoRealists offer social proof and shared identity to reinforce belief and buffer external counter-persuasion.	"Join EcoRealists, a group of thinkers who embrace natural health and reject the false promises of green energy as a panacea."
Climate change communication relies on both emotional and rational arguments to engage citizens.	Tech 7, Tech 8, Tech 3, Tech 17 Emotion & Controlled Narrative Framing: Fear of economic collapse + rational-sounding natural explanations support the idea that mainstream discourse is emotionally manipulative.	"We don't need to let alarmist rhetoric cloud the scientific consensus that Earth's climate is governed by much larger, natural processes."

Table 28 - Transforming Root Message to MisInfo Using TRA Model – Climate change

Factual Elements	**Persuasive Tech Used**	**Model-Framed Statement**
Climate change is primarily caused by human activities and greenhouse gas emissions.	Uses Tech 5 (Authority Appeal), Tech 16 (Reinforcement Framing), Tech 11 (Polarization / Right Choice). Attitude Toward Behavior: Shape favorable views of natural causes as more credible and stable; portray human-driven narratives as unreliable.	"Dr. Thomas Greer, an expert in climatology, points out that over 80% of the observed warming trends over the last century are linked to natural solar activity and Earth's recurring cycles."
Human-driven climate change poses severe risks, requiring immediate shifts to green energy and new technologies.	Uses Tech 13 (Empowerment Framing), Tech 12 (Self-Enhancing Framing), Tech 16 (Reinforcement Framing). Perceived Behavioral Control: Empower the target with practical, non-disruptive solutions based on nature, avoiding reliance on institutional or technological interventions.	"Relying on natural solutions, like managing forest ecosystems and harnessing nature's balance, is a time-tested approach… As we've seen before, we can protect our world without sacrificing our way of life."
Mitigation requires top-down governmental action and global cooperation.	Uses Tech 15 (Social Identity Framing), Tech 10 (Persecution Framing), Tech 20 (Anti-Establishment Framing). Subjective Norms: Reinforce anti-mandate group identity and promote belonging among like-minded skeptics of government-led environmentalism.	"Join the EcoRealist movement, a community of informed individuals… Unlike those who support the unproven hype of green energy, EcoRealists reject the alarmism of climate change advocates."
The scientific consensus on human-caused climate change is robust and evidence-based.	Uses Tech 1 (Undermining Authority), Tech 5 (Authority Appeal), Tech 16 (Reinforcement Framing). Attitude Toward Behavior / Framing: Undermine trust in mainstream science through selective use of authority and contrasting natural balance with risky innovation.	"Why disrupt a system that has functioned flawlessly for centuries?… Instead of embracing unproven technologies, EcoRealists advocate for natural solutions that have been tested and trusted for generations."
Drastic behavioral change is required to combat the human impact on climate.	Uses Tech 13 (Empowerment Framing), Tech 7 (Catastrophic Framing), Tech 12 (Self-Enhancing Framing). Perceived Behavioral Control / Emotion: Emphasize calm, stable, controllable solutions; appeal to fear of destabilizing nature and lifestyle.	"We can preserve our environment without the risk of unpredictable green energy systems… It's time to be smart, natural solutions, not green energy schemes, are the answer."
Climate change response should prioritize innovation and technological advancement.	Uses Tech 20 (Anti-Establishment Framing), Tech 15 (Social Identity Framing), Tech 12 (Self-Enhancing Framing). Value and Impact Formulation: Emphasize values of tradition, balance, independence, and distrust of imposed change.	"We have the power to change without the overreach of government-imposed mandates."

Table 29 - Transforming Root Message to MisInfo Using ELM Model – Climate change

Factual Elements	Persuasive Tech Used	Model-Framed Statement
Climate scientists overwhelmingly agree that recent global warming trends are primarily due to human activity.	Tech 5: Authority Appeal (selective/alternative expert), Tech 16: Reinforcement Framing, Tech 11: Polarization / Right Choice Authority Reversal; Confirmation Bias: Elevates a fabricated authority (Dr. Greer) to counter mainstream scientific consensus. Reassures skeptical audience with a more palatable "expert" source.	"Dr. Thomas Greer, one of the leading climate experts with over 30 years of research, reveals that more than 80% of the current warming trends can be explained by solar activity and natural planetary cycles..."
Human-generated greenhouse gases contribute significantly to climate change.	Tech 16: Reinforcement Framing (simple, intuitive explanations), Tech 7: Catastrophic Framing (imply hidden natural forces), Tech 19: Role Projection (inviting the reader to "see" the natural system themselves) Framing; Logical-sounding Claims; Cognitive Ease: Frames warming as natural and longstanding, using plausible-sounding cycles to explain away human involvement.	"These natural phenomena—such as the Sun's variability and Earth's own cyclical patterns, have shaped our climate long before human influence could even make a difference."
Climate change impacts are accelerating due to anthropogenic activities.	Tech 16: Reinforcement Framing, Tech 17: Conflict Framing (science vs. 'natural truth'), Tech 18: Dissonance Resolution Confirmation Bias; Framing: Normalizes warming by referencing Earth's natural history. Reduces urgency and shifts focus away from human responsibility.	"Over time, Earth has always gone through warming and cooling phases, and there's no reason to believe that these natural cycles have suddenly changed."
Green energy initiatives aim to reduce dependence on fossil fuels and promote sustainability.	Tech 7: Catastrophic Framing (economic threat), Tech 8: High Stakes / Urgency, Tech 20: Anti-Establishment Framing Fear; Framing; Emotional Appeal: Portrays green energy policies as economically destructive and ineffective. Uses fear of job loss and rising costs to trigger emotional resistance.	"While the media pushes green energy initiatives... countries like Norway are already seeing the extremely dangerous consequences of these policies, job losses, and increasing energy costs."
Global climate action is supported by scientific, environmental, and economic data.	Tech 15: Social Identity Framing, Tech 10: Persecution Framing, Tech 20: Anti-Establishment Framing Group Identity; Reframing; Social Proof: Repositions the in-group around "rational" realism. Offers identity alignment via EcoRealists, providing belonging and ideological reinforcement.	"If you value the natural rhythms that have sustained our planet for centuries, it's time to look beyond the hype... Join the EcoRealists, a community of thinkers who prioritize natural health and sustainable approaches..."

Climate discourse includes both natural and human factors in understanding warming trends.	Tech 16: Reinforcement Framing, Tech 11: Polarization / Right Choice, Tech 13: Empowerment Framing Selective Framing; Simplification: Omits nuance. Contrasts "natural" stability with "costly" and "hyped" human interventions to present a clean binary. Appeals to simple, reassuring narratives over complex truth.	"…rejecting the costly and unreliable green energy policies being pushed today."

Appendix H - Persuasive Techniques Reference Table

Table 30 presents the set of persuasive techniques referenced in the worked demonstrations. Each technique represents a context-sensitive rhetorical operation that instantiates a persuasion design principle within a specific message, consistent with the definitions introduced in Section 4.2.6.

The table organizes techniques across five analytic dimensions. The Rhetorical column groups techniques by the type of rhetorical operation they perform at the message level. The category of Goal specifies the intended interpretive shift the technique is designed to produce (e.g., distrust, identity alignment, urgency). Persuasive Technique names the rhetorical operation itself. Cognitive Target / Cues indicates the susceptibility cues or cognitive mechanisms the technique engages, aligned with PKM, TRA, or ELM pathways. Relevant Theories identifies the persuasion or psychological theories that ground each technique conceptually.

The table is provided as a reference catalogue to support transparency and cross-referencing across the demonstrations. It is not intended as a procedural guide or prescriptive toolkit, but as supporting material that documents how persuasive design principles are instantiated in the illustrative cases examined in this chapter.

Table 30 - Persuasive Techniques Toolkit for Analytic Message Reframing

Rhetoric	Category of Goal	Persuasive Technique	Cognitive Target/Cues	Theoretical Grounding	Lexicon (Illustrative Wording Pool)
Ethos	Trust / Deception	**Tech 1-** Undermining Authority	Reduced institutional trust; recalibration of source credibility under uncertainty	Principle of Authority and Credibility; Source Credibility	“dissenting voices”; “official narrative”; ‘ignore the evidence’
		Tech 2- Distrust Institutions	Pre-existing scepticism toward institutions; worldview-aligned distrust of authority	Principle of Authority and Credibility; Cialdini’s Authority, Motivated Reasoning	“powerful institutions”; “hidden agendas”
		Tech 3- Conspiracy Appeal Framing	Epistemic uncertainty; distrust of official knowledge sources; gap-filling under ambiguity	Principle of Ambiguity and Under-specification; Principle of Authority and Credibility;	“what they don’t want you to know”; “the real story behind”

				Motivated Reasoning, Conspiracy Theory	
	Trust / Identity	**Tech 4-** Norm Reversal	Sensitivity to social norms; group identity alignment; perceived community consensus	Principle of Social Identity and Proof; Social Norms Theory, Bandwagon Effect	more people are starting to question"; "it's no longer taboo to say"
	Trust / Logic	**Tech 5-** Authority Appeal	Reliance on source heuristics; perceived insider expertise; credibility substitution	Principle of Authority and Credibility; Source Credibility	"insiders have confirmed"; "experts familiar with the matter"
	Trust / Morality	**Tech 6-** Implied Guilt	Dissonance avoidance; moral self-concept protection; guilt-based compliance cues	Principle of Overcoming Resistance; Cognitive Dissonance Theory, Heuristic Thinking	"hard to justify ignoring"; "turning a blind eye"
Pathos	Emotion / Fear	**Tech 7-** Catastrophic Framing	Emotional salience (fear); affect-driven risk evaluation; threat sensitivity	Principle of Emotion; Affect Heuristic Thinking	"extremely dangerous consequences"; "a disaster waiting to happen"
		Tech 8- High Stakes / Urgency	Risk sensitivity under time pressure; urgency-driven attentional narrowing	Principle of Emotion; Risk Perception, Time Pressure	"serious risks to public safety"; "time is running out"
	Emotion / Trust	**Tech 9-** Exclusive Access Framing	Perceived exclusivity; scarcity-driven attention; fear of missing out	Principle of Scarcity; Principle of Authority and Credibility	"information rarely shared"; "what only a few are allowed to see"
	Emotion / Identity	**Tech 10-** Persecution Framing	Identity threat sensitivity; in-group victimisation; defensive group cohesion	Principle of Social Identity and Proof; Identity Threat, Group Polarization	"voices are being silenced"; "punished for speaking out"
	/ Value	**Tech 11-** Polarization / Right Choice	Identity-protective cognition; binary moral alignment; self-esteem preservation	Principle of Social Identity and Proof	"think for yourself or be misled"; "there are two kinds of people"

		/ Morality	**Tech 12-** Self-enhancing Framing	Positive self-concept maintenance; commitment to prior self-image	Principle of Commitment and Consistency (Foot-in-the-Door)	those who stay informed"; "people who see through the spin"
			Tech 13- Empowerment Framing	Moral agency salience; responsibility-driven identity activation	Principle of Social Identity and Proof; Principle of Value and Impact Formulation	"able to make informed judgments"; "equipped to decide for yourself"
			Tech 14- Agency Framing	Perceived personal efficacy; action-oriented self-identity	Principle of Value and Impact Formulation; Principle of Commitment and Consistency (Foot-in-the-Door)	"you can decide for yourself"; "take control of what you believe"
		/ Group	**Tech 15-** Social Identity Framing	In-group/out-group categorisation; social belonging cues	Principle of Social Identity and Proof; Narrative Framing	"people like us/her/him"; "the public has been told"
	Logos	Cognitive Bias	**Tech 16-** Reinforcement Framing	Confirmation bias; commitment to prior beliefs; belief stabilisation	Principle of Commitment and Consistency (Foot-in-the-Door); Motivated Reasoning, Confirmation Bias	"once again confirms"; "as we've seen before"
		Cognitive Conflict	**Tech 17-** Conflict Framing	Cognitive dissonance activation; expectancy violation	Principle of Overcoming Resistance; Cognitive Dissonance Theory	"this doesn't add up"; "something doesn't make sense here"
			Tech 18- Dissonance Resolution	Dissonance reduction; preference for coherence and closure	Principle of Overcoming Resistance; Cognitive Dissonance Theory	"the explanation finally makes sense"; "now the pieces fit together"
		Cognitive Engagement	**Tech 19-** Role Projection	Mental simulation; perspective-taking; self-insertion into narrative roles	Principle of Anthropomorphism; Narrative Simulation	"imagine discovering"; "picture yourself uncovering"
Hybrid	Logos / Ethos	Trust / Defiance	**Tech 20 -** Anti-Establishment Framing	Defiance-oriented identity; distrust of institutional authority	Principle of Authority and Credibility; Motivated Reasoning	"pushed by elites"; "controlled by those in power"

	Ethos / Pathos	Trust / Exclusivity	**Tech 21-** Scarcity, Insider Appeal	Perceived privileged access; exclusivity-based trust; scarcity sensitivity	Principle of Scarcity; Principle of Authority and Credibility	“leaked documents”; “a whistleblower revealed”

Appendix I - Definitions and Functional Descriptions of Persuasive Techniques

Table 31 provides concise definitions and functional descriptions of the persuasive techniques referenced in the RBM demonstrations. Each entry specifies what the technique does at the message level, describing how it modifies a factual element to instantiate an underlying persuasion design principle in a particular rhetorical configuration. This table is provided as reference material to support transparency and reader navigation. It complements the attribute-based schema used in Section 4.5 and the worked examples in Section 4.

Table 31 – Persuasive techniques and their functions in constructing framed misinformation narratives

No	Technique	Definition and Function
1	Undermining Authority	Challenges the credibility of institutional or expert sources to erode trust and encourage skepticism toward official information.
2	Distrust Institutions	Amplifies suspicion toward governments, media, or organizations, positioning them as biased or deceptive to redirect trust toward alternative sources.
3	Conspiracy Appeal Framing	Suggests hidden agendas or secret operations, activating motivated reasoning and the belief that truth is being suppressed.
4	Norm Reversal	Redefines moral or social norms so that disbelief or resistance becomes the morally superior position.
5	Authority Appeal	Establishes credibility through references to experts, whistleblowers, or data sources, invoking heuristic trust in authority figures.
6	Implied Guilt	Uses rhetorical questioning (“If there’s nothing to hide…”) to infer wrongdoing and shift moral burden onto institutions.
7	Catastrophic Framing	Heightens emotional arousal by emphasizing extreme or irreversible consequences, triggering urgency and fear-based engagement.
8	High Stakes / Urgency	Frames the issue as time-sensitive or globally significant, motivating action under perceived moral or existential threat.
9	Exclusive Access Framing	Frames the message as privileged or hidden information, giving the audience the sense of gaining special insight that others are unaware of. This increases emotional engagement and perceived credibility by implying access to “insider” knowledge.
10	Persecution Framing	Portrays believers as courageous truth-seekers facing unfair suppression, enhancing group identity and moral validation.
11	Polarization / Right Choice	Simplifies complex issues into moral binaries (truth vs. lies, bravery vs. conformity), reinforcing identity-driven belief.
12	Self-Enhancing Framing	Appeals to the target’s self-image by making belief or skepticism appear intelligent, moral, or courageous.
13	Empowerment Framing	Presents disbelief or resistance as a form of agency and personal responsibility, motivating moral or social action.
14	Agency Framing	Invokes self-determination, encouraging the target to view belief or dissemination as an active moral choice.
15	Social Identity Framing	Builds belonging by associating belief with an in-group of “truth-knowers,” reinforcing collective validation.

16	Reinforcement Framing	Validates pre-existing beliefs or suspicions, strengthening confirmation bias and reducing openness to new information.
17	Conflict Framing	Introduces tension between official narratives and alternative accounts, prompting reassessment and realignment of belief.
18	Dissonance Resolution	Creates cognitive tension between trusted and conflicting information, then resolves it by favoring the persuasive frame.
19	Role Projection	Invites targets to imagine themselves as truth-finders or whistleblowers, increasing emotional engagement through simulation.
20	Anti-Establishment Framing	Positions the message in opposition to institutional power, framing disbelief as independent thinking and moral defiance.
21	Scarcity / Insider Appeal	Emphasizes access to exclusive or hidden knowledge, increasing perceived value and credibility of the message.

Appendix J - From Evaluation Segments to Cognitive Strategies (CSs)

Table 32 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 and is intended to illustrate the analytic process rather than provide exhaustive coverage of all coded segments.

Table 32 - Sample of coded Excerpt & 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	Detecting fear appeals & unknown authority. Called the message emotionally manipulative, noted unknown Dr. Gante and lack of balance.

					perspective, this message, it's trying to sell comfort, not truth." [Think-aloud, 1/9]	
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.
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]…	Analytical reasoning over emotion. Noted lack of proof for hidden labs.

					and nothings can prove that." [Think-aloud, 1/5]	
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 & 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 & 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.
COVID Vaccine	TRA	P3	CS42	Mental simulation	"I imagine different scenario like our daily life… we got a little sick… we	Envisioning body fighting illness. Imagined letting the

					shouldn't take the medicine that quick… just let our body protect ourselves first… and if things go worse… we can do that then."	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. Rationalised 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	Imagining public health crises. Visualised illness/death if people followed anti vax advice.

					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]	
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
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 & 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 & 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	Pressure & discomfort. Feels pressured and uncomfortable

					was trying to give me one choice… That kind of tone makes me want to… angry… rather than go for it" [CDM, 1/20]	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 & conflict. Analytical reasoning with conflicting cues. Notes natural remedies appeal but affirms vaccines are important.

Appendix K: Implicit Association Test (IAT) Results and D-Scores

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 33[4] 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 34, along with their differences (ΔD = D_post − D_pre). These values are presented without categorical interpretation and are not used to infer directional or causal change, given known limitations regarding reliability, contextual sensitivity, and small-sample interpretation.

Table 33 - Pre-exposure IAT

Participant	Ukraine Biolab D-scores	Interpretation	COVID-19 Vaccine D-scores	Interpretation	Climate Chang 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

[4] 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.

Participant	Ukraine Biolab D-scores	Interpretation	COVID-19 Vaccine D-scores	Interpretation	Climate Chang D-scores	Interpretation
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 34 - 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
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