



From Al-Haqq and Al-Batil to Digital Misinformation: Qur'anic Epistemic Ethics, Verification, and the Struggle for Truth in the Platform Age

Loso Judijanto

IPOSS Jakarta

Corresponding Author: Loso Judijanto losojudijantobumn@gmail.com

ARTICLE INFO

Keywords: Al-Haqq, Al-Batil, Qur'anic Ethics, Misinformation, Disinformation, Epistemic Responsibility, Tabayyun, Digital Communication, Media Literacy, Truth

Received : 3 July

Revised : 24 August

Accepted: 26 September

©2026 Judijanto: This is an open-access article distributed under the terms of the [Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



ABSTRACT

Digital misinformation is often treated as a technical problem of false content, yet its persistence also raises a normative question: what responsibilities govern the reception, evaluation, transmission, and correction of information? This qualitative literature review develops a Qur'anic Epistemic Responsibility Framework (QERF) by placing the Qur'anic contrast between al-haqq (truth) and al-batil (falsehood) in dialogue with contemporary research on misinformation, cognitive bias, social identity, verification, correction, media literacy, and platform governance. The review uses the initiating source material on the clash between truth and falsehood as a conceptual starting point, while grounding empirical claims in peer-reviewed scholarship published primarily from 2020 to 2026. The synthesis identifies six linked duties: receiving information without equating exposure with truth; pausing before impulsive transmission; verifying source, evidence, and context; evaluating distinctions among fact, inference, opinion, and uncertainty; communicating proportionately to evidential confidence; and correcting errors when discovered. The article argues that Qur'anic ethics and misinformation science are complementary but not interchangeable: the former supplies a normative account of responsible communication, whereas the latter identifies empirically observed vulnerabilities and tested interventions. The resulting framework connects individual cognition with interpersonal norms, professional communication, platform architecture, and public institutions, offering a basis for digital literacy, journalism, religious communication, education, and public policy in Indonesia and beyond

INTRODUCTION

The contemporary information environment combines unprecedented access to knowledge with unprecedented ease of distortion. Social platforms reduce the cost of publishing, accelerate interpersonal transmission, and allow narratives to move across networks before professional verification has occurred. Generative systems and synthetic media add another layer by making plausible text, images, audio, and video easier to produce at scale. The result is not simply a larger quantity of false information, but a more difficult epistemic environment in which users must judge source quality, context, motive, evidence, and uncertainty under conditions of speed and attention scarcity (Ecker et al., 2022; Pennycook & Rand, 2021; Harris, 2022; Fallis, 2021).

Misinformation should therefore be understood as more than a collection of factually incorrect statements. Contemporary scholarship distinguishes false or misleading information from intentional deception, propaganda, satire, partisan framing, and other forms of information disorder, while recognizing that real-world messages may combine accurate elements with misleading implications. This is especially important because epistemic harm can arise from omission, decontextualization, authentic media paired with false captions, selective statistics, or true claims used to support an unwarranted conclusion. A bounded definition anchored in the best available evidence is more analytically useful than treating any contested claim as misinformation (Vraga & Bode, 2020; Thomson et al., 2022; Harris, 2022).

The source material provides a different but potentially complementary starting point. It frames *al-haqq* as a central Qur'anic concept associated with truth, correspondence, stability, and moral legitimacy, while *al-batil* is treated as its opposite: what is false, transient, deceptive, or ultimately without enduring value. The presentation repeatedly highlights Qur'anic injunctions not to mix truth with falsehood or conceal truth knowingly, and it gives particular prominence to the water-and-foam metaphor in Qur'an 13:17, where what benefits people remains while the foam disappears. Rather than importing these theological claims directly into empirical misinformation research, this article treats them as a normative vocabulary for examining epistemic responsibility.

That distinction matters. A religious text can articulate duties about truthfulness, verification, restraint, and moral accountability, but it cannot by itself establish the causal mechanisms through which people misperceive content on social media. Conversely, experiments on attention, cognitive reflection, identity, inoculation, or correction can identify reliable behavioral patterns without answering why truthfulness should be regarded as a moral obligation. The contribution of this article is therefore dialogical: Qur'anic ethics provides a normative architecture, while contemporary misinformation science provides empirical knowledge about vulnerabilities, interventions, and unintended consequences (Pennycook et al., 2021; Mosleh et al., 2021; van der Linden, 2022; Rathje et al., 2023).

The urgency of such an integration is particularly visible in Indonesia. The country's large, highly networked online population encounters misinformation through mainstream platforms, messaging services, religious networks, electoral

communication, and health discourse. Indonesian studies show that media literacy interventions can improve misinformation recognition and reduce sharing intentions, while research on Islamic communication emphasizes *tabayyun*, valid information, and careful speech as responses to hoaxes. Religious narratives themselves can also be incorporated into misleading discourse, demonstrating why religious authority must be part of the solution rather than assumed to be automatically protective (Thomas et al., 2022; Ford et al., 2023; Parhan et al., 2021; Jenuri et al., 2021; Triana et al., 2023).

This article asks six linked questions. How should *al-haqq* and *al-batil* be translated into an epistemic-ethical vocabulary suitable for contemporary communication analysis? How do fabrication, concealment, and mixtures of true and false material relate to modern misinformation? Why can unreliable information become visible and persuasive despite weak evidential foundations? What responsibilities follow from verification-oriented Qur'anic ethics before belief, before sharing, and after error? Which contemporary interventions correspond with these responsibilities, and where do they differ? Finally, how can an integrated framework contribute to digital communication ethics without conflating religious truth claims with empirical verification? The article answers these questions through an integrative qualitative literature review and develops the Qur'anic Epistemic Responsibility Framework (QERF).

LITERATURE REVIEW

The first conceptual foundation concerns the meaning of truth itself. In the initiating material, *al-haqq* is described as objective in the sense that its validity does not depend on whether people accept or reject it, and truthfulness is associated with correspondence between a statement and reality. That framing resonates with an epistemic concern central to misinformation research: veracity cannot be reduced to popularity, repetition, emotional appeal, or social endorsement. Contemporary studies show that people can judge accuracy relatively well when explicitly asked to do so, even though the same people may share content for reasons unrelated to accuracy, indicating a gap between evaluative capacity and communicative behavior (Pennycook et al., 2020; Pennycook et al., 2021; Pennycook & Rand, 2021).

A second foundation concerns *al-batil*. The source material does not portray falsehood only as an explicit lie; it also emphasizes mixing truth with falsehood, concealing known truth, and allowing appearances to obscure substance. This is analytically significant because many influential forms of misinformation contain genuine components. Misleading content may use a real image with a false description, a correct number without its denominator, a genuine quotation stripped of context, or a true event framed to imply a false causal story. Research on visual misinformation and disinformation and online fakes confirms that manipulation frequently operates by altering context and inference rather than fabricating every informational component (Thomson et al., 2022; Harris, 2022; Diakopoulos & Johnson, 2021).

Figure 1 positions the Qur'anic concepts of *al-haqq* and *al-batil* around a digital epistemic environment rather than treating them as isolated theological abstractions. It shows that the contemporary struggle over truth is mediated by

algorithms, virality, repetition, identity, trust, and synthetic media, all of which can influence visibility without determining veracity.

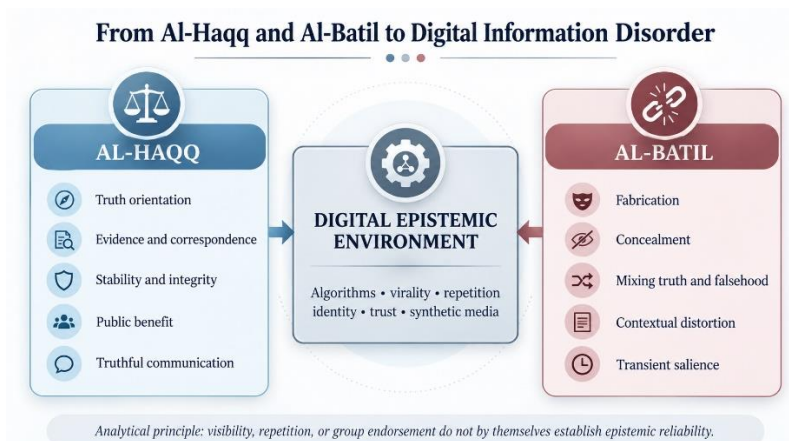


Figure 1. From Al-Haqq and Al-Batil to Digital Information Disorder

Figure 1 should be read as an interpretive map rather than a causal model. The left side identifies truth orientation, evidence, integrity, and public benefit as normative features associated with al-haqq in the source material; the right side identifies fabrication, concealment, mixing, contextual distortion, and transient salience as mechanisms associated with al-batil. The central digital environment indicates that these normative categories are filtered through technological and psychological systems that may reward attention independently of epistemic quality (Allen et al., 2020; Anderau, 2023; Fallis, 2021).

Psychological research helps explain why this mediation matters. People do not approach every claim through deliberate evidence evaluation. Attention is limited, familiar claims become easier to process, and social or partisan motivations can influence judgments. Cognitive reflection is associated with more discerning information behavior, but reflective capacity does not eliminate identity-based motivation; in political contexts, antagonism toward out-groups can predict sharing of unreliable material. Misinformation susceptibility is therefore better understood as a family of processes involving attention, prior beliefs, source cues, identity, emotion, reflective reasoning, and sharing motivations rather than as a single deficit in intelligence. A recent meta-analysis likewise shows that misinformation sharing is associated with multiple motivational, trust, and literacy-related traits rather than one uniform individual profile (Sun & Xie, 2024). Research also links social-media use to conspiracy and misinformation beliefs in ways that depend on prior predispositions and political orientations (Mosleh et al., 2021; Osmundsen et al., 2021; Rathje et al., 2023; Maertens et al., 2024; Enders et al., 2023).

The literature on conspiracy belief reinforces this multi-level perspective. Conspiracy theories can satisfy explanatory, existential, and social-identity motives, yet evidence does not support simplistic claims that conspiracism is uniformly increasing everywhere. Individual, intergroup, and national conditions shape susceptibility, and different conspiracy narratives mobilize different grievances and identities. For an epistemic-ethical framework, the implication is that truth-seeking requires more than condemning falsehood; it requires understanding why certain

claims become psychologically or socially attractive and designing practices that reduce those incentives without pathologizing entire groups (Douglas & Sutton, 2023; Hornsey et al., 2023; Uscinski et al., 2022).

The source material's water-and-foam metaphor offers a useful interpretive analogy for this problem. Foam is conspicuous at the surface but lacks the enduring usefulness attributed to the water beneath it, suggesting a distinction between salience and substance. Digital systems can reproduce an analogous separation: emotionally provocative material can receive disproportionate attention, while reliable but less spectacular information remains comparatively under-amplified. The analogy must be used carefully, because empirical popularity cannot be treated as evidence of falsity and low visibility cannot be treated as proof of truth; its value lies in reminding analysts not to confuse attention metrics with epistemic quality (Allen et al., 2020; Anderau, 2023).

Source trust adds another layer. Correction studies show that a retraction can be less effective when the source is perceived as untrustworthy, even when the correcting information is substantively sound. Trust is therefore neither a substitute for evidence nor an irrelevant social variable; it is part of the channel through which evidence becomes usable to audiences. This insight fits the ethical logic of *tabayyun*, which requires attention to who transmits information, what evidence supports it, and what consequences may follow from acting on it, while also suggesting that verification cannot be reduced to source reputation alone (Ecker & Antonio, 2021; Xiao, 2022).

The rise of deepfakes sharpens these questions by weakening the evidential status traditionally attached to recordings. Synthetic media can make false visual evidence more plausible, but the deeper problem is epistemic: once audiences know realistic fabrication is possible, genuine recordings may also become easier to dismiss. Philosophical work on deepfakes therefore emphasizes both direct deception and the broader erosion of evidential trust, while experimental communication research shows that synthetic political video can increase uncertainty and reduce trust even when viewers are not directly deceived. Media research consequently highlights the need for provenance, verification, and institutional practices capable of authenticating digital material (Fallis, 2021; de Ruiter, 2021; Diakopoulos & Johnson, 2021; Vaccari & Chadwick, 2020).

Corrective interventions form a major second body of literature. Fact-checks generally improve factual accuracy, but effects vary across topics, audiences, formats, and contexts. Global experiments show that fact-checking can work across diverse national settings, while meta-analytic evidence indicates that corrections rarely erase all influence of previously accepted misinformation. Effective correction must therefore be conceived as epistemic repair rather than a one-time deletion of false belief (Walter et al., 2020; Walter & Tukachinsky, 2020; Porter & Wood, 2021).

Prebunking and inoculation extend this logic upstream. Instead of waiting for a false claim to spread, these approaches teach people to recognize manipulation techniques such as emotional language, false dichotomies, impersonation, conspiratorial framing, or discrediting tactics. Large-scale studies and technique-based experiments show that inoculation can improve resilience, although

intervention durability is not unlimited and can depend on reinforcement and testing. This suggests a close functional parallel with ethical formation: durable epistemic responsibility requires habits, not isolated warnings (Roozenbeek et al., 2022a; Roozenbeek et al., 2022b; Capewell et al., 2024).

Media literacy is similarly promising but not universally effective. Experimental work in the United States and India found that a short digital literacy intervention improved discernment, while an intensive field intervention in India produced more mixed results and even revealed motivated-reasoning effects among strong political supporters. Recent meta-analyses nevertheless find that media literacy interventions, on average, improve misinformation discernment and credibility assessment, particularly when interventions are repeated or well targeted. Broader synthesis also positions media and information literacy as a societal capacity that links individual discernment with institutional accountability rather than as a stand-alone classroom technique (Mansoor, 2024). Reviews of psychological countermeasures similarly emphasize that debunking and prebunking work best when they are treated as complementary rather than mutually exclusive strategies. These findings support a layered approach in which skills training is combined with motivational, social, and institutional safeguards (Guess et al., 2020; Badrinathan, 2021; Huang et al., 2024; Lu et al., 2024; van der Linden et al., 2021).

Indonesia-specific studies provide evidence that such interventions can travel beyond Western settings. Online training has improved identification of misinformation and reduced intentions to share false headlines among Indonesian participants. At the same time, Islamic communication research repeatedly identifies *tabayyun*, honest communication, validation of information, and restraint in speech as relevant normative resources. These studies are valuable, but many remain primarily descriptive or theological; the analytical opportunity is to connect their normative insights to experimentally studied processes such as attention, verification, correction, and inoculation (Thomas et al., 2022; Ford et al., 2023; Parhan et al., 2021; Jenuri et al., 2021; Fauziyah, 2020).

The literature also warns against an overcorrection problem. Messages that repeatedly warn audiences about misinformation can increase generalized skepticism and reduce credibility judgments for accurate information. Recent experiments show that prominent anti-misinformation interventions can reduce belief in false information while simultaneously lowering confidence in true information, creating a risk that critical thinking becomes undirected cynicism. Newer evidence also suggests that repeated media attention to misinformation can contribute to epistemic uncertainty and, in some settings, lower trust in scientists, reinforcing the need to distinguish warranted skepticism from generalized distrust (Hoes et al., 2025). Any ethical framework centered on verification therefore needs a positive account of warranted trust, not merely suspicion (Hoes et al., 2024; van der Meer et al., 2023; Hameleers & van der Meer, 2024).

These findings suggest that the Qur'anic prohibition on mixing or concealing truth can be reconstructed as a broader responsibility to preserve the relationship between claims, evidence, context, and consequences. Such a reconstruction does not claim that contemporary misinformation categories are hidden inside scripture.

Instead, it treats the source material as a normative tradition that can illuminate questions empirical research often leaves open: what kinds of information conduct deserve moral praise or blame, what duties accompany communicative power, and what counts as responsible uncertainty when evidence is incomplete. This normative dimension is particularly important for religious communicators whose audiences may treat moral authority as a credibility shortcut (Hashmi et al., 2021; Triana et al., 2023; Fathurrohman et al., 2024).

METHODOLOGY

This study uses an integrative qualitative literature review rather than a systematic literature review. Its objective is conceptual synthesis across fields that use different vocabularies and research designs, including Qur'anic ethics, communication studies, social epistemology, cognitive psychology, political communication, media literacy, journalism, and information science. Because the purpose is theory development rather than exhaustive retrieval or pooled effect estimation, the review does not use PRISMA and does not claim that every relevant publication has been identified.

The initiating literature source is treated as contextual primary material rather than peer-reviewed evidence. Its recurring themes include the centrality of al-haqq, the opposition between truth and al-batil, the prohibition on mixing or concealing truth, the duty to speak truth even when difficult, and the water-and-foam metaphor of Qur'an 13:17. These themes were used to construct initial sensitizing concepts, but all empirical claims about misinformation, cognition, media effects, and interventions were grounded in peer-reviewed literature.

The academic evidence base prioritizes journal articles published from 2020 through 2026. Searches were organized around combinations of terms including misinformation, fake news, disinformation, epistemology, fact-checking, correction, inoculation, prebunking, media literacy, accuracy prompts, deepfakes, source credibility, conspiracy beliefs, Islam, Qur'an, tabayyun, Islamic communication, Indonesia, and digital ethics. Priority was given to established international journals and to Indonesia-specific studies that provided relevant contextual detail. The final synthesis emphasizes recent scholarship while using the Qur'an and the literature material as normative primary sources rather than counting them toward the journal-reference requirement.

Thematic coding proceeded in three stages. Initial coding identified mechanisms and duties such as truth orientation, fabrication, concealment, verification, source trust, attention, repetition, identity, correction, and public benefit. Axial coding then grouped these into six responsibility functions: receive, pause, verify, evaluate, communicate, and correct. Integrative coding extended those functions across five levels of the information environment: individual cognition, interpersonal networks, platform architecture, professional communicators, and public institutions.

A methodological safeguard was applied to prevent category errors. Qur'anic statements were not used as empirical proof of psychological mechanisms, and experimental findings were not treated as validation or invalidation of theological truth claims. Contemporary political or religious actors were not

classified as embodiments of *al-haqq* or *al-batil*. The review instead asks whether principles such as verification, non-concealment, proportional confidence, and correction can be articulated in a way that is normatively rooted yet empirically informed.

A second safeguard concerned confirmation bias. The review retained evidence that complicates intervention optimism, including field experiments with weak or heterogeneous effects and studies showing that anti-misinformation warnings can create generalized skepticism. It also retained evidence that misinformation exposure may sometimes prompt verification rather than passive acceptance. This balance is necessary because an epistemic-responsibility framework should increase accuracy without encouraging blanket distrust (Badrinathan, 2021; Xiao, 2022; Hoes et al., 2024; van der Meer et al., 2023).

RESULTS

The first thematic finding is that truth in the source material is simultaneously epistemic and ethical. *Al-haqq* is not merely a proposition that happens to be correct; it is connected to truthful speech, justice, stability, and conduct that should align with reality. Contemporary misinformation research makes a similar distinction between private belief accuracy and public communicative behavior. People may be capable of recognizing implausible content when prompted to judge accuracy yet still share it when attention is redirected toward novelty, identity, humour, or social reward (Pennycook et al., 2020; Pennycook et al., 2021).

This finding supports a two-part account of epistemic responsibility. The first part concerns belief formation: agents should seek evidence and calibrate confidence to what the evidence supports. The second concerns transmission: agents should consider whether they have sufficient grounds to amplify a claim and whether their framing preserves relevant context. The distinction is important because misinformation can spread through people who do not fully believe it, and sharing can function as identity performance rather than testimony (Pennycook & Rand, 2021; Osmundsen et al., 2021).

The second finding is that falsehood operates through more than fabrication. The literature repeatedly emphasizes the wrong of mixing truth with falsehood and concealing truth knowingly, which maps closely onto contemporary problems of selective framing, omission, false context, and misleading implication. Visual misinformation research likewise shows that authentic images can become deceptive when relabelled or decontextualized, while deepfake research shows that wholly synthetic content adds a new layer of manipulation. The ethical target is therefore the integrity of the evidence-to-claim relationship, not merely whether every sentence contains a literal falsehood (Thomson et al., 2022; Fallis, 2021; de Ruiter, 2021).

The third finding concerns visibility. False or misleading content can be attention-efficient because it may be novel, emotionally arousing, identity-affirming, or conflict-oriented. Yet information-ecosystem research cautions against assuming that fake news dominates all information exposure; its absolute share can be smaller than public rhetoric suggests, even while concentrated exposure among

certain users creates serious harms. The water-and-foam metaphor is useful precisely because it discourages an equation between surface prominence and durable value, but it should not be transformed into the simplistic claim that whatever is popular is false (Allen et al., 2020; Uscinski et al., 2022).

Figure 2 translates these findings into a responsibility cycle. The cycle begins before belief formation and continues after communication, reflecting the fact that epistemic responsibility is ongoing rather than exhausted by an initial verification decision.

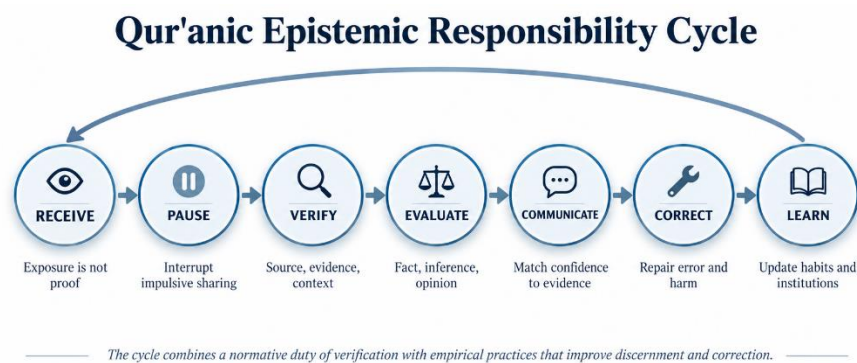


Figure 2. Qur'anic Epistemic Responsibility Cycle

The receive-pause-verify-evaluate-communicate-correct sequence combines normative and behavioral insights. The pause stage responds to evidence that attention can be shifted toward accuracy before sharing; verification connects with source and evidence checking; evaluation requires separating fact from inference and opinion; communication requires proportional confidence; and correction responds to evidence that false claims can continue influencing reasoning after retraction. The final learning stage emphasizes that habits and institutions should be updated after recurring failures (Pennycook et al., 2021; Ecker et al., 2022; Walter & Tukachinsky, 2020).

The fourth finding is that human vulnerability is predictable but heterogeneous. Reflective thinking, numeracy, trust in credible science, and information literacy are associated with better discernment in many studies, while identity conflict and polarized motivation can increase selective acceptance and sharing. Health-misinformation experiments also show that exposure to false vaccine claims can reduce vaccination intentions, illustrating that epistemic errors can translate into consequential behavioral choices. However, there is no single psychological profile of the misinformed person, and interventions that work in one population may not transfer unchanged to another. This argues against moralizing explanations that reduce misinformation to stupidity or bad character (Roozenbeek et al., 2020; Loomba et al., 2021; Mosleh et al., 2021; Osmundsen et al., 2021; Maertens et al., 2024).

The fifth finding is that verification is necessary but not sufficient. Verification can fail if audiences distrust the corrector, lack the skills to assess evidence, or encounter corrections after a false narrative has become integrated into memory and identity. It can also fail when platforms make verification costly relative to sharing. Studies of correction effectiveness, source credibility, and online

verification therefore point to the importance of both individual skill and environmental design (Ecker & Antonio, 2021; Zhang et al., 2022; Xiao, 2022).

The sixth finding is that correction must be treated as a process of repair. Fact-checks can reduce misperceptions across different national contexts, but the continued influence effect shows that people may still reason from retracted misinformation. Corrections are stronger when they provide a coherent alternative account and when trusted sources communicate them. Social correction is also common in digital environments, suggesting that ordinary users are not only recipients of misinformation but participants in epistemic repair (Walter et al., 2020; Walter & Tukachinsky, 2020; Porter & Wood, 2021; Bode & Vraga, 2021).

Prebunking contributes a preventive dimension. Inoculation interventions teach recognition of manipulation techniques before a specific false claim is encountered, and experimental evidence shows improved resilience across social-media contexts. However, recent longitudinal work indicates that effects can decay unless reinforced, making repeated practice important. The ethical analogue is the cultivation of stable habits of caution rather than occasional reactive skepticism (Roozenbeek et al., 2022a; Roozenbeek et al., 2022b; Capewell et al., 2024).

Media literacy emerges as a complementary but context-dependent strategy. Information literacy has been associated with better identification of fake news, digital literacy interventions have improved discernment in experimental settings, and meta-analyses report overall positive effects. Yet field evidence from India shows that political motivation can override training effects, and intervention design matters substantially. The practical lesson is that knowledge and motivation must be addressed together (Jones-Jang et al., 2021; Guess et al., 2020; Badrinathan, 2021; Huang et al., 2024; Lu et al., 2024).

The seventh finding is that religious communication requires special attention to authority. In Muslim contexts, appeals to Qur'anic verses, hadith, religious identity, or respected figures can provide moral motivation for verification, but they can also be used to increase the perceived credibility of misleading content. Indonesian research documents both Islamic ethical responses to hoaxes and the presence of religious narratives within misinformation. A responsible framework must therefore combine reverence for normative sources with verification of empirical claims made in their name (Parhan et al., 2021; Jenuri et al., 2021; Triana et al., 2023; Fathurrohman et al., 2024).

The eighth finding is that epistemic responsibility is distributed. Individuals decide what to believe and share, but platforms shape visibility, journalists and religious communicators shape interpretation, educators shape skills, and public institutions shape incentives and standards. The problem of misinformation is therefore neither purely personal nor purely technological. Effective responses require aligned responsibilities across the whole information environment (van der Linden, 2022; Diakopoulos & Johnson, 2021; Li & Chang, 2022).

DISCUSSION

The synthesis suggests that the most productive way to translate the literature into an academic contribution is to move from a binary label to a process theory. Al-haqq and al-batil remain normative categories, but the key scholarly question becomes how information practices move claims toward or away from

evidential integrity. Such a process approach is more useful than assigning moral labels to competing political camps because it focuses on observable duties: disclose evidence, preserve context, verify before amplification, distinguish knowledge from inference, and correct errors transparently.

This shift also protects the framework from epistemic arrogance. Commitment to truth is not identical to certainty about every contested empirical issue. A person can affirm a strong moral duty toward truth while acknowledging incomplete evidence, revising beliefs, or deferring to qualified expertise. Indeed, the contemporary misinformation literature shows that overconfidence and identity-protective cognition can be as damaging as ignorance, while robust truth-seeking often requires calibrated uncertainty and willingness to update (Ecker et al., 2022; Rathje et al., 2023; Hornsey et al., 2023).

The prohibition on mixing truth and falsehood becomes especially powerful when applied to hybrid messages. Many digital claims are neither wholly true nor wholly fabricated. An authentic photograph can be paired with a false date, an accurate statistic can omit a relevant baseline, and a real quotation can be edited so that its pragmatic meaning is reversed. The Qur'anic ethical concern with mixture and concealment therefore provides a richer normative vocabulary than a narrow focus on lying, while modern verification research supplies the methods needed to evaluate these hybrid forms (Vraga & Bode, 2020; Thomson et al., 2022; Harris, 2022).

Figure 3 develops this argument by separating visibility from epistemic substance. The vertical dimension captures attention, while the horizontal dimension captures evidential reliability; the resulting matrix prevents popularity from functioning as a proxy for truth.

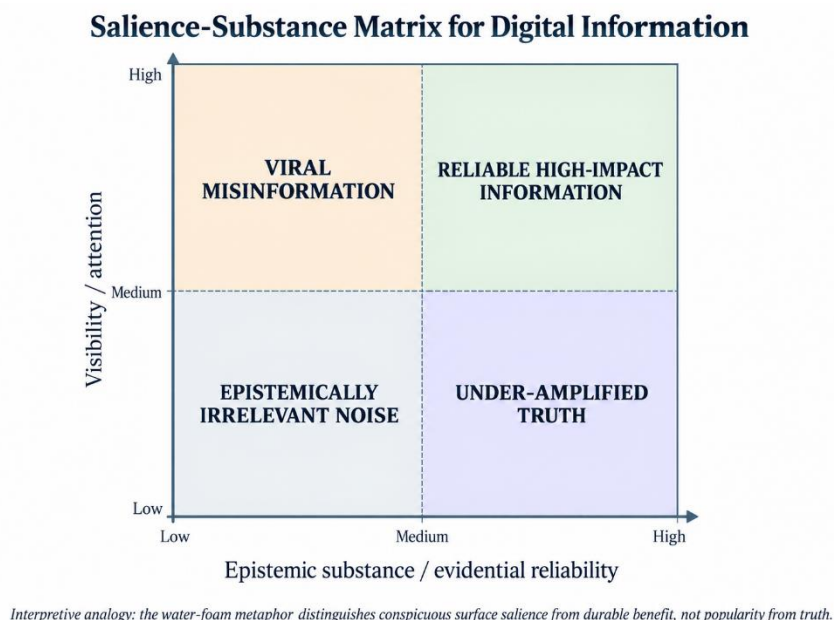


Figure 3. Salience-Substance Matrix for Digital Information

The matrix clarifies four possibilities. Reliable high-impact information combines visibility with substance and is normatively desirable when public

importance warrants amplification. Viral misinformation is visible but weakly grounded; under-amplified truth is reliable but receives insufficient attention; and low-value noise lacks both reliability and meaningful public salience. The model interprets the source material's water-and-foam metaphor as a warning against surface-level inference, not as an empirical law that all visible content is false (Allen et al., 2020; Anderau, 2023).

This distinction has consequences for platform governance. If ranking systems reward engagement without sufficient attention to provenance or manipulation, highly salient low-substance content can gain an advantage. Yet simply suppressing all controversial content would create other harms, including overreach and erosion of legitimate debate. A stewardship approach should instead increase friction for high-risk sharing, improve provenance signals, support independent verification, and preserve transparent appeal mechanisms, particularly as synthetic media make authenticity harder to infer from appearance alone (Fallis, 2021; Diakopoulos & Johnson, 2021; Hameleers, 2022). Recent governance research further indicates that platform self-governance can be insufficient when commercial operating principles and societal responsibilities are poorly aligned, strengthening the case for transparent, multi-stakeholder governance arrangements (Warnke et al., 2024; Huang & Krafft, 2024).

The QERF developed here has six operational stages. Receiving requires epistemic openness without assuming that exposure equals truth. Pausing interrupts the speed advantage of impulsive sharing. Verification examines source, evidence, corroboration, and context. Evaluation distinguishes fact, inference, probability, value judgment, and speculation. Communication aligns wording and confidence with evidential support. Correction requires acknowledging mistakes and repairing downstream harm rather than quietly deleting an inaccurate post.

The empirical literature provides support for each stage. Accuracy prompts demonstrate the value of redirecting attention before sharing, information-literacy studies support verification skills, and correction research shows that repair is more effective when it is credible and explanatory. At the same time, studies of spillover skepticism warn that a verification culture should not teach people to distrust everything. The goal is disciplined trust: confidence proportionate to evidence, track record, transparency, and institutional safeguards (Pennycook et al., 2021; Guess et al., 2020; Ecker & Antonio, 2021; Hoes et al., 2024).

The relationship between *tabayyun* and fact-checking deserves particular precision. *Tabayyun* is a moral-epistemic orientation toward clarification and verification before acting on consequential information; fact-checking is a modern professional practice using documentary evidence, expert sources, data, and transparent methods. They overlap in function but are not identical. Treating *tabayyun* simply as an ancient word for fact-checking would flatten both traditions, whereas treating them as complementary allows religious motivation to support rigorous contemporary verification practices (Fauziyah, 2020; Parhan et al., 2021; Jenuri et al., 2021).

A similar distinction applies to prebunking. Inoculation theory seeks to build resistance by exposing people to weakened examples of manipulation techniques and explaining how those techniques work. Qur'anic moral formation seeks to

cultivate character and responsibility, not merely pattern recognition. Yet both approaches are preventive and habit-oriented, suggesting a productive educational synthesis in which students learn manipulation techniques while also reflecting on the moral consequences of transmitting unverified claims (Roozenbeek et al., 2022a; Roozenbeek et al., 2022b; Huang et al., 2024).

For Indonesia, this synthesis has direct educational implications. Digital-literacy curricula can incorporate locally meaningful ethical language while retaining universal verification practices such as lateral reading, source tracing, reverse-image search, claim decomposition, and comparison with primary documents. The positive findings from Indonesian media literacy interventions suggest that context-sensitive programs can reduce misinformation vulnerability, but the mixed evidence from other Global South settings indicates that sustained reinforcement and attention to political identity are necessary (Thomas et al., 2022; Ford et al., 2023; Badrinathan, 2021).

Religious education can add a second layer by treating information conduct as part of moral formation. Students can be taught that repeating an unverified claim is not ethically neutral merely because the claim benefits their preferred group, and that correcting one's own error is a form of integrity rather than humiliation. This is especially important in polarized environments where misinformation can become a badge of loyalty. Research on partisan sharing and conspiracy beliefs shows how group identity can compete with accuracy motives, reinforcing the need for norms that reward correction and epistemic humility (Osmundsen et al., 2021; Douglas & Sutton, 2023; Hornsey et al., 2023).

Journalists and professional communicators also have responsibilities beyond fact-checking individual claims. Headline design, visual context, corrections policy, source disclosure, and uncertainty communication can either protect or damage epistemic trust. Visual-verification research emphasizes the practical difficulty of authenticating images, while studies of correction show that explanatory detail and source credibility matter. News organizations should therefore make verification methods visible enough to create warranted trust rather than relying only on institutional authority (Thomson et al., 2022; Vraga et al., 2022; Zhang et al., 2022).

Religious leaders and influencers face an analogous challenge. Because followers may grant them high moral credibility, they carry a heightened duty to distinguish scriptural interpretation from empirical claims about politics, health, economics, or science. The misuse of religious narratives in hoax discourse demonstrates the danger of allowing moral authority to substitute for empirical validation. Responsible religious communication should therefore disclose uncertainty, consult relevant expertise, and correct factual errors even when the correction is socially uncomfortable (Triana et al., 2023; Fathurrohman et al., 2024; Hashmi et al., 2021).

Figure 4 extends QERF beyond the individual. It places cognitive responsibility inside wider interpersonal, technological, professional, and institutional layers, emphasizing that responsible communication depends on environments as well as personal virtue.

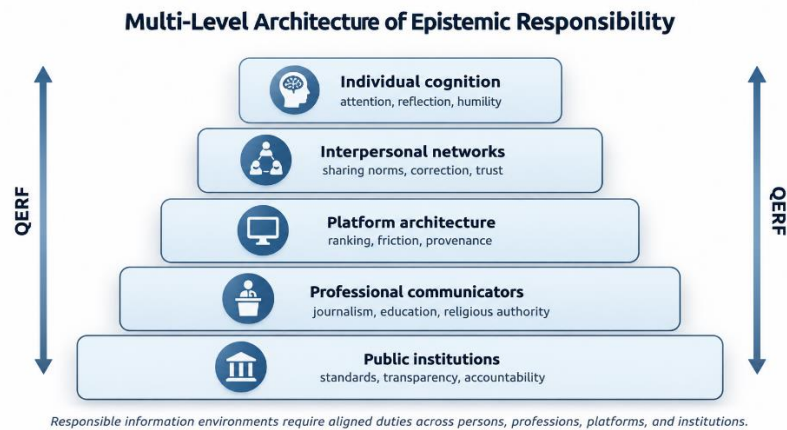


Figure 4. Multi-Level Architecture of Epistemic Responsibility

At the individual level, QERF emphasizes attention, reflection, and humility; at the network level, it emphasizes sharing norms and correction; at the platform level, it emphasizes friction, provenance, and ranking incentives; at the professional level, it applies to journalism, education, and religious authority; and at the institutional level, it concerns standards, transparency, and accountability. This multi-level design reflects evidence that misinformation exposure and correction are shaped by social context and infrastructure, not merely by isolated cognitive errors (van der Linden, 2022; Bode & Vraga, 2021; Li & Chang, 2022).

The framework also helps explain why anti-misinformation policy should not be reduced to censorship. Some false claims can be countered through correction, provenance, friction, or media literacy without removing content, while more harmful cases may require stronger action under clearly defined rules. From an epistemic perspective, legitimacy depends on whether interventions improve access to reliable evidence while preserving contestability and avoiding the suppression of genuine uncertainty. Overbroad interventions can generate mistrust, making transparency and proportionality essential (Hoes et al., 2024; van der Meer et al., 2023; Hameleers & van der Meer, 2024). Platform-governance scholarship likewise cautions that legitimacy depends not only on formal rules but also on who sets them, how data and moderation practices are made visible, and whether external stakeholders can scrutinize those arrangements (Huang & Krafft, 2024; Warnke et al., 2024).

The stakes extend beyond individual error because democratic decision-making depends on a sufficiently shared base of reliable public knowledge. When organized disinformation systematically discredits evidence-producing institutions or makes factual agreement impossible, the harm concerns the epistemic infrastructure of collective self-government rather than only isolated false beliefs. This broader perspective reinforces the need to protect both critical scrutiny and warranted institutional trust (Lewandowsky et al., 2023).

The framework's strongest contribution may therefore be its concept of epistemic stewardship. Stewardship differs from control: it does not presume that one institution should decide all truth claims. Instead, it requires actors with informational power to preserve conditions under which claims can be checked, sources traced, uncertainty communicated, and errors corrected. This is compatible

with pluralistic public debate because it regulates information practices rather than enforcing ideological agreement.

Several limitations should guide interpretation. The review is qualitative and integrative rather than systematic, so it does not estimate the prevalence or average causal effect of misinformation interventions. The Islamic communication literature used here is smaller and methodologically more heterogeneous than the psychological literature, and the initiating source material is a teaching document rather than a peer-reviewed scholarly source. Future research should therefore test QERF empirically, including experiments that compare purely technical media literacy interventions with interventions that add ethical motivation, especially in Indonesia and other Muslim-majority societies.

Future work should also examine whether moral framing improves correction willingness without increasing moralized polarization. A promising design would measure not only truth discernment but also confidence calibration, willingness to delay sharing, readiness to correct one's own post, and trust in reliable institutions. Such studies could determine whether Qur'anic ethical language functions as a motivational complement to media literacy, whether effects vary by religiosity or political identity, and whether repeated practice produces durable epistemic habits rather than short-lived compliance (Capewell et al., 2024; Maertens et al., 2024).

CONCLUSION

The contemporary struggle between truth and falsehood is not only a contest between propositions; it is a contest between information practices. The initiating source material's emphasis on *al-haqq*, *al-batil*, non-concealment, truthful speech, and the water-and-foam metaphor can be reconstructed as a normative vocabulary for responsible digital communication, provided that theological claims are not confused with empirical evidence. The resulting perspective asks how claims are received, verified, contextualized, transmitted, and corrected rather than merely who wins a rhetorical contest.

The qualitative synthesis shows that misinformation succeeds through multiple mechanisms, including inattention, repetition, source heuristics, identity, emotional salience, contextual distortion, and technological amplification. It also shows that no single intervention is sufficient. Fact-checking, correction, inoculation, accuracy prompts, and media literacy can all improve epistemic outcomes, but their effects depend on trust, motivation, reinforcement, context, and the possibility of unintended skepticism toward accurate information.

The Qur'anic Epistemic Responsibility Framework therefore proposes six linked duties: receive without equating exposure with truth, pause before impulsive transmission, verify source and evidence, evaluate the distinction between fact and inference, communicate proportionately to evidential confidence, and correct errors when they are discovered. These duties operate across individual cognition, interpersonal networks, platforms, professional communicators, and public institutions. Their purpose is not to create a single authority over truth, but to strengthen the conditions under which truthful claims can be distinguished from distortion.

For Indonesia, the framework offers a bridge between culturally resonant moral language and contemporary digital-literacy practice. It supports a model of *tabayyun* that is neither purely devotional nor merely technical: verification becomes a disciplined moral practice informed by evidence, humility, and responsibility for downstream harm. The broader proposition is that commitment to truth requires more than confidence in one's own beliefs. It requires institutions and habits that reward verification before belief, restraint before transmission, transparency about uncertainty, correction after error, and the preservation of public trust through accountable communication.

REFERENCES

- Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14), eaay3539. <https://doi.org/10.1126/sciadv.aay3539>
- Anderau, G. (2023). Fake news and epistemic flooding. *Synthese*, 202, 106. <https://doi.org/10.1007/s11229-023-04336-7>
- Badrinathan, S. (2021). Educative interventions to combat misinformation: evidence from a field experiment in India. *American Political Science Review*, 115(4), 1325-1341. <https://doi.org/10.1017/S0003055421000459>
- Bode, L., & Vraga, E. K. (2021). Correction experiences on social media during COVID-19. *Social Media + Society*, 7(2). <https://doi.org/10.1177/20563051211008829>
- Capewell, G., Maertens, R., Remshard, M., van der Linden, S., Compton, J., Lewandowsky, S., & Roozenbeek, J. (2024). Misinformation interventions decay rapidly without an immediate posttest. *Journal of Applied Social Psychology*, 54, 441-454. <https://doi.org/10.1111/jasp.13049>
- de Ruiter, A. (2021). The distinct wrong of deepfakes. *Philosophy & Technology*, 34(4), 1311-1332. <https://doi.org/10.1007/s13347-021-00459-2>
- Diakopoulos, N., & Johnson, D. (2021). Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society*, 23(7), 2072-2098. <https://doi.org/10.1177/1461444820925811>
- Douglas, K. M., & Sutton, R. M. (2023). What are conspiracy theories? A definitional approach to their correlates, consequences, and communication. *Annual Review of Psychology*, 74, 271-298. <https://doi.org/10.1146/annurev-psych-032420-031329>
- Ecker, U. K. H., & Antonio, L. M. (2021). Can you believe it? An investigation into the impact of retraction source credibility on the continued influence effect. *Memory & Cognition*, 49(4), 631-644. <https://doi.org/10.3758/s13421-020-01129-y>
- Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1, 13-29. <https://doi.org/10.1038/s44159-021-00006-y>
- Enders, A. M., Uscinski, J. E., Seelig, M. I., Klostad, C. A., Wuchty, S., Funchion, J. R., Murthi, M. N., Premaratne, K., & Stoler, J. (2023). The relationship between social media use and beliefs in conspiracy theories and misinformation.

- Political Behavior, 45(2), 781-804. <https://doi.org/10.1007/s11109-021-09734-6>
- Fallis, D. (2021). The epistemic threat of deepfakes. *Philosophy & Technology*, 34(4), 623-643. <https://doi.org/10.1007/s13347-020-00419-2>
- Fathurrohman, R., Anna, D. N., Sibawaihi, & Fahmi, M. Z. (2024). Framing religious hoaxes: logical, religious, and legal perspectives in Indonesia. *Religious: Jurnal Studi Agama-Agama dan Lintas Budaya*, 8(2). <https://doi.org/10.15575/rjsalb.v8i2.15545>
- Fauziyah, U. (2020). Tabayyun dan hukumnya sebagai penanggulangan berita hoax di era digital dalam perspektif fiqih. *Al Yasini: Jurnal Keislaman, Sosial, Hukum dan Pendidikan*, 5(1), 114-125. <https://doi.org/10.55102/alyasini.v5i1.3775>
- Ford, T., Yankoski, M., Facciani, M., & Weninger, T. (2023). Online media literacy intervention in Indonesia reduces misinformation sharing intention. *Journal of Media Literacy Education*, 15(2), 99-123. <https://doi.org/10.23860/JMLE-2023-15-2-8>
- Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences*, 117(27), 15536-15545. <https://doi.org/10.1073/pnas.1920498117>
- Hameleers, M. (2022). Separating truth from lies: comparing the effects of news media literacy interventions and fact-checkers in response to political misinformation in the US and Netherlands. *Information, Communication & Society*, 25(1), 110-126. <https://doi.org/10.1080/1369118X.2020.1764603>
- Hameleers, M., & van der Meer, T. G. L. A. (2024). Striking the balance between fake and real: under what conditions can media literacy messages that warn about misinformation maintain trust in accurate information?. *Behaviour & Information Technology*, 43, 1209-1221. <https://doi.org/10.1080/0144929X.2023.2267700>
- Harris, K. R. (2022). Real fakes: the epistemology of online misinformation. *Philosophy & Technology*, 35, 83. <https://doi.org/10.1007/s13347-022-00581-9>
- Hashmi, U. M., Rashid, R. A., & Ahmad, M. K. (2021). The representation of Islam within social media: a systematic review. *Information, Communication & Society*, 24(13), 1962-1981. <https://doi.org/10.1080/1369118X.2020.1847165>
- Hoes, E., Aitken, B., Zhang, J., Gackowski, T., & Wojcieszak, M. (2024). Prominent misinformation interventions reduce misperceptions but increase scepticism. *Nature Human Behaviour*, 8, 1545-1553. <https://doi.org/10.1038/s41562-024-01884-x>
- Hoes, E., Clemm, B., Gessler, T., Qian, S., & Wojcieszak, M. (2025). (Media attention to) misinformation can undermine trust in scientists. *Political Behavior*. <https://doi.org/10.1007/s11109-025-10090-y>
- Hornsey, M. J., Bierwaczzonek, K., Sassenberg, K., & Douglas, K. M. (2023). Individual, intergroup and nation-level influences on belief in conspiracy

- theories. *Nature Reviews Psychology*, 2, 85-97. <https://doi.org/10.1038/s44159-022-00133-0>
- Huang, G., Jia, W., & Yu, W. (2024). Media literacy interventions improve resilience to misinformation: a meta-analytic investigation of overall effect and moderating factors. *Communication Research*. <https://doi.org/10.1177/00936502241288103>
- Huang, K., & Krafft, P. M. (2024). Performing platform governance: Facebook and the stage management of data relations. *Science and Engineering Ethics*, 30, 13. <https://doi.org/10.1007/s11948-024-00473-5>
- Jenuri, Suwarma, D. M., Parhan, M., Sartika, A., Ramdani, A. D., & Rahmah, F. A. (2021). Islamic communication ethics towards hoax phenomenon on social media. *Dialogia*, 19(2), 515-534. <https://doi.org/10.21154/dialogia.v19i2.3446>
- Jones-Jang, S. M., Mortensen, T., & Liu, J. (2021). Does media literacy help identification of fake news? Information literacy helps, but other literacies do not. *American Behavioral Scientist*, 65(2), 371-388. <https://doi.org/10.1177/0002764219869406>
- Lewandowsky, S., Ecker, U. K. H., Cook, J., van der Linden, S., Roozenbeek, J., & Oreskes, N. (2023). Misinformation and the epistemic integrity of democracy. *Current Opinion in Psychology*, 54, 101711. <https://doi.org/10.1016/j.copsyc.2023.101711>
- Li, J., & Chang, X. (2022). Combating misinformation by sharing the truth: a study on the spread of fact-checks on social media. *Information Systems Frontiers*. <https://doi.org/10.1007/s10796-022-10296-z>
- Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K., & Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour*, 5, 337-348. <https://doi.org/10.1038/s41562-021-01056-1>
- Lu, C., Hu, B., Bao, M.-M., Wang, C., Bi, C., & Ju, X.-D. (2024). Can media literacy intervention improve fake news credibility assessment? A meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, 27(4), 240-252. <https://doi.org/10.1089/cyber.2023.0324>
- Maertens, R., Götz, F. M., Golino, H. F., Roozenbeek, J., Schneider, C. R., Kyrychenko, Y., Kerr, J. R., Stieger, S., McClanahan, W. P., Drabot, K., He, J., & van der Linden, S. (2024). The Misinformation Susceptibility Test (MIST): a psychometrically validated measure of news veracity discernment. *Behavior Research Methods*, 56(3), 1863-1899. <https://doi.org/10.3758/s13428-023-02124-2>
- Mansoor, H. M. H. (2024). Media and information literacy as a model of societal balance: a grounded meta-synthesis. *Heliyon*, 10(3), e25380. <https://doi.org/10.1016/j.heliyon.2024.e25380>
- Mosleh, M., Pennycook, G., Arechar, A. A., & Rand, D. G. (2021). Cognitive reflection correlates with behavior on Twitter. *Nature Communications*, 12, 921. <https://doi.org/10.1038/s41467-020-20043-0>
- Osmundsen, M., Bor, A., Vahlstrup, P. B., Bechmann, A., & Petersen, M. B. (2021). Partisan polarization is the primary psychological motivation behind political

- fake news sharing on Twitter. *American Political Science Review*, 115(3), 999-1015. <https://doi.org/10.1017/S0003055421000290>
- Parhan, M., Jenuri, & Islamy, M. R. F. (2021). Media sosial dan fenomena hoax: tinjauan Islam dalam etika berkomunikasi. *Communicatus: Jurnal Ilmu Komunikasi*, 5(1), 59-80. <https://doi.org/10.15575/cjik.v5i1.12887>
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388-402. <https://doi.org/10.1016/j.tics.2021.02.007>
- 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, 590-595. <https://doi.org/10.1038/s41586-021-03344-2>
- Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31(7), 770-780. <https://doi.org/10.1177/0956797620939054>
- Porter, E., & Wood, T. J. (2021). The global effectiveness of fact-checking: evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom. *Proceedings of the National Academy of Sciences*, 118(37), e2104235118. <https://doi.org/10.1073/pnas.2104235118>
- Rathje, S., Roozenbeek, J., Van Bavel, J. J., & van der Linden, S. (2023). Accuracy and social motivations shape judgements of (mis)information. *Nature Human Behaviour*, 7, 892-903. <https://doi.org/10.1038/s41562-023-01540-w>
- Roozenbeek, J., Schneider, C. R., Dryhurst, S., Kerr, J., Freeman, A. L. J., Recchia, G., van der Bles, A. M., & van der Linden, S. (2020). Susceptibility to misinformation about COVID-19 around the world. *Royal Society Open Science*, 7, 201199. <https://doi.org/10.1098/rsos.201199>
- Roozenbeek, J., Traberger, C. S., & van der Linden, S. (2022b). Technique-based inoculation against real-world misinformation. *Royal Society Open Science*, 9(5), 211719. <https://doi.org/10.1098/rsos.211719>
- Roozenbeek, J., van der Linden, S., Goldberg, B., Rathje, S., & Lewandowsky, S. (2022a). Psychological inoculation improves resilience against misinformation on social media. *Science Advances*, 8(34), eabo6254. <https://doi.org/10.1126/sciadv.abo6254>
- Sun, Y., & Xie, J. (2024). Who shares misinformation on social media? A meta-analysis of individual traits related to misinformation sharing. *Computers in Human Behavior*, 158, 108271. <https://doi.org/10.1016/j.chb.2024.108271>
- Thomas, P., Hogan-Taylor, C., Yankoski, M., & Weninger, T. (2022). Pilot study suggests online media literacy programming reduces belief in false news in Indonesia. *First Monday*, 27(1). <https://doi.org/10.5210/fm.v27i1.11683>
- Thomson, T. J., Angus, D., Dootson, P., Hurcombe, E., & Smith, A. (2022). Visual mis/disinformation in journalism and public communications: current verification practices, challenges, and future opportunities. *Journalism Practice*, 16(5), 938-962. <https://doi.org/10.1080/17512786.2020.1832139>
- Triana, H. W., Oktavianus, Ilham, M., & Khairunnisa, A. (2023). Religious narratives during the pandemic: transitivity analysis of Covid-19 hoax discourse on Indonesian social media. *Jurnal Arbitrer*, 10(1), 66-75. <https://doi.org/10.25077/ar.10.1.66-75.2023>

- Uscinski, J. E., Enders, A. M., Klofstad, C., Seelig, M., Drochon, H., Premaratne, K., & Murthi, M. (2022). Have beliefs in conspiracy theories increased over time?. *PLOS ONE*, 17(7), e0270429. <https://doi.org/10.1371/journal.pone.0270429>
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1). <https://doi.org/10.1177/2056305120903408>
- van der Linden, S. (2022). Misinformation: susceptibility, spread, and interventions to immunize the public. *Nature Medicine*, 28, 460-467. <https://doi.org/10.1038/s41591-022-01713-6>
- van der Linden, S., Roozenbeek, J., Maertens, R., Basol, M., Kacha, O., Rathje, S., & Traberg, C. S. (2021). How can psychological science help counter the spread of fake news?. *The Spanish Journal of Psychology*, 24, e25. <https://doi.org/10.1017/SJP.2021.23>
- van der Meer, T. G. L. A., Hameleers, M., & Ohme, J. (2023). Can fighting misinformation have a negative spillover effect? How warnings for the threat of misinformation can decrease general news credibility. *Journalism Studies*, 24(6), 803-823. <https://doi.org/10.1080/1461670X.2023.2187652>
- Vraga, E. K., & Bode, L. (2020). Defining misinformation and understanding its bounded nature: using expertise and evidence for describing misinformation. *Political Communication*, 37(1), 136-144. <https://doi.org/10.1080/10584609.2020.1716500>
- Vraga, E. K., & Bode, L. (2021). Addressing COVID-19 misinformation on social media preemptively and responsively. *Emerging Infectious Diseases*, 27(2), 396-403. <https://doi.org/10.3201/eid2702.203139>
- Vraga, E. K., Bode, L., & Tully, M. (2022). Creating news literacy messages to enhance expert corrections of misinformation on Twitter. *Communication Research*, 49(2), 245-267. <https://doi.org/10.1177/0093650219898094>
- Walter, N., & Tukachinsky, R. (2020). A meta-analytic examination of the continued influence of misinformation in the face of correction: how powerful is it, why does it happen, and how to stop it?. *Communication Research*, 47(2), 155-177. <https://doi.org/10.1177/0093650219854600>
- Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2020). Fact-checking: a meta-analysis of what works and for whom. *Political Communication*, 37(3), 350-375. <https://doi.org/10.1080/10584609.2019.1668894>
- Warnke, L., Maier, A.-L., & Gilbert, D. U. (2024). Social media platforms' responses to COVID-19-related mis- and disinformation: the insufficiency of self-governance. *Journal of Management and Governance*, 28, 1079-1115. <https://doi.org/10.1007/s10997-023-09694-5>
- Xiao, X. (2022). Not doomed: examining the path from misinformation exposure to verification and correction in the context of COVID-19 pandemic. *Telematics and Informatics*, 74, 101890. <https://doi.org/10.1016/j.tele.2022.101890>
- Zhang, Y., Guo, B., Ding, Y., Liu, J., Qiu, C., Liu, S., & Yu, Z. (2022). Investigation of the determinants for misinformation correction effectiveness on social media during COVID-19 pandemic. *Information Processing & Management*, 59(3), 102935. <https://doi.org/10.1016/j.ipm.2022.102935>