

Generative AI Disclosure in Advertising: From Intent- to Production-Based Disclosure

Denitsa Dineva & Oguz A. Acar

To cite this article: Denitsa Dineva & Oguz A. Acar (24 Sep 2026): Generative AI Disclosure in Advertising: From Intent- to Production-Based Disclosure, Journal of Advertising, DOI: [10.1080/00913367.2026.2731565](https://doi.org/10.1080/00913367.2026.2731565)

To link to this article: <https://doi.org/10.1080/00913367.2026.2731565>



© 2026 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 24 Sep 2026.



Submit your article to this journal [↗](#)



Article views: 197



View related articles [↗](#)



View Crossmark data [↗](#)

Generative AI Disclosure in Advertising: From Intent- to Production-Based Disclosure

Denitsa Dineva^a and Oguz A. Acar^b

^aCardiff University, Cardiff, UK; ^bKing's College London, London, UK

ABSTRACT

Generative artificial intelligence (GenAI) is transforming advertising production and introducing new forms of disclosure at the point of exposure. Existing research treats disclosure primarily as a cue signaling persuasive intent, yet GenAI disclosures increasingly communicate *how* content was produced rather than *why* it was created. This paper reconceptualizes disclosure in AI-mediated advertising by positioning GenAI disclosure as a signal of production conditions. Integrating signaling theory, source credibility theory, the persuasion knowledge model, and the heuristic-systematic model, the paper develops a conceptual framework that differentiates declarative, regulatory, and infrastructural disclosure as distinct rendering conditions. The framework proposes that GenAI disclosure influences consumer evaluations through two complementary inferential pathways: perceived source attribution and production-based agent knowledge, defined as consumers' beliefs about how persuasive content was generated and what those production processes imply for evaluating the persuasion agent. It further specifies how consumers' motivation and ability shape the likelihood of heuristic versus systematic processing. The paper advances advertising theory by explaining how disclosure operates when its primary referent shifts from persuasive intent to production conditions and by providing a framework for understanding why similar disclosure cues may generate divergent evaluative responses across consumers, contexts, and forms of GenAI use.

Introduction

Generative artificial intelligence (GenAI) is rapidly being integrated into advertising production (Campbell et al. 2022; Hartmann, Exner, and Domdey 2025; Huh, Nelson, and Russell 2023), expanding the focus of disclosure from communicating persuasive intent to communicating the conditions under which advertising content was produced (Wojdyski and Evans 2016). A now common manifestation of this shift is the use of labels such as “AI-generated,” “synthetic content,” or “AI info” attached directly to content at the point of exposure (see Figure 1 for an example; Adobe 2025; Meta 2024; YouTube 2024). Specifically, platform-level implementations make this explicit by identifying whether content has been generated or materially altered by GenAI systems, while

provenance standards embed production metadata within the content itself (Adobe 2025; Coalition for Content Provenance and Authenticity 2023; YouTube 2024).

Public controversies surrounding AI-generated campaigns further illustrate major implications of this shift. In several cases, brands disclosed their use of generative tools, yet disclosure did not consistently mitigate skepticism or backlash (BBC. 2026; To et al. 2025). Gucci's AI-generated promotional imagery ahead of its 2026 Milan Fashion Week show, for example, prompted criticism centered not on the brand's persuasive intent, but on the production process itself, with debate focusing on authenticity, craftsmanship, heritage, and the role of human creative labor (Shahidi 2026). Rather than altering perceptions of why the communication was intended to persuade,

CONTACT Denitsa Dineva  DinevaD@cardiff.ac.uk  Cardiff Business School, Cardiff University, Cardiff CF10 3EU, UK.

Denitsa Dineva (Ph.D., Aberystwyth University) is a Senior Lecturer (Associate Professor) in Marketing and Strategy, Cardiff Business School, Cardiff University.

Oguz A. Acar (Ph.D., Erasmus University Rotterdam) is a Professor of Marketing & Innovation, King's Business School, King's College London.

© 2026 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

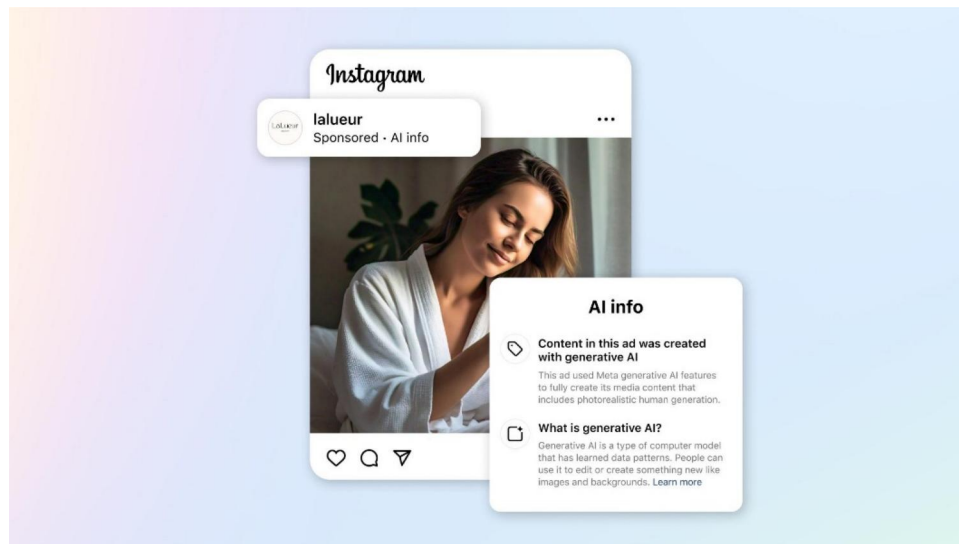


Figure 1. Example of GenAI disclosure (source: Meta 2025).

disclosure shifted attention to how the communication had been produced. This indicates that production disclosure introduces production conditions as an additional object of consumer evaluation.

Current advertising knowledge does not adequately address this phenomenon. Disclosure research has developed around a *stable* referent: the commercial relationship, whereby transparency in advertising has largely been conceptualized as the visibility of commercial intent (Boerman, Willemsen, and Van Der Aa 2017; Kim, Barasz, and John 2019; Wojdyski and Evans 2016). Disclosure mechanisms such as sponsorship statements, influencer tags, and “#ad” labels have long been used to make persuasive relationships explicit and enable audiences to calibrate trust (Boerman, Willemsen, and Van Der Aa 2017; Giuffredi-Kähr, Petrova, and Malär 2022). Within this tradition, disclosure activates persuasion knowledge by signaling advertiser intent (Friestad and Wright 1994), while platforms and advertisers may structure, standardize, and enforce how such information appears (Kerr et al. 2015).

GenAI disclosure does not fit this referent because it signals *how* content was produced rather than why it was created. Instead of indicating a persuasive relationship, these cues point to production conditions, including the configuration of human and machine inputs (Castelo, Bos, and Lehmann 2019; Longoni, Bonezzi, and Morewedge 2019; Molina and Sundar 2022). As a result, audiences infer not only who is attempting to persuade them, but also how the message was generated, often under conditions where source attribution is ambiguous or distributed (Luo

et al. 2019; Wang, Lopez, and Okazaki 2025). This shifts the object of disclosure from persuasive intent to production conditions, rendering authorship hybrid, distributing communicative agency across actors, and making production conditions increasingly transparent through technological systems (Shi and Jiang 2026).

Recent empirical evidence reinforces this mismatch. Disclosure can increase perceived transparency or honesty while simultaneously reducing authenticity, creativity, sincerity, or relational warmth (e.g., Baek, Kim, and Kim 2026; Brüns and Meißner 2024; Wortel, Vanwesenbeeck, and Tomas 2024). These mixed findings suggest that disclosures revealing production processes introduce evaluative considerations that are not fully captured by existing advertising disclosure models and point to production conditions as an additional basis for consumer evaluation.

Against this backdrop, we develop a production-based account of GenAI disclosure and make four theoretical contributions. First, we define GenAI disclosure in contemporary advertising, showing that the referent of disclosure shifts from signaling persuasive intent to signaling production conditions, extending prior conceptualizations of disclosure as a mechanism for making commercial relationships visible (Boerman, Willemsen, and Van Der Aa 2017; Friestad and Wright 1994; Wojdyski and Evans 2016).

Second, drawing on signaling theory, source credibility theory, the persuasion knowledge model (PKM), and the heuristic-systematic model (HSM), we develop a framework that conceptualizes disclosure as production-related information rather than a purely

intent-driven practice. In doing so, we advance production-based agent knowledge as an extension of the agent knowledge component of the PKM, capturing how beliefs about generativity, human oversight, provenance reliability, and normative appropriateness shape evaluations of persuasion agents (Castelo, Bos, and Lehmann 2019; Kirk and Givi 2025; Longoni, Bonezzi, and Morewedge 2019).

Third, we extend signaling theory by introducing three disclosure rendering conditions: declarative, regulatory, and infrastructural, which specify how production-related information is rendered and how communicative agency is distributed across advertising ecosystems (Giuffredi-Kähr, Petrova, and Malär 2022; Spence 1973). In particular, infrastructural disclosure embeds production information within technical systems such as provenance metadata and verification mechanisms rather than relying solely on externally displayed cues (Adobe 2025; Coalition for Content Provenance and Authenticity 2023).

Fourth and finally, we integrate these elements into a unified framework in which GenAI disclosure influences evaluation through two complementary inferential pathways: perceived source attribution and production-based agent knowledge. The framework explains how identical disclosure cues may generate divergent responses depending on how communicative agency is attributed, what production-related beliefs become salient, and whether those beliefs are processed heuristically or systematically.

Practically, the paper speaks to an emerging environment in which GenAI disclosure is increasingly institutionalized through regulation and platform governance, yet without convergence on a single best disclosure practice. As provenance systems and platform-level verification cues become more established (Adobe 2025; Coalition for Content Provenance and Authenticity 2023; European Commission 2024; YouTube 2024), disclosure is no longer solely a voluntary signaling device but an expected feature of communicative infrastructure. Consequently, advertisers, platforms, and regulators must make disclosure decisions in an environment where advertiser-led statements, mandated labels, and embedded provenance signals increasingly coexist, creating uncertainty regarding how production-related information will be interpreted and evaluated by consumers. By explaining how such disclosure conditions shape attribution and evaluation, the framework offers some guidance for designing more effective GenAI disclosure practices (ASA 2025; Federal Trade Commission 2009).

From Commercial to Production Disclosure

Advertising transparency has long been operationalized through disclosure, but disclosure processes have rarely been fully individual or linear because persuasive communications are often produced, distributed, amplified, and endorsed through multiple interconnected actors rather than a single identifiable source (Campbell, Mohr, and Verlegh 2013). Early regulatory and scholarly work, shaped by concerns over covert persuasion, established disclosure as a mechanism for signaling persuasive intent and protecting consumer autonomy, particularly in contexts where commercial influence may otherwise be obscured (Boerman, Willemsen, and Van Der Aa 2017; Campbell, Mohr, and Verlegh 2013). Subsequent research on sponsorship, influencer marketing, and native advertising shows that disclosure practices are structured by regulatory frameworks and platform governance as much as by advertiser intent (Evans et al. 2017; Giuffredi-Kähr, Petrova, and Malär 2022; Kerr et al. 2015; Wojdyski and Evans 2016).

Within this literature, sponsorship disclosure remains the dominant operationalization of transparency (Montecchi et al. 2024), typically realized through explicit cues such as labels or hashtags that make commercial relationships visible and activate persuasion knowledge (Boerman, Willemsen, and Van Der Aa 2017; Friestad and Wright 1994). The distributed character of disclosure across advertisers, regulators, and platforms has been recognized; however, this distribution is primarily coordination around a fixed referent (the commercial relationship) and a declarative cue format attached to the message.

What changes in GenAI contexts is not distribution per se but its locus and referent (Chen, Wang, and Hao 2025). Distribution becomes constitutive rather than coordinative: it is embedded in the production and verification pipeline and redefines what disclosure is about. Whereas sponsorship disclosure centers on commercial relationships and persuasive intent, GenAI transparency extends to production processes, content provenance, and machine involvement, altering (1) what is disclosed; (2) how disclosure is rendered; and (3) which actors are implicated in audience inference (Campbell et al. 2022; Castelo, Bos, and Lehmann 2019; Huh, Nelson, and Russell 2023; Longoni, Bonezzi, and Morewedge 2019). Disclosure is no longer confined to declarative cues attached to messages but increasingly emerges through platform-level labeling and infrastructural signals linked to content generation (Rough and Clift 2026), shifting it from a discrete communicative act to a system-level condition.

Table 1. Key dimensions of disclosure in advertising.

Dimension	Commercial disclosure (intent-based)	GenAI disclosure (production-based)
Object of disclosure	Commercial relationships and persuasive intent	Production conditions: how content is produced and the role of human and AI involvement
Core inference	Why is this message trying to persuade?	How this message is produced, by whom, and under what level of human oversight
Evaluative focus	Motives and persuasive intent of the advertiser	Production conditions, including generativity, human oversight, provenance, and appropriateness
Mechanism of disclosure	Predominantly explicit labels and sponsorship cues, often standardized or enforced by platforms and regulators (e.g., “#ad”)	Disclosure rendered through declarative, regulatory, and infrastructural conditions, including labels, verification cues, metadata, and provenance systems
Rendering mode	Usually through visible sponsorship cues, whether brand-provided, publisher-applied, or platform-standardized	Disclosure links content generation, platform governance, verification systems, and provenance infrastructures
Structure of accountability	Primarily tied to sponsor–publisher relations, with regulatory and platform oversight	Responsibility distributed across brands, platforms, model providers, and technical infrastructures

This shift has three interrelated implications (see Table 1). First, the object of disclosure shifts. Traditional sponsorship cues signal commercial relationships and prompt inferences about persuasive intent. In GenAI contexts, disclosure signals how content was produced. Labels such as “AI-generated” or provenance credentials indicate machine involvement and shift inference from motive to production, including questions of authorship, oversight, and autonomy (Adobe 2025; Meta 2024; YouTube 2024). Empirical evidence shows that such cues affect perceptions of authenticity, creativity, effort, and fairness even when persuasive intent is already clear (Baek, Kim, and Kim 2026; Kirkby, Baumgarth, and Henseler 2023; Wortel, Vanwesenbeeck, and Tomas 2024; Zhang and Hur 2025), indicating that production constitutes a distinct object of evaluation.

Second, the mechanism of disclosure changes. In sponsorship contexts, disclosure is typically rendered as a visible, declarative cue, whether voluntarily provided or enforced through platform rules (Montecchi et al. 2024). Under GenAI, disclosure is increasingly generated through socio-technical infrastructures. Generative systems may embed metadata, provenance standards attach credentials, platforms may apply detection algorithms, and governance rules determine how labels are displayed (YouTube 2024). Disclosure is therefore not only declared; in some settings, it is also computed or technically surfaced through integrated pipelines that link content production, embedded signals, platform verification, and policy enforcement.

Third, the structure of accountability shifts. Sponsorship disclosure primarily assigns responsibility to the advertiser, with platforms and regulators enforcing compliance (Montecchi et al. 2024). In GenAI contexts, accountability is distributed across brands, platforms, model providers, and technical

infrastructures. Errors or omissions in AI disclosure may therefore be attributed to platforms, automated systems, or governance processes rather than to advertisers alone, reflecting the distribution of perceived agency across multiple actors (Molina and Sundar 2022; Wang, Lopez, and Okazaki 2025). This aligns with research in human–machine communication showing that audiences may attribute agency and credibility to algorithmic systems (Lee 2024).

Conceptualizing GenAI Disclosure

Defining GenAI Disclosure in Advertising

GenAI disclosure is defined herein as follows: *the presence of a signal indicating the nature and extent of generative artificial intelligence’s involvement in the production of advertising or promotional content*. As a result, the framework is concerned specifically with disclosures of *generative* AI involvement rather than AI disclosures more broadly. The definition is anchored in cue availability within the message environment, such that disclosure refers to the presence of a production-related cue when audiences encounter the communication, irrespective of whether it is consciously noticed, accurately interpreted, or correctly distinguished from AI disclosures more generally. This treatment follows prior work in advertising disclosure, which distinguishes the objective presence of a disclosure cue from its subsequent recognition and processing (Boerman, Willemsen, and Van Der Aa 2017; Wojdyski and Evans 2016).

The definition is further bounded to cases in which GenAI involvement materially shapes the expressive or persuasive content of the message (ASA 2025). This aligns with emerging research on AI-generated content, which shows that audience responses are contingent on perceived authorship and the extent to

which machine involvement alters the final output (Castelo, Bos, and Lehmann 2019; Kirk and Givi 2025). GenAI involvement may range from minor assistance to fully generative production (Hanson, Carlson, and Pressler 2025; Hartmann, Exner, and Domdey 2025). The degree to which GenAI contributes to the final message is therefore likely to influence the strength, direction, and type of production-related inferences that disclosure activates (Longoni, Bonezzi, and Morewedge 2019). However, we argue that the presence of disclosure itself is not contingent on the degree of generative involvement. Rather, disclosure exists whenever production-related information concerning GenAI involvement is communicated at the point of exposure, while subsequent recognition, interpretation, and evaluation remain matters of consumer processing (Boerman, Willemsen, and Van Der Aa 2017; Wojdyski and Evans 2016).

Furthermore, GenAI disclosure is distinct from related constructs often treated interchangeably, namely transparency and awareness (Chaiken, Liberman, and Eagly 1989; Friestad and Wright 1994), and separating these constructs isolates the locus of effect across signal presence, perceptual uptake, and inferential processing. Disclosure refers to the existence of a signal indicating third-party involvement (Boerman, Willemsen, and Van Der Aa 2017), such as the use of GenAI. Transparency, by contrast, concerns the extent to which audiences perceive production processes as visible, understandable, or accountable (Kim, Barasz, and John 2019). Awareness refers to whether the disclosure cue is attended to, recognized, and correctly interpreted. In reality, a disclosure cue may be present without generating perceived transparency if it is ignored, misunderstood, or lacks credibility, and awareness, while necessary, is not always sufficient for downstream effects. Indeed, related empirical work shows that disclosure cues are frequently overlooked or misattributed, and that psychological responses arise only when such cues are processed (Luo et al. 2019).

Finally, GenAI disclosure should not be conceptualized as a unitary communicative act but as a disclosure condition that may be rendered through multiple agents. This shift requires reconsideration of *communicative agency*. Building on work in human-machine communication that recognizes technological actors as potential sources of communication (Lee 2024; Sundar 2008; Sundar and Nass 2001), we describe communicative agency herein as the locus from which audiences infer that production-related information originates and where responsibility for communicating

that information is perceived to reside. In conventional advertising disclosure, communicative agency is typically located with the advertiser or publisher responsible for the message (Boerman, Willemsen, and Van Der Aa 2017; Campbell, Mohr, and Verlegh 2013; Wojdyski and Evans 2016). In GenAI contexts, however, agency becomes distributed across brands, platforms, regulators, and technical infrastructures involved in generating, verifying, and surfacing production-related information (Lee 2024; Sundar 2008).

GenAI Disclosure Rendering Conditions

Drawing on emerging governance practices and platform implementations, we distinguish three disclosure rendering conditions (so termed because they specify how production-related information is rendered or made observable at the point of exposure): *declarative* (advertiser-controlled signaling; ASA 2025), *regulatory* (mandated disclosure obligations; European Commission 2024), and *infrastructural* (platform- or system-level disclosure embedded through technical infrastructures; Coalition for Content Provenance and Authenticity 2023; YouTube 2024). These rendering conditions specify how production-related information becomes visible to audiences and how communicative agency is distributed across the GenAI advertising ecosystem. Before discussing each in more detail, it is also worthwhile noting that their prevalence is contingent on platform governance and regulatory regimes, and any one condition may dominate, coexist with, or be absent without altering the underlying explanatory logic of the framework.

Declarative disclosure refers to signals initiated by the advertiser or content creator as part of the message itself. These disclosures are voluntarily included and are typically expressed through visible textual or verbal cues, such as statements indicating that GenAI tools were used in content creation. This rendering condition aligns with established sponsorship disclosure practices in which advertisers disclose commercial relationships through labels or statements (Boerman, Willemsen, and Van Der Aa 2017; Wojdyski and Evans 2016). Communicative agency remains primarily located with the advertiser, and the disclosure is therefore structurally attributable to brand-level communicative intent. Because declarative disclosure is discretionary, it may function as a costly signal of transparency and honesty (Giuffredi-Kähr, Petrova, and Malär 2022), while simultaneously exposing the brand to negative inferences concerning authenticity, creativity, or human effort in GenAI

contexts (Brüns and Meißner 2024; Kirkby, Baumgarth, and Henseler 2023).

Regulatory disclosure refers to signals imposed through platform rules, industry standards, or policy requirements requiring the identification of AI-generated or AI-manipulated content (e.g., ASA 2025; European Commission 2024). These disclosures are not fully controlled by advertisers at the point of rendering, even where compliance is required upstream. Instead, they are standardized in format and placement, reflecting institutional governance rather than brand-level communicative choice (Kerr et al. 2015). Contemporary examples include platform-applied labels such as “AI-generated,” “AI info,” or “altered content” implemented by Meta and YouTube (Meta 2024; YouTube 2024). Under this rendering condition, communicative agency is largely relocated from the advertiser to the institutional actors responsible for implementing disclosure. Consequently, audiences may attribute the disclosure to the platform or regulatory system rather than the advertiser, altering how responsibility, intentionality, and credibility are inferred.

Infrastructural disclosure refers to signals generated through technical systems embedded within the production and distribution pipeline. These include provenance metadata, watermarking, content credentials, and automated verification systems that identify or authenticate AI-generated content (Adobe 2025; Coalition for Content Provenance and Authenticity 2023; European Commission 2024). Unlike declarative disclosure, these signals are not appended during message composition but emerge from socio-technical infrastructures linking content generation, embedded provenance information, platform verification, and governance protocols (Coalition for Content Provenance and Authenticity 2023). Communicative agency is therefore no longer attributable to a single organizational actor but distributed across the technical architectures responsible for generating, authenticating, and surfacing production information.

Importantly, neither regulatory nor infrastructural disclosure should be assumed to provide a perfectly accurate representation of production conditions. Platform labeling systems may produce false positives or false negatives, and provenance and verification infrastructures may be incomplete, unavailable, or incorrectly interpreted (Adobe 2025; Coalition for Content Provenance and Authenticity 2023; Meta 2024). For the purposes of the present framework, however, disclosure operates as a signal at the point of exposure. Its effects therefore depend primarily on

the production-related cue encountered by audiences rather than on the objective accuracy of the underlying classification. Even when disclosure inaccurately represents production conditions, it may still activate inferences concerning authorship, oversight, provenance, authenticity, and accountability. The framework therefore focuses on the interpretive consequences of disclosure cues rather than on the technical validity of the systems generating them.

Moreover, the three rendering conditions are analytically distinct but not mutually exclusive. A single communication may simultaneously contain advertiser-declared disclosures, platform-applied labels, and technically generated provenance information (Schaefer 2024). Rather than representing alternative disclosure types, these conditions constitute distinct structural mechanisms through which production-related information becomes available to audiences. Their theoretical significance lies not in producing different disclosure effects directly, but in shaping the attributional environment within which disclosure is interpreted (Campbell and Marks 2015).

Specifically, each rendering condition distributes communicative agency differently: declarative disclosure locates agency primarily with advertisers, regulatory disclosure shifts agency toward platforms and regulators, and infrastructural disclosure distributes agency across socio-technical systems responsible for generating and verifying production information. Consequently, what varies across rendering conditions is where audiences perceive responsibility, intentionality, accountability, and credibility to reside (Schatzki 2025). The rendering conditions are therefore conceptualized as structural antecedents to attribution processes rather than independent psychological constructs (Weiner 1985). They do not require consumers to accurately identify their structural origin and are not theorized as having direct effects on evaluation. Instead, they shape the signaling environment from which perceived source attribution emerges, explaining their foundational role in the framework despite the propositions beginning at the level of consumer inference.

The Production-Based GenAI Disclosure Framework: Theory, Evidence, and Propositions

Having established the nature of GenAI disclosure and the conditions governing its rendering, we now turn to our central theoretical question: *How does such disclosure influence consumer evaluations?* In

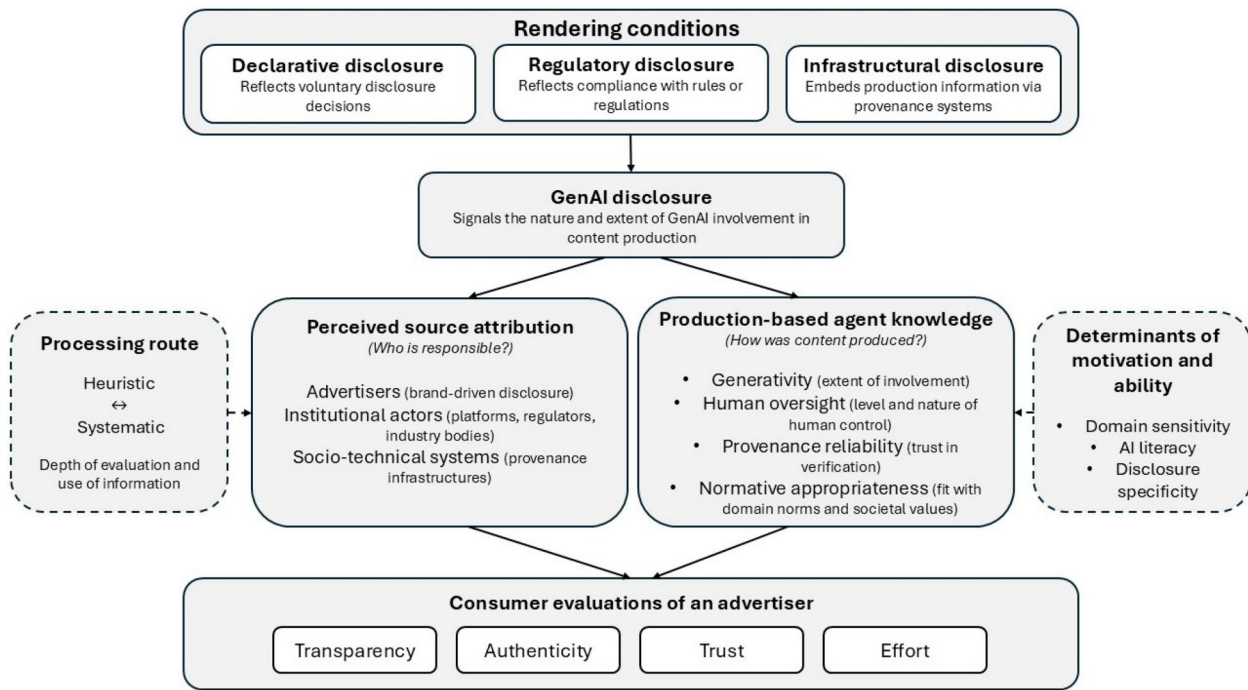


Figure 2. Production-based GenAI disclosure framework.

Figure 2, we present a production-based GenAI disclosure framework that explains this process. The framework integrates signaling theory, source credibility theory, the PKM, and the HSM into a unified explanatory framework linking disclosure as a socio-technical signal to downstream evaluations through source attribution and production-based agent knowledge. The sections that follow develop each component in turn and formalize the relationships between them through a set of propositions, and Table 2 summarizes the theoretical foundations and their role within the framework.

Signaling Production Disclosure

Signaling theory conceptualizes communication as the transmission of observable cues that reduce information asymmetry, with their informational value depending on factors such as signaling cost, credibility, intentionality, and context (Connelly et al. 2011; Spence 1973). In advertising, disclosure has traditionally been interpreted as a signal of honesty, transparency, or regulatory compliance, with its credibility depending partly on whether it is perceived as voluntary and potentially costly for the advertiser (Giuffredi-Kähr, Petrova, and Malär 2022; Kirmani and Rao 2000).

In GenAI contexts, however, disclosure cannot be understood as a single signal form because production-related information may be rendered through

distinct structural conditions that distribute communicative agency across different actors. Declarative disclosure is primarily attributable to advertisers and therefore reflects voluntary communicative intent. Regulatory disclosure is attributable to institutional governance and reflects compliance with platform rules, industry standards, or legal obligations rather than discretionary transparency (European Commission 2024; Federal Trade Commission 2009; Kerr et al. 2015). Infrastructural disclosure, in turn, is attributable primarily to socio-technical systems responsible for generating, verifying, and surfacing production information through mechanisms such as provenance metadata, content credentials, and automated verification systems (Adobe 2025; Coalition for Content Provenance and Authenticity 2023).

These rendering conditions alter the informational meaning of disclosure. Importantly, the proposed distinction does not imply that observable disclosure cues must differ across rendering conditions. Rather, visually similar disclosure cues may communicate different information because they originate from different structural conditions. Where disclosure is attributable to advertisers, it may function as a costly signal reflecting voluntary transparency and reputational commitment (Giuffredi-Kähr, Petrova, and Malär 2022; Kirmani and Rao 2000). In contrast, where disclosure is institutionally imposed, it primarily signals compliance with external governance regimes rather than voluntary communicative choice

Table 2. Theoretical foundations and proposed framework extensions.

Theory	Core assumption	Role in traditional disclosure	Extension in GenAI disclosure	Function in framework
Source credibility theory (Hovland and Weiss 1951)	Persuasion depends on the perceived trustworthiness and expertise of the source.	Source is typically treated as stable (e.g., advertiser or endorser); disclosure clarifies the commercial relationship.	GenAI disclosure weakens source stability by introducing multiple potential loci of communicative agency (advertiser, institutional actors, socio-technical systems), making source attribution an inferential process.	Mediator: Perceived source attribution determines where communicative agency, credibility, and responsibility are assigned.
Persuasion knowledge model (PKM) (Friestad and Wright 1994)	Consumers draw on knowledge about persuasion agents, persuasion tactics, and persuasion attempts to evaluate persuasion episodes.	Disclosure activates recognition of persuasive intent and associated coping responses.	GenAI disclosure extends the agent knowledge component of the PKM by making production conditions observable and activating beliefs about generativity, human oversight, provenance reliability, and normative appropriateness.	Mediator: Production-based knowledge shapes evaluations of agents beyond persuasive intent alone.
Signaling theory (Connelly et al. 2011; Spence 1973)	Signals reduce information asymmetry; their informational value depends on signaling cost, credibility, intentionality, and context.	Disclosure functions as a voluntary or mandated signal of honesty, transparency, or compliance.	The informational meaning of disclosure depends on its rendering condition (declarative, regulatory, or infrastructural), which distributes communicative agency, accountability, and signaling cost differently.	Structural antecedent: Rendering conditions shape the signaling environment within which source attribution emerges.
Heuristic-systematic model (Chaiken 1980; Chaiken and Maheswaran 1994)	Information is processed through heuristic or systematic routes depending on consumers' motivation and ability.	Disclosure often functions as a heuristic cue activating persuasion knowledge.	GenAI disclosure activates production-related heuristics or more systematic evaluation of production conditions, depending on consumers' motivation and ability.	Processing mechanism: Heuristic or systematic processing translates production-based agent knowledge into consumer evaluations of the advertiser; AI literacy, domain sensitivity, and disclosure specificity determine the predominant processing route.

(European Commission 2024; Kerr et al. 2015). Finally, where disclosure is generated through socio-technical infrastructures, it primarily signals content provenance, verification, or classification rather than discretionary advertiser intent (Adobe 2025; Coalition for Content Provenance and Authenticity 2023; Lee 2024).

Although audiences may not always accurately identify which rendering condition generated a particular disclosure cue, we suggest that these structural differences remain theoretically important because they shape the attributional environment within which consumers infer communicative agency, responsibility, credibility, and production involvement (Connelly et al. 2011; Wang, Lopez, and Okazaki 2025). As declarative, regulatory, and infrastructural disclosures increasingly coexist within the same communication environment, GenAI disclosure is better understood as a set of socio-technical signaling conditions rather than as a single, homogeneous disclosure signal. This in turn extends signaling theory by showing that the

informational meaning of GenAI disclosure depends not only on the disclosure cue itself but also on the structural rendering condition through which it is communicated, thereby shaping the inferential processes through which consumers evaluate AI-mediated communications (Connelly et al. 2011; Wang, Lopez, and Okazaki 2025).

Perceived Source Attribution under Production Ambiguity

Source credibility theory holds that persuasion depends on the perceived trustworthiness and expertise of a communicator (Hovland and Weiss 1951). In advertising, the communicator is typically treated as a stable and identifiable source, most often the brand or an endorsed spokesperson, with disclosure serving to clarify the nature of the commercial relationship (Boerman, Willemsen, and Van Der Aa 2017). Because credibility depends on the source that audiences perceive rather than on the objectively originating

source (Hovland and Weiss 1951), understanding how audiences infer the source of GenAI disclosure becomes central to explaining its persuasive effects. Under GenAI disclosure, however, this assumption of source stability is weakened. Disclosure cues foreground production processes and introduce multiple potential loci of communicative agency, including advertisers, institutional actors, and socio-technical systems, none of which is necessarily dominant at the point of exposure.

Research in computer-mediated communication demonstrates that source attribution is not fixed but is constructed through interaction with technological affordances (Sundar 2008; Sundar and Nass 2001). The MAIN model, for example, shows that modality, agency, interactivity, and navigability cues function as heuristic source cues, allowing interfaces and technological systems to be treated as communicative agents (Sundar 2008). Evidence from human–AI communication indicates that transparency and AI-related cues influence trust and evaluation (Molina and Sundar 2022). Extending this logic, we argue that disclosure cues may shape evaluations even when consumers cannot accurately identify the structural conditions through which those cues were rendered. Similarly, emerging research on GenAI disclosure demonstrates that production-related cues influence perceptions of authenticity, competence, sincerity, and trust (Schilke and Reimann 2025), suggesting that audiences form source-related inferences even under conditions of production ambiguity (Brüns and Meißner 2024; Kirkby, Baumgarth, and Henseler 2023; Wang, Lopez, and Okazaki 2025). Collectively, this evidence suggests that disclosure cues can function as source cues even when their structural origin remains uncertain or only partially understood.

Accordingly, source attribution is better conceptualized as an inferential process rather than a direct mapping between disclosure cues and objective sources. Structural origin and perceived source may diverge, with audiences attributing communicative agency to advertisers, institutional actors, or socio-technical systems irrespective of the actual rendering condition (Epley, Waytz, and Cacioppo 2007; Sundar 2008). For example, platform-generated disclosure may be interpreted as evidence of advertiser transparency, whereas advertiser-declared disclosure may be attributed to regulatory pressure or automated platform enforcement (Wojdyski and Evans 2016). Consequently, GenAI disclosure is not expected to influence consumer evaluations of the advertiser directly. Rather, its influence is mediated by perceived

source attribution, which determines whether disclosure is interpreted as primarily brand-, institution-, or system-driven.

Drawing on source credibility theory, these attributional outcomes shape perceptions of both trustworthiness and expertise (Hovland and Weiss 1951). Advertiser attribution may enhance perceptions of transparency and accountability; institutional attribution may signal independent oversight and verification capability; and attribution to socio-technical systems may increase perceptions of procedural expertise and reliability, even as it raises questions regarding authenticity and human judgment (Lee 2024). Consequently, different attributional outcomes may give rise to different evaluative responses even when consumers encounter identical disclosure cues. Integrating the preceding signaling and source attribution arguments, we propose the following:

Proposition 1: The effect of GenAI disclosure on consumer evaluations of an advertiser is mediated by perceived source attribution, such that the attribution of communicative agency to advertisers, institutional actors, or socio-technical systems influences perceptions of source trustworthiness and expertise, thereby shaping evaluations of authenticity, sincerity, and trust.

Production-Based Agent Knowledge within the PKM

The PKM explains how individuals draw on knowledge about persuasion agents, persuasion tactics, and persuasion attempts to interpret and respond to marketing communications (Friestad and Wright 1994). In sponsorship and native advertising contexts, disclosure typically supports this process by making persuasive intent explicit and enabling audiences to recalibrate trust or skepticism accordingly (Boerman, Willemsen, and Van Der Aa 2017). GenAI disclosure, however, expands the information available for evaluating persuasion agents. Rather than communicating persuasive motive alone, disclosures such as “AI-generated” foreground the conditions under which promotional content was produced by indicating that generative AI contributed to message creation.

Importantly, unlike conventional production tools such as Photoshop, for which editing generally remains attributable to a human creator (Castelo, Bos, and Lehmann 2019; Longoni, Bonezzi, and Morewedge 2019), GenAI disclosure can introduce ambiguity concerning authorship and the respective roles of humans and computational systems in content generation (Molina and Sundar 2022; Sundar

2008). Consequently, consumers evaluate not only why a persuasion attempt is made but also how persuasive content was produced and what that production implies for evaluating the persuasion agent (Zhang and Hur 2025).

For instance, emerging empirical evidence demonstrates that these production-related evaluations operate even when persuasive intent is already recognized (Grigsby, Michelsen, and Zamudio 2025). Disclosure of AI use can increase perceived transparency while simultaneously increasing skepticism or perceived manipulateness (Baek, Kim, and Kim 2026); amplify skepticism without increasing perceived deception (Wortel, Vanwesenbeeck, and Tomas 2024); and reduce perceived creativity through inferences about diminished human effort (De Freitas, Nave, and Puntoni 2025; Zhang and Hur 2025). In specific contexts, AI disclosure has also been shown to evoke concerns regarding authenticity and moral appropriateness (Kirk and Givi 2025; Schilke and Reimann 2025). Related evidence further demonstrates that AI-generated content disclosure can reduce consumer engagement by weakening parasocial relationships through reduced perceptions of human effort, even after controlling for content quality and general AI aversion (Carney, Riveros, and Tully 2026). Collectively, these findings suggest that consumers evaluate persuasion agents not only in terms of persuasive intent but also in terms of the production processes underlying persuasive communication.

These findings raise an important theoretical question concerning the scope of agent knowledge within the PKM. The model traditionally emphasizes consumers' beliefs about persuasion agents, including their motives, capabilities, strategies, and intentions (Friestad and Wright 1994). GenAI disclosure, however, makes visible an additional set of agent-related beliefs concerning how advertising content was produced. We therefore argue that GenAI disclosure activates *production-based agent knowledge*, defined as consumers' beliefs about how persuasive content was generated and what those production processes imply for evaluating the persuasion agent. Rather than replacing beliefs about agent motives and intentions, production-based agent knowledge extends the agent knowledge component of the PKM (Friestad and Wright 1994) by incorporating production-related beliefs that become salient when GenAI disclosure renders production processes observable.

Production-based agent knowledge encompasses four interrelated dimensions: the degree of generativity involved in producing the message, the level of

human oversight exercised during content creation, the reliability of provenance and verification systems, and the normative appropriateness of AI use within a given context. These beliefs may inform broader evaluations of persuasion agents by shaping perceptions of authenticity, effort, legitimacy, accountability, and fairness (Koning and Voorveld 2025; Voorveld, Meppelink, and Boerman 2024). Although production-related beliefs may inform judgements about agent motives and persuasive intent, they are not reducible to motive attribution alone. Instead, they broaden the agent knowledge consumers draw upon when evaluating persuasive communications involving GenAI.

Thus, we expect production-based agent knowledge to become particularly influential when persuasive intent is already explicit, held constant, or weakly diagnostic, such that variation in evaluation is more likely to arise from production-related beliefs than from differences in motive attribution. This expectation is consistent with prior research demonstrating that AI disclosure can alter consumer evaluations without corresponding changes in perceived persuasive intent (Baek, Kim, and Kim 2026; Wortel, Vanwesenbeeck, and Tomas 2024), suggesting that production-related beliefs contribute to persuasion outcomes beyond traditional intent-based evaluations (Friestad and Wright 1994; Zhang and Hur 2025). We therefore suggest:

Proposition 2: GenAI disclosure activates production-based agent knowledge concerning generativity, human oversight, provenance reliability, and normative appropriateness, which mediates the effect of disclosure on consumer evaluations of the advertiser, including perceptions of authenticity, creativity, and trust.

Processing Route: Heuristic Cues and Systematic Evaluation

The HSM distinguishes between two modes of information processing: heuristic processing, which relies on readily accessible heuristics under relatively low cognitive effort, and systematic processing, which involves more effortful evaluation of available information depending on individuals' motivation and ability (Chaiken 1980; Chaiken and Maheswaran 1994). This distinction also requires separating heuristics from heuristic processing. Rather than representing a processing mode, heuristics are learned decision rules or cognitive shortcuts that may be relied upon under heuristic processing or critically evaluated during systematic processing (Gigerenzer and Gaissmaier

2011; Todorov, Chaiken, and Henderson 2002; West, Acar, and Caruana 2020). In the context of GenAI disclosure, the presence of a cue such as an “AI-generated” label does not require consumers to identify the actor or system responsible for rendering the disclosure in order to influence evaluation. Instead, it activates readily accessible production-related heuristics concerning machine involvement, human effort, authorship, authenticity, and oversight. Consistent with the HSM, such heuristics may serve as direct judgment shortcuts under heuristic processing or be evaluated alongside other information under systematic processing (Todorov, Chaiken, and Henderson 2002).

Consistent with the HAI-TIME framework (Sundar 2020), disclosure functions as an AI cue that activates machine heuristics, triggering learned associations about automation, authorship, and creative effort. These associations operate as direct cue-judgment mappings. For instance, disclosure may prompt inferences of reduced human effort, which in turn lowers perceived creativity, or it may signal automation, which diminishes perceived authenticity (Carney, Riveros, and Tully 2026). Platform-applied labels can similarly function as cues that content has been subjected to platform governance or rule-based classification. Rather than requiring detailed evaluation of the underlying production or verification process, such cues may shape judgements through heuristic inference, consistent with evidence that transparency cues can influence trust and evaluation in human-AI interaction contexts (Molina and Sundar 2022). In this way, disclosure functions as a cognitively efficient input into evaluation, and its effects do not depend on accurate identification of its structural origin.

When consumers engage in more systematic processing, however, disclosure cues are no longer treated as sufficient evidence in themselves (Chaiken and Maheswaran 1994). Instead, these cues direct attention toward production information and encourage more deliberate evaluation of its implications. Judgements become contingent on contextual considerations, including normative expectations within the domain, the perceived appropriateness of AI use, and the extent of human oversight. In domains where human input is considered central, such as creative or moral appeals, AI involvement may reduce perceived authenticity or sincerity (Arango, Singaraju, and Niininen 2023; Kirk and Givi 2025; Wu, Dodoo, and Wen 2025). Conversely, in contexts where technological competence is valued, the same disclosure cue may enhance perceptions of innovation or capability.

Consumers may also distinguish between GenAI used for augmentation and GenAI used for substitution, evaluating the former more favorably when meaningful human involvement remains evident (Voorveld, Meppelink, and Boerman 2024). Thus, GenAI disclosure increases the salience of production-related information, but its evaluative consequences depend on whether the resulting source attributions and production-based beliefs are processed through rapid heuristic associations or through more effortful, context-sensitive reasoning. In the following section, we discuss the key factors determining which processing route predominates.

Motivation and Ability as Determinants of Processing

Consistent with the HSM, whether consumers rely primarily on heuristic or systematic processing depends on their motivation and ability to evaluate available information (Chaiken 1980; Chaiken and Maheswaran 1994). Within the proposed framework, we put forward that *domain sensitivity*, *AI literacy*, and *disclosure specificity* jointly determine these processing conditions and, consequently, how GenAI disclosure influences evaluation. Domain sensitivity reflects the extent to which a context places value on authenticity, human authorship, and process visibility rather than performance or efficiency alone (Mousavi and Villanova 2026; Napoli et al. 2014). In highly sensitive domains, such as creative production, luxury, journalism, or personalized communication, production information is evaluatively diagnostic because authenticity and human effort constitute central judgment criteria. In less sensitive domains, including routine information services or functional outputs, evaluations are less contingent on how content was produced.

Consumers' AI literacy reflects their understanding of GenAI capabilities, limitations, and appropriate use and shapes how production-related information is interpreted (Long and Magerko 2020; Tully, Longoni, and Appel 2025). Lower AI literacy encourages reliance on heuristics concerning automation, effort, and risk (Chaiken 1980; Sundar 2008), whereas higher AI literacy supports more differentiated evaluation of production conditions, including distinctions between levels of human oversight, forms of AI augmentation, and degrees of generativity (Castelo, Bos, and Lehmann 2019; Longoni, Bonezzi, and Morewedge 2019). Consequently, identical disclosure cues are more likely to produce simplified or categorical

evaluations among consumers with limited AI literacy and more contingent, context-sensitive evaluations among those with greater literacy.

Disclosure specificity determines the amount of diagnostic information available for evaluation. General disclosures, such as “AI-generated,” provide relatively limited information and therefore encourage reliance on broad heuristic associations (Boerman, Willemsen, and Van Der Aa 2017; Wojdyski and Evans 2016). More specific disclosures explaining the nature of AI involvement, particularly the extent of human oversight or the role of GenAI within the production process, provide richer diagnostic information that supports more systematic evaluation. Such disclosures can shift interpretation from AI substitution toward AI augmentation, thereby altering downstream evaluations (Longoni, Bonezzi, and Morewedge 2019; Voorveld, Meppelink, and Boerman 2024).

Together, these factors determine consumers’ motivation and ability to evaluate production-related information and, consequently, whether heuristic or systematic processing is more likely to predominate. Domain sensitivity increases motivation by determining the evaluative importance of authenticity, human authorship, and production processes. Because motivation and ability are distinct determinants of processing, consumers may be highly motivated yet lack the knowledge required to interpret production-related information, or possess substantial AI knowledge while regarding such information as relatively unimportant for evaluation. Consumers’ AI literacy increases their ability to interpret production-related information, whereas disclosure specificity determines whether sufficient diagnostic information is available to support systematic evaluation or whether consumers rely primarily on heuristic inference (Chaiken and Maheswaran 1994; Sundar 2008). As a result, we expect the following:

Proposition 3a: Higher domain sensitivity, greater AI literacy, and higher disclosure specificity increase systematic processing of production-related information.

Proposition 3b: Lower domain sensitivity, lower AI literacy, and lower disclosure specificity increase reliance on heuristic processing.

Taken together, the proposed framework explains why GenAI disclosure often produces inconsistent and seemingly contradictory effects. Existing disclosure theories largely assume that disclosure operates by making persuasive intent visible, leading to relatively predictable attributional and evaluative responses

(Boerman, Willemsen, and Van Der Aa 2017; Friestad and Wright 1994). GenAI disclosure differs because it renders production conditions visible, introducing an additional basis for evaluation alongside persuasive intent. Consequently, identical disclosure cues may generate divergent responses depending on how communicative agency is attributed, what production-related beliefs become salient, and whether those beliefs are processed heuristically or systematically.

The framework therefore reframes disclosure not simply as a mechanism for reducing information asymmetry, but as a production-related signal capable of activating multiple evaluative pathways simultaneously. In doing so, it helps explain why GenAI disclosure may increase perceptions of transparency, accountability, or honesty while simultaneously diminishing authenticity, creativity, relational warmth, or perceived human effort (Baek, Kim, and Kim 2026; Kirk and Givi 2025; Wortel, Vanwesenbeeck, and Tomas 2024; Zhang and Hur 2025). Rather than producing uniformly positive or negative effects, disclosure redistributes attention across multiple evaluative criteria, including trustworthiness, authenticity, creativity, fairness, effort, and accountability. This provides a theoretical explanation for the mixed empirical findings that characterize the emerging GenAI disclosure literature and suggests that disclosure effects are inherently multidimensional rather than unidirectional.

Implications for Advertising Theory and Practice

Theoretical Contributions

This article advances advertising theory by defining GenAI disclosure and showing that GenAI disclosure shifts the primary referent of disclosure from persuasive intent to production conditions. Prior work has conceptualized advertising disclosure primarily as a mechanism for making persuasive intent visible, with disclosure functioning to signal commercial relationships and activate coping mechanisms within the PKM (Boerman, Willemsen, and Van Der Aa 2017; Friestad and Wright 1994; Wojdyski and Evans 2016). In GenAI contexts, however, disclosure increasingly communicates how content was generated, including the degree of machine involvement, human oversight, and content provenance. Emerging empirical work demonstrates that GenAI disclosure affects perceptions of authenticity, creativity, effort, and fairness even when persuasive intent is already recognized (Baek, Kim, and Kim 2026; Kirk and Givi 2025;

Zhang and Hur 2025). These effects are not fully explained by motive recognition alone, indicating that production conditions constitute a distinct object of consumer evaluation.

Building on this shift, the article extends the PKM by conceptualizing production-based agent knowledge as an extension of the agent knowledge component of the model, operating together with traditional intention-based agent inferences (Friestad and Wright 1994). Existing applications of the PKM primarily explain how consumers interpret the intentions and tactics of persuasion agents. In contrast, GenAI disclosure activates beliefs about the conditions under which content is produced, including perceptions of generativity, human oversight, provenance reliability, and the normative appropriateness of AI use within a given context. Research showing that consumers evaluate AI-generated outputs in terms of effort, authenticity, and moral acceptability independently of persuasive intent supports this extension (Castelo, Bos, and Lehmann 2019; Kirk and Givi 2025; Longoni, Bonezzi, and Morewedge 2019). The framework therefore clarifies that production-based agent knowledge is not simply a proxy for motive attribution but a distinct evaluative mechanism that becomes salient when production conditions are rendered visible.

The article also extends signaling theory by conceptualizing disclosure as a structurally heterogeneous and multi-origin signal rather than a uniform communicative cue. Disclosure may originate from multiple loci of control, including brands, regulators, platforms, and technical infrastructures, each of which shapes perceived intentionality, accountability, credibility, and signaling cost differently (Giuffredi-Kähr, Petrova, and Malär 2022; Kerr et al. 2015; Spence 1973). Most notably, infrastructural disclosure represents a conceptual shift in which production-related information becomes embedded within the technical substrate of content through metadata, provenance credentials, and automated verification systems (Adobe 2025; Coalition for Content Provenance and Authenticity 2023). Unlike conventional disclosure, the visibility of production information is no longer dependent solely on an advertiser's communicative choice but may persist through technical infrastructures embedded within the content itself. Building on research demonstrating that technological systems can function as communicative actors in digital environments (Molina and Sundar 2022; Sundar 2008), we extend this perspective by arguing that infrastructural disclosure redistributes

perceived communicative agency across advertisers, platforms, regulators, and socio-technical systems.

Finally, the article proposes a unified explanatory framework in which GenAI disclosure influences evaluation through two complementary inferential pathways: perceived source attribution and production-based agent knowledge. Drawing on source credibility theory, the PKM, signaling theory, and the HSM (Chaiken and Maheswaran 1994; Friestad and Wright 1994; Hovland and Weiss 1951; Spence 1973), disclosure is modeled as a socio-technical cue whose effects depend on how communicative agency is attributed, what production-related beliefs become salient, and whether those beliefs are processed heuristically or systematically. Because communicative agency is distributed across human and technical actors, disclosure effects do not depend on precise source identification; even unresolved attribution can shape evaluative responses (Sundar 2008). The framework therefore provides a mechanism for explaining why identical GenAI disclosure cues may produce divergent outcomes across consumers, contexts, and evaluative dimensions, helping reconcile the mixed findings emerging within the GenAI disclosure literature.

Practical Implications

Our framework suggests that GenAI disclosure should not be treated as a simple question of whether to disclose AI involvement. Its effects depend on how disclosure is rendered, who is perceived to be responsible for it, and what audiences infer from it. In practice, disclosure is not a single design choice, but a condition shaped by multiple actors and interpreted in different ways. Table 3 summarizes key managerial implications of the framework.

For advertisers, the framework suggests that disclosure strategy cannot be separated from production strategy. When brands voluntarily disclose GenAI use, they make visible a production decision that may simultaneously increase perceptions of transparency while reducing perceptions of authenticity, creativity, or human effort. Declarative disclosure therefore presents both opportunity and risk. Rather than viewing disclosure as inherently beneficial or detrimental, firms should frame GenAI as augmenting rather than replacing human expertise wherever this framing accurately reflects production practice, while making meaningful human oversight visible where it exists. Disclosure strategy should therefore be aligned with broader brand positioning rather than implemented as a uniform best practice.

Table 3. Practical implications of the proposed framework.

Stakeholder	Framework insight	Implication for practice
Advertisers	Declarative disclosure attributes communicative agency to the brand and makes production decisions visible.	Frame GenAI as augmentation rather than substitution where appropriate; communicate meaningful human oversight; align disclosure with brand positioning and category expectations.
Platforms	Regulatory disclosure attributes communicative agency to institutional governance and standardizes disclosure practices.	Design disclosure labels that are consistent, prominent, and sufficiently informative; recognize that disclosure design influences consumer attribution and evaluation, not simply regulatory compliance.
Regulators	Disclosure effectiveness depends on interpretation rather than visibility alone.	Evaluate disclosure policies according to behavioral outcomes as well as compliance; recognize that identical disclosure labels may produce different consumer responses across contexts.
Infrastructure providers	Infrastructural disclosure renders production-related information visible through provenance and verification technologies.	Develop provenance systems that produce meaningful and interpretable production information; recognize that verification infrastructures function as communication mechanisms as well as technical standards.

The framework further indicates that managerial control over disclosure is increasingly constrained by platform governance and regulatory requirements. Platform-applied disclosures attribute communicative agency primarily to institutional actors rather than advertisers and may therefore influence evaluations independently of brand intentions. Organizations should anticipate platform-generated disclosure, understand evolving regulatory requirements, and ensure that creative executions remain appropriate under likely disclosure conditions rather than assuming that messaging alone determines consumer responses.

Beyond advertiser- or platform-generated disclosure, production information is increasingly communicated through technical infrastructures such as provenance credentials, content credentials, watermarking, and verification systems. As these infrastructures become more widely adopted, production-related information becomes increasingly embedded within digital content and less dependent on voluntary disclosure alone. Organizations should therefore assume that production-related information may become observable regardless of whether they voluntarily disclose GenAI use and ensure that production workflows, documentation, and human oversight practices are consistent with the information that technical verification systems may reveal.

The framework also has implications for platform designers and regulators. Disclosure labels do more than satisfy disclosure and governance requirements; they shape source attribution, activate production-related beliefs, and influence subsequent evaluation. Consequently, disclosure policies should be designed not only to maximize visibility or regulatory compliance but also to facilitate accurate consumer interpretation. Similarly, developers of provenance infrastructures should recognize that technical verification systems function simultaneously as

communication systems. The way production information is generated, verified, and surfaced influences how communicative agency is attributed and how brands are subsequently evaluated.

Finally, the framework suggests that GenAI disclosure is unlikely to produce uniformly positive or negative effects. Instead, its consequences depend on how production information is rendered, the attributional inferences consumers draw, the production-related beliefs that are activated, and the extent to which information is processed heuristically or systematically. Disclosure decisions should therefore be aligned with category norms, audience characteristics, and strategic positioning rather than treated as universal best practice. More broadly, organizations should integrate disclosure strategy with production governance, platform requirements, and provenance management, recognizing that meaningful transparency increasingly depends on the coordination of communication, technology, and governance rather than on disclosure cues alone.

Beyond effectiveness considerations, the framework highlights that production disclosure raises ethical questions distinct from those traditionally associated with persuasive intent disclosure. If disclosure increasingly communicates how content was created rather than solely why it was created, organizations face ethical responsibilities concerning the transparency of production practices, human oversight, and the appropriate use of generative systems. As a result, disclosure decisions should be evaluated not only in terms of their impact on trust or persuasion but also in terms of accountability and consumers' ability to make informed judgements about AI-mediated communications.

Conclusion

This paper reconceptualizes disclosure in AI-mediated advertising by shifting the analytical focus from persuasive intent to production conditions. In doing so,

we develop a framework that explains how GenAI disclosure influences consumer evaluations through source attribution, production-based agent knowledge, and the processing mechanisms through which production-related information is interpreted. While we have sought to define these concepts and relationships as clearly and precisely as possible, the conceptual work presented herein is subject to the following few limitations.

The distinction between rendering conditions is specified analytically and may not be perceptually separable in practice because audiences may not reliably identify or attend to their structural origin (Lee 2024). In addition, the framework is developed in the context of current platform and regulatory conditions, which are evolving rapidly and may alter how disclosure is implemented and interpreted over time. Because the framework is conditional on cue presence rather than on specific institutional arrangements, its empirical salience will vary with platform governance and regulatory regimes. Changes in these regimes may increase, reduce, or eliminate observable disclosure, thereby affecting when and how production-based agent knowledge becomes salient without rendering the underlying process obsolete. Moreover, infrastructural disclosure mechanisms depend on system adoption and integrity and may be incomplete or circumvented (Rough and Clift 2026), potentially affecting the availability and interpretation of production-related information. The effects specified are also contingent on processing route and may differ depending on whether disclosure is interpreted through heuristic associations or more systematic evaluation of production conditions.

These limitations point to several directions for future research. Empirical work is needed to test the proposed relationships between disclosure rendering conditions, source attribution, production-based agent knowledge, and consumer evaluations, including the extent to which audiences distinguish or conflate different forms of disclosure. Research should also examine how processing varies under different levels of attention and involvement, and how production-based agent knowledge interacts with established forms of persuasion knowledge. Lastly, further work is needed to explore boundary conditions (e.g., Arango, Singaraju, and Niininen 2023; Chen, Wang, and Hao 2025), including variation across product categories, cultural contexts, and platform environments, as well as the role of emerging provenance and labeling systems in shaping attribution, trust, and evaluation.

As GenAI becomes embedded in everyday marketing practice, the central challenge will no longer be whether AI use is disclosed, but how production transparency can be implemented in ways that support understanding, accountability, and trust.

Acknowledgements

The authors thank the three anonymous reviewers for their thoughtful and constructive feedback, which helped shape and strengthen the manuscript. We also thank the Guest Editors, Colin Campbell and Jisu Huh, for a rigorous and constructive revision process, and for their flexibility throughout.

Author Contributions

CRedit: **Denitsa Dineva**: Conceptualization, Data curation, Investigation, Methodology, Resources, Writing – original draft, Writing – review & editing; **Oguz A. Acar**: Conceptualization, Resources, Writing – original draft, Writing – review & editing.

Declaration of Generative AI Use

In the spirit of the phenomenon examined in this paper, the authors disclose that generative AI tools were used during manuscript preparation for copyediting and to generate visual mockups of the conceptual framework. These outputs were subsequently reviewed and revised by the authors, who take full responsibility for the final content of the manuscript.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

References

- Adobe. 2025. "Content Credentials Overview." <https://helpx.adobe.com/creative-cloud/apps/adobe-content-authenticity/content-credentials/overview.html>.
- Arango, L., S. P. Singaraju, and O. Niininen. 2023. "Consumer Responses to AI-Generated Charitable Giving Ads." *Journal of Advertising* 52 (4): 486–503. <https://doi.org/10.1080/00913367.2023.2183285>.
- ASA. 2025. "Disclosure of AI in Advertising: Striking the Balance Between Creativity and Responsibility." <https://www.asa.org.uk/news/disclosure-of-ai-in-advertising-striking-the-balance-between-creativity-and-responsibility.html>.
- Baek, T. H., J. Kim, and J. H. Kim. 2026. "Effect of Disclosing AI-Generated Content on Prosocial Advertising Evaluation." *International Journal of Advertising* 45 (1): 171–192. <https://doi.org/10.1080/02650487.2024.2401319>.

- BBC. 2026. "AI 'Slop' Is Transforming Social Media—and a Backlash Is Brewing." <https://www.bbc.co.uk/news/articles/c9wx2dz2v44o>.
- Boerman, S. C., L. M. Willemsen, and E. P. Van Der Aa. 2017. "'This Post Is Sponsored' Effects of Sponsorship Disclosure on Persuasion Knowledge and Electronic Word of Mouth in the Context of Facebook." *Journal of Interactive Marketing* 38 (1): 82–92. <https://doi.org/10.1016/j.intmar.2016.12.002>.
- Brüns, J. D., and M. Meißner. 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>.
- Campbell, C., and L. J. Marks. 2015. "Good Native Advertising Isn't a Secret." *Business Horizons* 58 (6): 599–606. <https://doi.org/10.1016/j.bushor.2015.06.003>.
- Campbell, C., K. Plangger, S. Sands, and J. Kietzmann. 2022. "Preparing for an Era of Deepfakes and AI-Generated Ads: A Framework for Understanding Responses to Manipulated Advertising." *Journal of Advertising* 51 (1): 22–38. <https://doi.org/10.1080/00913367.2021.1909515>.
- Campbell, M. C., G. S. Mohr, and P. W. Verlegh. 2013. "Can Disclosures Lead Consumers to Resist Covert Persuasion? The Important Roles of Disclosure Timing and Type of Response." *Journal of Consumer Psychology* 23 (4): 483–495. <https://doi.org/10.1016/j.jcps.2012.10.012>.
- Carney, S., I. Riveros, and S. Tully. 2026. "Made with AI: Consumer Engagement with Social Media Containing AI Disclosures." *Journal of Consumer Research: ucag013*. <https://doi.org/10.1093/jcr/ucag013>.
- Castelo, N., M. W. Bos, and D. R. Lehmann. 2019. "Task-Dependent Algorithm Aversion." *Journal of Marketing Research* 56 (5): 809–825. <https://doi.org/10.1177/0022243719851788>.
- Chaiken, S. 1980. "Heuristic versus Systematic Information Processing and the Use of Source versus Message Cues in Persuasion." *Journal of Personality and Social Psychology* 39 (5): 752–766. <https://doi.org/10.1037/0022-3514.39.5.752>.
- Chaiken, S., A. Liberman, and A. H. Eagly. 1989. "Heuristic and Systematic Information Processing within and Beyond the Persuasion Context." In *Unintended Thought*, edited by J. S. Uleman and J. A. Bargh, 212–252. New York: Guilford Press.
- Chaiken, S., and D. Maheswaran. 1994. "Heuristic Processing Can Bias Systematic Processing: Effects of Source Credibility, Argument Ambiguity, and Task Importance on Attitude Judgment." *Journal of Personality and Social Psychology* 66 (3): 460–473. <https://doi.org/10.1037/0022-3514.66.3.460>.
- Chen, H., P. Wang, and S. Hao. 2025. "AI in the Spotlight: The Impact of Artificial Intelligence Disclosure on User Engagement in Short-Form Videos." *Computers in Human Behavior* 162: 108448. <https://doi.org/10.1016/j.chb.2024.108448>.
- Coalition for Content Provenance and Authenticity. 2023. "C2PA Specification (Content Provenance and Authenticity)." <https://c2pa.org>.
- Connelly, B. L., S. T. Certo, R. D. Ireland, and C. R. Reutzel. 2011. "Signaling Theory: A Review and Assessment." *Journal of Management* 37 (1): 39–67. <https://doi.org/10.1177/0149206310388419>.
- De Freitas, J., G. Nave, and S. Puntoni. 2025. "Ideation with Generative AI—in Consumer Research and beyond." *Journal of Consumer Research* 52 (1): 18–31. <https://doi.org/10.1093/jcr/ucaf012>.
- European Commission. 2024. "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence and Amending." <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>.
- Evans, N. J., J. Phua, J. Lim, and H. Jun. 2017. "Disclosing Instagram Influencer Advertising: The Effects of Disclosure Language on Advertising Recognition, Attitudes, and Behavioral Intent." *Journal of Interactive Advertising* 17 (2): 138–149. <https://doi.org/10.1080/15252019.2017.1366885>.
- Epley, N., A. Waytz, and J. T. Cacioppo. 2007. "On Seeing Human: A Three-Factor Theory of Anthropomorphism." *Psychological Review* 114 (4): 864–886. <https://doi.org/10.1037/0033-295X.114.4.864>.
- Federal Trade Commission. 2009. "FTC Publishes Final Guides Governing Endorsements." Testimonials. <https://www.ftc.gov/sites/default/files/attachments/press-releases/ftc-publishes-final-guides-governing-endorsements-testimonials/091005revisedendorsementguides.pdf>.
- Friestad, M., and P. Wright. 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>.
- Gigerenzer, G., and W. Gaissmaier. 2011. "Heuristic Decision Making." *Annual Review of Psychology* 62: 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346>.
- Giuffredi-Kähr, A., A. Petrova, and L. Malär. 2022. "Sponsorship Disclosure of Influencers—a Curse or a Blessing?" *Journal of Interactive Marketing* 57 (1): 18–34. <https://doi.org/10.1177/10949968221075686>.
- Grigsby, J. L., M. Michelsen, and C. Zamudio. 2025. "Service Ads in the Era of Generative AI: Disclosures, Trust, and Intangibility." *Journal of Retailing and Consumer Services* 84: 104231. <https://doi.org/10.1016/j.jretconser.2025.104231>.
- Hanson, S., J. Carlson, and H. Pressler. 2025. "The Differential Impact of AI Salience on Advertising Engagement and Attitude: Scary Good AI Advertising." *Journal of Advertising Research* 65 (2): 190–201. <https://doi.org/10.1080/00218499.2025.2464307>.
- Hartmann, J., Y. Exner, and S. Domdey. 2025. "The Power of Generative Marketing: Can Generative AI Create Superhuman Visual Marketing Content?" *International Journal of Research in Marketing* 42 (1): 13–31. <https://doi.org/10.1016/j.ijresmar.2024.09.002>.
- Hovland, C. I., and W. Weiss. 1951. "The Influence of Source Credibility on Communication Effectiveness." *Public Opinion Quarterly* 15 (4): 635–650. <https://doi.org/10.1086/266350>.
- Huh, J., M. R. Nelson, and C. A. Russell. 2023. "ChatGPT, AI Advertising, and Advertising Research and

- Education.” *Journal of Advertising* 52 (4): 477–482. <https://doi.org/10.1080/00913367.2023.2227013>.
- Kerr, G., D. E. Schultz, P. J. Kitchen, F. J. Mulhern, and P. Beede. 2015. “Does Traditional Advertising Theory Apply to the Digital World? A Replication Analysis Questions the Relevance of the Elaboration Likelihood Model.” *Journal of Advertising Research* 55 (4): 390–400. <https://doi.org/10.2501/JAR-2015-001>.
- Kim, T., K. Barasz, and L. K. John. 2019. “Why am I Seeing This Ad? The Effect of Ad Transparency on Ad Effectiveness.” *Journal of Consumer Research* 45 (5): 906–932. <https://doi.org/10.1093/jcr/ucy039>.
- Kirmani, A., and A. R. Rao. 2000. “No Pain, no Gain: A Critical Review of the Literature on Signaling Unobservable Product Quality.” *Journal of Marketing* 64 (2): 66–79. <https://doi.org/10.1509/jmkg.64.2.66.18000>.
- Kirk, C. P., and J. Givi. 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>.
- Kirkby, A., C. Baumgarth, and J. Henseler. 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 (7): 1108–1122. <https://doi.org/10.1108/JBPM-02-2022-3864>.
- Koning, B., and H. A. Voorveld. 2025. “Disclaimer! This Content Is AI-Generated: How AI-Disclosures Influence Trust in Advertisements and Organizations.” *Journal of Interactive Advertising* 25 (3): 240–253. <https://doi.org/10.1080/15252019.2025.2554149>.
- Lee, E. J. 2024. “Minding the Source: Toward an Integrative Theory of Human–Machine Communication.” *Human Communication Research* 50 (2): 184–193. <https://doi.org/10.1093/hcr/hqad034>.
- Long, D., and B. Magerko. 2020. “What Is AI Literacy? Competencies and Design Considerations.” In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, April, 1–16.
- Longoni, C., A. Bonezzi, and C. K. Morewedge. 2019. “Resistance to Medical Artificial Intelligence.” *Journal of Consumer Research* 46 (4): 629–650. <https://doi.org/10.1093/jcr/ucz013>.
- Luo, X., S. Tong, Z. Fang, and Z. Qu. 2019. “Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases.” *Marketing Science* 38 (6): 937–947. <https://doi.org/10.1287/mksc.2019.1192>.
- Meta. 2024. “Our Approach to Labeling AI-Generated Content and Manipulated Media.” <https://about.fb.com/news/2024/04/metasp-approach-to-labeling-ai-generated-content-and-manipulated-media/>.
- Meta. 2025. “Expanding GenAI Transparency for Meta’s Ads Products.” <https://about.fb.com/news/2025/02/gen-ai-transparency-metas-ads-products/>.
- Molina, M. D., and S. S. Sundar. 2022. “When AI Moderates Online Content: Effects of Human Collaboration and Interactive Transparency on User Trust.” *Journal of Computer-Mediated Communication* 27 (4): zmac010. <https://doi.org/10.1093/jcmc/zmac010>.
- Montecchi, M., K. Plangger, D. West, and K. de Ruyter. 2024. “Perceived Brand Transparency: A Conceptualization and Measurement Scale.” *Psychology & Marketing* 41 (10): 2274–2297. <https://doi.org/10.1002/mar.22048>.
- Mousavi, J., and D. Villanova. 2026. “Consumer Evaluations of Generative Artificial Intelligence Disclosures in Responses to Health Care Reviews.” *Journal of Public Policy & Marketing* 45 (1): 1–17. <https://doi.org/10.1177/07439156251367762>.
- Napoli, J., S. J. Dickinson, M. B. Beverland, and F. Farrelly. 2014. “Measuring Consumer-Based Brand Authenticity.” *Journal of Business Research* 67 (6): 1090–1098. <https://doi.org/10.1016/j.jbusres.2013.06.001>.
- Rough, E., and J. Clift. 2026. “AI Content Labelling.” *Research Briefing*. <https://researchbriefings.files.parliament.uk/documents/CBP-10467/CBP-10467.pdf>.
- Schaefer, A. 2024. “AI Disclaimers: What Marketing Leaders Need to Know.” *SproutSocial*. <https://sproutsocial.com/insights/ai-disclaimer/>.
- Schatzki, T. R. 2025. “Agency.” *Information and Organization* 35 (1): 100553. <https://doi.org/10.1016/j.infoandorg.2024.100553>.
- Schilke, O., and M. Reimann. 2025. “The Transparency Dilemma: How AI Disclosure Erodes Trust.” *Organizational Behavior and Human Decision Processes* 188: 104405. <https://doi.org/10.1016/j.obhdp.2025.104405>.
- Shahidi, R. 2026. “Gucci’s New AI-Generated Ads Have Hit a Nerve Online.” *Business Insider*. <https://www.businessinsider.com/gucci-ai-ads-backlash-branding-experts-culturally-relevant-2026-2>.
- Shi, Y., and Z. Jiang. 2026. “Consumer Responses to AI Disclosure Labels: The Role of Novelty and Authenticity.” *Sage Open* 16 (1): 21582440261417793. <https://doi.org/10.1177/21582440261417793>.
- Spence, M. 1973. “Job Market.” *The Quarterly Journal of Economics* 87 (3): 355–374.
- Sundar, S. S. 2008. *The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility*, 73–100. Cambridge, MA: MacArthur Foundation Digital Media and Learning Initiative.
- Sundar, S. S. 2020. “Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction (HAI).” *Journal of Computer-Mediated Communication* 25 (1): 74–88. <https://doi.org/10.1093/jcmc/zmz026>.
- Sundar, S. S., and C. Nass. 2001. “Conceptualizing Sources in Online News.” *Journal of Communication* 51 (1): 52–72. <https://doi.org/10.1111/j.1460-2466.2001.tb02872.x>.
- To, R. N., Y. C. Wu, P. Kianian, and Z. Zhang. 2025. “When AI Doesn’t Sell Prada: Why Using AI-Generated Advertisements Backfires for Luxury Brands.” *Journal of Advertising Research* 65 (2): 202–236. <https://doi.org/10.1080/00218499.2025.2454120>.
- Todorov, A., S. Chaiken, and M. D. Henderson. 2002. “The Heuristic-Systematic Model of Social Information Processing.” In *The Persuasion Handbook: Developments in Theory and Practice*, 23, 195–212. Thousand Oaks, CA: SAGE Publications, Inc.
- Tully, S. M., C. Longoni, and G. Appel. 2025. “Lower Artificial Intelligence Literacy Predicts Greater AI Receptivity.” *Journal of Marketing* 89 (5): 1–20. <https://doi.org/10.1177/00222429251314491>.

- Voorveld, H. A., C. S. Meppelink, and S. C. Boerman. 2024. "Consumers' Persuasion Knowledge of Algorithms in Social Media Advertising: Identifying Consumer Groups Based on Awareness, Appropriateness, and Coping Ability." *International Journal of Advertising* 43 (6): 960–986. <https://doi.org/10.1080/02650487.2023.2264045>.
- Wang, F., C. Lopez, and S. Okazaki. 2025. "Signaling Transparency in the Era of Artificial Intelligence." *Internet Research* 35 (2): 1–25. <https://doi.org/10.1108/INTR-11-2023-1041>.
- Weiner, B. 1985. "An Attributional Theory of Achievement Motivation and Emotion." *Psychological Review* 92 (4): 548–573. <https://doi.org/10.1037/0033-295X.92.4.548>.
- West, D. C., O. A. Acar, and A. Caruana. 2020. "Choosing among Alternative New Product Development Projects: The Role of Heuristics." *Psychology & Marketing* 37 (11): 1511–1524. <https://doi.org/10.1002/mar.21397>.
- Wojdyski, B. W., and N. J. Evans. 2016. "Going Native: Effects of Disclosure Position and Language on the Recognition and Evaluation of Online Native Advertising." *Journal of Advertising* 45 (2): 157–168. <https://doi.org/10.1080/00913367.2015.1115380>.
- Wortel, C., I. Vanwesenbeeck, and F. Tomas. 2024. "Made with Artificial Intelligence: The Effect of Artificial Intelligence Disclosures in Instagram Advertisements on Consumer Attitudes." *Emerging Media* 2 (3): 547–570. <https://doi.org/10.1177/27523543241292096>.
- Wu, L., N. A. Dodoo, and T. J. Wen. 2025. "Disclosing AI's Involvement in Advertising to Consumers: A Task-Dependent Perspective." *Journal of Advertising* 54 (1): 20–38. <https://doi.org/10.1080/00913367.2024.2309929>.
- YouTube. 2024. "Updating Our Policies on AI-Altered Content." *YouTube Official Blog*. <https://blog.youtube/news-and-events/disclosing-ai-generated-content/>.
- Zhang, L., and C. Hur. 2025. "The Impact of Generative AI Images on Consumer Attitudes in Advertising." *Administrative Sciences* 15 (10): 395. <https://doi.org/10.3390/admsci15100395>.