

The Deepfake Dilemma: Can Brands Stay Authentic in the Synthetic Media Age?

Muhammad Imadudin Septiono^{1*}

¹ Sekolah Tinggi Ilmu Ekonomi Widya Wiwaha, Yogyakarta, Indonesia

Abstract

Article history:

Received: July 24, 2025

Revised: August 7, 2025

Accepted: September 29, 2025

Published: December 30, 2025

Keywords:

Brand Authenticity, Consumer Trust, Deepfake Advertising, Digital Advertising, Synthetic Media.

Identifier:

Nawala

Page: 64-79

<https://nawala.io/index.php/iraim>

Brands now operate in a context where deepfakes and other synthetic media make highly realistic fabrications easy to produce, raising new questions about whether authenticity can be sustained. As marketers experiment with deepfake campaigns, AI-generated content, and virtual influencers, they gain creative efficiency and personalization possibilities, but also face heightened risks to perceived sincerity, trust, and ethical acceptability. This article undertakes a systematic literature review of peer-reviewed studies published between 2019 and 2024 that examine deepfakes and synthetic brand communications. The included studies are coded by type of synthetic content, theoretical framing, methodological design, and brand-related outcomes, with a focus on authenticity, trust, and consumer response. The synthesis shows that synthetic media often acts as a double-edged stimulus: it can improve engagement and flexibility, yet frequently undermines authenticity and trust when disclosures are unclear or executions conflict with brand values. The article closes by proposing a research agenda centred on disclosure practices, governance mechanisms, and long-term brand equity in increasingly synthetic media environments.

*Corresponding author:
(Muhammad Imadudin Septiono)

©2025 The Author(s).

This is an open-access article under CC-BY-SA license (<https://creativecommons.org/licence/by-sa/4.0/>)



1. Introduction

Deepfake and other synthetic media technologies now make it possible to generate hyper-realistic but fabricated audio-visual content that is increasingly difficult to distinguish from authentic recordings (Westerlund, 2019). As these tools move from niche experimentation into mainstream platforms, they disrupt long-held assumptions about visual evidence, identity, and trust. Brands are not merely bystanders in this shift. Synthetic advertising that relies on deepfakes and generative adversarial networks allows marketers to create synthetic ads in which celebrity endorsements, spokespersons, or even ordinary consumers can be convincingly simulated, potentially transforming creative production, personalization, and media efficiency (Campbell et al., 2021).

From a branding perspective, this capacity is double-edged. On one hand, deepfake techniques promise hyper-personalized storytelling, scalable talent, and agile content iteration that may strengthen differentiation and relevance in crowded digital markets (Kietzmann et al., 2021). On the other hand, the same realism amplifies concerns about deception, consent, and the erosion of seeing is believing, challenging core pillars of brand authenticity and stakeholder trust (Kietzmann et al., 2020; de Ruiters, 2021). Recent conceptual work in advertising and communication argues that synthetic media requires rethinking how falsity, originality, and consumer response are theorized in manipulated advertising, yet empirical evidence in brand-specific contexts remains fragmented (Campbell et al., 2021).

Emerging studies illustrate that the deepfake dilemma for brands is not merely technical but psychological and ethical. Experimental research shows that AI-

generated or deepfaked ads can both stimulate curiosity and impair intentions, as consumers struggle with feelings of eeriness, manipulation, or loss of control (Sivathanu & Pillai, 2023; Gu et al., 2024). Work on deepfake information more broadly highlights that ethical concerns such as informed consent, privacy, and non-deception significantly shape social acceptance, and that entertainment value can weaken the influence of ethical judgments (Li & Wan, 2023). In parallel, research on AI-generated models and brand inclusivity suggests that using synthetic representations may backfire when audiences perceive them as tokenistic or inauthentic rather than genuinely value-driven (Sands et al., 2024). Together, these findings suggest that synthetic media can simultaneously enhance and undermine perceived brand authenticity, depending on how it is designed, disclosed, and governed (Sivathanu et al., 2023).

Despite growing attention, the marketing literature still lacks an integrated understanding of how deepfake and synthetic media practices intersect with brand authenticity as a strategic resource in the synthetic media age. Existing work is dispersed across technology reviews, ethical analyses, and context-specific advertising studies, with limited synthesis of how different brand uses of deepfakes affect trust, authenticity, and long-term brand equity. To address this gap, this article conducts a systematic literature review of peer-reviewed studies on deepfakes, synthetic advertising, and AI-generated brand communications published between 2019 and 2024. The review aims to map how brands are currently using or encountering deepfake technologies, identify the mechanisms through which these practices influence perceived authenticity and stakeholder trust, and derive a research

agenda for managing strategic brand authenticity when reality itself becomes malleable.

2. Literature Review

Early work on deepfake and synthetic media technologies has primarily focused on defining the phenomenon, mapping technical developments, and outlining broad social risks. Westerlund (2019) highlights how generative adversarial networks enable realistic but fabricated audio-visual content that undermines traditional cues of authenticity. Subsequent conceptual analyses argue that deepfakes destabilize core assumptions about evidence, identity, and responsibility in mediated communication, raising distinctive ethical concerns around deception, consent, and manipulation (Kietzmann et al., 2020; de Ruiter, 2021). Within advertising, Campbell et al. (2022) frame deepfakes and AI fakes as a structural shift in the advertising ecosystem, lowering production costs while amplifying concerns about misrepresentation, regulation, and consumer protection. In parallel, research on deepfake information more broadly shows that ethical concerns, norms, and perceived enjoyment jointly shape attitudes toward regulation and acceptance (Li & Wan, 2023).

Empirical marketing studies have begun to examine how consumers respond to AI-generated and deepfake advertising. Campbell et al. (2021) propose a framework suggesting that manipulated ads evoke complex appraisals of novelty, deception, and fairness, which in turn drive behavioral intentions. Experimental work on deepfake video ads indicates that such content can increase curiosity and

engagement but also trigger discomfort, reactance, and reduced behavioral intentions when viewers experience heightened eeriness or loss of control (Sivathanu & Pillai, 2023; Gu et al., 2024). Related research on AI-based deepfake ads in e-commerce contexts finds that perceived usefulness and entertainment can coexist with concerns over manipulation, with purchase intentions depending on how viewers balance these competing appraisals (Sivathanu et al., 2023). Together, these studies suggest that deepfake and AI-generated ads should be understood as ambivalent stimuli that can both enhance and erode persuasion outcomes, depending on design, disclosure, and audience characteristics.

A parallel stream of literature examines virtual and AI-generated influencers as a neighboring form of synthetic brand communication. Conti et al. (2022) describe virtual influencers as computer-generated imagery characters that simulate human influencers while offering brands high control, scalability, and cross-market adaptability. Choudhry et al. (2022) show that such influencers can generate real influence on followers, particularly when parasocial relationships and perceived humanness are strong, yet they also surface questions about authenticity and transparency. More recent work on AI-generated content in social media marketing indicates that generative tools can weaken perceived brand authenticity when audiences discover that brands rely on AI rather than human creators, especially in contexts where originality and personal expression are valued (Brüns & Meißner, 2024; Sands et al., 2024).

Across these streams, brand authenticity appears as a recurring but often secondary construct, treated as one outcome among many rather than as the central

lens for interpreting synthetic media practices. Existing studies typically isolate specific formats, such as deepfake video ads or virtual influencers, and examine short-term responses rather than long-term brand relationship dynamics (Campbell et al., 2021; Conti et al., 2022; Sivathanu & Pillai, 2023). There is still limited synthesis of how different types of synthetic media, governance choices, and disclosure strategies jointly shape perceived authenticity, trust, and brand equity over time. This fragmentation motivates a systematic literature review that integrates insights from deepfake advertising, virtual influencer research, and AI-generated content studies to clarify how brands can remain or become authentic in an environment where mediated reality itself is increasingly synthetic.

3. Methods

This study uses a systematic literature review approach to synthesize current knowledge on deepfakes, synthetic media, and brand authenticity. The review focuses on peer-reviewed journal articles published between 2019 and 2024 to capture the period in which deepfake technologies and AI-generated brand communications became salient in marketing and communication research. Searches were conducted in major databases, including Scopus, Web of Science, ScienceDirect, and Google Scholar, using combinations of keywords related to deepfakes, synthetic media, AI-generated advertising, virtual influencers, brand authenticity, trust, and consumer response. Articles were included if they were written in English, published in peer-reviewed journals, and addressed synthetic or AI-generated media in a branding, advertising, or consumer context. Studies that

focused only on technical detection, purely legal analysis without consumer or brand implications, or non-branded political deepfakes were excluded.

The screening process proceeded in two stages. First, titles and abstracts were reviewed to remove clearly irrelevant records and to ensure alignment with the core themes of deepfakes or synthetic media and their implications for brands and consumers. Second, full-text screening was used to confirm that each study engaged with brand-related outcomes such as authenticity, trust, credibility, or brand equity, or with closely related constructs such as perceived deception, eeriness, and ethical evaluation. The final set of articles was coded using a structured template that captured the type of synthetic media examined, theoretical lens, methodological approach, brand and consumer constructs, and key findings. Coding focused in particular on how authenticity was defined and measured, how deepfakes and synthetic content were framed as opportunities or risks, and what governance or disclosure strategies were proposed. The resulting synthesis informs the thematic analysis presented in the results and discussion sections.

4. Results and Discussion

The review shows that scholarship on deepfakes and synthetic media in branding is still emerging, yet it converges on a central tension between technological possibility and the preservation of authenticity. Conceptual and review work describes deepfakes as part of a broader wave of synthetic media that undermines traditional cues of reality and opens new avenues for manipulation (Westerlund, 2019; Kietzmann et al., 2020; Kietzmann et al., 2021). These studies argue that when

any image, voice, or video can be fabricated with high realism, brands can no longer rely on visual evidence or human presence as straightforward signals of genuineness. Ethical analysis reinforces this view, suggesting that deepfakes involve a distinct wrong because they appropriate identity and agency in ways that go beyond conventional misrepresentation, which raises serious normative questions for any brand that uses the technology (de Ruiter, 2021).

Work on brand authenticity provides an important theoretical anchor for interpreting these developments. Authenticity is consistently identified as a key driver of positive brand evaluations, perceived quality, and willingness to buy, particularly in categories where identity and values are salient (Cinelli & LeBoeuf, 2020). From this perspective, deepfakes and synthetic media are best understood not only as creative tools but also as potential threats to the perceived sincerity, continuity, and integrity of a brand's identity. The review indicates that many empirical studies treat authenticity as one outcome among several, yet across contexts it systematically mediates the relationship between AI-based communication and downstream attitudes or intentions. Synthetic media therefore does not simply add a layer of novelty; it can reconfigure how consumers interpret what is real about the brand.

Empirical work on AI-generated and deepfake-like advertising points to an ambivalent pattern in consumer responses. Studies on AI-generated ads find that such content can be perceived as intelligent, novel, and entertaining, but that it simultaneously evokes feelings of eeriness and manipulation that dampen acceptance (Gu et al., 2024). Research on deepfake video advertisements in tourism shows that

while these ads can draw attention and stimulate interest, they may lower booking intentions when viewers experience them as deceptive or uncanny (Sivathanu & Pillai, 2023). In e-commerce settings, AI-based deepfake advertising influences online shopping intention through a combination of perceived usefulness, entertainment, and concerns about manipulation, again highlighting mixed effects (Sivathanu et al., 2023). Evidence from social media marketing similarly indicates that when users discover that brands rely on generative AI for content creation, perceived brand authenticity can decline even if the content is aesthetically appealing (Brüns & Meißner, 2024). Taken together, these findings suggest that synthetic content often functions as a double-edged stimulus: it enhances efficiency and engagement while simultaneously eroding authenticity and trust if not carefully framed.

A related stream examines virtual and AI-generated influencers as a form of synthetic brand persona. Conti et al. (2022) describe virtual influencers as fully controlled, computer-generated figures that allow brands to scale presence and messaging across platforms. Choudhry et al. (2022) show that these virtual influencers can exert real influence on followers, shaping attitudes and intentions when parasocial relationships and perceived credibility are strong. However, their synthetic nature raises questions about transparency and authenticity, especially when audiences are uncertain about whether the influencer is human or AI-generated. In parallel, research on AI-generated models in brand communication indicates that attempts to use synthetic faces for inclusive representation can backfire if audiences interpret them as inauthentic or opportunistic rather than sincere efforts

at diversity (Sands et al., 2024). Bui et al. (2024) further suggest that perceived authenticity of AI-generated visuals can itself become a distinct evaluative dimension, with some consumers accepting AI-based imagery as long as the synthetic nature is clear and aligned with the context. These findings highlight that synthetic characters and images do not inevitably undermine authenticity; instead, perceptions depend on disclosure, coherence with brand identity, and the role that synthetic elements play in the broader narrative.

Governance and disclosure emerge as crucial, yet underdeveloped, levers in managing the deepfake dilemma. Studies on deepfake information show that ethical concerns and perceived enjoyment jointly shape support for regulation, implying that purely technical or legal solutions are unlikely to be sufficient without attention to user experience and norms (Li & Wan, 2023). In branding contexts, work on AI-disclosed brand voices finds that openly stating that a message is AI-generated can reduce perceived authenticity and brand attitude if the disclosure is not framed in a way that fits the brand's identity and category expectations (Kirkby et al., 2023). At the same time, disclosure and transparency are central in frameworks on deepfake advertising, which argue that brands must develop clear principles for consent, labelling, and permissible uses if they want to avoid accusations of deception (Campbell et al., 2021; Campbell et al., 2022). Collectively, these studies suggest that disclosure is necessary but not automatically sufficient: how and why a brand uses synthetic media, and how it explains this to audiences, matters as much as the mere fact of disclosure.

Viewed together, these streams contribute to a more nuanced picture of authenticity in the synthetic media age. The evidence indicates that authenticity is not simply destroyed by the presence of deepfakes or AI-generated content; rather, it is renegotiated. Brands that use synthetic media as a hidden substitute for genuine human expression, or that deploy deepfakes in ways that blur lines of identity and consent, are likely to face authenticity penalties and ethical backlash (Kietzmann et al., 2020; de Ruiter, 2021; Sands et al., 2024). In contrast, brands that position synthetic media as an explicit creative device, align it with their established values, and maintain a credible human core can sometimes integrate these tools without severe authenticity losses and, in some cases, generate new forms of perceived openness and innovation (Conti et al., 2022; Bui et al., 2024; Brüns & Meißner, 2024). This implies that authenticity in a synthetic context is less about avoiding technology and more about how convincingly brands reconcile synthetic techniques with honest, value-consistent self-presentation.

The review also highlights several gaps and directions for future research. Most studies focus on short-term responses to specific executions, while long-term effects on brand equity, relationship quality, and trust remain under-examined. Measurement of authenticity is often treated as a secondary variable, using relatively simple scales, rather than as a rich, multi-dimensional construct that can capture how deepfakes affect perceived continuity, integrity, originality, and symbolism over time (Cinelli & LeBoeuf, 2020). There is also limited comparative work across categories, cultures, and regulatory environments, even though ethical norms and expectations around synthetic media are likely to vary. Future research would benefit from

longitudinal and cross-cultural designs, as well as experimental work that systematically manipulates disclosure formats, consent cues, and degrees of synthetic manipulation. Integrating these designs with richer authenticity measures and brand-level outcomes would provide a stronger basis for answering the central question of this article: under what conditions can brands remain or even become more authentic in an age where media reality itself is increasingly synthetic.

In this section not only relate the findings found in the results, or report additional findings that have not been discussed. This section emphasizes more on the broader implications of research findings and relates them to previous research. Make sure that the conclusions you reach follow logically from and be strengthened by the evidence presented in your research.

5. Conclusion

This review concludes that deepfakes and synthetic media create a genuine authenticity dilemma for brands rather than a purely technical challenge. Across conceptual and empirical studies, synthetic content is shown to offer efficiency, creative flexibility, and sometimes stronger engagement, yet it also introduces substantial risks to perceived sincerity, trust, and ethical standing. Authenticity consistently appears as a central mechanism that links AI-based communication to consumer attitudes and intentions, which means that synthetic media cannot be treated as a neutral upgrade in production, but as a strategic choice with reputational consequences.

At the same time, the evidence indicates that authenticity is not automatically destroyed by the presence of synthetic content. Brand outcomes depend heavily on how deepfakes and AI-generated media are framed, disclosed, and integrated with existing brand meanings. Brands that treat synthetic elements as hidden substitutes for human expression, or that blur boundaries around consent and identity, are more likely to face authenticity penalties. In contrast, brands that are transparent, values consistent, and deliberate about their use of synthetic media may be able to maintain, or even reconfigure, authenticity in ways that audiences accept.

This review also highlights important shortcomings and limitations in the underlying studies that reduce the strength and generalizability of current conclusions. Most research relies on short-term experiments, narrow categories, or single-country samples, which makes it difficult to know whether observed effects persist over time, hold in other cultural or regulatory environments, or scale to full brand portfolios. Measures of authenticity and trust are often simplified, and many studies treat governance choices such as disclosure and consent as fixed rather than systematically manipulated. These limits should prompt readers to question how far existing findings can be extended and in what ways unmeasured factors might have influenced the reported results.

Future research should therefore develop more rigorous and context-sensitive examinations of how brands can remain authentic in the synthetic media age. Longitudinal designs, cross-cultural comparisons, and studies that integrate richer authenticity scales with brand-level outcomes would help clarify long-term effects

on equity and relationships. Experimental work that systematically varies disclosure formats, degrees of manipulation, and alignment with brand values would also strengthen causal claims. By directly addressing these gaps and limitations, subsequent studies can provide more valid evidence and more actionable guidance for brands navigating the deepfake dilemma.

References

- 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.
- 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.
- Campbell, C. L., Plangger, K., Sands, S., & Kietzmann, J. (2021). Preparing for an era of deepfakes and AI-generated ads: A framework for understanding responses to manipulated advertising. *Journal of Advertising*, 50(1), 22-38.
- Campbell, C., Plangger, K., Sands, S., Kietzmann, J., & Bates, K. (2022). How deepfakes and artificial intelligence could reshape the advertising industry: The coming reality of AI fakes and their potential impact on consumer behavior. *Journal of Advertising Research*, 62(3), 241-251.

- Choudhry, A., Han, J., Xu, X., & Huang, Y. (2022). “I felt a little crazy following a ‘doll’”: Investigating real influence of virtual influencers on their followers. *Proceedings of the ACM on Human-Computer Interaction*, 6(GROUP), 1-28.
- Cinelli, M. D., & LeBoeuf, R. A. (2020). Keeping it real: How perceived brand authenticity affects product perceptions. *Journal of Consumer Psychology*, 30(1), 40-59.
- Conti, M., Gathani, J., & Tricomi, P. P. (2022). Virtual influencers in online social media. *IEEE Communications Magazine*, 60(8), 86-91.
- de Ruiter, A. (2021). The distinct wrong of deepfakes. *Philosophy & Technology*, 34(4), 1311-1332.
- Gu, C., Jia, S., Lai, J., Chen, R., & Chang, X. (2024). Exploring consumer acceptance of AI-generated advertisements: From the perspectives of perceived eeriness and perceived intelligence. *Journal of Theoretical and Applied Electronic Commerce Research*, 19(3), 2218-2238.
- Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons*, 63(2), 135-146.
- Kietzmann, J., Mills, A. J., & Plangger, K. (2021). Deepfakes: Perspectives on the future reality of advertising and branding. *International Journal of Advertising*, 40(3), 473-485.
- Kirkby, A., Baumgarth, C., & Henseler, J. (2023). To disclose or not disclose, is no longer the question: Effect of AI disclosed brand voice on brand authenticity and attitude. *Journal of Product & Brand Management*, 32(8), 1108-1122.

- Li, M., & Wan, Y. (2023). Norms or fun? The influence of ethical concerns and perceived enjoyment on the regulation of deepfake information. *Internet Research*, 33(5), 1750-1773.
- Sands, S., Demsar, V., Ferraro, C., Campbell, C., & Cohen, J. (2024). Inauthentic inclusion: Exploring how intention to use AI-generated diverse models can backfire. *Psychology & Marketing*, 41(6), 1396-1413.
- Sivathanu, B., & Pillai, R. (2023). The effect of deepfake video advertisements on the hotel booking intention of tourists. *Journal of Hospitality and Tourism Insights*, 6(5), 1669-1687.
- Sivathanu, B., Pillai, R., & Metri, B. (2023). Customers online shopping intention by watching AI-based deepfake advertisements. *International Journal of Retail & Distribution Management*, 51(1), 124-145.
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 39-52.