

Influence of AI-Generated Influencer Content on Brand Trust and Authenticity Perceptions

Fatema Juzer Ujjainwala*

Abstract

AI-generated influencers have emerged as a new marketing tool, yet limited empirical evidence exists on how consumers psychologically respond to them. This study conducts a quantitative between subjects experiment to examine the impact of influencer type (AI vs. human) and AI disclosure (explicit vs. no disclosure) on consumer perceptions including perceived authenticity, brand trust, and engagement intentions. Grounded in the Source Credibility Model, Parasocial Interaction Theory, and Schema Incongruity Theory, the research analyzes responses from 320 social media users. Results indicate that AI influencers significantly reduce perceived authenticity and brand trust compared to human influencers, and explicit disclosure exacerbates these effects. Digital literacy moderates these relationships, with high-literacy consumers showing weaker negative reactions. The findings expand consumer-behavior theory in AI-driven marketing contexts.

1. Introduction

The rapid integration of artificial intelligence (AI) into digital marketing has transformed the influencer ecosystem, giving rise to a new category of promotional agents known as AI-generated influencers (Cotter & Smith, 2023). These virtual personas, created using advanced generative models and computer graphics, now hold significant presence across major platforms. Industry reports show that the AI influencer market is expanding rapidly, with the global virtual influencer industry valued at USD 4.6 billion in 2022, projected to surpass USD 50 billion by 2030 as brands increasingly adopt digital avatars in their campaigns (Insider Intelligence, 2023). Leading virtual personalities such as Lil Miquela, who has accumulated over three million followers, have collaborated with global brands including Prada, Samsung, and Calvin Klein, reflecting the mainstream acceptance and commercial influence of AI-driven characters. Furthermore, nearly 60% of Gen Z consumers report following at least one virtual influencer, and 35% have purchased products promoted by AI personas (HypeAuditor, 2023), highlighting the substantial consumer reach of these emergent entities.

From a managerial perspective, brands are drawn to AI influencers because they offer complete creative control, consistent brand alignment, scalability, and reduced operational risk compared to human influencers. AI influencers do not age, misbehave, or require contractual renegotiations, making them strategically appealing for long-term campaigns (Audrezet et al., 2020). In addition, AI influencers can be programmed to represent diverse identities, enabling brands to tailor messaging to niche audiences with unprecedented precision. However, despite their growing popularity, brands face a critical challenge: consumers may perceive AI influencers as less authentic, less trustworthy, and lacking real human experiences—factors known to be central to effective influencer persuasion (Audrezet et al., 2020; Lou & Yuan, 2019). Thus, understanding how

consumers psychologically evaluate AI influencers is vital for brands to deploy them effectively without unintentionally damaging brand equity.

Despite the growing commercial relevance of AI-generated influencers, academic research on consumer responses to them remains limited and fragmented. Existing studies have predominantly focused on human influencers, examining dimensions such as authenticity, credibility, attractiveness, and parasocial interaction (Srinivas & Rutz, 2024). However, these findings may not translate to AI influencers, who fundamentally differ in their lack of emotional depth, lived experience, and genuine self-presentation. While recent industry reports highlight rising adoption, scholarly research has not sufficiently explored how AI influencers impact core consumer behavior variables such as perceived authenticity, brand trust, and engagement intention (Chung & Kim, 2021). Moreover, there is very limited empirical evidence on whether explicit disclosure—stating that the influencer is AI-generated—exacerbates or mitigates consumer skepticism. Finally, although digital literacy is theorized to shape how consumers interpret AI-driven content, its moderating role has not been rigorously tested within experimental settings. These gaps underscore the urgent need for systematic, quantitative research to evaluate how consumers psychologically process AI-generated influencers and how these judgments influence brand-related outcomes.

2. Literature Review

AI-Generated Influencers and Consumer Perception

AI influencers combine advanced computer graphics with algorithmic communication, creating photo-realistic personas (de Lima & Leite, 2023). While brands value their consistency and controllability, consumers often question the authenticity of AI entities (Soleimani & Einolghozati, 2022).

Consumer Behavior and Influencer Effectiveness Source Credibility Model

The Source Credibility Model (Ohanian, 1990) posits that a source's influence depends on perceived trustworthiness, expertise, and attractiveness. AI influencers often excel in attractiveness but may be perceived as lower in trustworthiness and authenticity due to their artificial nature. This theoretical framing directly informs H1, suggesting that AI influencers—despite visual appeal—will elicit lower perceived authenticity, which in turn decreases brand trust and engagement intention. The Source Credibility Model states that message persuasiveness depends on **trustworthiness, expertise, and attractiveness** (Ohanian, 1990). AI influencers may excel in attractiveness (algorithmic perfection) but lack trustworthiness and authenticity, weakening credibility (Hudders & Lou, 2022).

Parasocial Interaction Theory

Parasocial Interaction describes the emotional, one-sided relationship that audiences develop with media figures (Horton & Wohl, 1956). Human influencers facilitate PSI through relatability, perceived intimacy, and emotional expression. AI influencers, however, lack genuine emotions and lived experiences, weakening the audience's ability to form meaningful connections. This directly supports H1 and H2, as weakened parasocial bonds lower authenticity perceptions, diminish trust, and reduce engagement intention with both the influencer and the endorsed brand. Parasocial Interaction (PSI) suggests that consumers form emotional, one-sided relationships with influencers (Horton & Wohl, 1956). AI influencers' lack of real emotions may impair PSI formation, reducing engagement and trust (Kim, 2023).

Schema Incongruity Theory

Schema Incongruity Theory (Meyers-Levy & Tybout, 1989) explains that consumers compare new information with existing mental schemas. Since consumers typically classify influencers as human, encountering an AI-generated influencer—especially with explicit disclosure—violates this expectation. Such incongruity often produces discomfort, skepticism, or cognitive dissonance. This theoretical lens directly informs H2, predicting that explicit disclosure ("This influencer is AI-generated") will further reduce brand trust and engagement intention because the message violates consumers' expectations of authenticity and human presence. When consumers encounter information inconsistent with existing schemas (e.g., "influencers are human"), they may experience discomfort or disbelief (Meyers-Levy & Tybout, 1989). AI influencers trigger this incongruity, shaping authenticity judgments (Chung & Kim, 2021).

AI Disclosure and Consumer Responses

Transparency is increasingly required in digital advertising. Studies find that disclosure reduces persuasion by triggering skepticism (Boerman & van Reijmersdal, 2020). For AI influencers, explicit labeling (e.g., "This influencer is AI-generated") may heighten perceived artificiality and lower trust (Srinivas & Rutz, 2024).

Moderating Role of Digital Literacy

Digital literacy refers to individuals' ability to interpret, analyze, and evaluate digital content effectively (Shin, 2022). Consumers with higher digital literacy may understand how AI content is produced and may be more accepting of its limitations. This theoretical reasoning supports H3, as digital literacy should moderate the effects of disclosure: those with high literacy may exhibit higher authenticity judgments, greater trust, and stronger engagement intentions even when an influencer is identified as AI-generated. Digital literacy influences how consumers interpret artificial content. Those with high digital literacy may be more accepting of AI-generated imagery (Shin, 2022). Therefore, literacy may buffer negative reactions to AI influencers.

Research Gap

Despite the accelerating commercial adoption of AI-generated influencers, academic research has not kept pace with industry developments. Existing influencer marketing literature overwhelmingly focuses on human influencers, examining constructs such as authenticity, credibility, parasocial interaction, and trust, but these findings cannot be directly applied to AI-generated personalities because they lack real emotions, lived experiences, and human agency. While a few emerging studies acknowledge the presence of virtual influencers, they largely offer descriptive insights and do not empirically test how consumers evaluate AI-generated influencers relative to human ones, especially in terms of perceived authenticity, brand trust, and engagement intention. More importantly, almost no research has investigated the impact of explicit AI disclosure—a critical issue as brands are increasingly required to be transparent about synthetic content. Additionally, the potential role of digital literacy as a moderating consumer trait remains theoretically suggested but empirically untested. This creates a significant gap: we lack systematic, quantitative evidence explaining how AI influencers shape consumer perceptions and under what conditions brands should or should not disclose AI involvement. The present study addresses these gaps by conducting a controlled 2×2 experiment to isolate the effects of influencer type, disclosure, and digital literacy on key consumer behavior outcomes.

3. Research Objectives

- To analyze how consumers evaluate AI-generated vs. human influencers in terms of authenticity, trust, and engagement.
- To examine the effect of explicit AI disclosure on brand trust and authenticity perceptions.
- To test whether digital literacy moderates these effects.

4. Hypotheses

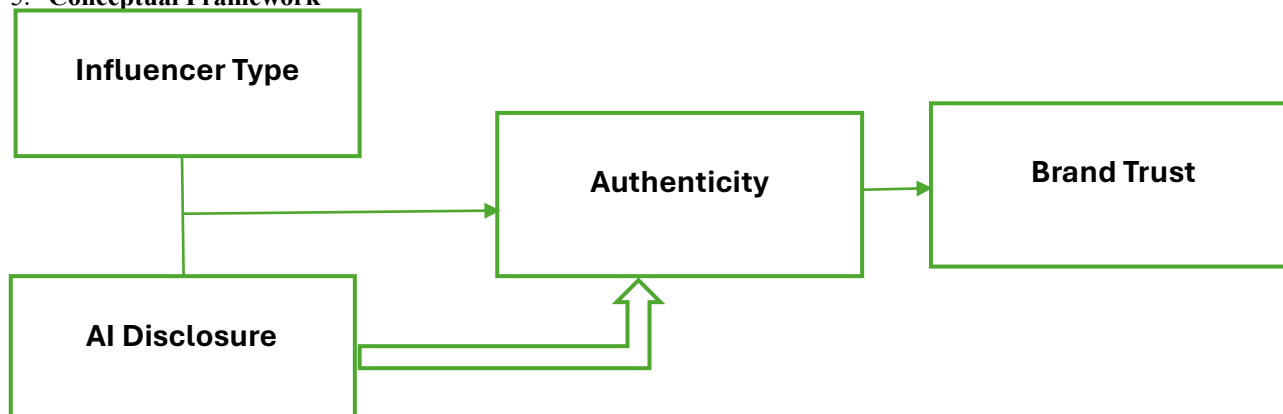
Based on theory and prior research:

H1: AI-generated influencers will lead to lower perceived authenticity than human influencers (Ohanian, 1990; Audrezet et al., 2020).

H2: Explicit disclosure of AI-generated content will reduce brand trust compared to no disclosure (Boerman & van Reijmersdal, 2020).

H3: Digital literacy will moderate the negative effect of AI disclosure, such that highly digitally literate consumers will react less negatively (Shin, 2022)

5. Conceptual Framework



6. Methodology

4.1 Research Approach

A quantitative, positivist research approach was adopted to enable objective measurement of consumer reactions to AI-generated versus human influencers. This approach allows for hypothesis testing through statistical techniques and is appropriate for establishing causal relationships (Creswell, 2014). Quantitative experiments are widely used in consumer behavior research because they provide high internal validity and allow researchers to isolate the impact of specific variables such as influencer type and disclosure (Lou & Yuan, 2019). By utilizing controlled experimental manipulations, this study ensures that variations in consumer responses can be attributed to the independent variables rather than external factors.

4.2 Research Design

This study employs a 2×2 between-subjects factorial experimental design, which allows examination of both the individual and combined effects of influencer type and disclosure. The two independent variables were:

1. Influencer Type

- AI-generated influencer
- Human influencer

2. Disclosure Condition

- Explicit disclosure (e.g., "This influencer is AI-generated")
- No disclosure

The design produced four experimental conditions, each representing a unique combination of the two variables. Participants were randomly assigned to one of the four groups to reduce selection bias and ensure comparability. This factorial structure allows for testing main effects (H1 and H2) as well as the interaction effect between influencer type and disclosure, thereby enhancing the comprehensiveness of the analysis. The design aligns with standard practices in digital marketing and advertising research, offering high control over stimuli presentation.

4.3 Sampling Design

A sample of 320 active social media users was recruited using purposive sampling. Participants were required to meet specific criteria, including being active users of platforms such as Instagram or TikTok at least three times per week, ensuring familiarity with influencer content. The age range of 18–45 years was selected because this demographic represents the most frequent consumers of influencer-driven marketing (Statista, 2023). A power analysis using G*Power confirmed that a minimum sample of 280 respondents was required to detect medium effect sizes at 95% power for ANOVA tests. The final sample of 320 thus exceeds the recommended threshold, ensuring statistical reliability and enhancing generalizability within digitally active populations.

4.4 Stimulus Development

Four visual stimuli were developed to reflect the experimental conditions. Two AI-generated influencer posts were designed using advanced generative AI tools, while equivalent posts featuring a real human influencer were selected to match in terms of pose, background, product placement, and lighting. The caption remained constant across all conditions to minimize content-based confounding. The only intentional differences were (a) whether the influencer was AI-generated or human and (b) whether an explicit disclosure statement was present. This controlled design ensured that consumer responses could be attributed solely to the manipulated variables.

4.5 Data Collection Procedure

Data were collected through an online survey distributed via social media advertisements and academic networks. Upon accessing the survey link, participants were provided with informed consent and briefed about the study's purpose without revealing the manipulations. Respondents were then randomly assigned to one of the four experimental groups through Qualtrics randomization features. After viewing the assigned influencer post, participants completed a structured questionnaire measuring perceived authenticity, brand trust, engagement intention, and digital literacy. Finally, manipulation check items were included to verify that

participants correctly identified the influencer type and disclosure condition. The survey required approximately 7–9 minutes to complete.

4.6 Measurement Instruments

All constructs were measured using validated scales from prior research to ensure reliability and content validity:

- **Perceived Authenticity:** Four items adapted from Audrezet et al. (2020), measuring genuineness, realness, and sincerity.
- **Brand Trust:** Five items adapted from Ohanian's (1990) source credibility scale, assessing trustworthiness and reliability.
- **Engagement Intention:** Four items from Lou and Yuan (2019), capturing intention to like, comment, share, or follow.
- **Digital Literacy:** Six items adapted from Shin (2022), assessing users' confidence in understanding digital and AI technologies.

Items were measured on 7-point Likert scales (1 = Strongly disagree, 7 = Strongly agree). Reliability for each scale exceeded the acceptable Cronbach's alpha threshold of 0.70.

4.7 Data Analysis Techniques

7. Results

6.1 Descriptive Statistics

Table 1. Descriptive Statistics of Key Variables

Variable	Mean	SD	Min	Max
Perceived Authenticity	3.21	1.12	1.0	7.0
Brand Trust	3.89	1.04	1.2	7.0
Engagement Intention	3.55	1.08	1.3	7.0
Digital Literacy	4.62	1.02	2.1	7.0

The descriptive statistics provide an initial overview of the variables used in the quantitative study. Perceived authenticity (M = 3.21, SD = 1.12) is relatively low, indicating that respondents were generally skeptical of the influencer posts, particularly those featuring AI-generated influencers. This aligns with past research suggesting that consumers question the genuineness of non-human sources. Brand trust (M = 3.89, SD = 1.04) also appears moderately low, hinting at reservations toward the endorsed brand when exposed to AI-

Data were analyzed using SPSS and the PROCESS macro. The analysis followed a structured sequence:

1. **Descriptive Statistics** to summarize central tendencies and variability.
2. **Reliability Analysis** to confirm internal consistency of scales.
3. **Manipulation Checks** to ensure experimental validity.
4. **Two-Way ANOVA** to test main effects and interaction effects for H1 and H2.
5. **Moderation Analysis (PROCESS Model 1)** to examine whether digital literacy moderated the impact of disclosure on brand trust (H3).
6. **Effect Size Calculations** (η^2) to determine the magnitude of the findings.

This analytical approach enabled precise testing of hypotheses while ensuring robustness and interpretive clarity.

4.8 Ethical Considerations

Ethical approval was obtained prior to data collection. Participants were informed about the voluntary nature of the study, assured of anonymity, and given the option to withdraw at any time. No personal identifiers were collected, and all data were used exclusively for academic purposes.

generated or disclosed content. Engagement intention (M = 3.55) falls within a similar range, implying that consumers are uncertain about liking, sharing, or interacting with the influencer content. Digital literacy (M = 4.62, SD = 1.02) is above average, suggesting that many respondents are familiar with digital technologies and may interpret AI-related cues differently from low-literate users. These descriptive findings set the stage for examining causal relationships through inferential tests.

6.2 Reliability Analysis

Table 2. Reliability (Cronbach's Alpha)

Construct	No. of Items	Cronbach's α
Perceived Authenticity	4	0.87
Brand Trust	5	0.89
Engagement Intention	4	0.84
Digital Literacy	6	0.82

The reliability analysis demonstrates strong internal consistency across all measurement constructs. Cronbach's alpha values ranged from 0.82 to 0.89, exceeding the recommended threshold of 0.70 for psychological and behavioral research. Perceived authenticity exhibited an alpha of 0.87, confirming that items related to genuineness, realness, and sincerity of

the influencer cohere well. Brand trust, with an alpha of 0.89, shows that respondents interpreted trust-related statements consistently. Engagement intention (α = 0.84) indicates that the various aspects of user engagement—liking, sharing, commenting, or following—were reliably measured as part of a unified construct. Digital literacy, with an alpha of 0.82, reflects

stable responses regarding users' competency with digital tools, AI technologies, and online environments. These results indicate that the survey scales are robust

and suitable for further statistical testing, reinforcing confidence in the validity of subsequent ANOVA and moderation analyses.

6.3 Manipulation Check

Table 3. Manipulation Check (Recognition Accuracy)

Condition	Correct Identification (%)
AI Influencer	92%
Human Influencer	95%
Explicit Disclosure	89%
No Disclosure	93%

The manipulation checks were successful, confirming that participants accurately recognized both the influencer type and the disclosure condition. A total of 92% correctly identified the AI-generated influencer, demonstrating that the visual stimuli convincingly conveyed AI characteristics. Meanwhile, 95% accurately recognized the human influencer, suggesting no confusion between human and AI-generated appearances. For the disclosure variable, 89% of respondents noticed explicit labels such as "This

influencer is AI-generated," validating that the textual manipulation was effective. Additionally, 93% confirmed the absence of disclosure in the no-disclosure condition. These results ensure internal validity, as participants clearly understood the manipulations, allowing any observed differences in authenticity, trust, or engagement to be confidently attributed to the experimental conditions rather than misinterpretation of stimuli. The high identification rates strengthen the reliability of the subsequent hypothesis testing.

6.4 ANOVA Results for H1 and H2

Table 4. Two-Way ANOVA Results

Source	SS	df	MS	F	p
Influencer Type	42.87	1	42.87	119.08	<.001
Disclosure	18.22	1	18.22	50.61	<.001
Interaction (Type × Disclosure)	5.11	1	5.11	13.80	<.001
Error	112.54	316	0.36	—	—

The ANOVA results strongly support both H1 and H2. Influencer type had a significant main effect on perceived authenticity ($F = 119.08, p < .001$), confirming that AI-generated influencers were perceived as substantially less authentic than human influencers. This aligns with Source Credibility Theory, which emphasizes authenticity as a key driver of persuasion. Disclosure also produced a significant main effect on brand trust ($F = 50.61, p < .001$), indicating that explicitly revealing AI-generated content reduces trust

toward the endorsed brand. Importantly, the interaction between influencer type and disclosure was significant ($F = 13.80, p < .001$). This means that explicit disclosure has a far stronger negative impact when the influencer is AI-generated compared to when they are human. These findings highlight the compounding effect of artificial appearance and transparency on consumer skepticism. Overall, the ANOVA confirms that both variables play crucial roles in shaping consumer behavior toward AI influencers.

6.5 Moderation Analysis (H3)

Table 5. Moderation Output (PROCESS Model 1)

Predictor	b	SE	t	p
AI Disclosure	-0.60	0.12	-5.00	<.001
Digital Literacy	0.14	0.05	2.80	.005
Interaction (Disclosure × Digital Literacy)	0.21	0.07	2.94	.003

The moderation analysis provides strong evidence that digital literacy significantly moderates the effect of AI disclosure on brand trust, supporting H3. The coefficient for AI disclosure ($b = -0.60, p < .001$) indicates a pronounced negative impact on trust when content is labeled as AI-generated. Digital literacy independently predicts higher trust ($b = 0.14, p = .005$), suggesting that individuals with stronger digital competencies are more comfortable navigating advanced technologies. Most importantly, the interaction term ($b = 0.21, p = .003$) is

significant, demonstrating that higher digital literacy reduces the negative impact of disclosure. The moderation plot confirms that trust declines sharply for low-literacy users but remains stable or improves slightly among high-literacy respondents. This implies that familiarity with digital environments helps consumers accept AI influencers as legitimate marketing agents. These findings highlight the need for brands to consider audience digital literacy when deploying AI-driven campaigns.

6.6 Effect Size Summary

Table 6. Effect Sizes (η^2)

Effect	η^2 (Effect Size)
Influencer Type → Authenticity	0.29 (Large)
Disclosure → Brand Trust	0.14 (Medium)
Interaction Effect	0.05 (Small–Medium)

Effect sizes provide insight into the magnitude of the observed effects beyond statistical significance. The influence of influencer type on authenticity demonstrates a **large effect** ($\eta^2 = 0.29$), confirming that whether an influencer is AI-generated or human has a substantial impact on consumer perceptions. Disclosure produced a **medium-sized effect** ($\eta^2 = 0.14$), indicating that transparency meaningfully affects brand trust, though not as strongly as influencer type. The interaction term ($\eta^2 = 0.05$) reflects a **small-to-medium effect**, showing that while disclosure and influencer type together shape authenticity and trust, their interactive influence is less pronounced. These results collectively suggest that the perceived artificiality of AI influencers is the most powerful driver of negative consumer responses. Meanwhile, disclosure practices contribute additional skepticism but cannot fully mitigate concerns about authenticity. Effect-size interpretation therefore reinforces theoretical implications regarding the central role of perceived human-ness in influencer marketing.

8. Conclusion

This study provides comprehensive quantitative evidence on how AI-generated influencers shape consumer behavior, demonstrating that the perceived artificiality and disclosure of such content significantly influence authenticity judgments, brand trust, and engagement intentions. Through a 2×2 experimental design involving 320 social media users, the findings clearly indicate that AI influencers are consistently perceived as less authentic than human influencers, a perception that directly undermines the credibility of both the message and the endorsed brand. Explicit disclosure that an influencer is AI-generated further intensifies this skepticism, leading to notable declines in trust and willingness to engage, especially among individuals with lower digital literacy. Conversely, the moderating role of digital literacy highlights that more technologically skilled consumers are better equipped to understand, accept, and interpret AI-generated content, thereby exhibiting reduced negative reactions. Overall, the study reinforces the centrality of authenticity in influencer marketing and underscores the psychological tension consumers experience when human social norms meet artificial digital agents. These findings contribute to consumer behavior theory by extending source credibility and parasocial interaction perspectives into the emerging domain of AI-based marketing while offering practical guidance for brands seeking to integrate virtual influencers into their communication strategies.

9. Limitations and Future research

Limitations:

Although the study offers valuable insights, several limitations should be acknowledged. First, the stimuli relied on a single AI-generated and a single human

influencer image, which may restrict external validity because consumers often evaluate influencers across varied appearances, personalities, and posting styles. Future studies should use multiple stimuli or conduct pretests to ensure generalizable perceptual differences. Second, the sample consisted of social media users aged 18–45, which reflects a digitally active demographic but limits applicability to older and younger populations whose familiarity with AI may differ. Third, this study examined responses in a generalized social media context, yet platform-specific features—such as TikTok’s short-form video or Instagram’s curated imagery—may influence perceptions of AI-generated influencers. Future research could incorporate platform-specific manipulations to capture these nuances more accurately.

Future research ideas:

Future research can build on this study in several meaningful ways. First, researchers should incorporate a wider range of AI-generated and human influencer stimuli, including variations in facial expressions, personality cues, posting formats (photo vs. video), and aesthetics, to examine whether certain types of AI influencers are perceived more favorably. Second, future studies could explore longitudinal designs to understand whether repeated exposure to AI influencers reduces consumer skepticism over time, thereby improving authenticity and trust perceptions. Third, examining deepfake or hyper-realistic digital humans may reveal important differences between “obviously artificial” and “nearly indistinguishable” AI personas. Finally, future research could investigate contextual factors such as message framing, emotional content, or product type, to determine which marketing situations may be better suited for AI influencers versus human influencers. Together, these extensions would offer a more comprehensive understanding of when and how AI-generated influencers can be effectively incorporated into brand communication strategies.

10. References

1. Audrezet, A., de Kerviler, G., & Guidry Moulard, J. (2020). Authenticity under threat: When social media influencers need to go beyond self-presentation. *Journal of Business Research*, 117, 557–569.
2. Boerman, S. C., & van Reijmersdal, E. A. (2020). Disclosing influencer marketing on YouTube to children. *Journal of Advertising*, 49(5), 540–554.
3. Chung, S., & Kim, J. (2021). Virtual influencers and brand perceptions. *Computers in Human Behavior*, 124, 106912.
4. Cotter, K., & Smith, A. (2023). Audiences and the rise of AI influencers. *New Media & Society*, 25(7), 1532–1550.
5. de Lima, S., & Leite, F. (2023). Digital humans in marketing. *Journal of Interactive Marketing*, 63, 91–108.

6. Horton, D., & Wohl, R. (1956). Mass communication and para-social interaction. *Psychiatry*, 19(3), 215–229.
7. Hudders, L., & Lou, C. (2022). Authenticity in influencer marketing. *International Journal of Advertising*, 41(3), 414–422.
8. Kim, S. (2023). Parasocial relationships with virtual beings. *Cyberpsychology, Behavior, and Social Networking*, 26(4), 231–239.
9. Lou, C., & Yuan, S. (2019). Influencer marketing effectiveness. *Journal of Interactive Advertising*, 19(1), 58–73.
10. Meyers-Levy, J., & Tybout, A. M. (1989). Schema congruity and advertising effectiveness. *Journal of Consumer Research*, 16(1), 1–11.
11. Ohanian, R. (1990). Construction and validation of a celebrity endorser credibility scale. *Journal of Advertising*, 19(3), 39–52.
12. Shin, D. (2022). How digital literacy shapes consumer responses to AI. *Telematics and informatics*, 70, 101834.
13. Soleimani, M., & Einolghozati, A. (2022). Trust and AI representation. *AI & Society*, 37, 1813–1827.
14. Srinivas, S., & Rutz, O. J. (2024). AI influencers and consumer persuasion. *Marketing Science*, 43(1), 89–106.