

Blurred Authenticity in Social Media Marketing: A Literature Review on Its Impact on Consumer Trust

Mukhamad Rinda Rifa'i ^{*1}

¹ Management Study Program, Faculty of Economics and Business, University of Cahaya Surya (UCS),
Kediri, Indonesia.

*e-mail : mukhamad.rinda@gmail.com

Abstract

Background: The proliferation of social media platforms has created an environment in which branded content and genuine personal expression increasingly coexist within the same communicative space, generating conditions of systematic authenticity ambiguity that threaten consumer trust formation.

Objective: This study introduces and theorizes blurred authenticity as a structural condition of social media marketing and examines the mechanisms through which it affects consumer trust, with the aim of advancing an integrated conceptual framework applicable to influencer marketing practice.

Method: A conceptual narrative literature review synthesizing five anchor studies was complemented by an exploratory qualitative inquiry involving seventeen purposively selected informants from University of Cahaya Surya, encompassing students, lecturers, content creators, and public relations practitioners. Data were analyzed using thematic analysis following Braun and Clarke's six-phase framework.

Results: Findings identify narrative-niche consistency as the dominant authenticity heuristic among consumers, cited by 82.4% of informants. Three trust maintenance mechanisms were empirically validated: perceived transparency, narrative coherence, and value alignment. Transparent commercial disclosure enhances trust only when accompanied by substantively candid content. Perceived value misalignment between creators and endorsed brands generates durable trust erosion extending beyond the immediate interaction.

Conclusion: Blurred authenticity is a systemic rather than individual phenomenon requiring structural responses from brands, creators, and platforms. Brands that prioritize narrative fit in influencer selection and support genuine communicative candor in branded content are best positioned to sustain consumer trust in increasingly saturated digital markets. Future research should pursue longitudinal, experimental, and cross-cultural designs to validate and extend the conceptual model advanced here.

Keywords: Blurred Authenticity, Social Media Marketing, Consumer Trust, Influencer Marketing, Brand Transparency

INTRODUCTION

The accelerating growth of social media platforms over the past decade has fundamentally reconfigured the relationship between brands, communicators, and consumers. Platforms such as Instagram, TikTok, YouTube, and Facebook have evolved far beyond their original function as spaces for interpersonal connection; today, they constitute a dense commercial ecosystem in which branded content, influencer partnerships, and organic user expression coexist, often indistinguishably, within the same digital feed. This structural convergence has made social media marketing one of the most consequential and most contested terrains in contemporary marketing practice. Scholars have consistently demonstrated that social media marketing exerts significant influence on consumer attitudes and behavior, operating primarily through the formation and reinforcement of brand trust rather than through direct persuasive appeals (Arshad, Jannat, Imran, Meer, & Shabbir Ali, 2026; Jain, Verma, Kumar Agarwal, Rajkumar, & Professor, 2026). Trust, in this context, is not merely a peripheral outcome, it is the central psychological mechanism through which social media communication is converted into durable consumer commitment. Without it, the persuasive architecture of digital marketing collapses.

Within this landscape, authenticity has emerged as the dominant currency of credibility. Consumers increasingly evaluate social media content not only on the basis of its informational value but on the degree to which it feels genuine, unmanipulated, and reflective of the

communicator's actual experience (Amini, 2025). The demand for authentic communication is particularly acute among younger demographics, Millennials and Generation Z, who have developed sophisticated skepticism toward overt promotional messaging and who actively seek out content that resonates with lived experience rather than commercial strategy (Huang, 2025). This shift has accelerated the mainstreaming of influencer marketing as a preferred mode of branded communication, precisely because influencers appear to offer something traditional advertising cannot: a personal voice, a consistent persona, and a relationship of familiarity with their audience. Yet the rapid commercialization of this space has introduced a paradox. As influencer marketing has matured into a multibillion-dollar industry, the very authenticity that makes it effective has become increasingly difficult to sustain, let alone verify. The result is a growing deficit of authentic communication at precisely the moment when consumer demand for it is at its highest.

The central problem animating this study is what we term blurred authenticity, a structural condition in which the boundary between genuine consumer-oriented communication and commercially motivated promotional content becomes systematically indistinct. Unlike deliberate deception, which involves the conscious misrepresentation of facts, blurred authenticity is a more pervasive and arguably more consequential phenomenon. It emerges not from individual dishonesty but from the structural dynamics of the influencer marketing ecosystem: the financial incentives that shape which products creators endorse, the narrative strategies that embed branded messages within personal life stories, the algorithmic pressures that reward emotional engagement regardless of its commercial origins, and the misaligned authenticity priorities of the multiple stakeholders, creators, brand managers, and marketing agencies, who co-produce branded content (Duffek, Eisingerich, Merlo, & Lee, 2025). In this environment, consumers are left to make trust judgments on the basis of signals that are inherently ambiguous, and the cognitive mechanisms through which trust is ordinarily constructed are systematically disrupted.

The consequences of this disruption extend well beyond individual consumer decisions. Trust, as established in relationship marketing theory, is the cornerstone of long-term brand relationships; it is the mechanism through which repeated positive interactions are converted into commitment, advocacy, and loyalty (Morgan & Hunt, 1994).

When authenticity is blurred, trust formation becomes unreliable: consumers may extend trust to strategically constructed authenticity while withholding it from genuinely authentic communicators who appear promotional. Either outcome constitutes a failure of the trust formation process. At the aggregate level, the cumulative effect is a progressive erosion of consumer confidence in digital brand communication, a dynamic that (Jain et al., 2026) document empirically, showing that content saturation in digital markets has significantly reduced consumers' willingness to grant trust to branded content. Understanding the mechanisms through which blurred authenticity operates, and the conditions under which its trust-eroding effects can be mitigated, is therefore a question of considerable theoretical and practical urgency.

Prior scholarship has identified several mechanisms that may attenuate the negative effects of blurred authenticity on consumer trust. The most consistently supported is perceived transparency, the degree to which a communicator makes their commercial arrangements visible and their endorsements balanced and candid. (Duffek et al., 2025) demonstrate through a large-scale qualitative study spanning 185 in-depth interviews that consumers assign substantially higher authenticity to influencers who disclose the terms of their brand partnerships and who include both positive and negative dimensions in their product assessments. Critically, these scholars find evidence for what they term transparent opportunism: influencers who openly acknowledge their financial motivation, sometimes with humor or self-deprecation, can paradoxically strengthen rather than undermine consumer trust. This counterintuitive dynamic suggests that the relationship between commercial disclosure and trust is not linear but

conditional, shaped by the manner and context of the disclosure rather than its mere presence or absence.

A second mechanism concerns the role of narrative coherence in sustaining trust under conditions of authenticity ambiguity. (Huang, 2025) identifies three distinct modes of influencer storytelling, personal-experience narratives, brand-centric narratives, and lifestyle-integrated narratives, and demonstrates that each activates different trust-relevant psychological processes. Personal-experience storytelling generates the strongest trust signals because it embeds branded content within intimate life narratives that followers are already emotionally invested in, drawing on Narrative Transportation Theory to explain how immersive stories reduce consumers' epistemic vigilance toward commercial intent (Green & Brock, 2000). Lifestyle-integrated storytelling, by contrast, represents the most structurally blurred mode: products appear as natural elements of aspirational daily life, making their commercial placement nearly invisible. The implication is that narrative coherence, the degree to which branded content is consistent with the influencer's established personal narrative, functions as a proxy for authenticity that consumers can assess even when they cannot directly verify the communicator's genuine experience. A third mechanism, value alignment, refers to the degree to which consumers perceive a genuine correspondence between the influencer's expressed values and the brand they endorse (Duffek et al., 2025). Together, these three mechanisms, perceived transparency, narrative coherence, and value alignment, constitute the primary levers through which trust can be maintained or restored in blurred authenticity contexts.

Taken collectively, the existing literature offers valuable but fragmented insights into the authenticity-trust relationship in social media marketing. (Amini, 2025) provides a rigorous taxonomy of authenticity dimensions but does not account for how these dimensions interact with commercial pressures to produce blurring. (Duffek et al., 2025) illuminate the multi-stakeholder dynamics of authenticity construction but focus primarily on influencer-brand misalignments rather than on consumer trust outcomes.

(Arshad et al., 2026) and (Jain et al., 2026) establish robust empirical links between authenticity, trust, and loyalty, but treat authenticity as a measured construct rather than as a dynamic, context-dependent condition subject to blurring. (Huang, 2025) illuminates the narrative mechanics of authenticity but does not integrate these findings into a coherent account of trust formation under ambiguity. The result is a significant gap in the literature: there is, as yet, no unified conceptual framework that defines blurred authenticity as a structural phenomenon, maps the mechanisms through which it affects consumer trust, and specifies the conditions under which its effects can be moderated. This gap is the central motivation for the present study.

This study addresses the identified gap through a conceptual narrative literature review complemented by an exploratory qualitative inquiry involving seventeen informants drawn from the academic and administrative community of Cahaya Surya University, encompassing students, lecturers, public relations practitioners, and content creators. Three primary objectives guide the investigation. First, the study introduces and theorizes blurred authenticity as a distinct conceptual construct, differentiating it from related concepts such as perceived inauthenticity and deliberate deception. Second, it synthesizes evidence from across the social media marketing, influencer authenticity, storytelling, and consumer trust literatures to map the mechanisms through which blurred authenticity operates and affects trust formation. Third, it advances a conceptual model specifying the conditions, perceived transparency, narrative coherence, and value alignment, under which blurred authenticity's trust-eroding effects are attenuated or reversed. The novelty of this study lies in its integrative conceptualization of blurred authenticity as a structural rather than an individual phenomenon, its multi-stakeholder grounding in real communicative contexts, and its derivation of actionable guidance for brand managers, content creators, and policymakers operating in increasingly ambiguous digital marketing environments. The scope is bounded to social media marketing contexts in which influencer-mediated branded content constitutes the primary communicative mode, with particular attention to platforms such

as Instagram and TikTok where the convergence of personal and commercial expression is most structurally pronounced.

Literature Review

Social Media Marketing and Consumer Trust

Social media marketing has repositioned itself from a supplementary communication channel into a primary driver of brand perception and consumer decision-making. Unlike traditional advertising, which operates through one-directional message delivery, social media marketing functions through a relational logic, brands and their representatives engage consumers in ongoing, dialogic exchanges that accumulate into perceptions of credibility and trustworthiness over time (Dwivedi et al., 2020). This relational architecture means that the persuasive value of social media marketing is not carried by any single message but by the cumulative consistency of communication across multiple interactions and touchpoints.

Consumer trust, in this context, operates as the central mediating construct between marketing exposure and behavioral outcome. (Arshad et al., 2026) confirm through structural equation modeling that social media marketing does not exert a direct significant effect on purchase intention; its influence is fully channeled through brand trust, which functions as the psychological gateway through which consumers translate positive brand communication into commitment and action. This finding aligns with broader evidence from digital marketing scholarship demonstrating that trust mediates the relationship between perceived content quality and long-term consumer loyalty (Jain et al., 2026).

The implication is structurally significant: any disruption to the trust-formation process, including disruptions introduced by authenticity ambiguity, will have cascading consequences for the entire downstream pathway from marketing exposure to behavioral loyalty.

Authenticity in the Context of Social Media Influencers

Defining Authenticity as a Multi-Dimensional Construct

Authenticity in marketing communication is not a unitary concept. Through a systematic review of 88 peer-reviewed articles, (Amini, 2025) identifies five key dimensions through which authenticity is constructed and assessed in influencer marketing contexts: emphasis on origin (the degree to which content reflects genuine, original experience rather than derivative or manufactured material); alignment with societal expectations (the degree to which the communicator's behavior corresponds to culturally shared norms of honest communication); expressions of inner integrity (the degree to which content reflects the communicator's actual values rather than externally imposed commercial scripts); individual interpretation (the recognition that authenticity is not an objective quality but a subjective judgment shaped by each consumer's own experiential lens); and intrapersonal connection (the degree to which content generates a felt sense of personal resonance in the audience).

This multidimensional conceptualization is consequential for understanding blurred authenticity. Because authenticity is assessed simultaneously across multiple dimensions, a communicator may appear authentic on some dimensions while appearing inauthentic on others, and consumers will weight these dimensions differently depending on their prior relationship with the influencer, their product category involvement, and their general disposition toward digital skepticism. Authenticity, in other words, is never simply present or absent; it is always a complex, weighted judgment made under conditions of incomplete information.

Authenticity as a Contextual and Stakeholder-Dependent Outcome

(Duffek et al., 2025) advance this understanding further by reconceptualizing influencer authenticity not as an inherent property of the communicator but as a contextual outcome that emerges from the dynamic interplay among multiple stakeholders. Drawing on assemblage theory (Delanda, 2006), these scholars argue that influencer authenticity is produced through the ongoing interactions of five key components, the influencer, consumers, brand managers, influencer marketing agencies, and the branded content itself, each of which interacts with the others through distinct material and expressive capacities. The authenticity of any given piece of content is therefore not determined by the influencer's personal integrity alone but by the alignment or misalignment of authenticity priorities across all stakeholders simultaneously.

This framework identifies five properties through which authenticity is evaluated and negotiated: expertise, connectedness, integrity, originality, and transparency. Critically, different stakeholders prioritize these properties differently, consumers place highest priority on integrity and transparency, influencers on originality and connectedness, and brand managers on expertise and engagement metrics, generating systematic misalignments that create the structural conditions for blurred authenticity. The study's key theoretical contribution is its demonstration that authenticity is not stable but situational: it can be built, eroded, and rebuilt through specific communicative choices and relational configurations.

Storytelling as a Mechanism of Authenticity Construction Narrative Transportation Theory

The primary mechanism through which influencers construct perceived authenticity is narrative, the embedding of branded messages within personal stories that draw audiences into an experiential world and reduce their critical resistance to commercial content. This process is theorized through Narrative Transportation Theory, originally developed by (Green & Brock, 2000), which posits that compelling stories transport individuals into an imagined narrative world where identification with characters, emotional engagement with situations, and suspension of analytical judgment collectively enhance message acceptance and attitude change.

(Huang, 2025) applies this framework to influencer marketing, demonstrating that narrative transportation operates across three distinct storytelling modes. Personal-experience storytelling, in which creators embed branded products within intimate accounts of their own lives, generates the strongest transportation effects and the highest authenticity perceptions, precisely because the narrative stakes are personal and the emotional investment is observable. Brand-centric storytelling, which foregrounds product features and performance, activates source credibility processes rather than transportation, generating trust through perceived expertise rather than emotional immersion. Lifestyle-integrated storytelling, in which products appear as natural props within aspirational daily-life narratives, produces the most structurally ambiguous authenticity signal: consumers may experience transportation into the aspirational lifestyle while simultaneously failing to register the commercial dimension of the content.

Source Credibility Theory

Complementing Narrative Transportation Theory, Source Credibility Theory (Ohanian, 1990), establishes that the persuasive effectiveness of any message depends critically on the audience's perception of the source's expertise, trustworthiness, and attractiveness. In influencer marketing contexts, these three dimensions map onto distinct communicative behaviors: expertise is signaled through consistent, knowledgeable content creation within a defined niche; trustworthiness is signaled through candid, balanced endorsements that include acknowledgment of product limitations; and attractiveness is signaled through parasocial intimacy, the sense of familiarity and personal connection that followers develop through sustained engagement with a creator's content. (Huang, 2025) demonstrates that these credibility dimensions interact differently with each storytelling mode, such that the persuasive

architecture of influencer content is always a product of both narrative immersion and source credibility simultaneously.

Content Authenticity, Trust, and Long-Term Consumer Loyalty

(Jain et al., 2026) provide the most direct empirical evidence for the downstream consequences of content authenticity on consumer trust and loyalty. Their structural equation model, estimated from a sample of 135 digital consumers, demonstrates that content authenticity, operationalized through dimensions of transparency, consistency, credibility, and relatability, exerts a strong positive effect on consumer trust ($\beta = 0.58$, $p < .001$), which in turn significantly predicts long-term consumer loyalty ($\beta = 0.42$, $p < .01$). Consumer trust partially mediates this relationship, indicating that authenticity builds loyalty both directly and indirectly, directly by generating favorable brand attitudes, and indirectly by fostering the psychological confidence in brand honesty and reliability that sustains commitment over time.

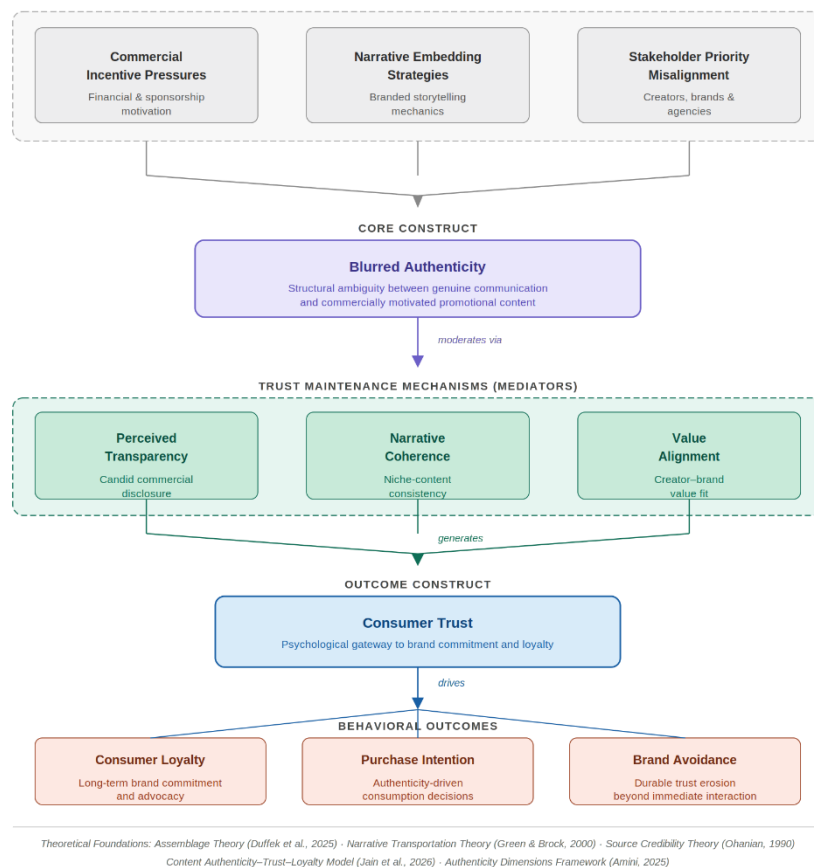
These findings carry particular weight in the context of content saturation. As (Jain et al., 2026) document, the proliferation of branded content across digital platforms has progressively eroded consumers' baseline willingness to trust digital communication, creating conditions in which only content that consistently signals authenticity retains persuasive effectiveness. This dynamic positions authenticity not merely as a desirable attribute of marketing communication but as a structural requirement for brand survival in saturated digital markets.

Theoretical Integration: Toward a Framework of Blurred Authenticity

The theories and empirical findings reviewed across the preceding sections converge on a set of shared insights that collectively motivate the present study's conceptual contribution. First, consumer trust in social media marketing is not produced by exposure alone but by the quality of authenticity signals embedded in communication, signals that are assessed across multiple dimensions simultaneously and that are shaped by the interplay of multiple stakeholders. Second, storytelling is the primary mechanism through which these signals are constructed, but its effects on authenticity perception are highly contingent on narrative mode, source credibility, and the degree to which commercial intent is visible or concealed. Third, the downstream consequences of authenticity, or its blurring, extend well beyond individual purchase decisions to encompass long-term loyalty, advocacy, and the durability of the consumer-brand relationship.

What remains theoretically underdeveloped across this literature is a unified account of the condition that emerges when these dynamics interact systemically, when commercialization pressures, narrative transportation strategies, and stakeholder misalignments combine to produce content in which authenticity signals are not merely imperfect but structurally ambiguous. This is the conceptual space that the construct of blurred authenticity is designed to occupy, and it is in relation to this gap that the present study advances its primary theoretical contribution.

Figure 1. Conceptual Framework of Blurred Authenticity and Consumer Trust



Note. The framework illustrates the proposed relationships among the key constructs examined in this study. Structural antecedents, commercial incentive pressures, narrative embedding strategies, and stakeholder priority misalignment, collectively produce conditions of blurred authenticity (Duffek et al., 2025). Three trust maintenance mechanisms moderate its effects : perceived transparency, narrative coherence, and value alignment (Amini, 2025; Huang, 2025). Consumer trust functions as the central outcome construct (Arshad et al., 2026; Jain et al., 2026), subsequently driving behavioral outcomes of consumer loyalty, purchase intention, and brand avoidance. Arrows indicate directional relationships grounded in Narrative Transportation Theory (Green & Brock, 2000) and Source Credibility Theory (Ohanian, 1990).

Method

Sample Participants

This study employs a purposive sampling strategy to identify informants capable of providing rich, experience-grounded perspectives on authenticity perception in social media marketing contexts. Purposive sampling is particularly appropriate for exploratory qualitative inquiry, as it allows researchers to deliberately select participants whose backgrounds and lived experiences are most likely to illuminate the phenomenon under investigation. Rather than seeking statistical representativeness, this approach prioritizes informational depth and stakeholder diversity, a principle directly informed by (Duffek et al., 2025) multi-stakeholder framework, which demonstrates that authenticity in social media marketing is best understood by triangulating the perspectives of those who consume, produce, and manage branded content simultaneously.

A total of seventeen informants were recruited from the academic and administrative community of Cahaya Surya University, representing a deliberately heterogeneous sample in terms of occupational role, age, and gender. The participant pool encompasses students, lecturers, public relations practitioners, content creators, and administrative staff, reflecting a range of relationships with social media content, from passive consumption to active production and institutional management. This occupational diversity is theoretically significant: as (Amini, 2025) observes, authenticity judgments are not uniform across audiences but are shaped by individuals' prior experience with content creation and their familiarity with the commercial dynamics of influencer marketing.

By sampling across roles that carry different levels of media literacy and platform engagement, this study is positioned to capture the variability in authenticity perception that a homogeneous sample would obscure. The age and gender diversity of the participant pool further enriches this variability, given documented differences in social media usage patterns and trust orientations across demographic groups (Dwivedi et al., 2020).

Instrument

Data were collected using a semi-structured interview protocol consisting of five open-ended descriptive questions, each designed to elicit narrative accounts of informants' lived experience with authenticity and trust in social media marketing contexts. The decision to employ semi-structured rather than fully structured questioning reflects the exploratory orientation of the study: while the five questions provide a consistent thematic framework across all informants, the open-ended format allows for follow-up probing and the emergence of unanticipated perspectives that a fixed-response instrument would preclude (Braun & Clarke, 2006).

The five questions address the following thematic domains in sequence. The first question invites informants to describe the signals through which they assess whether social media content feels genuine or commercially constructed, anchoring the inquiry in the construct of authenticity signal recognition. The second question explores informants' direct experience with authenticity ambiguity, asking them to recount occasions on which they were uncertain whether content reflected genuine recommendation or paid promotion, and to describe how they navigated that uncertainty. The third question examines the conditions under which trust is maintained even when commercial involvement is known, targeting the trust maintenance mechanisms of perceived transparency, narrative coherence, and value alignment theorized in the conceptual framework. The fourth question addresses the specific role of commercial disclosure labels in shaping authenticity perception and trust, directly engaging the paradox of transparent opportunism identified by (Duffek et al., 2025). The fifth and final question elicits accounts of behavioral outcomes driven by authenticity perception, asking informants to describe occasions on which their assessment of content genuineness influenced a purchase decision or shaped a lasting attitude toward a brand, thereby grounding the authenticity-trust-loyalty pathway established by (Jain et al., 2026) in concrete experiential terms.

All questions were formulated in Indonesian to ensure natural and fluent communication with informants, then translated into English for reporting purposes. The instrument was reviewed for clarity and accessibility prior to data collection, with particular attention to ensuring that the language used was comprehensible across the diverse educational and professional backgrounds represented in the sample.

Data Collection Procedures

Interviews were conducted individually with each of the seventeen participants in a setting that ensured privacy and conversational comfort, consistent with ethical standards for qualitative research involving human participants (Creswell & David Creswell, 2018). Each session was conducted in the informant's preferred language, Indonesian, and lasted approximately twenty to forty minutes depending on the depth and spontaneity of responses. Participants were

informed of the academic purpose of the study prior to commencement, and verbal consent was obtained before documentation began. Confidentiality of responses was assured, and participants were made aware of their right to withdraw from the study at any point without consequence.

Responses were documented in written form during the interview and subsequently verified for accuracy. Given the descriptive and narrative nature of the questions, written documentation was deemed sufficient for capturing the substantive content of informants' accounts. To facilitate systematic analysis, all responses were transcribed into a structured format organized by informant and by question. The raw data were retained in their original Indonesian form throughout the analysis process, with English translation applied only at the point of reporting in order to minimize interpretive distortion introduced by premature translation. Participant identities are protected in reporting through role-based attribution, for example, "student informant" or "content creator informant", rather than by name, consistent with the confidentiality commitments made at the point of data collection.

Data Analysis

Qualitative data were analyzed using thematic analysis following the six-phase framework established by (Braun & Clarke, 2006), which comprises familiarization with the data, generation of initial codes, development of candidate themes, review and refinement of themes, definition and naming of final themes, and production of the written report. This approach was selected because it is simultaneously flexible enough to accommodate the exploratory nature of the inquiry and sufficiently rigorous to produce transparent, auditable findings that can be meaningfully integrated with the theoretical framework derived from the literature review (Nowell, Norris, White, & Moules, 2017).

In the familiarization phase, all seventeen interview transcripts were read in their entirety multiple times to develop a holistic understanding of the dataset before any coding was applied. Initial codes were then generated inductively from the data, capturing recurring words, phrases, and experiential patterns without imposing a priori theoretical categories. These initial codes were subsequently organized into candidate themes aligned with the five conceptual domains established by the interview protocol: authenticity signal recognition, navigation of blurred authenticity, trust maintenance mechanisms, the role of transparency disclosure, and authenticity-driven behavioral outcomes. Each candidate theme was reviewed against the full dataset to verify that it was coherently supported across multiple informants and clearly distinguishable from adjacent themes.

Final themes were defined with explicit reference to both the empirical data and the theoretical framework, ensuring that the qualitative findings functioned as genuine validation and contextualization of the literature-derived conceptual model rather than as a parallel or disconnected narrative. This integration strategy reflects the hybrid nature of the study's methodology, a conceptual narrative literature review enriched by exploratory qualitative grounding, and is consistent with established conventions for theory-building research in marketing (Amini, 2025; Duffek et al., 2025). Throughout the analysis, direct quotations from informants were retained as illustrative evidence and selected on the basis of their capacity to crystallize a theme's core meaning with clarity and specificity, ensuring that individual voices remain audible and analytically productive within the broader conceptual argument.

RESULTS AND DISCUSSION

Authenticity Signal Recognition in Social Media Content

The first research question examined how informants identify and assess authenticity signals in social media content involving branded products or services. Thematic analysis of responses to the first interview question revealed that informants across all occupational groups

consistently relied on a combination of narrative consistency, communicative naturalness, and content-niche alignment as primary indicators of genuine communication.

Rather than depending solely on formal disclosure mechanisms, informants described an intuitive but experientially grounded evaluative process in which the perceived coherence between a creator's established persona and their branded content functioned as the dominant authenticity cue. This finding resonates directly with (Amini, 2025) identification of origin emphasis and inner integrity as foundational dimensions of authenticity assessment, as well as with (Duffek et al., 2025) conceptualization of originality, defined as the degree to which an influencer's content reflects their authentic personal style and positioning, as a high-priority authenticity property among consumers. Informants demonstrated a sophisticated awareness of the distinction between content that emerges organically from a creator's established narrative identity and content that appears to have been externally scripted or commercially imposed. A student informant articulated this distinction with particular clarity: *"When the content flows naturally with what they usually talk about, I feel it is honest. But when someone suddenly promotes something completely different from their usual topics, I immediately feel that something is off, it feels like they are just being paid to say it."* A lecturer informant similarly noted: *"Authenticity for me is about consistency. If a person has always been straightforward in sharing their experiences, including the negatives, I tend to believe them more than someone who only says good things about everything they promote."*

Table 1 summarizes the distribution of primary authenticity signals identified across the seventeen informants, organized by the frequency with which each signal type was mentioned spontaneously during the interview.

Table 1. Primary Authenticity Signals Identified by Informants (N = 17)

No	Authenticity Signal	Number of Informants Mentioning	Percentage
1	Narrative-niche consistency	14	82.4%
2	Naturalness of delivery	13	76.5%
3	Acknowledgment of product limitations	11	64.7%
4	Prior organic usage before sponsorship	9	52.9%
5	Frequency and selectivity of promotions	8	47.1%

Note. Informants could mention multiple signal types; percentages reflect proportion of total sample.

As Table 1 illustrates, narrative-niche consistency emerged as the most frequently cited authenticity signal, mentioned by fourteen of seventeen informants. This pattern corroborates (Huang, 2025) finding that personal-experience storytelling generates the strongest authenticity perceptions precisely because it embeds branded content within narrative frameworks that audiences already find credible and emotionally engaging.

Navigating Blurred Authenticity: Informant Experiences with Authenticity Ambiguity

The second research question explored informants' direct experiences with blurred authenticity, occasions on which they were genuinely uncertain whether content reflected voluntary recommendation or commercial obligation, and the coping strategies they employed in response to that uncertainty. All seventeen informants reported having encountered such

situations, confirming that blurred authenticity is not a peripheral or occasional experience but a routine feature of social media consumption for this population.

The coping strategies described by informants fell into three broad patterns: cross-platform verification, community consultation through reading comment sections, and conditional trust extension based on prior relationship with the creator. These strategies reflect the cognitive labor that blurred authenticity imposes on consumers, a dynamic consistent (Jain et al., 2026) observation that content saturation has trained digital consumers to approach branded communication with heightened skepticism and to seek corroborating signals before extending trust. A content creator informant, whose perspective carries particular analytical weight given their insider familiarity with the production side of branded content, described the experience of navigating this ambiguity from a dual position: *"As someone who creates content myself, I know how sponsorships work. So when I watch other creators, I always ask, does this person actually use this product, or did they only receive it yesterday? I look at whether they mention the same product in their non-sponsored content. If they do, I believe them. If they don't, I assume it's purely commercial."* A public relations informant offered a complementary perspective grounded in institutional experience: *"In my work, I understand that every piece of communication has a purpose behind it. So when I see sponsored content on social media, I always read the comments first. The audience's reactions often tell me more about whether the product is genuinely good than the creator's own words."*

Trust Maintenance Mechanisms Under Commercial Involvement

The third research question addressed the conditions under which consumer trust is sustained even after commercial involvement becomes known. Informants' responses converged on three recurring factors: the influencer's demonstrated selectivity in accepting brand partnerships, the perceived alignment between the influencer's personal values and the brand being promoted, and the influencer's track record of candid communication including acknowledgment of product shortcomings. These three factors map precisely onto the trust maintenance mechanisms, value alignment, narrative coherence, and perceived transparency, theorized in the conceptual framework and empirically supported by (Duffek et al., 2025) and (Jain et al., 2026).

Informants were notably consistent in distinguishing between influencers who appeared to endorse indiscriminately and those who communicated a discernible standard of selectivity. An administrative staff informant captured this distinction succinctly: *"I do not stop trusting someone just because they work with a brand. What matters to me is whether they seem to have chosen that brand for a reason, whether it fits who they are. If someone suddenly promotes a brand that has nothing to do with their usual content or their values, that is when I start to question everything they say."* A female student informant added a dimension related to candor: *"The creators I trust most are the ones who are honest about what a product cannot do, not just what it can. That kind of honesty makes me think they are genuinely trying to help me make a good decision, not just trying to get me to buy something."*

The Role of Transparency Disclosure in Trust Formation

The fourth research question examined how informants responded to explicit commercial disclosure labels such as "paid partnership" or "sponsored content," and whether such transparency strengthened or weakened their trust in the content. Responses revealed a nuanced and largely positive orientation toward disclosure, with twelve of seventeen informants indicating that transparent labeling either maintained or enhanced their trust in the creator, provided the content itself remained substantively balanced and credible. Five informants expressed ambivalence, noting that disclosure alone was insufficient if the content itself appeared excessively promotional or inconsistent with the creator's established voice.

This pattern constitutes empirical support for the paradox of transparent opportunism theorized by (Duffek et al., 2025) the counterintuitive finding that explicit acknowledgment of commercial motivation can strengthen rather than undermine authenticity perception when it is accompanied by other stabilizing properties of the authenticity assemblage. A male lecturer informant articulated the underlying logic with precision: *"When a creator says openly that this content is sponsored, I actually respect that more than when there is no label but everything feels like an advertisement. The label tells me they are being honest about the situation. What I then evaluate is whether their actual opinion still seems genuine, or whether the sponsorship has clearly changed what they are willing to say."* A content creator informant, drawing on personal professional experience, reinforced this position: *"In my own content, I always disclose when something is sponsored. I think audiences can tell anyway, so hiding it only makes things worse. Being upfront actually gives me more freedom to be honest about the product, which in the end protects my relationship with my audience more than any single brand deal."*

Authenticity-Driven Behavioral Outcomes

The fifth research question explored the downstream behavioral consequences of authenticity perception, examining whether and how informants' assessments of content genuineness translated into purchase decisions or lasting attitudinal orientations toward brands. All seventeen informants reported at least one instance in which perceived authenticity had directly influenced a consumption-related decision, either by motivating a purchase they would not otherwise have made or by generating a sustained avoidance of a brand whose promotional content they had found inauthentic or misleading.

The behavioral patterns described by informants confirm the authenticity-trust-loyalty pathway established empirically by (Jain et al., 2026), and extend it by illustrating the mechanisms through which that pathway operates in lived experience. A student informant described a purchase motivated by perceived authenticity in terms that directly reflect the narrative transportation mechanism theorized by (Huang, 2025): *"I bought a skincare product because the creator shared her experience over several weeks, including the first two weeks when her skin actually got worse before it improved. That kind of honesty made me feel like she was genuinely documenting her experience, not performing for a camera. I trusted her enough to try it myself."* Conversely, a female administrative staff informant described a lasting brand avoidance rooted in perceived inauthenticity: *"There was a creator I followed for a long time who suddenly started promoting a financial product very aggressively. Everything about the way she talked about it felt rehearsed and unlike her usual style. I have not bought anything she recommended since then, and I actually became skeptical of the brand she was promoting as well."* These accounts illustrate that the consequences of blurred authenticity extend well beyond the immediate interaction, shaping consumer orientations toward both the communicator and the brand across an extended temporal horizon, a dynamic that cross-sectional research designs are structurally ill-equipped to capture and that underscores the need for longitudinal investigation into authenticity's cumulative trust effects.

Discussion

Authenticity as a Structural Condition, Not an Individual Trait

The findings of this study collectively affirm and extend the central theoretical proposition that authenticity in social media marketing is not a stable, inherent quality of any individual communicator but a dynamic, context-dependent condition produced through the interaction of multiple structural forces. This conclusion, which echoes (Duffek et al., 2025) assemblage-theoretic reconceptualization of influencer authenticity, carries significant implications for how both scholars and practitioners understand the phenomenon of blurred authenticity. The near-universal experience of authenticity ambiguity reported by informants, all seventeen participants

described having encountered situations in which they could not reliably distinguish genuine recommendation from commercial obligation, confirms that blurred authenticity is not an occasional failure of individual integrity but a routine, systemic feature of the contemporary social media marketing environment.

What is particularly noteworthy is that informants demonstrated a sophisticated awareness of the structural forces generating this ambiguity. Rather than attributing blurred authenticity exclusively to dishonest creators, informants consistently described it as an emergent condition arising from the intersection of commercial pressures, narrative conventions, and platform dynamics, a characterization that aligns closely with the structural account developed in the conceptual framework. This level of consumer awareness has important theoretical consequences: it suggests that the blurring of authenticity is not simply a problem of disclosure or individual ethics but a deeper architectural challenge that requires systemic responses from brands, platforms, and regulatory bodies alike.

The Primacy of Narrative Coherence as an Authenticity Heuristic

Among the most substantively significant findings of this study is the identification of narrative-niche consistency as the dominant authenticity signal across the informant population, cited by 82.4% of participants. This pattern extends (Huang, 2025) theoretical account of personal-experience storytelling by demonstrating empirically that consumers do not evaluate authenticity primarily through formal disclosure mechanisms but through narrative coherence, the perceived fit between a creator's established personal narrative and the branded content they produce. When that fit is disrupted, when a creator whose identity is associated with one domain suddenly endorses products from an unrelated category, consumers interpret the discontinuity as a signal of commercial motivation overriding genuine experience, triggering skepticism that extends to both the content and the creator's broader communicative credibility.

This finding has a practical implication that the literature has not previously articulated with sufficient directness: the most consequential authenticity decision a brand makes is not whether to disclose a commercial partnership but which creator to partner with in the first place. A partnership that is narratively coherent, that positions the brand as a natural element within the creator's existing content world, generates authenticity perceptions that no amount of post-hoc disclosure engineering can replicate. Conversely, a partnership that disrupts narrative coherence generates skepticism that disclosure cannot fully repair. The strategic implication is that narrative fit should be treated as a primary criterion in influencer selection, weighted at least as heavily as follower count or engagement metrics, a reorientation that current industry practice, with its predominant focus on quantitative performance indicators, has been slow to adopt.

Transparent Opportunism as a Trust-Building Paradox

The finding that twelve of seventeen informants responded positively to explicit commercial disclosure, reporting that transparent labeling either maintained or enhanced their trust in the creator, constitutes one of the study's most theoretically generative results, and one that deserves careful interpretive attention.

On the surface, this pattern might appear to suggest that disclosure is straightforwardly trust-enhancing, a conclusion that would license a simple policy recommendation: require more disclosure, generate more trust. The qualitative texture of informants' responses, however, reveals a more conditional dynamic. Trust enhancement through disclosure was consistently contingent on the perceived genuineness of the content that accompanied the disclosure. Informants distinguished clearly between creators who disclosed a commercial arrangement and then proceeded to offer a candid, balanced assessment, including acknowledgment of limitation, and those who disclosed the arrangement but then delivered what appeared to be a scripted, uniformly positive endorsement. In the former case, disclosure amplified trust by functioning as

a signal of the creator's broader communicative integrity. In the latter case, disclosure was insufficient to counteract the inauthenticity signals carried by the content itself.

This nuanced pattern is consistent with (Duffek et al., 2025) concept of transparent opportunism, but extends it by specifying the conditions under which that paradox operates. Transparent acknowledgment of commercial motivation builds trust not because transparency is inherently valued in the abstract, but because it activates consumers' perception of inner integrity, the sense that the creator is prioritizing honest communication with their audience over the financial incentives of the brand relationship. When transparency is accompanied by content that contradicts this signal, when the disclosed partnership is followed by an evidently scripted endorsement, the disclosure loses its trust-building function and may even amplify skepticism by making the commercial compromise more visible.

Value Alignment as a Long-Term Trust Architecture

The results pertaining to trust maintenance under conditions of known commercial involvement reveal a consistent and theoretically coherent pattern: consumers sustain trust in commercially active creators not despite their commercial involvement but through their perception that such involvement reflects genuine value alignment rather than indiscriminate financial opportunism. This finding directly validates the value alignment mechanism theorized in the conceptual framework and empirically supported by (Jain et al., 2026) who demonstrate that content perceived as consistent with the creator's expressed values generates significantly stronger trust and loyalty outcomes than content that appears commercially motivated without corresponding value coherence.

The behavioral outcome data reported in Section 4.5 extend this insight by illustrating the temporal durability of value alignment's effects on trust. The brand avoidance behavior described by several informants, in which a single perceived authenticity violation generated a lasting negative orientation toward both the creator and the associated brand, suggests that value alignment functions not merely as a momentary trust signal but as a foundational component of the consumer-creator relationship whose disruption carries consequences that extend well beyond the immediate interaction. This temporal dimension of trust erosion is a dimension that the existing literature, with its predominant reliance on cross-sectional designs, has been structurally unable to capture, and it represents one of the most compelling arguments for the longitudinal research agenda proposed in this study.

Limitations and Interpretive Boundaries

Several limitations of the present study warrant explicit acknowledgment in the interest of interpretive transparency. The sample, while occupationally and demographically diverse within its institutional context, is bounded to the community of a single university in Indonesia, which necessarily constrains the generalizability of the qualitative findings.

Cultural factors, including norms around commercial communication, trust orientations, and social media usage patterns that are specific to the Indonesian context, may shape authenticity perceptions in ways that differ from other cultural environments, and the findings should be interpreted with this boundary condition in mind. Future research employing cross-cultural comparative designs would be well-positioned to examine the degree to which the mechanisms identified here operate consistently across different social and regulatory contexts.

Additionally, the exploratory nature of the qualitative inquiry means that the findings are best understood as theoretical illustrations that enrich and validate the conceptual framework rather than as definitive empirical claims. The sample size of seventeen informants, while appropriate for a purposive exploratory design, does not support the kind of saturation analysis that larger qualitative studies employ to establish the comprehensiveness of their thematic findings. The conceptual model advanced in this study therefore represents a theoretically grounded but empirically preliminary account of blurred authenticity's mechanisms and effects, one that is

explicitly designed to generate testable hypotheses for subsequent quantitative and longitudinal investigation rather than to constitute a final empirical settlement of the questions it raises.

CONCLUSION

This study introduced and theorized blurred authenticity as a structural condition of contemporary social media marketing, one in which the boundary between genuine communication and commercial promotion becomes systematically indistinct, disrupting the trust formation processes upon which durable consumer-brand relationships depend. Through the integration of a conceptual narrative literature review with exploratory qualitative inquiry involving seventeen informants from University of Cahaya Surya, the study identified three primary mechanisms through which trust is maintained or eroded under conditions of authenticity ambiguity: perceived transparency, narrative coherence, and value alignment.

The findings demonstrate that consumers do not evaluate authenticity through disclosure mechanisms alone but through a sophisticated, experientially grounded assessment of narrative consistency, communicative candor, and the perceived alignment between a creator's expressed values and their commercial partnerships. Transparent acknowledgment of commercial involvement enhances trust only when accompanied by substantively honest content, a conditionality that refines existing theoretical accounts of transparent opportunism and carries direct strategic implications for brand-influencer partnership design.

This study contributes to the marketing literature by establishing blurred authenticity as a theoretically distinct construct, differentiating it from deliberate deception and perceived inauthenticity, and by mapping the specific mechanisms through which its trust consequences are moderated. Future research should pursue longitudinal designs to examine how trust violations accumulate over time, experimental studies to isolate individual mechanism effects, and cross-cultural investigations to establish the boundary conditions of the present findings across different social and regulatory environments.

REFERENCES

- Amini, M. (2025). Authenticity in the Realm of Influencer Marketing: A Systematic Review. In *Journal of Promotional Communications* (Vol. 11). Retrieved from <http://promotionalcommunications.org/index.php/pc/about/submissions>
- Arshad, L., Jannat, A., Imran, T., Meer, W., & Shabbir Ali, K. (2026). The Impact of Social Media Marketing on Consumer Purchase Intention: The Mediating Role of Brand Trust. *Policy Journal of Social Science Review*, 4(1).
- Audrezet, A., De Kerviler, G., & Moulard, J. G. (2020). Authenticity under threat: When social media influencers need to go beyond self-presentation. *Journal of Business Research*, 117, 557–569. <https://doi.org/10.1016/j.jbusres.2018.07.008>
- Beverland, M. B., & Farrelly, F. J. (2010). The quest for authenticity in consumption: Consumers' purposive choice of authentic cues to shape experienced outcomes. *Journal of Consumer Research*, 36(5), 838–856. <https://doi.org/10.1086/615047>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Creswell, J. W., & David Creswell, J. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. New Delhi.
- Campbell, C., & Farrell, J. R. (2020). More than meets the eye: The functional components underlying influencer marketing. *Business Horizons*, 63(4), 469–479. <https://doi.org/10.1016/j.bushor.2020.03.003>
- Delanda, M. (2006). *A New Philosophy of Society*. Chennai.
- Duffek, B., Eisingerich, A. B., Merlo, O., & Lee, G. (2025). Authenticity in Influencer Marketing: How Can Influencers and Brands Work Together to Build and Maintain Influencer Authenticity? *Journal of Marketing*, 89(5), 21–46. <https://doi.org/10.1177/00222429251319786>

- Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., ... Wang, Y. (2020). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*, 59. <https://doi.org/10.1016/j.ijinfomgt.2020.102168>
- Green, M. C., & Brock, T. C. (2000). The Role of Transportation in the Persuasiveness of Public Narratives. *Journal of Personality and Social Psychology*, 79(5), 701–721. <https://doi.org/10.1037/0022-3514.79.5.701>
- Huang, K. (2025). Storytelling as a Strategic Tool in Influencer Marketing: Impact on Engagement, Authenticity and Purchase Intent. *Studies in Media and Communication*, 14(1), 291. <https://doi.org/10.11114/smc.v14i1.7935>
- Jain, V., Verma, C., Kumar Agarwal, M., Rajkumar, A., & Professor, A. (2026). Influence of Content Authenticity on Long-Term Consumer Loyalty in Digital Markets. *Conference Proceedings: "Sustainability Through Digital Transformation: Insights from Management, Humanities, and Sciences,"* 14(S1), 30–31. <https://doi.org/10.64882/ijrt.v14.iS1.1070>
- Kim, D. Y., & Kim, H. Y. (2021). Trust me, trust me not: A nuanced view of influencer marketing on social media. *Journal of Business Research*, 134, 223–232. <https://doi.org/10.1016/j.jbusres.2021.05.024>
- Lee, J. A., & Eastin, M. S. (2021). Perceived authenticity of social media influencers: Scale development and validation. *Journal of Research in Interactive Marketing*, 15(4), 822–841. <https://doi.org/10.1108/JRIM-12-2020-0253>
- Lou, C., & Yuan, S. (2019). Influencer marketing: How message value and credibility affect consumer trust of branded content on social media. *Journal of Interactive Advertising*, 19(1), 58–73. <https://doi.org/10.1080/15252019.2018.1533501>
- Morgan, R. M., & Hunt, S. D. (1994). The Commitment-Trust Theory of Relationship Marketing. In *Source: Journal of Marketing* (Vol. 58).
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, 16, 1–13. <https://doi.org/10.1177/1609406917733847>
- Nunes, J. C., Ordanini, A., & Giambastiani, G. (2021). The concept of authenticity: What it means to consumers. *Journal of Marketing*, 85(4), 1–20. <https://doi.org/10.1177/0022242921997081>
- Ohanian, R. (1990). Construction and Validation of a Scale to Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness. In *Source: Journal of Advertising* (Vol. 19).
- Patton, M. Q. (2014). *Qualitative research and evaluation methods* (4th ed.). SAGE Publications.
- Paul, J., & Criado, A. R. (2020). The art of writing literature review: What do we know and what do we need to know? *International Business Review*, 29(4), 101717. <https://doi.org/10.1016/j.ibusrev.2020.101717>
- Sokolova, K., & Kefi, H. (2020). Instagram and YouTube bloggers promote it, why should I buy? How credibility and parasocial interaction influence purchase intentions. *Journal of Retailing and Consumer Services*, 53, 101742. <https://doi.org/10.1016/j.jretconser.2019.01.010>
- Wellman, M. L., Stoldt, R., Tully, M., & Ekdale, B. (2020). Ethics of authenticity: Social media influencers and the production of sponsored content. *Journal of Media Ethics*, 35(2), 68–82. <https://doi.org/10.1080/23736992.2020.1736078>