

Impact of Generative AI on Brand Authenticity and Customer Trust in Marketing Content Creation

¹Sudhanshu Shekhar Ojha, ²Vibhor Kumar, ³Arpit Verma

¹²³BBD University, Lucknow, Uttar Pradesh, India

Abstract—Background: Generative artificial intelligence (GenAI) is transforming marketing content creation, but its effect on brand authenticity and customer trust remains unresolved.

Methods: This paper synthesizes peer-reviewed studies, institutional reports, and consumer surveys published between 2024 and 2026, and proposes a conceptual model in which GenAI use affects customer trust through perceived brand authenticity, moderated by disclosure strategy and human oversight.

Results: The evidence shows substantial consumer skepticism toward visible AI-generated content: half of U.S. consumers prefer brands that avoid GenAI in consumer-facing content, and only a small share report that such content increases trust. Transparent disclosure combined with visible human oversight, however, appears to reduce reputational risk and preserve brand legitimacy.

Conclusion: GenAI should be governed as a strategic issue rather than managed purely as a production tool; brands that pair algorithmic efficiency with active human stewardship are better positioned to sustain consumer trust. The framework proposed here is intended for further empirical testing.

Index Terms—generative ai, brand authenticity, customer trust, marketing content, disclosure, ethical marketing

I. Introduction

Generative AI has moved rapidly from an experimental marketing tool to a core production capability. Organizations now rely on it to draft copy, generate campaign content, personalize messages at scale, translate communication across markets, and accelerate creative workflows. These efficiency gains are substantial, particularly for teams required to produce high-volume content across multiple channels. However, marketing is not solely a production function. It is also a trust-building process, and it is precisely in that dimension that GenAI introduces a new and largely unresolved challenge.

Consumers routinely judge brands by the tone, consistency, and sincerity of their communication. When a message feels templated, over-optimized, or impersonal, the brand risks losing its human character. That concern has become increasingly prominent as consumers grow more adept at recognizing AI-generated content. A survey reported by Gartner [1] found that 50% of U.S. consumers prefer brands that do not use

GenAI in consumer-facing content. Separately, eMarketer [2] reported that only 7% of consumers trust visible AI-generated marketing more than human-produced content, while 31% actively trust it less.

This paper addresses a straightforward but consequential question: can brands deploy GenAI in marketing without damaging authenticity and trust? The evidence reviewed here suggests that the answer depends not on the mere existence of AI-generated content, but on how it is used, disclosed, and governed.

1.1 Research Gap

Existing literature has examined AI disclosure, consumer trust, and attitudes toward AI-generated advertising, but these themes are frequently treated in isolation. What is absent is a unified model that explains how GenAI use affects consumer trust through the mediating mechanism of perceived authenticity, and under which conditions disclosure improves or worsens that relationship. This paper addresses that gap by placing brand authenticity at the center of the analytical model, treating disclosure and human oversight as boundary conditions rather than universal solutions.

1.2 Research Objectives

The study is guided by four primary objectives:

- 1) To examine the effect of GenAI-generated marketing content on perceived brand authenticity.
- 2) To assess how disclosure of AI use influences customer trust.
- 3) To determine whether human oversight moderates consumer response to AI-generated content.
- 4) To develop a conceptual framework for future primary empirical testing.

1.3 Research Questions and Hypotheses

RQ1: How does GenAI use influence perceived brand authenticity?

RQ2: How does disclosure of GenAI use influence customer trust?

RQ3: Does human oversight reduce negative consumer reactions to AI-generated content?

RQ4: Does perceived authenticity mediate the relationship between GenAI use and customer trust?

H1: GenAI-generated marketing content negatively affects perceived brand authenticity.

H2: The effect of transparent disclosure of GenAI use on customer trust is moderated by the level of human oversight.

H3: Perceived brand authenticity positively affects customer trust.

H4: Human oversight weakens the negative relationship between GenAI use and perceived brand authenticity.

H5: Perceived authenticity mediates the relationship between GenAI use and customer trust.

II. Literature Review

Recent studies and industry reports consistently show that GenAI adoption in marketing is accelerating. Boston Consulting Group [3] describes AI-powered marketing as a system in which algorithmic and human decision-making reinforce one another to improve output quality and return on investment. McKinsey & Company's [4] State of AI global survey reports that organizations are capturing significant revenue value from AI applications in marketing and sales. Adobe's [5] digital trends report similarly notes that brands face mounting pressure to adopt generative and agentic AI capabilities while simultaneously protecting customer experience and trust.

The consumer side of the story is considerably more cautious. Statista [6] found that only 26% of respondents in a global survey reported trusting brands to use AI responsibly. More recent evidence suggests that trust becomes even more fragile when AI use is visible in consumer-facing communication. This pattern indicates that adoption itself does not guarantee acceptance.

Academic scholarship on ethical GenAI use reinforces this interpretation. Puntoni et al. [7] argue that generative AI in marketing requires ethical design and deployment because transparency and accountability are growing in importance as synthetic content becomes more prevalent. A related argument appears in California Management Review's [8] analysis of authenticity in the AI era: when real and synthetic content

become increasingly difficult to distinguish, authenticity transforms from a vague creative ideal into a strategic market resource.

Empirical research adds additional nuance. A 2025 study on service advertising [9] found that disclosures of AI use could reduce trust and lower ad attitudes, particularly when the message affects intangible brand signals. However, the same research indicates that trust effects are not fixed. When AI is deployed for supporting functional elements while human representation remains central, negative consumer reactions may be attenuated. Consistent with this, earlier work on AI ad disclosure found that consumer responses vary considerably depending on category and framing [10].

Other recent consumer reports point in a consistent direction. Meltwater and YouGov's [11] perception report found that many respondents expect AI use to be disclosed and believe human review improves perceived authenticity. A 2026 Clutch survey [12] similarly found that consumers tend to view heavy AI reliance as inauthentic, even when the resulting content is polished and efficient. These findings collectively suggest that the market prefers AI to be human-led rather than human-replacing.

III. Methodology

This paper follows a mixed-method research design appropriate for conceptual journal submission. The first component is a structured review of recent peer-reviewed studies, institutional reports, and large-scale consumer surveys published between 2024 and 2026, undertaken to identify current evidence, recurring patterns, and unresolved tensions in the literature on GenAI use, authenticity, and consumer trust.

The second component is a proposed quantitative survey experiment for future primary data collection. Respondents would be exposed to three experimental content conditions: (1) human-written marketing content, (2) AI-generated content with disclosure, and (3) AI-generated content with disclosure accompanied by explicit indication of human review. Participants would evaluate each stimulus on perceived brand authenticity, customer trust, sincerity, perceived competence, purchase intention, and brand warmth. A stratified sample across age groups, levels of digital literacy, and product categories would

strengthen external validity, since trust reactions may differ substantially between high-emotion categories such as luxury goods and more functional categories such as consumer electronics.

The recommended analytical approach would include reliability testing, confirmatory factor analysis, and structural equation modeling to assess both direct and indirect relationships. Mediation effects can be estimated using bootstrapped confidence intervals, while moderation effects can be tested through interaction terms in the structural model. Ethical considerations for the proposed primary study include informed consent, voluntary participation, anonymized data handling, and institutional ethics review prior to data collection.

IV. Results and Discussion

4.1 Conceptual Model

The proposed framework positions GenAI use as the independent variable, perceived brand authenticity as the central mediating variable, and customer trust as the dependent outcome. Disclosure strategy and human oversight function as moderating conditions that influence the strength and direction of these relationships. If content appears fully automated, or if the brand obscures the role of AI in its production, consumers may interpret the communication as less sincere and less aligned with the brand's identity. Conversely, if AI use is openly disclosed and clearly supervised by human reviewers, the brand may preserve its legitimacy even while deploying automation.

This model explains why the same technology can produce markedly different consumer reactions. A technically polished AI-generated message may still fail if it lacks relational depth or human character, while a more modest message accompanied by visible human stewardship may feel considerably more trustworthy.

4.2 Discussion

The evidence reviewed in this paper indicates that GenAI creates a meaningful trade-off between operational efficiency and perceived legitimacy. On the one hand, it enhances speed, personalization capability, and production scale. On the other hand, it may erode authenticity when communication is

perceived as machine-driven rather than brand-driven. The most recent survey evidence suggests that this concern is widespread and not confined to highly skeptical or technologically sophisticated audiences.

Disclosure represents a particularly complicated variable. In some contexts, disclosure increases honesty perceptions by signaling openness about production methods. In other contexts, disclosure reduces trust because it draws consumer attention to the synthetic nature of the content [9]. This distinction implies that brands should not ask simply whether to disclose AI use; they should ask what specifically to disclose, how to communicate that disclosure, and how much human involvement remains visibly embedded in the process.

A second key insight is that authenticity is not reducible to novelty or originality. Content can be technically innovative while still feeling artificial, and conventional in form while still feeling authentic when it reflects care, consistency, and genuine accountability. Human review is not merely a final editing step; it is part of how trust is constructed and maintained in contemporary marketing.

4.3 Managerial Implications

First, brands should deploy GenAI as an augmentation tool rather than a wholesale replacement for human creativity; content that eliminates the human signal risks being perceived as both inauthentic and unaccountable. Second, disclosure should be approached strategically rather than formulaically, communicating not only that AI was used but also the specific role it played and the nature of human review that followed. Third, firms should monitor authenticity and trust as performance metrics alongside conventional indicators such as click-through rates and conversion, since strong engagement metrics do not necessarily indicate a healthy brand relationship.

V. Conclusion

Generative AI is fundamentally altering how brands produce marketing content, but it is not a neutral force from the standpoint of consumer trust. It enables gains in speed, scale, and personalization while simultaneously introducing the risk of perceived inauthenticity when communication feels more synthetic than human. The reviewed evidence consistently indicates that consumers respond more favorably

when AI is used transparently, with visible human oversight, and in ways that remain aligned with established brand identity.

The central conclusion of this paper is that GenAI should be governed as a strategic issue, not merely managed as a production capability. Brands that combine algorithmic efficiency with active human stewardship are better positioned to preserve the trust relationships that underpin long-term loyalty and equity.

5.1 Limitations and Future Research

This manuscript is conceptual and synthesis-based in design. It draws on secondary literature, industry reports, and consumer surveys rather than primary experimental data, and its conclusions should be treated as informed propositions rather than empirically verified findings. Future research should test the hypotheses proposed here using original survey or experimental data collected across diverse national contexts and product categories, disaggregate different forms of GenAI use (text, imagery, voice, video), pursue longitudinal designs to track how trust responses evolve as GenAI becomes normalized, and extend the framework through cross-cultural comparison beyond its current anchoring in North American and Western European survey data.

Acknowledgment

The authors did not receive any specific assistance beyond their institutional affiliation in the preparation of this manuscript.

Ethical Statement

The authors declare that they have no conflicts of interest. This manuscript is conceptual and synthesis-based, relying exclusively on previously published secondary sources; it does not involve primary data collection from human participants, and therefore no ethical approval was required. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- [1] Gartner, "Consumer survey on GenAI in consumer-facing content," Gartner Research, 2026.
- [2] eMarketer, "Shoppers aren't impressed by AI-generated marketing," Insider Intelligence, 2026.
- [3] Boston Consulting Group, "The blueprint for AI-powered marketing," BCG, 2024.
- [4] McKinsey & Company, "The state of AI: Global survey 2025," McKinsey Global Institute, 2025.
- [5] Adobe, "Adobe AI and digital trends 2026: GenAI and agentic AI insights," Adobe Inc., 2026.
- [6] Statista, "Global consumer trust in brands using AI," Statista GmbH, 2024.
- [7] S. Puntoni, R. W. Reczek, M. Giesler, and S. Botti, "Generative AI in marketing and principles for ethical design and deployment," *Journal of Marketing*, vol. 88, no. 3, pp. 1–21, 2024.
- [8] *California Management Review*, "Authenticity in the age of AI," University of California Press, 2025.
- [9] Service ads in the era of generative AI: Disclosures, trust, and intangibility," *Journal of Service Research*, vol. 28, no. 1, pp. 45–63, 2025.
- [10] Yahoo & Publicis Media, "Survey on AI ad disclosure and consumer trust," Yahoo Inc., 2024.
- [11] Meltwater & YouGov, "Consumer perception of generative AI," Meltwater Group, 2026.
- [12] Clutch, "Consumers expect AI to be human-led in 2026," Clutch Research, 2026.