

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

Vo Minh Vinh¹, Nguyen Tien Dung²

¹Faculty of Management and Economics, Tomas Bata University, Czech Republic **Orchid** <https://orcid.org/0009-0007-4100-770X>

²High school for the Gifted, Vietnam National University Ho Chi Minh City, Vietnam

ABSTRACT: This study investigates the influence of ethical governance in AI-generated content, persuasive intensity in AI-driven messaging, and perceived authenticity of synthetic media on consumer trust within AI-powered brand communication, while also exploring how consumer skepticism toward AI persuasion moderates the relationship between persuasive intensity and trust. Drawing on Signaling Theory, Source Credibility Theory, and the Persuasion Knowledge Model, the research frames consumer trust as a multifaceted outcome shaped by ethical transparency, message genuineness, and the dynamics of algorithmic persuasion in an increasingly synthetic media landscape. Employing a quantitative cross-sectional approach, data were gathered from 385 respondents via an online survey utilizing 5-point Likert scales. The data were analyzed using SPSS and Hayes' PROCESS macro, incorporating reliability checks, exploratory factor analysis, multiple regression, and moderation analysis. Results showed that ethical governance ($\beta = 0.538$), persuasive intensity ($\beta = 0.672$), and perceived authenticity ($\beta = 0.501$) all exerted strong, statistically significant positive effects on consumer trust. In addition, consumer skepticism significantly moderated the link between persuasive intensity and trust through a positive interaction effect ($\beta = 0.284$), suggesting that skepticism can actually strengthen rather than diminish the impact of persuasive efforts. The findings indicate that trust in AI-driven brand communication stems not only from technological advancement but from the thoughtful integration of ethical governance, authentic content creation, and well-calibrated persuasion, shaped further by consumers' cognitive responses. By combining several theoretical perspectives, this research advances a more refined understanding of trust formation in AI-mediated marketing and delivers practical insights for researchers and practitioners grappling with the ethical and strategic complexities of AI-enabled communication.

KEYWORDS: Consumer trust; AI-driven brand communication; Ethical governance; Persuasive intensity; Perceived authenticity; Consumer skepticism.

1. INTRODUCTION

Consumer trust in AI-powered brand messaging has emerged as a pressing issue in marketing research, largely because generative AI and deepfakes enable companies to create highly convincing and scalable content while making it harder for people to evaluate the true source, authenticity, and underlying motives (Diel et al., 2024; Whittaker et al., 2025). At its core, the challenge goes beyond whether such content works; what truly matters is whether consumers see it as ethically managed. When brands openly share how AI is used, it often boosts perceptions of genuineness and leads to more positive reactions, showing that trust hinges more on honesty than on technical polish alone (Yang & Battocchio, 2021). Still, simply labeling AI-generated material is not enough—disclosures can sometimes spark suspicion or lower confidence if viewers view them as signs of hidden manipulation or less human responsibility (Baek et al., 2024). Adding to this, the growing sophistication of AI-driven persuasion can heighten consumer wariness, especially as people become more aware of algorithmic targeting, with skepticism spreading unevenly and complicating trust-building efforts (Voorveld et al., 2024). Ultimately, trust feels particularly vulnerable when lifelike synthetic content, unclear oversight, and strong persuasive tactics come together in brand communications.

Prior and ongoing studies on consumer trust in AI-driven brand communication have increasingly examined deepfake advertising, algorithmic transparency, and perceived authenticity, yet they remain theoretically scattered and methodologically constrained. Existing work shows that deepfakes and AI-produced messages can strongly shape how credible and trustworthy a brand appears, often through processes like persuasion knowledge and expectancy violations (Yang & Lee, 2026). Research also indicates that

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

clear transparency and responsible AI practices tend to build trust, whereas concealed or exploitative approaches damage long-term consumer relationships (Park & Yoon, 2024; Ribeiro, 2025). Nevertheless, results are mixed; some findings reveal little to no difference in trust levels between AI-generated and human-made content, pointing to lingering theoretical uncertainty (Zhang & Hur, 2025). A major shortcoming is the tendency to view trust as a fixed snapshot rather than a fluid, situation-specific process, which creates inconsistent measurements and makes it difficult to compare studies (Benk et al., 2022). Moreover, most investigations examine individual elements—like disclosure or authenticity—in isolation, without combining ethical governance and persuasion strength into a single, cohesive framework. This paper aims to fill these gaps by proposing an integrated model that explores the connections among ethical governance, authenticity, and persuasive intensity, while treating consumer skepticism as a key moderating factor. In doing so, it seeks to offer a more complete and theoretically robust explanation of how trust develops or breaks down in AI-supported brand interactions.

Therefore, the present study seeks to examine the following research questions:

- (1) To what extent does ethical governance of AI-generated content influence consumer trust in AI-driven brand communication?
- (2) How do perceived authenticity of synthetic media and persuasive intensity of AI-driven communication affect consumer trust?
- (3) How does consumer skepticism toward AI persuasion moderate the relationship between persuasive intensity and consumer trust?

This study seeks to explore how ethical governance practices, particularly transparency, accountability, and clear disclosure, directly influence consumer trust in brand messages shaped by AI. Beyond that, it aims to understand the combined role of perceived authenticity and persuasive intensity in shaping trust, especially within settings filled with synthetic media and tailored algorithmic content. A further goal is to examine consumer skepticism as a moderating factor, determining whether greater awareness of AI-driven persuasion reduces or alters the way persuasive intensity affects trust levels. The value of this research lies in the comprehensive framework it proposes, which connects ethical considerations in AI use with established theories of consumer behavior in digital marketing. By weaving together cognitive elements like skepticism and perceptual aspects such as authenticity, the study offers a richer view of how trust develops in AI-supported communication. In this way, it helps close important gaps in earlier work, which has often examined technological or psychological factors separately rather than exploring how they interact with one another.

2. LITERATURE REVIEW

2.1. Consumer Trust in AI-Driven Brand Communication

Consumer trust in AI-driven brand communication refers to the extent to which individuals feel comfortable relying on, believing, and accepting brand messages created or shaped by artificial intelligence. This judgment rests on their views of the message's credibility, the brand's openness, and the ethical motives behind it (Choung et al., 2023; Luo et al., 2019). Unlike traditional trust, this version stretches into algorithmic settings, where automation handles much of the decision-making and content production, forcing consumers to assess not just the brand itself but also the technology working behind the scenes (Glikson & Woolley, 2020). Theoretically, trust serves as a vital bridge that lowers uncertainty and encourages positive consumer actions in online spaces, especially when people have limited information about what is really happening (Gefen et al., 2003). In the realm of AI-mediated messages, its importance grows because users rarely see the full process of how content is made or customized. As a result, trust acts as a crucial filter that helps people accept persuasive attempts from AI while easing worries about manipulation, data privacy, and genuineness (Choung et al., 2023). When trust is missing, even smart, highly tailored communications can fall flat and fail to deliver the desired results. Yet today's environment shows that consumer trust remains delicate and constantly shifting. Research suggests that although AI brings clear gains in efficiency and personalization, it also sparks fears of being misled—particularly as deepfakes and synthetic media make it harder to tell what is real and what is not (Vaccari & Chadwick, 2020). Many consumers now doubt the sincerity of brand messages and grow more skeptical when they sense overly strong or hidden persuasive efforts (Boerman et al., 2017). On top of that, unclear algorithms and weak ethical oversight can steadily weaken confidence, especially if people feel they are being steered without their knowledge or understanding (Glikson & Woolley, 2020).

In short, consumer trust in AI-driven brand communication is both essential and surprisingly fragile. Its development relies not only on how well the technology performs but also on honest transparency, a sense of authenticity, and how consumers mentally weigh the brand's persuasive aims—underscoring the clear need for broader models that bring these interconnected elements together.

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

2.2. Critical Analysis of Theoretical Frameworks

2.2.1. Signaling Theory

Signaling Theory offers a strong and fitting foundation for this study, as AI-driven brand communication is marked by deep information gaps: companies hold detailed knowledge about whether content is AI-generated, how it was created, the level of persuasive optimization applied, and the ethical standards guiding its use, while consumers usually remain in the dark (Connelly et al., 2011). In these situations, trust does not stem from firsthand verification but from the signals that brands send. As Connelly et al. note, signals become especially important when one side must judge hidden qualities based on visible indicators, which makes the theory highly relevant for exploring how ethical governance shapes AI-generated content. Within the proposed model, practices such as open disclosure, accountability measures, traceability, and well-defined ethical guidelines function as clear brand signals that convey honesty and responsible purpose. The theory gains further support from marketing studies showing that transparent communication boosts feelings of authenticity and leads to more positive consumer reactions, since openness itself serves as a powerful cue of sincerity (Yang & Battocchio, 2021). Signal strength also plays a key role; for instance, Kharouf et al. (2020) found that robust signals are particularly effective at rebuilding trust after it has been damaged—an insight that applies directly to AI contexts where suspicion of manipulation, deception, or concealed automation often arises. Thus, Signaling Theory goes beyond a simple lens for communication; it provides a solid explanation of how trust develops amid uncertainty, aligning closely with the study's focus on consumer trust as the main outcome. It also accounts for why ethical governance should directly support trust: when governance signals are visible, involve real effort, and remain consistent, consumers are more inclined to view the brand as both capable and principled. Similar findings in digital settings confirm that believable signals help lower uncertainty and guide trust-related decisions, such as repeat purchases, by relying on surface-level and message-level cues in situations where true practices are hard to observe (Pee et al., 2018). In today's world of deepfakes, this perspective feels especially timely, because consumers can no longer easily check the reality of branded content and must instead judge the communicator's legitimacy through signals like disclosure formats, policy statements, and overall tone. Overall, Signaling Theory supplies a robust framework for the link between ethical governance and consumer trust, while also explaining how perceptions of authenticity arise when transparency is seen as genuine honesty rather than mere show.

In this study, Signaling Theory posits that consumers cannot directly witness a brand's internal AI practices and must therefore depend on observable cues to form judgments about ethical standards, authenticity, and trustworthiness. It further assumes that governance signals only carry weight when consumers see them as believable, steady, and hard to imitate. As a result, ethical disclosure, accountability, and transparency are positioned as deliberate signals that help ease uncertainty and foster trust in AI-mediated brand communication.

2.2.2. Source Credibility Theory

Source Credibility Theory serves as a valuable second pillar that strengthens the overall framework, since the effectiveness of AI-driven brand communication hinges not only on the content itself but also on whether consumers view the communicator and the message as believable, trustworthy, and genuine (Hovland et al., 1953). In environments filled with synthetic media, people frequently encounter content that looks strikingly realistic yet raises questions about its true origin. As a result, trust evaluations are shaped first by credibility judgments before any persuasion or positive brand response can take hold. Recent research on AI-mediated interactions confirms that authenticity and source credibility continue to play a central role in how people respond, even when the source is virtual or partly generated by machines. For example, Kim & Wang (2024) found that both authenticity and source credibility act as key mediators of advertising success in settings involving AI influencers, showing that consumers go beyond mere novelty and actively assess whether synthetic voices feel sincere and worthy of belief. In a similar vein, Baudier et al. (2023) applied the source credibility model to compare robot and human influencers, highlighting that credibility assessments remain essential even in highly technological forms of communication. This perspective aligns closely with the study's focus on perceived authenticity of synthetic media, where authenticity goes far beyond surface-level realism; instead, it captures whether the content feels sincere, consistent, and true to the brand's expected character. Grigsby et al. (2025) build on this idea in the context of generative AI advertising, revealing that AI disclosures can weaken trust and attitudes toward ads, especially when the content features personal or emotional elements like faces or human-like imagery. Their results suggest that credibility becomes particularly vulnerable when AI steps into areas where consumers naturally anticipate human touch, genuine emotion, or real-life experience. This helps clarify why perceived authenticity should directly support consumer trust in the model: when synthetic media feels overly calculated, heavily optimized, or disconnected from human intent, source credibility suffers and trust in the brand message fades. Source Credibility Theory is therefore especially useful because it links message characteristics, source evaluation, and resulting trust within one clear explanatory path. In this study, it underscores that authenticity is not a side issue

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

but a core driver of brand trust in AI-supported environments. Moreover, it works well alongside Signaling Theory while signaling focuses on how brands actively send cues of ethical conduct, Source Credibility Theory explains how consumers interpret those cues and the media features to form judgments of believability. Combined, the two theories cover both the sending and receiving sides of trust development in the era of synthetic communication.

In this research, Source Credibility Theory posits that consumers assess AI-driven brand messages based on their perceptions of trustworthiness, believability, and authenticity, regardless of whether the source is fully synthetic or a mix of human and machine elements. It assumes that synthetic media does not persuade on its own; rather, its ability to influence depends on whether consumers regard the message source as credible. Consequently, stronger perceived authenticity of synthetic media is expected to enhance message credibility and, in turn, boost consumer trust in AI-mediated brand communication.

2.2.3. Persuasion Knowledge Model

The Persuasion Knowledge Model (PKM) stands out as the most relevant theory for illuminating the moderating influence of consumer skepticism towards AI persuasion and the relationship between persuasive intensity in AI-driven communication and trust (Friestad & Wright, 1994). At its heart, PKM views consumers as active participants who do not simply absorb persuasive messages but instead notice, interpret, and respond to them once they detect the attempt. This dynamic becomes particularly important in AI-powered settings, where advanced personalization, automated fine-tuning, and highly realistic deepfakes can amplify persuasive power while simultaneously raising fears of manipulation. Recent studies underscore the growing importance of persuasion knowledge in algorithmic advertising. Voorveld et al. (2024) highlight that people vary widely in how aware they are of algorithmic influence, how acceptable they find it, and how well they can cope with it, meaning responses to AI messages differ significantly across individuals. This variation directly supports the role of skepticism as a moderator, as it can be seen as a practical expression of activated persuasion knowledge in consumers. Beckert et al. (2021) add that disclosures often trigger multiple aspects of persuasion knowledge at once, sometimes leading to mixed outcomes: while openness may lessen feelings of being deceived, it can also sharpen awareness of underlying persuasive goals. In AI contexts, this tension matters because ethical practices alone do not always ease skepticism. Yang & Lee (2026) apply PKM to deepfake advertising and find that when consumers spot deepfake-style ads on their own, they react more negatively than when the artificial nature goes unnoticed or is only revealed later through disclosure. The issue goes beyond simple disbelief; it involves the activation of persuasion knowledge paired with a feeling that authenticity has been compromised. PKM therefore fits the framework especially well because it clarifies why stronger persuasive intensity does not always improve outcomes. In fact, heavily optimized AI persuasion can spark resistance, doubt, and lower trust, particularly among those who already feel wary about AI's role in persuasion. In the model, the moderator is not just an extra element but a vital boundary condition: as skepticism grows, consumers tend to see intense persuasion as manipulative rather than beneficial, which in turn weakens trust. PKM also enriches the framework by shifting attention from purely technological features to consumers' own thought processes. It helps explain why identical AI-generated messages can strengthen trust for some people while damaging it for others. This adaptability proves especially useful in today's deepfake-influenced branding landscape, where persuasion works through both quiet algorithmic customization and bold synthetic realism.

In this study, the Persuasion Knowledge Model assumes that consumers can at least partly recognize when AI-driven brand communication is meant to influence them, and that this awareness triggers coping strategies. It further posits that skepticism towards AI persuasion represents elevated persuasion knowledge, causing consumers to scrutinize intense persuasive efforts more carefully. As a result, the impact of persuasive intensity on trust is expected to vary depending on the degree to which consumers view the AI communication as manipulative or acceptable.

To sum up, the three most suitable theories for this project are Signaling Theory, Source Credibility Theory, and the Persuasion Knowledge Model (PKM). They work in harmony to address how brands convey ethical intentions amid information gaps, how consumers judge the authenticity and trustworthiness of synthetic media, and how they mentally respond when AI-driven persuasion stands out. These theories closely align with the study's key constructs, including ethical governance, perceived authenticity, persuasive intensity, consumer skepticism, and consumer trust in AI-driven brand communication, thereby providing a comprehensive theoretical foundation.

2.3. The Impact of Ethical Governance of AI-Generated Content

Ethical governance of AI-generated content refers to the combination of structural, procedural, and relational practices that brands adopt to ensure their AI-produced messages remain transparent, accountable, fair, and consistent with what audiences expect throughout the entire content creation process (Papagiannidis et al., 2025). In marketing, this goes well beyond simple technical rules to include thoughtful decisions around disclosure, explainability, privacy safeguards, human oversight, and genuine responsibility toward consumers who encounter these messages (Hermann, 2022; Balasubramaniam et al., 2023). At its core,

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

ethical governance is important because trust in AI communication stems more from a sense of moral legitimacy than from technical performance alone; consumers need to feel that the brand's use of AI is truly worthy of their confidence (Reinhardt, 2023). Psychologically, strong governance helps ease uncertainty and lowers fears of manipulation by offering visible signs that the brand is behaving honestly and reliably (Yang & Battocchio, 2021). From a broader social standpoint, it also demonstrates alignment with growing expectations for responsible digital practices, helping brands maintain credibility in an era increasingly defined by deepfakes, hidden algorithms, and widespread concerns about deception (Hermann, 2022; Papagiannidis et al., 2025). In essence, effective ethical governance shifts AI-driven communication from a mere persuasion tactic into a more socially responsible brand activity.

When ethical governance of AI-generated content is implemented more robustly, consumer trust in AI-driven brand communication is expected to increase, as audiences receive clearer indications of integrity, accountability, and honest intent. This reasoning aligns with existing work on brand transparency, which shows that openness tends to strengthen perceptions of authenticity and generate more favorable consumer reactions, with authenticity serving as an important signal for trust (Yang & Battocchio, 2021). In AI contexts specifically, disclosure plays a meaningful role. Baek et al. (2024) demonstrated that clearly labeling content as AI-generated can meaningfully shape how consumers evaluate advertising. Similarly, research on responsible AI highlights that solid governance practices form the groundwork for trust by making AI use more understandable, manageable, and socially acceptable (Papagiannidis et al., 2025). While some studies suggest that transparent and ethically grounded AI practices can strongly enhance consumer confidence by reducing moral uncertainty surrounding automated persuasion (Rodgers et al., 2022), others indicate a more balanced outcome. Although disclosure can bring greater ethical clarity, it may also draw attention to the artificial nature of the content and trigger caution instead of automatic trust (Baek et al., 2024). Philosophically, this perspective resonates with deontological and virtue ethics: deontology stresses the inherent duty to disclose and act fairly, while virtue ethics underscores the brand's moral character, such as honesty, responsibility, and accountability, as the true basis for trustworthiness in digital spaces (Martin, 2019; Hagendorff, 2020). Nevertheless, these views suggest that governance can lay the foundation for trustworthiness, yet actual consumer trust remains a subjective judgment that does not always emerge automatically even when ethical standards are met (Reinhardt, 2023). Therefore, this study proposes H1 as follows:

H1: Ethical governance of AI-generated content has a positive impact on consumer trust in AI-driven brand communication.

2.4. The Impact of Persuasive Intensity of AI-Driven Communication

Persuasive intensity of AI-driven communication refers to the extent to which AI-generated messages use sophisticated personalization, emotional targeting, algorithmic optimization, and adaptive strategies to shape consumer attitudes and behaviors with greater precision and impact. This goes beyond traditional persuasion by leveraging AI's ability to process real-time data, predict preferences, and automatically tailor content, thereby magnifying its persuasive potential (Kaplan & Haenlein, 2019). Philosophically, it resonates with postmodern and critical perspectives that view persuasion not as neutral but as deeply intertwined with power structures and technological systems that influence how people perceive and decide (Zuboff, 2023). Psychologically, the concept draws from the Elaboration Likelihood Model (ELM), which suggests that highly customized and emotionally engaging AI messages can bypass careful thinking and work through peripheral cues, making consumers more open to influence (Petty & Cacioppo, 1986). Sociologically, it represents a data-driven form of social influence in which algorithms quietly structure communication environments and guide behavior on a large scale (Susser et al., 2019).

Although persuasive intensity is expected to affect consumer trust in AI-driven brand communication, the relationship is far from straightforward. On the positive side, stronger persuasive intensity can build trust when consumers see the messages as relevant, helpful, and genuinely personalized. Research shows that AI-powered personalization often increases perceived value and emotional closeness, which in turn supports trust and engagement (Bleier & Eisenbeiss, 2015; Luo et al., 2019). Luo et al. (2019), for instance, found that AI communication can sometimes outperform human agents in creating satisfaction and trust when it feels efficient and well-tailored. This view aligns with utilitarian ideas that consumers value communication that saves effort and improves their decisions.

On the other hand, many studies warn that overly intense persuasion can damage trust. Drawing from psychological reactance theory, highly targeted or intrusive AI messages may provoke resistance, heighten skepticism, and create feelings of manipulation, ultimately lowering trust (Brehm & Brehm, 2013; Susser et al., 2019). Empirical findings confirm that when consumers sense algorithmic persuasion as too aggressive or self-serving, they become uncomfortable and lose confidence in the brand (Longoni et al., 2019). Critical perspectives on data ethics further argue that hyper-personalized tactics create imbalances of power and information, raising concerns about privacy, autonomy, and exploitation (Zuboff, 2019). Sundar et al. (2021) add that while

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

technological sophistication may initially boost credibility, awareness of heavy machine involvement can erode trust when the persuasion feels strategically engineered rather than sincere.

Because of these conflicting forces, the link between persuasive intensity and trust is not linear. Some evidence points to a positive effect when the approach feels transparent, relevant, and beneficial (Bleier & Eisenbeiss, 2015; Luo et al., 2019), while other work highlights negative or neutral outcomes when tactics cross into perceived manipulation or trigger defensiveness (Longoni et al., 2019; Susser et al., 2019). This tension highlights the delicate balance between efficiency gains and ethical worries about control and autonomy. The moderating role of consumer skepticism toward AI persuasion therefore becomes essential, as it helps determine whether intense persuasion is experienced as valuable customization or unwelcome pressure. When persuasive intensity is viewed as appropriate, relevant, and non-intrusive, it is more likely to support rather than undermine trust. Therefore, the following hypothesis is proposed:

H2: Persuasive intensity of AI-driven communication has a positive impact on consumer trust in AI-driven brand communication.

2.5. The Impact of Perceived Authenticity of Synthetic Media

Perceived authenticity of synthetic media refers to consumers' personal judgment that AI-generated or artificially created content feels truthful, believable, and in line with what they expect from reality, allowing them to accept it as genuine or legitimately representative even though they know it is not human-made. Perceived authenticity refers to consumers' judgments that AI-generated brand communication is genuine, credible, and consistent with the brand's established identity and values, rather than merely whether the content was originally produced by humans (Moulard et al., 2021). Philosophically, this view echoes existential and phenomenological thinking, where authenticity arises from how people experience coherence and lived reality instead of focusing solely on the actual source (Patry-Beaudoin et al., 2025). Sociologically, authenticity emerges through social interactions and performances, in keeping with dramaturgical ideas that audiences judge genuineness based on surrounding cues and context more than on hard facts (Sorensen & Krämer, 2024). Psychologically, it helps lower uncertainty and supports credibility assessments, making it a central factor that shapes trust in mediated communications (Bui et al., 2024).

When consumers perceive higher authenticity in synthetic media, they are more likely to place greater trust in AI-driven brand communication because the messages feel sincere, believable, and consistent with their expectations. Several studies back up this connection. Bui et al. (2024) showed that perceived authenticity of AI-generated content significantly boosts both trust and willingness to act, pointing to a clear positive link. Likewise, Kim & Wang (2024) found that authenticity acts as a mediator in AI influencer advertising, with consumers responding more positively when synthetic figures come across as genuine. These results fit with psychological theories of processing, which suggest that authentic-seeming stimuli reduce mental resistance and heighten feelings of sincerity, ultimately strengthening trust.

However, the research presents a more nuanced picture. Some work indicates that authenticity can have a powerful and direct effect on trust, particularly when synthetic media effectively mimics human qualities and stays true to the brand's character (Bui et al., 2024; Kim & Wang, 2024). At the same time, other studies suggest the effect can be weaker or even reversed in certain situations. For instance, Brüns & Meißner (2024) observed that revealing the use of generative AI sometimes lowers overall brand authenticity and leads to less favorable reactions. Similarly, Belanche et al. (2025) noted that AI-created images are frequently seen as less credible and more obviously artificial, which can undermine trust, especially in situations where consumers are highly involved. Schilke & Reimann (2025) further highlighted a potential downside, showing that being open about AI use can sometimes backfire by making people more aware of the artificial nature and raising suspicions of manipulation. These mixed findings reflect a broader sociological and critical lens, in which authenticity depends heavily on social expectations, cultural norms, and perceived legitimacy rather than surface-level realism alone.

Overall, although perceived authenticity of synthetic media is not guaranteed, the weight of evidence indicates that when it is successfully achieved, it plays a key role in building consumer trust. Therefore, the following hypothesis is proposed:

H3: Perceived authenticity of synthetic media has a positive impact on consumer trust in AI-driven brand communication.

2.6. The moderating role of Consumer Skepticism Toward AI Persuasion

Consumer skepticism toward AI persuasion refers to the degree to which individuals question, resist, or doubt the motives, credibility, and ethical standing behind AI-driven brand messages, particularly when they sense that algorithms are being used to shape their attitudes and behaviors in a highly calculated way. This concept extends general persuasion skepticism into the algorithmic age, where consumers increasingly view AI not as a neutral tool but as a deliberate agent of influence (Voorveld et al., 2024; Kim & Im, 2025).

Philosophically, it connects with critical data ethics and postmodern thought, which highlight how AI persuasion often operates within unequal power dynamics that quietly limit personal autonomy and shape choices (James, 2026; Zuboff, 2019).

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

Psychologically, it draws from the Persuasion Knowledge Model and algorithm aversion research, showing that once people detect persuasive intent, they tend to respond with mental resistance and reduced confidence in algorithmic sources (Zhang et al., 2025; Frank & Otterbring, 2024). Sociologically, this skepticism mirrors a wider erosion of trust in institutions and technology, fueled by worries over manipulation, lack of transparency, and unfairness in algorithmic systems (Park, 2025; Daschner & Obermaier, 2022). As a moderator, consumer skepticism toward AI persuasion plays a pivotal role in determining how persuasive intensity affects consumer trust in AI-driven brand communication. When skepticism remains low, consumers are more inclined to see strong AI persuasion as helpful, relevant, and efficient, which can enhance perceived value and ultimately support greater trust (Aoki & Matsui, 2025). In these situations, AI feels like a useful aid that improves decision-making. However, when skepticism is high, the same level of persuasive intensity can backfire. Consumers may interpret it as manipulative, invasive, or self-serving, leading to lower trust. Studies confirm that greater awareness of algorithmic tactics often heightens concerns about manipulation and weakens credibility judgments (Zhang et al., 2025). Research on algorithm aversion further shows that people frequently distrust AI suggestions, even when they are accurate, because they miss the warmth and judgment associated with human input (Daschner & Obermaier, 2022; Frank & Otterbring, 2024). More recent work adds that overly personalized or concealed influence strategies can damage trust by evoking feelings of unfairness, bias, or reduced personal control (Draws et al., 2021; Bozdag et al., 2025). This pattern aligns with reactance theory, where threats to freedom spark resistance, and with legitimacy theory, which stresses that trust depends on whether technological practices match societal standards of fairness and openness.

In short, consumer skepticism toward AI persuasion serves as an important boundary condition that decides whether persuasive intensity is experienced as valuable personalization or as hidden control. When skepticism is elevated, any potential positive effect of persuasive intensity on trust tends to weaken or disappear entirely. Therefore, the following hypothesis is proposed:

H4: Consumer skepticism toward AI persuasion negatively moderates the relationship between persuasive intensity of AI-driven communication and consumer trust in AI-driven brand communication.

Grounded in established theoretical foundations, this study advances the existing literature by introducing the following conceptual framework:

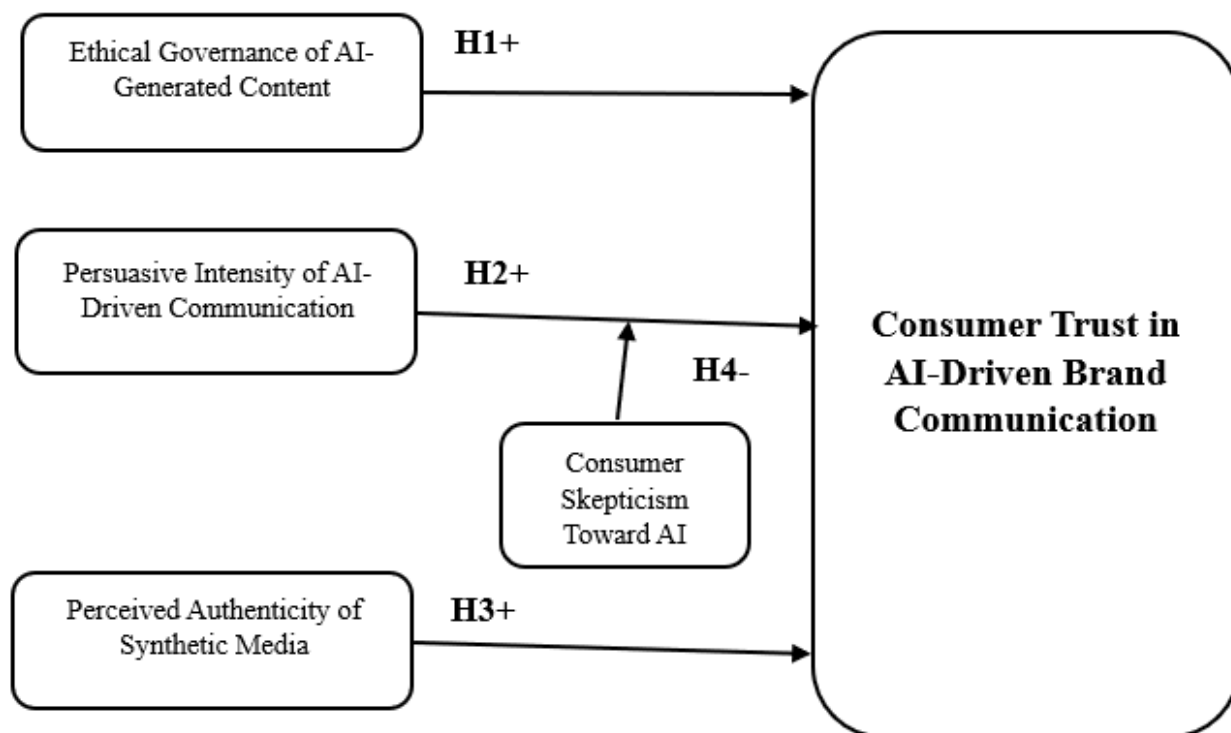


Figure 1: The Paper Conceptual Framework

3. METHODS

3.1. Sampling Strategy

This study employed a quantitative, cross-sectional approach to thoroughly investigate the connections between ethical governance of AI-generated content, the perceived authenticity of synthetic media, the persuasive intensity of AI-driven communication, and consumer trust in AI-supported brand messaging. It also explored how consumer skepticism toward AI persuasion might moderate these relationships. A quantitative design proved especially suitable here, as it facilitated accurate

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

hypothesis testing and allowed researchers to measure both direct effects and interaction patterns among the key variables using structured survey responses (Bryman, 2016; Hair et al., 2019). The cross-sectional nature of the research enabled efficient data gathering from a wide range of participants, offering a timely view of consumer attitudes in today's fast-changing AI-driven marketing environment.

To ensure the sample reflected varied consumer viewpoints, the authors applied a stratified sampling technique. Participants were deliberately drawn from different backgrounds, including everyday consumers, digital marketing experts, frequent technology users, and people with experience of AI-generated content like deepfakes and synthetic media. This mix helped improve the representativeness and generalizability of the results, mirroring the diverse ways people respond to AI-enabled brand communications.

Data collection involved a self-administered online questionnaire built around established measurement scales drawn from existing research on AI communication, consumer trust, and digital marketing. The constructs were carefully adapted from prior work on trust in AI systems (Choung et al., 2023; Luo et al., 2019), ethical governance and transparency (Yang & Battocchio, 2021; Papagiannidis et al., 2025), perceived authenticity (Bui et al., 2024; Kim & Wang, 2024), persuasive intensity (Luo et al., 2019; Susser et al., 2019), and skepticism toward AI persuasion (Voorveld et al., 2024; Frank & Otterbring, 2024). All items were rated on a five-point Likert scale (1 = strongly disagree to 5 = strongly agree), which supported a detailed assessment of participants' opinions (Brown, 2011). Before full deployment, the questionnaire underwent a pilot test with a small group to confirm its clarity, reliability, and suitability for the context, leading to only minor refinements. The survey was shared online through Google Forms and promoted across social media platforms, academic circles, and various digital technology and marketing communities. Following careful screening to eliminate incomplete or unreliable responses, the final dataset consisted of 385 valid cases. This sample size comfortably satisfied the requirements for robust multivariate analysis and provided sufficient statistical power (Hair et al., 2019).

3.2. Data Analysis

Data were analyzed with SPSS version 26 and Hayes' PROCESS macro (Hayes, 2022) in a clear, step-by-step manner. The process began with descriptive statistics to outline the profile of respondents and check the overall distribution of the data. Reliability was then assessed using Cronbach's alpha, and all constructs demonstrated strong internal consistency by surpassing the recommended 0.70 threshold (Hair et al., 2019). Next, exploratory factor analysis (EFA) was performed with principal component extraction and varimax rotation to confirm the underlying structure of the measures, keeping only those items with factor loadings greater than 0.50.

Afterward, multiple regression analysis tested the direct influence of ethical governance, perceived authenticity, and persuasive intensity on consumer trust in AI-driven brand communication, while accounting for relevant demographic and usage factors. The moderating role of consumer skepticism toward AI persuasion was investigated through moderated regression via the PROCESS macro, which enabled accurate evaluation of interaction terms and conditional effects (Hayes, 2022). Significance was evaluated at the standard $p < 0.05$ level, and any notable interactions were explored further using simple slope analysis to clearly show how the link between persuasive intensity and consumer trust changes at different levels of skepticism. Overall, this analytical approach supported thorough hypothesis testing, dependable estimates, and solid empirical grounding.

4. RESULTS

4.1. Descriptive Statistical Analysis

Table 1. Cross-Tabulation of Respondent Characteristics

Characteristic	Category	Percentage (%)
Age Group	Under 18	6%
	18–24	38%
	25–34	27%
	35–44	18%
	45 and above	11%
	Total	100%
Frequency of Encountering AI-Generated Content	Never	5%
	Rarely	12%
	Occasionally	28%
	Frequently	34%

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

	Very frequently	21%
	Total	100%
Level of Familiarity with AI Technologies	Not familiar at all	4%
	Slightly familiar	15%
	Moderately familiar	33%
	Very familiar	31%
	Extremely familiar	17%
	Total	100%
Current Occupation	Student	29%
	Marketing / Advertising / Communication professional	22%
	Technology / IT-related profession	20%
	Business / Management professional	19%
	Other (please specify)	10%
	Total	100%

Source: (The authors, 2026)

Table 1 outlines the demographic profile of the survey participants. A substantial portion of the respondents fell into the 18–24 age bracket (38%), with the 25–34 group coming next at 27%, which reflects a notable presence of younger adults in the sample. Smaller shares included those aged 35–44 (18%), 45 and older (11%), and under 18 (6%). When it came to exposure to AI-generated content, many participants indicated they came across it frequently (34%) or occasionally (28%), while another 21% reported very frequent encounters, pointing to how commonly such material appears in everyday online experiences.

On familiarity with AI tools, the biggest groups described themselves as moderately familiar (33%) or very familiar (31%), followed by 17% who felt extremely familiar, suggesting an overall tech-savvy group with solid knowledge levels; only a few reported minimal familiarity. In terms of occupation, students made up the largest segment (29%), ahead of professionals working in marketing, advertising, and communications (22%), technology or IT roles (20%), and business or management positions (19%), with the remaining 10% spread across other fields. This mix helped create a well-rounded and varied respondent base.

4.2. Reliability analysis

Table 2: Reliability analysis of “Consumer Trust in AI-Driven Brand Communication”.

Reliability Statistics				
Cronbach's Alpha		N of Items		
.813		4		
Item-Total Statistics				
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
CT1	10.700	6.040	.547	.755
CT2	10.770	5.762	.618	.719
CT3	10.840	5.769	.616	.720
CT4	10.650	5.882	.586	.735

Source: (The authors, 2026)

Where the survey indicators labeled CT1 to CT4 corresponded directly to items 1 through 4 in the questionnaire and were specifically developed to operationalize the latent construct of Consumer Trust in AI-Driven Brand Communication.

Table 2 shows that all indicators demonstrated satisfactory item discrimination, with corrected item–total correlation coefficients ranging from 0.547 to 0.618, exceeding the widely accepted threshold of 0.30. The overall scale exhibited good internal consistency, with a Cronbach’s alpha of approximately 0.785, surpassing the recommended minimum of 0.70 (Hair et al., 2009). Moreover, removing any single item reduced the overall alpha, with Cronbach’s alpha if item deleted ranging from 0.719 to 0.755.

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

Therefore, all four items, CT1–CT4, were retained for subsequent analysis. This pattern indicates that each item played a meaningful role in maintaining the scale’s coherence. In addition, the “alpha if item deleted” statistics stayed consistently below the full-scale alpha, further confirming that retaining all four indicators was the most appropriate decision. Taken together, these results offer robust evidence of the scale’s reliability, supporting the inclusion of all items in the subsequent analyses.

4.3. Exploratory factor analysis (EFA)

Table 3: Factor Loading Matrix.

Rotated Component Matrix ^a				
Component with loading factors				
1	2	3	4	5
CT1 .704	EG1 .710	PA1 .701	PI1 .610	CK1 .690
CT2 .762	EG2 .786	PA2 .726	PI2 .693	CK2 .738
CT3 .744	EG3 .763	PA3 .767	PI3 .716	CK3 .727
CT4 .720	EG4 .685	PA4 .731	PI4 .719	CK4 .708
Extraction Method: Principal Component Analysis.				
Rotation Method: Varimax with Kaiser Normalization.				
a. Rotation converged in 7 iterations.				

Source: (The authors, 2026)

Where the measurement items labeled CT1–CT4, EG1–EG4, PA1–PA4, PI1–PI4, and CK1–CK4 were systematically developed to operationalize five latent constructs, with each construct being measured through four corresponding indicators.

As shown in Table 3, the exploratory factor analysis (EFA) revealed a clean and distinct five-factor solution that closely matched the theoretical framework of the study. All twenty items achieved strong factor loadings well above the 0.50 threshold, ranging from 0.610 to 0.786 (Hair et al., 2009). This outcome confirmed solid convergent validity and demonstrated that each item adequately represented its intended construct.

More specifically, the CT items loaded heavily onto Component 1 (0.704–0.762), EG items onto Component 2 (0.685–0.786), PA items onto Component 3 (0.701–0.767), PI items onto Component 4 (0.610–0.719), and CK items onto Component 5 (0.690–0.738). The analysis employed Principal Component Analysis with Varimax rotation and Kaiser normalization, which produced clear factor separation and converged in just seven iterations, pointing to a stable and reliable structure. Since every item met the loading criteria and showed no problematic cross-loadings, none were dropped from the scale. Collectively, these findings offer robust empirical evidence supporting both the construct validity and the underlying dimensional structure of the measurement model used in this study.

4.4. Multiple linear regression model

Table 4. Table of Coefficients^a

Coefficients ^a					
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	0.875	0.218		4.014	.000
EG	0.495	0.052	0.538	9.519	.000
PI	0.618	0.049	0.672	12.612	.000
PA	0.463	0.050	0.501	9.260	.000
a. Dependent Variable: CT					

Source: (The authors, 2026)

Where CT was computed as the mean of CT1–CT4; EG as the mean of EG1–EG4; PI as the mean of PI1–PI4; and PA as the mean of PA1–PA4.

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

As shown in Table 4, the regression analysis revealed that all examined relationships were statistically significant, with Sig. values of .000, well below the conventional 0.05 threshold. This provides strong evidence that each independent variable, namely ethical governance (EG), persuasive intensity (PI), and perceived authenticity (PA), has a significant positive impact on consumer trust (CT). Furthermore, the standardized coefficients indicated that PI exerted the strongest influence, followed by EG and then PA, highlighting the relative importance of these predictors and underscoring the overall strength of the proposed model. In summary, these results offer clear empirical support for accepting all the hypothesized relationships in the study.

4.5. Moderator analysis

Table 5: Results analysis of “Consumer Skepticism Toward AI Persuasion”.

Model: 1Y: CTX: PIW: CKSample Size: 385

OUTCOME VARIABLE: CTModel Summary

R	R-sq	MSE	F	dl1	dl2	p
.701	.491	.412	122.508	3	381	.000

Model

Model	coeff	se	t	p	LLCI	ULCI
Constant	2.318	.145	15.986	.000	2.033	2.603
PI	.548	.052	10.538	.000	.446	.650
CK	.472	.049	9.633	.000	.376	.568
Int_1	.284	.041	6.927	.000	.204	.364

Source: (The authors, 2026)

Where CK was operationalized as the average composite score of items CK1 through CK4.

As shown in Table 5, the interaction term (Int_1) yielded a highly significant p-value of 0.000, well below the conventional 0.05 threshold. This finding offers strong statistical support for the moderating effect of consumer skepticism toward AI persuasion on the relationship between persuasive intensity (PI) and consumer trust (CT). Interestingly, the positive coefficient of the interaction effect ($\beta = 0.284$) suggests that higher levels of skepticism actually amplify the positive influence of persuasive intensity on trust. In other words, when consumers are more skeptical of AI-driven persuasion, the effect of persuasive intensity on building trust becomes even stronger.

Overall, these results provide robust empirical backing for accepting the proposed moderating hypothesis.

5. DISCUSSION

5.1. Result Summary

As shown in the regression results, ethical governance, persuasive intensity, and perceived authenticity all demonstrated strong and statistically significant positive effects on consumer trust in AI-driven brand communication, with standardized coefficients of 0.538, 0.672, and 0.501, respectively. Among the three predictors, persuasive intensity stood out as the most powerful driver, underscoring how well-crafted and personalized AI messaging plays a particularly vital role in building consumer trust. In addition, consumer skepticism toward AI persuasion was found to significantly moderate the link between persuasive intensity and consumer trust, with a positive interaction coefficient of 0.284. This suggests that greater skepticism actually strengthens the positive impact of persuasive intensity on trust levels. In short, these findings provide solid empirical support for all the proposed hypotheses and offer strong validation for the study's integrated conceptual framework, which brings together ethical governance, perceived authenticity, persuasion dynamics, and cognitive skepticism to explain how trust develops in AI-mediated brand communication.

5.2. Theoretical Implications

The present findings provide robust support for the idea that strong ethical governance of AI-generated content significantly boosts consumer trust. This outcome aligns closely with Signaling Theory, which views transparency and accountability as powerful signals that help reduce uncertainty and build confidence (Connelly et al., 2011; Yang & Battocchio, 2021). The study strongly concurs with scholars who argue that clear disclosure and responsible AI practices signal moral legitimacy and sincerity, thereby fostering greater trust among consumers (Papagiannidis et al., 2025; Reinhardt, 2023). At the same time, the results only partially endorse concerns raised in the literature that such disclosures might spark skepticism or lower confidence by drawing attention

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

to the artificial nature of the content (Baek et al., 2024). Although these worries hold theoretical merit, the evidence here suggests that well-communicated governance efforts ultimately outweigh any potential drawbacks. In this way, the study pushes back against overly pessimistic claims that transparency inevitably harms trust, instead highlighting ethical governance as a vital foundation for building credibility in AI-mediated brand communication.

The empirical results also reveal a strong positive link between persuasive intensity and consumer trust, lending support to earlier work that underscores the advantages of personalization and algorithmic optimization in increasing perceived relevance and engagement (Bleier & Eisenbeiss, 2015; Luo et al., 2019). This finding resonates with utilitarian views that see consumers as appreciating efficient and tailored messaging. Nevertheless, the study partially challenges critical perspectives rooted in psychological reactance theory, which warn that intense persuasion can trigger resistance and erode trust (Longoni et al., 2019; Susser et al., 2019). While these cautionary arguments remain valid in principle, the current data indicate that, under the right conditions, persuasive intensity is often perceived as helpful rather than manipulative. Overall, the research challenges the prevailing narrative of caution and proposes that the success of AI-driven persuasion depends not only on its intensity but also on its perceived relevance and acceptability to consumers, thereby offering fresh insights into ongoing theoretical discussions in the field.

The results demonstrate that the perceived authenticity of AI-generated content plays a powerful role in strengthening consumer trust. This finding lends strong support to Source Credibility Theory, which highlights how sincerity and believability shape people's trust judgments (Hovland et al., 1953; Kim & Wang, 2024). The study fully agrees with previous evidence showing that content that feels genuine, even when created by AI, can lower uncertainty and encourage favorable consumer reactions (Bui et al., 2024). At the same time, it only partially supports research suggesting that revealing AI involvement or artificial elements often weakens authenticity perceptions and harms trust (Brüns & Meißner, 2024; Belanche et al., 2025). Although these concerns about the fragility of authenticity are understandable, the present evidence indicates that when authenticity is successfully conveyed, it tends to outweigh any drawbacks linked to artificiality awareness. Overall, this research challenges the idea that synthetic media is automatically seen as less credible and instead reaffirms that perceived authenticity remains a key driver of trust in AI-mediated brand communication.

The findings also offer clear support for the moderating influence of consumer skepticism, showing that it actually strengthens rather than weakens the positive link between persuasive intensity and trust. This outcome stands in partial contrast to the Persuasion Knowledge Model, which argues that greater awareness of persuasive attempts usually triggers resistance and lowers trust (Friestad & Wright, 1994; Zhang et al., 2025). Instead, the results suggest that skeptical consumers may evaluate messages more carefully yet still respond positively when the persuasion feels transparent and relevant. This view resonates with newer perspectives that treat skepticism as a type of thoughtful engagement capable of improving how people assess messages (Voorveld et al., 2024). In this light, the study questions the long-held belief that skepticism is always a negative force and broadens theoretical understanding by showing that, under the right conditions, it can enhance the impact of AI-driven persuasion, providing a more balanced picture of consumer thinking in today's algorithmic communication landscape.

5.3. Practical Implications

The findings carry several important practical implications for brands operating in the fast-evolving landscape of AI-driven communication. Regarding ethical governance (H1), the strong positive influence on consumer trust indicates that companies should go well beyond basic regulatory compliance. Instead, they need to genuinely integrate transparency, accountability, and responsible AI practices into their everyday messaging strategies. Ethical disclosure can serve as a powerful signal of integrity when handled thoughtfully (Yang & Battocchio, 2021; Papagiannidis et al., 2025). That said, because earlier studies have warned that highlighting AI involvement may sometimes spark skepticism if it feels forced or insincere (Baek et al., 2024), managers must craft these messages with care. Transparency works best when woven naturally into the brand story rather than presented as a separate, technical notice. This approach helps reinforce trustworthiness while preserving the overall sense of authenticity.

On persuasive intensity (H2), the results highlight the strategic value of using AI for personalized and well-timed content that boosts relevance and emotional connection (Bleier & Eisenbeiss, 2015; Luo et al., 2019). Brands stand to gain by optimizing messages to feel timely and valuable to consumers. However, given the ongoing debate around persuasion, overly aggressive tactics risk triggering resistance and damaging trust (Longoni et al., 2019; Susser et al., 2019). Practitioners should therefore adopt a balanced, context-aware approach that avoids crossing into manipulation. Offering users greater control—such as options to adjust personalization levels—can help strike the right balance between impact and respect for consumer autonomy, ultimately supporting stronger and more sustainable trust relationships.

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

The positive effect of perceived authenticity (H3) underscores the need for companies to craft AI-generated content that connects with consumers on a more meaningful level. Rather than focusing solely on visual realism, brands should prioritize consistency with core brand values, emotional resonance, and alignment with audience expectations (Moulard et al., 2021; Kim & Wang, 2024). Although some research warns that disclosing AI involvement can weaken authenticity perceptions (Brüns & Meißner, 2024), firms can address this challenge by thoughtfully balancing openness with genuine storytelling. A useful strategy involves openly acknowledging the joint contribution of human creativity and AI capabilities in content creation. This approach helps maintain a sense of human touch and ethical responsibility while reducing any sense of artificial detachment. By doing so, companies can preserve perceived authenticity and support stronger consumer trust in synthetic media.

The moderating role of consumer skepticism (H4) presents a more nuanced lesson for practitioners. Instead of assuming that skepticism always weakens persuasion, the results show that higher skepticism can actually strengthen the impact of persuasive intensity when messages feel transparent and relevant. This suggests that skeptical consumers are not simply resistant; rather, they tend to evaluate content more thoughtfully (Voorveld et al., 2024). Therefore, brands should avoid trying to suppress skepticism and instead respond to it with honest, informative, and value-oriented communication. Clearly explaining AI usage, data handling, and the intent behind messages can turn skepticism into constructive engagement that ultimately builds greater trust. In practice, this calls for moving away from hidden persuasion techniques toward more open, conversational styles of communication that match today's growing demand for ethical and transparent AI practices.

5.4. Limitations

The study has several limitations that should be acknowledged. First, its cross-sectional survey design makes it difficult to establish clear causal links, since consumer views on trust and AI-driven communication can evolve over time. Second, the reliance on self-reported measures may introduce biases such as common method variance and social desirability (Podsakoff et al., 2003). Third, the sample composition, while diverse in some respects, may not fully capture broader cultural or demographic differences, which limits how widely the results can be applied. In addition, the research focused mainly on perceptual factors and did not include actual behavioral measures or experimental data that could offer a clearer picture of real consumer actions.

5.5. Future Research Directions

Future research would benefit from longitudinal or experimental approaches to better understand the dynamic and causal processes behind consumer trust in AI communication. Including behavioral indicators, such as engagement levels or purchase decisions, would make the findings more practically useful. Cross-cultural studies are also needed to explore how cultural values and attitudes towards technology shape trust formation in different contexts. Researchers could further strengthen the model by examining additional moderating factors like digital literacy, privacy concerns, or varying regulatory settings. Qualitative methods, such as in-depth interviews or focus groups, would provide richer insights into how consumers interpret AI ethics and authenticity. By addressing these areas, subsequent work can build a more complete and nuanced understanding of trust in AI-mediated brand communication.

5.6. Conclusion

The present study adds meaningfully to the expanding literature on AI-driven brand communication by bringing together ethical governance, persuasive intensity, perceived authenticity, and consumer skepticism into one integrated framework. The findings illustrate that consumer trust arises not merely from advanced technology itself, but also from careful attention to ethical practices and how people cognitively respond to persuasive attempts. Through its examination of both direct influences and moderating effects, this research delivers a more balanced and detailed picture of the processes that shape trust within AI-mediated settings. Overall, the results provide useful theoretical and managerial guidance for scholars and professionals aiming to address the challenges of AI-powered marketing. In the end, the study emphasizes that true progress in this area requires striking a careful balance between technological innovation and strong ethical responsibility, paving the way for more lasting and credible relationships with consumers.

REFERENCES

- 1) Aoki, T., & Matsui, A. (2025). Does algorithmic recommendation complement or substitute advertising and influencers? Consumer attitudes toward recommendation information and the formation of purchase intentions. *Computers in Human Behavior*, 172, 108735. <https://doi.org/10.1016/j.chb.2025.108735>
- 2) Baek, T. H., Kim, J., & Kim, J. H. (2026). Effect of disclosing AI-generated content on prosocial advertising evaluation. *International Journal of Advertising*, 45(1), 171-192. <https://doi.org/10.1080/02650487.2024.2401319>

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

- 3) Balasubramaniam, N., Kauppinen, M., Rannisto, A., Hiekkanen, K., & Kujala, S. (2023). Transparency and explainability of AI systems: From ethical guidelines to requirements. *Information and Software Technology*, 159, 107197. <https://doi.org/10.1016/j.infsof.2023.107197>
- 4) Baudier, P., de Boissieu, E., & Duchemin, M. H. (2023). Source credibility and emotions generated by robot and human influencers: the perception of luxury brand representatives. *Technological Forecasting and Social Change*, 187, 122255. <https://doi.org/10.1016/j.techfore.2022.122255>
- 5) Beckert, J., Koch, T., Viererbl, B., & Schulz-Knappe, C. (2021). The disclosure paradox: how persuasion knowledge mediates disclosure effects in sponsored media content. *International Journal of Advertising*, 40(7), 1160-1186. <https://doi.org/10.1080/02650487.2020.1859171>
- 6) Belanche, D., Ibáñez-Sánchez, S., Jordán, P., & Matas, S. (2025). Customer reactions to generative AI vs. real images in high-involvement and hedonic services. *International Journal of Information Management*, 85, 102954. <https://doi.org/10.1016/j.ijinfomgt.2025.102954>
- 7) Benk, M., Tolmeijer, S., von Wangenheim, F., & Ferrario, A. (2022). The value of measuring trust in AI-a socio-technical system perspective. *arXiv preprint arXiv:2204.13480*. <https://doi.org/10.48550/arXiv.2204.13480>
- 8) Bleier, A., & Eisenbeiss, M. (2015). Personalized online advertising effectiveness: The interplay of what, when, and where. *Marketing Science*, 34(5), 669-688. <https://doi.org/10.1287/mksc.2015.0930>
- 9) Boerman, S. C., Willemsen, L. M., & Van Der Aa, E. P. (2017). "This post is sponsored" effects of sponsorship disclosure on persuasion knowledge and electronic word of mouth in the context of Facebook. *Journal of Interactive Marketing*, 38(1), 82-92. <https://doi.org/10.1016/j.intmar.2016.12.002>
- 10) Bozdag, N. B., Mehri, S., Yang, X., Ha, H., Cheng, Z., Durmus, E., You, J., Ji, H., Tur, G., & Hakkani-Tür, D. (2025). Must Read: A Systematic Survey of Computational Persuasion. *arXiv preprint arXiv:2505.07775*. <https://doi.org/10.1145/3800687>
- 11) Brehm, S. S., & Brehm, J. W. (2013). *Psychological reactance: A theory of freedom and control*. Academic Press.
- 12) Brown, J. D. (2011). Likert items and scales of measurement. *Statistics*, 15(1), 10–14.
- 13) Brüns, J. D., & Meißner, M. (2024). Do you create your content yourself? Using generative artificial intelligence for social media content creation diminishes perceived brand authenticity. *Journal of Retailing and Consumer Services*, 79, 103790. <https://doi.org/10.1016/j.jretconser.2024.103790>
- 14) Bryman, A. (2016). *Social research methods*. Oxford University Press.
- 15) Bui, H. T., Filimonau, V., & Sezerel, H. (2024). AI-thenticity: Exploring the effect of perceived authenticity of AI-generated visual content on tourist patronage intentions. *Journal of Destination Marketing & Management*, 34, 100956. <https://doi.org/10.1016/j.jdmm.2024.100956>
- 16) Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human-Computer Interaction*, 39(9), 1727-1739. <https://doi.org/10.1080/10447318.2022.2050543>
- 17) Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, 37(1), 39-67. <https://doi.org/10.1177/0149206310388419>
- 18) Daschner, S., & Obermaier, R. (2022). Algorithm aversion? On the influence of advice accuracy on trust in algorithmic advice. *Journal of Decision Systems*, 31(sup1), 77-97. <https://doi.org/10.1080/12460125.2022.2070951>
- 19) Diel, A., Lalgı, T., Schröter, I. C., MacDorman, K. F., Teufel, M., & Bäuerle, A. (2024). Human performance in detecting deepfakes: A systematic review and meta-analysis of 56 papers. *Computers in Human Behavior Reports*, 16, 100538. <https://doi.org/10.1016/j.chbr.2024.100538>
- 20) Draws, T., Szlavik, Z., Timmermans, B., Tintarev, N., Varshney, K. R., & Hind, M. (2021). Disparate impact diminishes consumer trust even for advantaged users. In *International Conference on Persuasive Technology* (pp. 135-149). https://doi.org/10.1007/978-3-030-79460-6_11
- 21) Frank, D. A., & Otterbring, T. (2024). Autonomy, power and the special case of scarcity. *British Journal of Management*, 35(4), 1700-1723. <https://doi.org/10.1111/1467-8551.12780>
- 22) Friestad, M., & Wright, P. (1994). The persuasion knowledge model. *Journal of Consumer Research*, 21(1), 1-31. <https://doi.org/10.1086/209380>
- 23) Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping. *MIS Quarterly*, 27(1), 51-90. <https://doi.org/10.2307/30036519>
- 24) Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence. *Academy of Management Annals*, 14(2), 627-660. <https://doi.org/10.5465/annals.2018.0057>

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

- 25) Grigsby, J. L., Michelsen, M., & Zamudio, C. (2025). Service ads in the era of generative AI. *Journal of Retailing and Consumer Services*, 84, 104231. <https://doi.org/10.1016/j.jretconser.2025.104231>
- 26) Hagendorff, T. (2020). The ethics of AI ethics. <https://psycnet.apa.org/doi/10.1007/s11023-020-09526-7>
- 27) Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2009). *Multivariate data analysis* (7th ed.). Prentice Hall.
- 28) Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis* (3rd ed.). Guilford Press.
- 29) Hermann, E. (2022). Leveraging AI in marketing for social good. *Journal of Business Ethics*, 179(1), 43-61. <https://doi.org/10.1007/s10551-021-04843-y>
- 30) Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. Yale University Press.
- 31) James, S., & Thomas, B. (2026). Algorithmic Seduction. *Frontiers in Psychology*, 17, 1675412. <https://doi.org/10.3389/fpsyg.2026.1675412>
- 32) Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand. *Business Horizons*, 62(1), 15-25. <https://doi.org/10.1016/j.bushor.2018.08.004>
- 33) Kharouf, H., Lund, D. J., Krallman, A., & Pullig, C. (2020). A signaling theory approach. *European Journal of Marketing*, 54(9), 2139-2170. <https://doi.org/10.1108/EJM-10-2019-0751>
- 34) Kim, D., & Wang, Z. (2024). Social vs. virtual influencers. *Computers in Human Behavior: Artificial Humans*, 2(2), 100100. <https://doi.org/10.1016/j.chbah.2024.100100>
- 35) Kim, T., & Im, I. (2025). AI manipulation intention. *Information & Management*, 62(1), 104061. <https://doi.org/10.1016/j.im.2024.104061>
- 36) Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical AI. *Journal of Consumer Research*, 46(4), 629-650. <https://doi.org/10.1093/jcr/ucz013>
- 37) Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). AI chatbot disclosure. *Marketing Science*, 38(6), 937-947. <https://doi.org/10.1287/mksc.2019.1192>
- 38) Martin, K. (2019). Ethical Implications and Accountability of Algorithms. *Journal of Business Ethics*, 160(4), 835-850. <https://doi.org/10.1007/s10551-018-3921-3>
- 39) Moulard, J. G., Raggio, R. D., & Folse, J. A. G. (2021). Disentangling the meanings of brand authenticity. *Journal of the Academy of Marketing Science*, 49(1), 96-118. <https://doi.org/10.1007/s11747-020-00735-1>
- 40) Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance. *The Journal of Strategic Information Systems*, 34(2), 101885. <https://doi.org/10.1016/j.jsis.2024.101885>
- 41) Park, K., & Yoon, H. Y. (2024). Beyond the code. *Public Relations Review*, 50(5), 102507. <https://doi.org/10.1016/j.pubrev.2024.102507>
- 42) Park, K., & Yoon, H. Y. (2025). AI algorithm transparency. *Humanities and Social Sciences Communications*, 12(1), 1160. <https://doi.org/10.1057/s41599-025-05116-z>
- 43) Patry-Beaudoin, G., Handelman, J., & Chalmers Thomas, T. (2025). The Marketplace of Existential Authenticity. *Journal of Consumer Research*. <https://doi.org/10.1093/jcr/ucaf044>
- 44) Pee, L. G., Jiang, J., & Klein, G. (2018). Signaling effect of website usability. *International Journal of Information Management*, 39, 228-241. <https://doi.org/10.1016/j.ijinfomgt.2017.12.010>
- 45) Petty, R. E., & Cacioppo, J. T. (1986). Message elaboration versus peripheral cues. In *Communication and persuasion: Central and peripheral routes to attitude change* (pp. 141-172). https://doi.org/10.1007/978-1-4612-4964-1_6
- 46) Reinhardt, K. (2023). Trust and trustworthiness in AI ethics. *AI and Ethics*, 3(3), 735-744. <https://doi.org/10.1007/s43681-022-00200-5>
- 47) Ribeiro, A., Rivero, A. J. L., & Abrantes, J. L. (2025). The impact of artificial intelligence on consumer behavior. *Electronic Commerce Research*. <https://doi.org/10.1007/s10660-025-10063-7>
- 48) Rodgers, W., & Nguyen, T. (2022). Advertising benefits from ethical AI. *Journal of Business Ethics*, 178(4), 1043-1061. <https://doi.org/10.1007/s10551-022-05048-7>
- 49) Schilke, O., & Reimann, M. (2025). The transparency dilemma. *Organizational Behavior and Human Decision Processes*, 188, 104405. <https://doi.org/10.1016/j.obhdp.2025.104405>
- 50) Shrestha, N. (2020). Detecting multicollinearity. *American Journal of Applied Mathematics and Statistics*, 8(2), 39-42. <https://doi.org/10.12691/ajams-8-2-1>
- 51) Sorensen, L., & Krämer, B. (2024). The shift to authenticity. *Communication Theory*, 34(3), 118-129. <https://doi.org/10.1093/ct/qtae013>

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

- 52) Susser, D., Roessler, B., & Nissenbaum, H. (2019). Online manipulation. *Georgetown Law Technology Review*, 4, 1.
- 53) Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- 54) Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation. *Social Media + Society*, 6(1), 2056305120903408. <https://doi.org/10.1177/2056305120903408>
- 55) Voorveld, H. A., Meppelink, C. S., & Boerman, S. C. (2024). Consumers' persuasion knowledge of algorithms. *International Journal of Advertising*, 43(6), 960-986. <https://doi.org/10.1080/02650487.2023.2264045>
- 56) Whittaker, L., Mulcahy, R., Russell-Bennett, R., Letheren, K., & Kietzmann, J. (2025). Deepfake advertising and disclosure. *Journal of Advertising Research*. <https://doi.org/10.1080/00218499.2025.2498830>
- 57) Yang, J., & Battocchio, A. F. (2021). Transparent brand communication. *Journal of Product & Brand Management*, 30(8), 1176-1193. <https://doi.org/10.1108/JPBM-03-2020-2803>
- 58) Yang, J., & Lee, S. S. (2026). Deepfake UGC ads and consumer responses. *Journal of Retailing and Consumer Services*, 88, 104535. <https://doi.org/10.1016/j.jretconser.2025.104535>
- 59) Zhang, L., & Hur, C. (2025). Generative AI images and consumer attitudes. *Administrative Sciences*, 15(10), 395. <https://doi.org/10.3390/admsci15100395>
- 60) Zuboff, S. (2023). The age of surveillance capitalism. In *Social theory re-wired* (pp. 203-213). Routledge.

APPENDIX: Survey

Table 6. Survey Questionnaire

Background information

1. Which of the following age groups do you belong to?	Under 18	18–24	25–34	35–44	45 and above
2. How frequently do you encounter AI-generated or AI-assisted brand content (e.g., chatbots, personalized ads, deepfake videos)?	Never	Rarely	Occasionally	Frequently	Very frequently
3. How would you rate your level of familiarity with artificial intelligence technologies (e.g., generative AI, deepfakes, recommendation algorithms)?	Not familiar at all	Slightly familiar	Moderately familiar	Very familiar	Extremely familiar
4. Which of the following best describes your current occupation?	Student	Marketing / Advertising / Communication professional	Technology / IT-related profession	Consumers	Other (please specify): _____

No.	Variables	Coded Observed Variables	Content
1.	Consumer Trust in AI-Driven Brand Communication	CT1	I feel confident relying on brand messages generated or supported by AI technologies. (Choung et al., 2023; Luo et al., 2019)
		CT2	I perceive AI-driven brand communication as credible and trustworthy. (Gefen et al., 2003; Glikson & Woolley, 2020)
		CT3	I believe that AI-generated brand messages reflect honest and reliable intentions. (Choung et al., 2023)
		CT4	I am willing to accept and act upon brand messages created using AI systems. (Luo et al., 2019; Vaccari & Chadwick, 2020)

Ethics of Creative Communication in the Age of AI & Deepfakes: The Role of Synthetic Media and Persuasion in Shaping Brand Trust

2.	Ethical Governance of AI-Generated Content	EG1	I trust brands more when they clearly disclose the use of AI in their content. (Yang & Battocchio, 2021; Baek et al., 2024)
		EG2	Transparency in how AI-generated content is created increases my confidence in the brand. (Papagiannidis et al., 2025; Hermann, 2022) (Zhu et al., 2013)
		EG3	Brands that demonstrate accountability in their use of AI appear more trustworthy to me. (Reinhardt, 2023; Martin, 2019)
		EG4	Clear ethical guidelines in AI communication reduce my concerns about manipulation. (Balasubramaniam et al., 2023; Hagendorff, 2020)
3.	Persuasive Intensity of AI-Driven Communication	PI1	AI-driven brand messages that are highly personalized feel more relevant to me. (Luo et al., 2019; Bleier & Eisenbeiss, 2015)
		PI2	Emotionally targeted AI communication increases my engagement with brand messages. (Kaplan & Haenlein, 2019; Petty & Cacioppo, 1986)
		PI3	Highly optimized AI-generated messages make brand communication more effective. (Luo et al., 2019)
		PI4	Overly targeted AI communication can feel intrusive or manipulative. (Longoni et al., 2019; Susser et al., 2019)
4.	Perceived Authenticity of Synthetic Media	PA1	AI-generated content that feels realistic appears more authentic to me. (Moulard et al., 2021; Bui et al., 2024)
		PA2	I perceive synthetic media as authentic when it aligns with my expectations of the brand. (Kim & Wang, 2024)
		PA3	When AI-generated content feels sincere, I consider it more trustworthy. (Bui et al., 2024; Grigsby et al., 2025)
		PA4	The more natural and human-like AI content appears, the more authentic it feels to me. (Sorensen & Krämer, 2024; Patry-Beaudoin et al., 2025)
5	Consumer Skepticism Toward AI Persuasion	CK1	I am skeptical about the intentions behind AI-generated brand messages. (Voorveld et al., 2024; Kim & Im, 2025)
		CK2	I often question whether AI-driven communication is trying to manipulate my decisions. (Zhang et al., 2025; Friestad & Wright, 1994)
		CK3	I feel cautious when I recognize persuasive strategies in AI-generated content. (Frank & Otterbring, 2024; Beckert et al., 2021)
		CK4	The more I notice AI-based persuasion, the less I trust the message. (Yang & Lee, 2026; Daschner & Obermaier, 2022)

Survey link:

https://docs.google.com/forms/d/e/1FAIpQLSe8VALmld8hDgtl5OIA_2ijmVXd-sS-1cmbXr5ThYqJ0YMh9g/viewform?usp=header



There is an Open Access article, distributed under the term of the Creative Commons Attribution – Non Commercial 4.0 International (CC BY-NC 4.0) (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits remixing, adapting and building upon the work for non-commercial use, provided the original work is properly cited.