
The Impact of Perceived Authenticity of AI-Generated Marketing Content on Brand Trust and Purchase Intention: The Mediating Role of Perceived Brand Transparency

Reza Zeinali

Independent Researcher, Tehran, Iran, Marketing and Consumer Behavior Track

Abstract

Generative artificial intelligence (GenAI) is rapidly becoming part of the marketing communication infrastructure, yet its ability to create efficient and visually persuasive content has generated a parallel concern: consumers may question whether machine-assisted messages are genuine, accountable, and worthy of trust. This conceptual study examines how the perceived authenticity of AI-generated marketing content may shape brand trust and purchase intention, with perceived brand transparency proposed as a central mediating mechanism. Drawing on Signaling Theory and the Stimulus–Organism–Response (S–O–R) framework, the paper integrates recent evidence on AI authorship, disclosure, brand authenticity, transparency, algorithm aversion/appreciation, and consumer trust. The proposed model argues that authentic-seeming AI content can function as a credibility cue that reduces ambiguity surrounding algorithmic authorship. Consumers who interpret AI-generated content as coherent with a brand’s identity, honest in origin, and non-manipulative are more likely to infer that the brand communicates openly and voluntarily. Such perceived transparency should strengthen brand trust and, in turn, purchase intention. The framework also allows for direct effects of authenticity on trust and purchase intention, creating a serial mediation pathway from perceived authenticity to transparency, trust, and behavioral intention. The study contributes by separating content authenticity from brand transparency, explaining why transparency is not equivalent to mere AI disclosure, and reconciling conflicting findings showing that disclosure can either strengthen or weaken consumer responses. Managerially, the model suggests that firms should combine AI efficiency with human oversight, recognizable brand voice, clear but proportionate disclosure, verifiable claims, and accessible explanations of AI’s role. The article concludes with a detailed empirical validation agenda using experimental and structural-equation approaches.

Keywords: generative artificial intelligence; AI-generated marketing content; perceived authenticity; brand transparency; brand trust; purchase intention; signaling theory; S–O–R framework



. Introduction

Generative artificial intelligence has shifted from an experimental technology to a practical content-production infrastructure for brands. Marketing teams can now generate copy, images, video, product descriptions, customer-service scripts, and localized campaign variants at a speed and scale that were previously difficult to achieve. This transition extends the broader transformation described by Davenport et al. (2020) and Puntoni et al. (2021), who anticipated that AI would reshape both marketing strategy and consumer experience. However, GenAI differs from many earlier marketing technologies because it can become an apparent author of the message itself. The consumer is therefore not only evaluating a product claim or an advertisement; the consumer may also be evaluating the legitimacy of a non-human or human-AI source.

This authorship ambiguity creates a strategic paradox. On the one hand, AI can improve relevance, personalization, creative variation, and operating efficiency. On the other hand, the same technology can activate concerns about deception, synthetic imagery, hidden automation, emotional insincerity, and the absence of human intention. Recent research illustrates this tension. Brüns and Meißner (2024) found that GenAI use in social-media content can diminish perceived brand authenticity, particularly when automation appears to replace human creation. Kirk and Givi (2025) similarly showed that consumers can respond less favorably to emotional marketing communication when they believe AI authored the message, with perceived authenticity helping to explain the effect. Yet other evidence suggests that AI is not inherently distrusted. Logg et al. (2019) document algorithm appreciation under certain conditions, while recent studies of AI-generated advertising show that transparency cues, verification mechanisms, and positive attitudes toward AI can improve or stabilize trust (Bui, 2026; Kailani et al., 2026; Phan & Truong-Dinh, 2026).

The construct that can help reconcile these apparently conflicting findings is perceived authenticity. Authenticity is not identical to whether a message was physically produced by a human. In consumer research, authenticity is a perceptual judgment concerning whether an object, brand, or communication appears true to its identity, credible in its claims, consistent in its values, and sincere in its meaning (Morhart et al., 2015). AI-generated content can therefore be technically artificial while still being experienced as authentic if consumers believe it accurately represents the brand, does not manipulate them, and preserves a coherent brand voice. Conversely, human-written content can be experienced as inauthentic when it is exaggerated, opportunistic, or inconsistent with prior brand behavior.

A second construct is perceived brand transparency. Contemporary transparency research emphasizes that transparency is more than the amount of information a firm discloses. Montecchi et al. (2024) conceptualize perceived brand transparency in terms of observability, comprehensibility, and intentionality: consumers must be able to see relevant value-creation information, understand the communication, and perceive that the brand willingly provides relevant information. This distinction is crucial for AI marketing. A small label such as ‘AI-generated’ may satisfy a narrow disclosure requirement but still fail to create a broader perception that the brand is open, understandable, and accountable. Conversely, an appropriately designed explanation of what AI did, what humans reviewed, and how claims can be verified can operate as a stronger transparency signal.

The present study therefore asks: How does the perceived authenticity of AI-generated marketing content influence brand trust and purchase intention, and through what mechanism does perceived brand transparency transmit this influence? The article develops a conceptual model in which authenticity is the initiating perception, brand transparency is an interpretive mechanism, brand trust is the relational outcome, and purchase intention is the behavioral response. Signaling Theory explains how authenticity and transparency reduce information asymmetry surrounding algorithmic authorship, while the S-O-R framework explains the psychological sequence from marketing stimulus to internal evaluation and behavioral response.

The study makes four contributions. First, it separates the authenticity of the content from the transparency of the brand, avoiding the common assumption that disclosure and transparency are interchangeable. Second, it positions perceived brand transparency as a mediator linking authentic AI content to trust. Third, it integrates



the growing but fragmented research on AI disclosure, algorithm aversion, authenticity, and trust into one parsimonious model. Fourth, it offers a testable empirical agenda suitable for marketing research across different product categories, cultural contexts, and levels of AI involvement. The issue is especially timely as regulators, platforms, and consumers increasingly expect firms to identify or explain synthetic content while brands simultaneously seek to preserve persuasion effectiveness.

2. Conceptual Review Method

This article follows an integrative conceptual-review approach designed for theory building rather than statistical meta-analysis. The purpose is to synthesize adjacent streams of research—AI in marketing, algorithmic judgment, brand authenticity, transparency, trust, and purchase intention—and use their intersection to derive a coherent set of propositions. Integrative reviews are appropriate when an emerging phenomenon spans multiple theoretical traditions and when the central research problem is not merely to estimate an average effect but to clarify constructs and causal logic.

The review prioritizes peer-reviewed studies published in leading or established outlets in marketing, consumer behavior, psychology, information systems, retailing, and digital commerce, with emphasis on research published from 2020 through 2026. Foundational works were retained where they define core constructs or theories, including signaling, persuasion knowledge, algorithm aversion/appreciation, brand authenticity, and brand trust. Recent empirical work was included when it directly examined AI-generated advertising or content, AI disclosure, authenticity perceptions, consumer trust, transparency, brand evaluation, or purchase intention. The synthesis also uses scale-development research for brand authenticity, brand trust, and perceived brand transparency because the proposed model is intended to be empirically testable.

The review is intentionally theory driven. Studies were compared according to (a) the role assigned to AI authorship or disclosure, (b) the psychological mechanism proposed to explain consumer response, (c) the focal outcome, and (d) contextual conditions that could explain inconsistent findings. Particular attention was given to contradictory results. For example, disclosure can reduce authenticity or advertising value in some studies, while transparency and verification cues can increase trust in others. Rather than treating these findings as mutually exclusive, the present framework interprets them as evidence that the consumer's subjective assessment of authenticity and transparency—not the mere presence of AI—determines whether AI functions as a positive or negative market signal.

3. Literature Synthesis

3.1 AI-generated marketing content as a new authorship cue

Traditional advertising models usually assume a human organization as the intentional source of communication. GenAI complicates this assumption because creative execution can be produced through varying degrees of automation: fully human, AI-assisted human work, human-supervised AI generation, or largely autonomous generation. Consumers may not observe these distinctions directly. Instead, they infer authorship and accountability from labels, stylistic cues, platform notices, brand explanations, and prior knowledge. This makes AI involvement a cue whose meaning depends on interpretation.

Research increasingly shows that consumers respond not only to what AI produces but to how the brand frames AI's role. Brüns and Meißner (2024) found less negative response when GenAI assisted humans rather than replaced them. Kirk and Givi (2025) similarly reported that the adverse AI-authorship effect was weaker when AI was used for editing rather than as the apparent author of emotional communication. These findings suggest that human oversight can protect the perceived intentionality and moral accountability that consumers often associate with authentic brand expression.

At the same time, there is no universal preference for human judgment. Algorithm appreciation research demonstrates that consumers can value algorithmic advice, particularly when it is perceived as objective,



competent, or data driven (Logg et al., 2019). This is consistent with the idea that AI authorship is not intrinsically disqualifying. Its effect depends on whether the technology signals competence without simultaneously undermining sincerity, humanness, or accountability.

3.2 Perceived authenticity in the AI context

Brand authenticity has been conceptualized as a multidimensional perception involving credibility, integrity, symbolism, and continuity (Morhart et al., 2015). For AI-generated marketing content, these dimensions can be translated into four practical questions. Is the content credible and factually plausible? Is it consistent with the brand's values and ethical positioning? Does it express meaning that consumers can connect to the brand? And does it preserve continuity with the brand's established identity rather than appearing as opportunistic or generic automation?

Recent AI research supports the relevance of these questions. Bui, Filimonau, and Sezerel (2024) show that perceived authenticity of AI-generated visual content can positively influence trust and patronage intentions in destination marketing. Brüns and Meißner (2024), however, demonstrate that disclosure of GenAI use can reduce authenticity perceptions in social-media branding. Kirk and Givi (2025) add an important boundary condition: the authenticity penalty is particularly relevant when the message carries emotional meaning, because consumers may believe that machine authorship lacks the experiential basis expected for sincere emotional expression. Together, these findings imply that authenticity is best treated as a consumer inference rather than a technical property of the production process.

This inference is also sensitive to product category and communication goal. In utilitarian contexts, consumers may value accuracy, speed, and informativeness, making AI competence a positive cue. In identity-laden or hedonic contexts, authenticity and humanness may carry greater weight. Jung et al. (2025), for example, show that trust in AI and perceived humanness are important pathways in luxury advertising. Thus, authenticity can operate as the bridge between technological execution and relational brand outcomes.

3.3 Perceived brand transparency

Transparency is often reduced in practice to a compliance label, but perceived transparency is broader. Montecchi et al. (2024) define the construct through observability, comprehensibility, and intentionality. A transparent brand allows consumers to understand how value is created, communicates in an accessible manner, and voluntarily makes relevant information visible. This definition is particularly useful for AI marketing because consumers may care not only whether AI was involved but also what it did, how much control humans retained, and whether the output was verified.

Experimental evidence on traditional brand communication indicates that transparent communication can increase perceived authenticity, trust, and behavioral intentions (Yang et al., 2021). In the AI domain, recent work similarly indicates that transparency can become a trust-building mechanism when it is meaningful and supported by credible signals. Kailani et al. (2026) find that AI transparency contributes particularly to trust in social-media marketing. Research on verification signals further suggests that institutional or traceable cues can increase trust toward AI-generated advertising (Consumer Trust and Purchase Intentions in AI-Generated Advertising, 2026). These findings are compatible with signaling theory: transparency reduces uncertainty by making the brand's actions easier to interpret and verify.

However, transparency may also impose a cost when disclosure activates negative persuasion knowledge or algorithm aversion. Bui (2026) reports that AI-involvement disclosure can reduce advertising value and purchase intention, partly through lower advertising credibility. This apparent contradiction reinforces the present paper's core distinction. Disclosure is an objective message feature; transparency is a consumer perception. A disclosure that feels legalistic, unexplained, or threatening may increase awareness of AI without increasing transparency. Therefore, the effect of transparency depends on how the consumer interprets the brand's motives and the usefulness of the information provided.



3.4 Brand trust and purchase intention

Brand trust reflects consumers' confident expectations that a brand is reliable and acts with acceptable intentions in situations involving uncertainty or vulnerability. Delgado-Ballester, Munuera-Alemán, and Yagüe-Guillén (2003) provide a validated foundation for measuring this relational construct. In AI-mediated marketing, uncertainty is heightened because consumers may not know how content was generated, whether claims were hallucinated or embellished, or how responsibility is allocated between humans and systems. Trust therefore becomes a central mechanism translating communication perceptions into behavioral outcomes.

Recent AI advertising studies consistently place trust close to purchase intention. Bui et al. (2024) identify trust as a mechanism connecting authenticity to patronage intention. Israfilzade (2025) examines consumer trust and purchase intent in AI- versus human-created advertising. Kailani et al. (2026) identify trust as a central mediator between AI communication features and purchase intention. Phan and Truong-Dinh (2026) likewise show that trust formation depends on the resolution of competing credibility and eeriness signals. Collectively, this evidence supports a process view: consumers do not purchase because content is AI-generated or human-generated per se; they purchase when the communication environment produces sufficient confidence in the brand.

Table 1. Selected recent literature informing the conceptual model

Study	Context	Core mechanism	Key implication	Role in model
Brüns & Meißner (2024)	Social-media brand content	Algorithm authenticity; aversion;	GenAI can reduce brand authenticity, especially when replacing humans	Authenticity as key evaluator
Bui, Filimonau & Sezerel (2024)	AI-generated destination images	Authenticity; trust	Perceived authenticity supports trust and patronage intention	Authenticity → trust → intention
Kirk & Givi (2025)	AI-authored marketing communication	Authenticity; moral disgust	AI authorship can weaken response to emotional messages	Boundary condition: message type
Jung et al. (2025)	Luxury AI advertising	Trust in AI; humanness	Trust and humanness shape advertising effectiveness	Trust mechanism
Montecchi et al. (2024)	Brand transparency	Observability; comprehensibility; intentionality	Transparency predicts trust and purchase-related evaluations	Mediator construct
Bui (2026)	AI advertising disclosure	Persuasion knowledge; credibility	Disclosure can lower ad value and purchase intention	Disclosure ≠ transparency
Kailani et al. (2026)	AI social-media communication	S-O-R; trust	AI transparency primarily contributes to trust	Supports transparency → trust
Phan & Truong-Dinh (2026)	AI retail advertising	Signaling; credibility vs. eeriness	Trust depends on competing signals	Signaling logic

4. Theoretical Foundations

4.1 Signaling Theory

Signaling Theory originates from markets characterized by information asymmetry, where one party possesses information that another party cannot directly observe (Spence, 1973). Marketing communication is a signaling environment because consumers rarely observe the full process through which claims, prices, product quality, or creative messages are produced. AI intensifies this asymmetry. A consumer seeing a polished synthetic image may not know whether it accurately represents a real product, whether the brand intentionally created a fictional scene, whether a human verified the result, or whether the message is generated automatically from personal data.

Signals become useful when they are interpretable and sufficiently credible. In this paper, perceived authenticity functions as a signal quality judgment: the consumer evaluates whether the AI-generated content appears genuine, coherent, and congruent with the brand. Perceived brand transparency then represents a second-order inference about the sender: if the brand appears willing to make its process understandable and its claims observable, uncertainty falls. Trust is the relational consequence of successful signal resolution. Purchase intention follows when trust is sufficient to support action.



This logic also explains why disclosure can produce different outcomes. An AI label is a signal, but signals vary in meaning and strength. If the label communicates accountability and is paired with verifiable information, it can support transparency. If it merely highlights automation without explaining purpose or human oversight, it can activate skepticism. Therefore, the central theoretical question is not simply whether a brand discloses AI, but whether the total signal system makes the brand appear authentic and transparent.

4.2 Stimulus–Organism–Response framework

The S–O–R framework distinguishes environmental stimuli from internal psychological states and behavioral responses. Applied to digital marketing, the model is useful because the same technological stimulus can produce different behaviors depending on how consumers interpret it. In the present framework, the perceived authenticity of AI-generated content represents the focal stimulus-level evaluation, perceived brand transparency and brand trust represent organism-level cognitive and relational states, and purchase intention represents the response.

The model is consistent with recent AI-marketing research. Kailani et al. (2026) use S–O–R logic to explain how AI communication features influence trust, brand attitude, and purchase intention. The present study extends this approach by specifying a more focused sequence in which authenticity shapes transparency perceptions before trust is formed. This sequence is especially relevant because transparency is not purely an external stimulus; it is a subjective interpretation of the brand’s communication behavior.

4.3 Persuasion knowledge and algorithm aversion/appreciation as boundary logics

Consumers are active interpreters of persuasion attempts. The Persuasion Knowledge Model argues that consumers develop knowledge about marketers’ tactics and use that knowledge when evaluating influence attempts (Friestad & Wright, 1994). AI disclosure may therefore activate questions about why automation was used, whether personalization is excessive, or whether the brand is trying to lower costs while simulating human creativity. Such interpretations can reduce authenticity even when the output is technically high quality.

At the same time, algorithm aversion is not universal. Dietvorst, Simmons, and Massey (2015) show that people can lose confidence in algorithms quickly after observing errors, whereas Logg et al. (2019) demonstrate conditions under which algorithmic advice is preferred. The implication is that AI skepticism should be treated as a boundary condition rather than assumed as a constant. This paper therefore keeps AI attitudes, literacy, product involvement, and message emotionality outside the core parsimonious model but identifies them as priority moderators for empirical extension.

5. Conceptual Model and Research Propositions

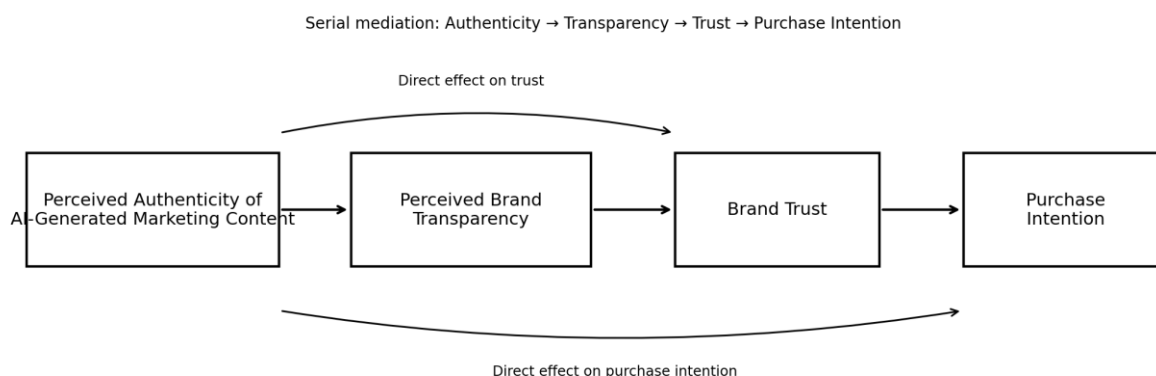


Figure 1. Proposed conceptual framework



5.1 Perceived authenticity and perceived brand transparency

Authentic-seeming content should encourage consumers to infer that the brand is communicating in a way that is faithful to its identity and not attempting to manipulate the audience through synthetic deception. This inference can spill over from the message to the sender. When consumers perceive AI-generated content as credible, value-consistent, and coherent with the brand's established voice, they have less reason to suspect hidden opportunism. In turn, they are more likely to believe that the brand is willing to make relevant aspects of value creation and communication visible.

The relationship is not equivalent to saying that authenticity automatically proves transparency. Rather, authenticity provides evidence from which consumers may infer the brand's communicative openness. The inference should be strongest when the brand's AI use is accompanied by understandable explanations, consistent visual identity, accessible factual verification, and clear human accountability.

Proposition 1 (P1). Perceived authenticity of AI-generated marketing content is positively associated with perceived brand transparency.

5.2 Perceived brand transparency and brand trust

Transparency reduces uncertainty by making the brand's intentions, processes, and claims easier to evaluate. Montecchi et al. (2024) provide evidence that perceived transparency supports brand trust, while Yang et al. (2021) show that transparent communication can strengthen authenticity and consumer responses. In AI settings, this function is especially important because the consumer cannot directly observe training data, prompting, editing, or human review. A transparent brand does not need to disclose technically sensitive details; it needs to provide sufficient observability, comprehensibility, and intentionality for consumers to feel that relevant information is not being strategically concealed.

Trust should therefore rise when consumers believe that the brand voluntarily explains AI involvement, distinguishes synthetic representation from factual product information, accepts responsibility for outputs, and offers a route for verification or correction. This mechanism is consistent with recent research showing that AI transparency and verification cues can support trust.

Proposition 2 (P2). Perceived brand transparency is positively associated with brand trust in AI-mediated marketing communication.

5.3 Brand trust and purchase intention

Purchase intention requires consumers to move from evaluation to anticipated action. This transition is risky when the source of communication is uncertain or when synthetic media may exaggerate product characteristics. Brand trust reduces perceived relational and performance risk by providing confidence that the brand will behave reliably and responsibly. The trust-intention relationship is well established in consumer research and is repeatedly observed in recent AI-marketing studies.

In the present model, trust is expected to be the most proximal relational predictor of purchase intention. Consumers may admire the creativity of an AI-generated advertisement yet remain unwilling to buy if they distrust the brand. Conversely, moderate creative quality may still lead to purchase when trust is high.

Proposition 3 (P3). Brand trust is positively associated with purchase intention toward brands using AI-generated marketing content.

5.4 Direct effect of authenticity on brand trust

Authenticity can influence trust independently of transparency because consumers may directly interpret authentic communication as evidence of integrity and reliability. Morhart et al. (2015) connect authenticity to favorable relational outcomes, and Bui et al. (2024) report a positive relationship between perceived



authenticity of AI-generated visual content and trust. In AI-mediated communication, a message that feels true to the brand can reassure consumers even when the production process is not fully salient.

This direct path is important because transparency is unlikely to account for all of authenticity's effect. Symbolic fit, emotional congruence, continuity, and credibility can strengthen trust without consumers consciously thinking about disclosure or production visibility.

Proposition 4 (P4). Perceived authenticity of AI-generated marketing content is positively associated with brand trust, above and beyond its indirect effect through perceived brand transparency.

5.5 Direct effect of authenticity on purchase intention

Authenticity can also influence purchase intention directly through perceived fit, reduced skepticism, and self-relevance. Consumers may interpret authentic brand expression as diagnostic evidence that the offering is congruent with the brand promise. Research in tourism and retail contexts indicates that authenticity can support patronage and favorable behavioral intention. Nonetheless, the present model expects much of this effect to operate through transparency and trust, particularly when AI authorship increases uncertainty.

A residual direct path remains theoretically plausible because purchase intention is not purely relational: consumers may respond to perceived aesthetic fit, creativity, novelty, or symbolic meaning even before trust is fully formed.

Proposition 5 (P5). Perceived authenticity of AI-generated marketing content is positively associated with purchase intention.

5.6 Mediation through perceived brand transparency

The core contribution of the model lies in the proposed mediating role of perceived brand transparency. If authentic AI content signals that the brand's communication is coherent and non-deceptive, consumers should infer greater openness and accountability. Transparency then reduces information asymmetry and supports trust. This mechanism explains why two advertisements with equally visible AI disclosure may produce different outcomes: one can appear authentic and therefore transparent, while the other can appear synthetic, opportunistic, or strategically evasive.

Mediation also offers a theoretical explanation for mixed disclosure findings. A disclosure that lowers authenticity may weaken transparency inference despite increasing factual awareness of AI use. By contrast, a disclosure embedded within a broader transparency strategy may strengthen trust. Thus, transparency is not the disclosure itself; it is the psychological state produced when disclosure, message quality, brand identity, and accountability cues are jointly interpreted as openness.

Proposition 6 (P6). Perceived brand transparency mediates the positive relationship between perceived authenticity of AI-generated marketing content and brand trust.

5.7 Serial mediation to purchase intention

Combining the preceding propositions yields a sequential process. Authenticity first shapes whether the brand appears open and accountable; transparency then strengthens trust; trust finally supports purchase intention. This serial mediation path integrates signaling theory with S-O-R logic and provides a process explanation rather than a simple correlation between AI content quality and consumer behavior.

The serial pathway should be especially important in contexts where consumers face high informational uncertainty—for example, when synthetic images depict product use, when personalized claims are generated automatically, or when the brand communicates emotionally significant messages. In these cases, authenticity alone may be insufficient unless it also supports a belief that the brand is transparent and trustworthy.



Proposition 7 (P7). Perceived authenticity has a positive indirect effect on purchase intention through the serial pathway of perceived brand transparency and brand trust.

Table 2. Summary of research propositions

Proposition	Relationship	Expected direction
P1	Authenticity → perceived brand transparency	Positive
P2	Perceived brand transparency → brand trust	Positive
P3	Brand trust → purchase intention	Positive
P4	Authenticity → brand trust	Positive
P5	Authenticity → purchase intention	Positive
P6	Authenticity → transparency → trust	Positive mediation
P7	Authenticity → transparency → trust → purchase intention	Positive serial mediation

6. Construct Clarification and Operationalization

A major risk in AI-marketing research is construct overlap. Authenticity, credibility, transparency, trust, and disclosure are related but theoretically distinct. Clear operationalization is therefore necessary if the model is to be tested without tautology. Table 3 distinguishes the focal constructs and provides measurement guidance derived from validated research traditions.

Table 3. Construct definitions and recommended operationalization

Construct	Working definition	Measurement focus	Primary source base
Perceived authenticity of AI content	Extent to which AI-generated marketing content is experienced as genuine, credible, value-consistent, and true to the brand	Credibility, integrity, symbolic fit, continuity; adapted to content level	Morhart et al. (2015); Bui et al. (2024)
Perceived brand transparency	Extent to which the brand appears to make value creation and AI use observable, understandable, and voluntarily disclosed	Observability, comprehensibility, intentionality	Montecchi et al. (2024)
Brand trust	Confidence that the brand is reliable and acts with acceptable intentions despite uncertainty	Reliability and intentions / benevolence dimensions	Delgado-Ballester et al. (2003)
Purchase intention	Consumer's stated likelihood or willingness to purchase the focal brand/product	Likelihood, willingness, consideration	Established consumer-intention measures; adapted to context
AI attitude / literacy (moderator)	General evaluation of AI and capacity to understand AI-mediated content	Positive/negative AI attitude; familiarity and literacy	Recommended extension based on Bui (2026); Kailani et al. (2026)

7. Recommended Empirical Validation Design

Although the present article is conceptual, the model is designed for rigorous empirical testing. A two-study program would provide stronger evidence than a single cross-sectional survey. Study 1 should use an experiment to establish causal effects of AI-content authenticity and transparency cues. Study 2 should use a larger survey or field sample to test the full structural model and assess external validity.

For Study 1, a 2×2 between-subjects design is recommended. The first factor would manipulate authenticity cues (high versus low) while holding the advertised product and basic claim constant. High-authenticity execution should use brand-consistent language, realistic product representation, coherent tone, and human-reviewed factual claims; low-authenticity execution should use generic wording, inconsistent imagery, or exaggerated synthetic cues without becoming obviously fraudulent. The second factor would manipulate transparency (high versus low). High transparency could state that AI assisted in content production, identify the role of human review, and provide a concise verification link or explanation. Low transparency could omit such process information. Manipulation checks should assess perceived authenticity and transparency rather than assuming that the experimental treatments worked.

A sample of approximately 300–500 adult consumers would normally provide adequate power for medium effects and allow comparisons across conditions, although an a priori power analysis should determine the final number. Participants should be randomly assigned and exposed to identical brand/product information



except for the manipulated cues. Brand familiarity can be controlled by using either a fictitious brand or a real brand with measured familiarity. Product involvement, AI attitude, AI literacy, age, and prior use of generative AI should be measured as controls or moderators.

For Study 2, the full model can be tested using covariance-based SEM or PLS-SEM depending on the research objective and distributional characteristics. A minimum sample in the range of 400–600 would be preferable for a model containing multiple latent variables, mediation paths, controls, and potential multi-group comparisons. Measurement validation should precede structural testing. Researchers should report reliability, convergent validity, discriminant validity, model fit where applicable, standardized path coefficients, confidence intervals, and indirect effects estimated using bootstrapping.

Common-method concerns should be addressed through both procedural and statistical remedies. Procedurally, predictor and criterion measures can be separated, item order randomized, scale anchors varied where appropriate, and anonymity emphasized. Statistically, researchers can assess a common latent factor or use marker-variable procedures rather than relying solely on Harman’s one-factor test. Because the topic concerns AI, attention checks should verify that participants noticed the source/manipulation without explicitly revealing hypotheses before outcome measures are completed.

Cross-cultural replication would be especially valuable. The meaning of transparency and the legitimacy of algorithmic authorship may vary across individualistic and collectivistic markets, high- and low-trust environments, and jurisdictions with different AI disclosure norms. Recent evidence from Vietnam and Romania indicates that cultural context can shape AI disclosure and trust relationships. Multi-group SEM or moderated mediation could test whether the serial pathway is stable across markets.

Table 4. Illustrative measurement indicators for future empirical testing (paraphrased, not verbatim scale items)

Construct	Illustrative indicators
AI-content authenticity	The content feels true to the brand; the content appears sincere rather than artificial for its own sake; the message is consistent with what I know about the brand; the claims and visuals feel credible.
Brand transparency	The brand makes relevant aspects of how it uses AI visible; its explanation is easy to understand; the brand appears willing rather than forced to share relevant information; I can understand how the marketing content was produced or checked.
Brand trust	I can rely on this brand; I expect the brand to keep its promises; I believe the brand would act responsibly if a problem occurred; I am confident in the brand’s intentions.
Purchase intention	I would consider buying from this brand; I am likely to choose this brand when purchasing in the category; the advertisement increases my willingness to purchase.

8. Discussion

The proposed framework reframes the managerial question from ‘Should brands use AI-generated content?’ to ‘Under what perceptual conditions can AI-generated content support trustworthy brand relationships?’ This shift is important because the empirical literature does not support a simple deterministic answer. AI can generate authenticity penalties when consumers interpret machine authorship as emotionally insincere, opportunistic, or replacement oriented. Yet AI can also be accepted—and sometimes valued—when it is associated with competence, relevance, verification, and responsible oversight.

The model explains this divergence through two linked judgments. The first is message authenticity: does the content feel genuine and consistent with the brand? The second is sender transparency: does the brand appear open and accountable about how the message was created and what it represents? These judgments are conceptually separable. A technically disclosed AI advertisement can feel inauthentic, and an undisclosed advertisement can feel authentic. Likewise, a brand can disclose AI use without being perceived as transparent if the disclosure is vague, difficult to interpret, or apparently motivated only by compliance.



Placing transparency between authenticity and trust advances current AI-marketing theory in three ways. First, it treats transparency as an interpretive state rather than a binary treatment. Second, it helps explain why disclosure effects are inconsistent across studies. Third, it connects authenticity research—often focused on brand meaning—with trust research focused on risk and relational confidence. In an environment where synthetic media increases the cost of verifying what is real, the ability of brands to transform authentic execution into understandable and voluntary transparency may become a core source of trust.

The serial mediation path also clarifies why purchase intention may not respond directly to disclosure. Disclosure can influence purchase only after consumers decide what it means. If disclosure signals honesty, it can improve transparency and trust. If it signals low effort, automation, job replacement, or manipulation, it can reduce authenticity and advertising credibility. The consumer's interpretive pathway therefore matters more than the label alone.

9. Theoretical Contributions

First, the study extends brand-authenticity research into a context where the apparent author may not be human. Classic authenticity dimensions remain useful, but AI creates a distinctive question of indexicality: content can be highly realistic while having no direct physical referent or lived human experience. By treating authenticity as a consumer judgment about credibility, integrity, symbolic fit, and continuity, the model avoids equating authenticity with human authorship.

Second, the study advances transparency research by emphasizing perceived transparency rather than disclosure. The distinction is theoretically consequential. Disclosure is an observable communication action; transparency is the consumer's integrated inference that relevant information is visible, comprehensible, and intentionally provided. This distinction helps explain why mandatory or minimal disclosures can have effects that differ from broader transparency strategies.

Third, the model integrates signaling theory with S-O-R. Signaling theory explains why authenticity and transparency matter under information asymmetry, while S-O-R specifies the process through which content perceptions become relational evaluations and behavioral intentions. The integration yields a parsimonious sequence from authenticity to transparency, trust, and purchase intention.

Fourth, the framework provides an explanation for mixed findings in AI-disclosure research without assuming that one set of studies is incorrect. Negative and positive disclosure effects can coexist because disclosure changes the signal environment, but consumers differ in how they interpret the signal. Authenticity, perceived transparency, AI attitudes, product type, message emotionality, and verification cues can all shift the meaning of the same disclosure.

10. Managerial Implications

For managers, the central implication is that AI governance and brand communication should be designed together. A firm cannot protect trust solely by adding an 'AI-generated' label after content has been produced. Trust protection begins earlier, with decisions about which tasks are delegated to AI, what claims require human verification, how closely generated content must follow brand identity, and how consumers can understand the production process.

Brands should preserve recognizable continuity. AI-generated copy and imagery should be trained, prompted, or edited to reflect established brand voice, values, visual codes, and factual product attributes. Generic AI aesthetics can create a sense that the brand is outsourcing its identity. Human oversight is particularly important for emotional messages, social causes, crisis communication, testimonials, and other contexts in which audiences expect lived experience or moral accountability.



Transparency should be proportionate and informative. Consumers do not need a technical description of the model architecture. They need information that helps them evaluate the message: whether AI assisted or generated the content, whether a human reviewed factual claims, whether an image is illustrative rather than documentary, and where further information can be verified. Verification devices such as traceable landing pages, provenance information, or third-party assurance can strengthen the credibility of transparency cues when the context justifies them.

Managers should also segment by context. Utilitarian categories may tolerate or reward AI competence, whereas luxury, creative, hospitality, health-adjacent, and emotionally expressive categories may require stronger authenticity safeguards. Likewise, experienced AI users may interpret disclosure differently from consumers who have low AI familiarity. Therefore, pretesting should measure perceived authenticity and transparency rather than only ad liking.

Finally, firms should treat transparency as a relationship asset rather than a compliance burden. A brand that repeatedly communicates clearly about synthetic content can develop a reputational expectation of openness. Such consistency may reduce the future cost of AI disclosure because consumers learn how the brand uses technology and where human accountability remains.

Table 5. Managerial design principles for trustworthy AI-generated marketing

Managerial objective	Recommended action	Risk if ignored	Expected consumer inference
Preserve authenticity	Maintain brand voice, realistic representation, human review	Generic or uncanny content; identity dilution	"This still feels like the brand."
Create meaningful transparency	Explain AI role, human oversight, and limits in plain language	Disclosure without understanding	"The brand is open with me."
Strengthen verifiability	Provide source/provenance or accessible verification where material	Synthetic claims appear unverifiable	"I can check what matters."
Protect emotional sincerity	Use humans in high-emotion communication and disclose assistance appropriately	Moral discomfort; emotional insincerity	"A responsible person stands behind this."
Build trust over time	Use consistent AI policy across touchpoints	Inconsistent disclosure and surprise	"I know how this brand uses AI."

11. Ethical and Governance Considerations

AI-generated marketing raises ethical questions beyond persuasion effectiveness. Synthetic content can blur the boundary between illustration and evidence, reproduce biased representations, imitate creative styles, or personalize messages using data that consumers do not expect to be used for content generation. A high-performing AI campaign may therefore be commercially effective while still damaging long-run legitimacy if consumers later discover that important aspects of the communication were misleading.

The proposed transparency mechanism offers a governance principle: information should be disclosed when it is material to consumer interpretation. Materiality depends on whether knowledge of AI involvement would reasonably change how a consumer evaluates authenticity, product evidence, endorsement, or accountability. For example, a synthetic lifestyle background may require less explanation than an AI-generated testimonial-like person, simulated product performance, or emotionally charged message presented as if written from lived experience.

Human accountability must remain identifiable. Consumers should not face a situation in which the brand attributes errors to 'the AI' while benefiting from the persuasive output when it performs well. Responsible deployment therefore requires documented review processes, correction pathways, and clear ownership of claims. Transparency is most credible when it is supported by internal governance rather than treated as a surface-level communication tactic.

12. Limitations and Future Research

The principal limitation is that the article is conceptual and therefore does not estimate the magnitude of the proposed relationships. The model should be tested experimentally and across field settings before causal



claims are generalized. In particular, the proposed path from authenticity to perceived transparency is theoretically plausible but less established than the transparency-to-trust relationship; this makes P1 and P6 especially important targets for empirical validation.

Future research should distinguish levels of AI contribution. Consumers may react differently when AI generates an initial draft, edits human content, creates imagery under human direction, or autonomously generates and publishes content. A continuous measure of perceived AI agency may be more informative than a binary AI/human label. Human-AI collaboration should also be studied as a signal in its own right.

Message type is another priority. Emotional, factual, humorous, luxury, social-cause, crisis, and service-recovery messages impose different expectations of authenticity and human intention. Kirk and Givi (2025) show that emotional communication is particularly sensitive to AI authorship. Future studies can test whether transparency mediates authenticity effects more strongly when the message invites empathy or moral identification.

AI literacy and general attitudes toward AI should be tested as moderators. Consumers with high familiarity may better distinguish AI assistance from full automation, but they may also be more alert to synthetic artifacts or privacy risks. Cultural orientation, regulatory environment, age, and occupational exposure to AI may similarly shape the model. Cross-national studies are especially valuable because transparency norms differ across markets.

Researchers should also examine actual behavior. Purchase intention is a useful conference-level outcome, but click-through, add-to-cart, conversion, repeat purchase, complaint behavior, and willingness to share personal data would provide stronger evidence of marketplace consequences. Field experiments in e-commerce or social platforms could test whether transparent AI labels affect real conversion differently for existing versus new customers.

Finally, future work should consider dynamic trust. A single exposure may not capture how consumers learn a brand's AI practices over time. Longitudinal research can examine whether consistent transparency creates a trust buffer that protects the brand when an AI error occurs, or whether repeated synthetic communication gradually erodes perceived humanness and authenticity.

13. Conclusion

Generative AI makes marketing faster and more scalable, but it also changes the meaning of authorship. Consumers increasingly evaluate not only what a brand says but how the message was produced and whether the brand remains accountable for it. This paper proposes that perceived authenticity of AI-generated marketing content can support brand trust and purchase intention when it contributes to a broader perception of brand transparency.

The central theoretical insight is that authenticity, transparency, and trust form a sequence rather than a set of interchangeable positive evaluations. Authenticity concerns whether the content feels genuine and true to the brand. Transparency concerns whether the brand's actions and communication process are visible, understandable, and voluntarily disclosed. Trust concerns whether the consumer is willing to rely on the brand under uncertainty. Purchase intention is the behavioral consequence. By integrating these constructs through signaling theory and S-O-R, the framework explains why AI disclosure can sometimes help and sometimes hurt: the outcome depends on the meaning consumers assign to the signal.

For practice, the message is straightforward. Brands should not choose between AI efficiency and authenticity. They should design AI systems, human oversight, disclosure, and verification so that technological efficiency remains compatible with a recognizable brand identity and an accountable relationship with consumers. In an era of synthetic media, the brands most likely to sustain trust will be those that make AI use understandable without making the brand itself feel artificial.

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. <https://doi.org/10.1016/j.jretconser.2024.103790>
- Bui, H. T. (2026). Examining the effect of AI advertising involvement disclosure on advertising value and purchase intentions. *Journal of Research in Interactive Marketing*, 20(5), 613–632. <https://doi.org/10.1108/JRIM-02-2025-0066>
- 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>
- Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, 48, 24–42. <https://doi.org/10.1007/s11747-019-00696-0>
- Delgado-Ballester, E., Munuera-Alemán, J. L., & Yagüe-Guillén, M. J. (2003). Development and validation of a brand trust scale. *International Journal of Market Research*, 45(1), 1–18. <https://doi.org/10.1177/147078530304500103>
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114–126. <https://doi.org/10.1037/xge0000033>
- Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, 21(1), 1–31. <https://doi.org/10.1086/209380>
- Israfilzade, K. (2025). AI-generated versus human-created advertising: Effects on consumer trust and purchase intent. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 20(4), 1301–1337. <https://doi.org/10.24136/eq.4038>
- Jung, T., Koghut, M., Lee, E., & Kwon, O. (2025). Artificial creativity in luxury advertising: How trust and perceived humanness drive consumer response to AI-generated content. *Journal of Retailing and Consumer Services*, 87, 104403. <https://doi.org/10.1016/j.jretconser.2025.104403>
- Kailani, C., Alniacik, U., Barbu, A. M., Ionescu, F. T., & Ploșteanu, M. G. (2026). How AI-driven marketing communication shapes consumer attitudes and purchase intentions: Evidence from social media users in Romania. *Journal of Theoretical and Applied Electronic Commerce Research*, 21(8), 281. <https://doi.org/10.3390/jtaer21080281>
- Kirk, C. P., & Givi, J. (2025). The AI-authorship effect: Understanding authenticity, moral disgust, and consumer responses to AI-generated marketing communications. *Journal of Business Research*, 186, 114984. <https://doi.org/10.1016/j.jbusres.2024.114984>
- Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90–103. <https://doi.org/10.1016/j.obhdp.2018.12.005>
- Montecchi, M., Plangger, K., West, D., & de Ruyter, K. (2024). Perceived brand transparency: A conceptualization and measurement scale. *Psychology & Marketing*, 41, 2274–2297. <https://doi.org/10.1002/mar.22048>
- Morhart, F., Malär, L., Guèvremont, A., Girardin, F., & Grohmann, B. (2015). Brand authenticity: An integrative framework and measurement scale. *Journal of Consumer Psychology*, 25(2), 200–218. <https://doi.org/10.1016/j.jcps.2014.11.006>
- Phan, K. D., & Truong-Dinh, B. Q. (2026). Signal competition in AI-generated retail advertising: Credibility, eeriness, and trust formation. *Acta Psychologica*, 269, 107561. <https://doi.org/10.1016/j.actpsy.2026.107561>
- Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. *Journal of Marketing*, 85(1), 131–151. <https://doi.org/10.1177/0022242920953847>
- Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355–374. <https://doi.org/10.2307/1882010>
- Yang, J., & Battocchio, A. F. (2021). Effects of transparent brand communication on perceived brand authenticity and consumer responses. *Journal of Product & Brand Management*, 30(8), 1176–1193. <https://doi.org/10.1108/JPB-03-2020-2803>
- Yao, Y. (2025). Transparent technology: Evaluating the impact of AI-generated ad disclosure on consumer trust, authenticity, and purchase intention. In *2025 Global Marketing Conference at Hong Kong Proceedings* (p. 289).
- Saran, D., & Datt, V. (2026). Consumer trust and purchase intentions in AI-generated advertising: The role of verification signals. *Computers in Human Behavior Reports*, 22, 101112. <https://doi.org/10.1016/j.chbr.2026.101112>

Appendix A. Suggested Empirical Questionnaire Architecture

A future empirical paper can operationalize the model using a seven-point Likert response format. To minimize construct contamination, authenticity items should focus on qualities of the content, transparency items on the brand's openness and comprehensibility, trust items on the relationship with the brand, and purchase-intention items on anticipated behavior. Exact validated wording should be obtained from and cited to the original scale publications before field administration.

Recommended sequence: (1) stimulus exposure; (2) manipulation/attention checks; (3) perceived authenticity; (4) perceived brand transparency; (5) brand trust; (6) purchase intention; (7) AI attitudes and AI literacy; (8) product involvement and brand familiarity; (9) demographics. If a real brand is used, prior brand attitude should be measured before exposure or modeled as a control.

For mediation, indirect effects should be estimated using bias-corrected bootstrap confidence intervals. The serial indirect effect of greatest theoretical interest is Authenticity → Perceived Brand Transparency → Brand Trust → Purchase Intention. Researchers should compare this pathway with alternative models, such as Transparency → Authenticity → Trust, to demonstrate that the proposed ordering offers better theoretical and empirical fit.