

The unpaid auditors: Trust and De-influencing in E-Commerce

Wiam ZRIMI, (PhD Student)

*Laboratory of economic sciences and public policies (LSEPP)
Faculty of economics and management
Ibn Tofail University, Kenitra, Morocco*

Fatima CHAREF, (Full Professor)

*Laboratory of economic sciences and public policies (LSEPP)
Faculty of economics and management
Ibn Tofail University, Kenitra, Morocco*

Correspondence address :	Laboratory of Economic sciences and public policies (LSEPP), Faculty Ibn Tofail University Avenue des Universités, Campus Universitaire, Kenitra, Morocco B.P 14000
Disclosure Statement :	The authors declare that they have not received any financial support that could have influenced the objectivity of this study. They take full responsibility for any potential plagiarism, the use of artificial intelligence in the writing process, as well as for the results presented in this article.
Conflict of Interest :	The authors report no conflicts of interest.
Cite this article :	ZRIMI, W., & CHAREF, F. (2026). The unpaid auditors: Trust and De-influencing in E-Commerce. <i>International Journal of Accounting, Finance, Auditing, Management and Economics</i> , 6(7), 535–554. https://doi.org/10.5281/zenodo.20753507
License	This is an open access article under the CC BY-NC-ND license

The unpaid auditors: Trust and De-influencing in E-Commerce

Abstract

In digital commerce, AI-driven recommendation systems increasingly decide which products, influencers, and stories a consumer sees, and that visibility carries a trust problem of its own. When consumers sense saturation, staged authenticity, or hidden commercial motives, suspicion sets in and they begin checking claims themselves, scrutinizing details, flagging what does not add up, warning others, and pointing to alternatives. Existing work usually discusses de-influencing, digital labor, and e-commerce adoption separately, leaving the unpaid evaluative work that consumers perform around influencer content largely untheorized. We address that gap by naming this unpaid checking audience correction labor and arguing that it has become one of the ways trust now gets built in algorithmic e-commerce. The construct works at two levels. For the individual, it appears as consumer due diligence, where influencer content is treated as evidence to be tested rather than advice to be followed. For the market, it becomes reverse influence, as audience reactions travel back and reshape how credible influencers, brands, products, and platform recommendations appear. The article makes three contributions, an operational definition of audience correction labor and its two dimensions, a four-layer framework linking AI-mediated visibility, trust conditions, correction labor, and trust recalibration to e-commerce adoption, and a set of eleven testable propositions. Read this way, de-influencing looks less like anti-consumption and more like trust-building work, a shift that changes what the behavior means in practice. We close with the Moroccan case, where checking that happens off-platform, in WhatsApp and Facebook groups, weighs heavily on online purchasing because trust in formal institutions remains low. AI does not only make e-commerce more convenient; it also raises credibility problems that consumers resolve by correcting one another, and this matters most where trust is the barrier to buying online.

Keywords: De-influencing; Audience Correction Labor; E-Commerce Adoption; Algorithmic Visibility; Trust Recalibration; Influencer Marketing.

Classification JEL: M31, M37, D12, L81.

Paper type: Theoretical Research.

1. Introduction

For more than ten years, influencer marketing ran on a fairly simple exchange. Brands reached engaged audiences, and those audiences got product tips from people they could relate to. As long as consumers kept responding, through their attention, their engagement, and their purchases, the model held together. The trouble is that the very mechanisms that once amplified influencer marketing are now wearing down its credibility.

Recommendation algorithms have done more than spread influencer content around. They have packed digital spaces full of it. Ten minutes of scrolling can now put dozens of product recommendations in front of a consumer, and most of them were placed there by an algorithm chasing engagement, not by anyone asking whether the influencer actually uses the product or believes in it. Reviewing what the influencer-marketing literature says about purchase intention, Sardar and Vijay (2025) point out that the field has hardly started to ask what all this saturation is doing to the way consumers decide. George et al. (2025), looking at influencer marketing in fashion, arrive at much the same conclusion from another direction.

Out of this, a consumer response has taken shape, namely de-influencing. A growing part of the audience now handles promotional content as evidence to be checked rather than guidance to be followed. Counter-recommendations, sponsorship call-outs, collections of negative reviews, warnings between peers, all of these now move around next to ordinary influencer content, and at times as a direct reply to it. The question is no longer whether influence works. It is whether influence still travels in only one direction.

We argue that this change plays out on two connected levels. At the individual level, de-influencing works as a kind of consumer due diligence suited to the conditions of digital commerce, in that the consumer stops reading influencer content as advice and starts treating it as evidence that has to be tested. At the collective level, de-influencing creates what we call reverse influence, where audience reactions travel back upstream and reshape the credibility, the visibility, and the commercial worth of the content that was first aimed at them. Put together, these two levels turn the older one-way picture of persuasion into something closer to a contested negotiation over credibility.

What makes this shift matter is that influencer content no longer reaches the consumer on its own. A sponsored post today turns up surrounded by comments, stitches, screenshots, counter-reviews, bad experiences, and rival recommendations. In other words, the promotional message arrives with its own contestation already attached. The audience is no longer the last stop in a chain of persuasion; it takes part in that chain, corrects it, and sometimes reverses it before anyone decides to buy. We name what audiences do here audience correction labor, and we treat it as part of how trust is now put together in AI-mediated commerce.

This argument bears directly on e-commerce adoption. Whether a consumer is willing to buy online rests on the platform being usable, on price, on convenience, and, more and more, on whether the recommendations around a product come across as credible. It is the algorithm that decides how much visibility a product or creator gets on a platform. When that visibility breeds suspicion, the consumer holds back, puts off the purchase, or leans on peer correction before taking up the channel at all. Correction work therefore becomes one of the trust conditions that adoption depends on. So the question we ask is not simply whether consumers trust influencers; it is whether, and under what conditions, they take up e-commerce in the first place.

Three contributions grow out of this argument. The first is the construct of audience correction labor itself. We situate it within the digital-labor literature (Terranova, 2000; Duffy, 2017; Abidin, 2016) while spelling out what sets it apart, namely its corrective, regulatory, and unpaid character. The second is a bridge between two ideas that the literature has mostly kept separate, consumer due diligence at the individual level (Friestad & Wright, 1994; Obermiller & Spangenberg, 1998), and reverse influence at the market level, which extends the electronic

word-of-mouth tradition (Hennig-Thurau et al., 2004; Cheung & Thadani, 2012). Our claim is that these are two sides of one activity. Third, we put forward a four-layer framework that ties AI-mediated visibility, trust conditions, correction labor, and trust recalibration to e-commerce adoption outcomes, and that reads de-influencing as trust-building work rather than as anti-consumption behavior.

Methodologically, the article is a work of conceptual theory-building rather than empirical testing, and it should be read in that light from the outset. The question it sets out to answer is how audience correction labor, understood as unpaid evaluative work that consumers perform around influencer content, contributes to trust recalibration and to e-commerce adoption in algorithmically mediated environments, and under what conditions its effect is strongest.

Section 2 sets out the theoretical background on which the framework draws. Section 3 explains the methodological position. Sections 4 and 5 trace the way algorithmic visibility creates the conditions for de-influencing. The two-level construct is built up in Sections 6 and 7. Section 8 lays out the framework and the propositions. Section 9 turns to implications, the Moroccan context among them, and Section 10 concludes.

2. Theoretical Background

The framework draws on six streams of theory that have, until now, been mobilized largely in isolation. Setting them out together at the start clarifies the conceptual lenses through which de-influencing is read in the sections that follow.

The first stream is source credibility theory, which holds that a persuasive message succeeds only insofar as its source is judged trustworthy and expert (Hovland & Weiss, 1951). The second is the persuasion knowledge model, which describes how consumers accumulate knowledge of marketers' tactics and draw on it to manage persuasion attempts (Friestad & Wright, 1994), together with the wider literature on consumer skepticism toward advertising (Obermiller & Spangenberg, 1998). These two streams explain why dense, commercially loaded influencer content invites scrutiny rather than acceptance.

The third stream is digital trust, which distinguishes trusting beliefs from trusting intentions and treats trust as something calibrated through experience (Rousseau et al., 1998; McKnight & Chervany, 2001). Because trust in online exchange is shaped by context as much as by individual disposition, the framework also draws on work showing that the antecedents of e-commerce trust vary across cultural and institutional settings (Gefen & Straub, 2000). This matters for the emerging-market case developed later, where interpersonal trust often carries more weight than institutional trust. The fourth stream is the technology adoption tradition, in particular the unified view that links perceived usefulness, social influence, and facilitating conditions to use (Venkatesh, Morris, Davis, & Davis, 2003), which situates correction labor among the antecedents of adoption rather than treating it as a stand-alone attitude.

The fifth stream is electronic word-of-mouth, which examines how consumer-generated evaluations circulate and shape the judgments of others (Bickart & Schindler, 2001; Hennig-Thurau, Gwinner, Walsh, & Gremler, 2004; Cheung & Thadani, 2012). Reverse influence, defined below, extends this tradition to the corrective content that now travels alongside promotion. The sixth stream is digital labor on platforms, which shows that much of the value circulating in the platform economy is produced by users for free (Terranova, 2000; Abidin, 2016; Duffy, 2017), and which frames the analytical step of treating de-influencing as unpaid evaluative work. Across all six streams, the architectures that decide what consumers see are themselves an object of analysis, a point developed in the literature on algorithmic and surveillance-driven visibility (Zuboff, 2019).

None of these streams, on its own, accounts for the corrective activity that consumers now perform around influencer content. The contribution of the framework is to bring them into a single structure, which the remaining sections build up.

3. Methodological Positioning and Conceptual Theory-Building

This is a conceptual article. It does not put propositions to the test against fresh data. What it sets out to do is build a theoretical framework that others can then test. Conceptual work of this sort has a long standing in marketing and consumer research, and it is especially useful when a phenomenon is still new enough that the empirical literature is scattered and no organizing frame has settled in yet (MacInnis, 2011; Jaakkola, 2020). De-influencing fits that picture. Practitioners and researchers alike have given it real attention, but the vocabulary used to analyze it has not caught up with what the phenomenon seems to be. The empirical base is also thin and recent, and what exists is fragmented across designs, platforms, and settings (Penttinen, Ciuchita, & Pedersen, 2026; Khalil et al., 2025; Wiratha & Kemalasari, 2025). Individual studies establish that de-influencing affects trust and purchase intention, but they do not yet share a common construct or a model that links the individual and market levels. A conceptual synthesis is therefore useful before further empirical testing.

The article pulls together the six theoretical streams set out in Section 2 and reads them against two fast-moving empirical domains, AI-mediated commerce and the emerging de-influencing literature (Terranova, 2000; Duffy, 2017; Abidin, 2016). We do not claim that the framework uncovers a new phenomenon. The claim is rather that it brings together what these literatures only glimpse in pieces, and that it gives a name to a construct, audience correction labor, which none of them on its own has named.

The literature was assembled through a structured narrative review rather than a systematic protocol, which suits a theory-building aim. We searched Scopus, Web of Science, and Google Scholar for work published mainly between 2015 and 2026, combining terms for de-influencing, influencer marketing, consumer skepticism, electronic word-of-mouth, digital and consumer trust, algorithmic recommendation, and digital labor. Sources were retained when they spoke to one of the streams identified above, or to the emerging de-influencing literature, and when they either reported relevant empirical findings or advanced a usable construct. Foundational earlier works were added through backward citation tracing. The retained material was then read thematically, grouping findings by the analytical role they play in the framework rather than by discipline, which is what allowed the individual-level and market-level facets of correction labor to be brought together.

An approach that reads across literatures can bring to light mechanisms that single-discipline accounts tend to overlook. Until it is tested, though, the framework stays a structured hypothesis rather than a verified explanation. The propositions in Section 8 are written so that they can be tested, and Section 9 points to empirical strategies that could either challenge the framework or sharpen it.

4. Influencer Marketing and AI-Mediated Visibility

4.1 Algorithmic visibility

Influencer marketing does not operate in some neutral pool of information. It operates inside recommendation architectures that were built to push engagement as high as possible. Yet engagement is not the same thing as trust or relevance. The broader turn toward social-media-mediated commerce, which Mangold and Faulds (2009) already flagged as a basic change in the marketing communications mix, has only sped up under platform algorithms that surface content according to interaction signals such as views, shares, saves, and comments. Products and creators that draw strong engagement early on are handed wider distribution. The outcome is a feedback loop in which visibility keeps reinforcing itself, whether or not the recommendation underneath is any good. Lou and Yuan (2019) show that influencer content persuades only to the degree that consumers see informational value and source trustworthiness

in it together, and that amplification logics rewarding reach over fit can wear both of these down.

This dynamic gives rise to what we will call algorithmic visibility, a state in which certain products, creators, and narratives are exposed far more than others, not because they have any special merit, but because they suit the platform's engagement-optimization logic. AI-mediated content distribution lays another layer on top of this. Studies of AI disclosure and brand credibility find that when the source of algorithmically generated or curated advertising is left unclear, consumers grow ambivalent and the content loses credibility (Jiang, Zheng, & Kong, 2025). Consumers are more and more aware, even if not in precise terms, that what reaches them is shaped by systems whose interests do not always line up with their own.

4.2 Influencer saturation and declining trust

Influencer-produced content has expanded sharply over the last ten years. Recent reviews of AI-mediated commerce tend to single out trust, privacy, and transparency as the main worries in consumer adoption, and they note how recommendation architectures can eat away at adoption whenever their workings stay opaque. For influencer marketing in particular, the sheer density of sponsored content is not merely irritating; it is corrosive to trust.

Penttinen, Ciuchita, and Pedersen (2026) give us a direct empirical anchor here. Drawing on real-world reporting about how influencers behave, their work shows that consumer skepticism turns up in recognizable circumstances, that is, when influencers hide brand partnerships, when they push products they do not use themselves, or when they sign up for sponsorships that look bad for consumers. Over time the result is a consumer base that has learned to read influencer recommendations with suspicion, not with the trusting attention that early influencer marketing took for granted. This fits the long-standing finding that consumers build up a general tendency to disbelieve advertising claims (Obermiller & Spangenberg, 1998), and it fits the wider Persuasion Knowledge Model too (Friestad & Wright, 1994; see Eisend & Tarrahi, 2022 for a meta-analytic review), which holds that consumers gradually learn about persuasion tactics and then draw on that knowledge to protect their own decisions. So skepticism itself is not what is new about algorithmic commerce. What is new is its scale, its speed, and the public spaces in which it now plays out.

Three structural conditions push the effect along. First, influencer work has become professionalized, producing a class of creators whose income hangs on brand partnerships, which keeps the disclosure of conflicts of interest a constant issue rather than an occasional one (Abidin, 2016; Boerman et al., 2017). Second, promotional formats have grown standardized, the unboxing video, the GRWM sequence, the morning routine, the haul, so that the outward features of promotional content have become easy to predict, and that in turn makes commercial intent easier for the audience to spot. Third, platforms have rolled out affiliate links, in-app shopping, and direct sponsorship tools, lowering the cost of turning content into commerce and ramping up the pressure to churn out promotional material at scale. All of this leaves authenticity structurally harder to keep up than it was in the early days of influencer marketing.

5. From Authenticity to Perceived Staged Authenticity

5.1 Authenticity as the foundation of influencer credibility

Across the influencer-marketing literature, authenticity comes up again and again as a central mediator of trust and purchase intention. Source Credibility Theory (Hovland & Weiss, 1951), which underpins much of this work, holds that a message lands only insofar as the communicator looks trustworthy and expert. Examining how de-influencing reshapes trust in online communities, Khalil et al. (2025) find that de-influencers tend to be seen as more credible than ordinary influencers, largely because their critical stance on overconsumption reads as a

sign that no obvious commercial motive is in play. In a parallel study on the future of influencer marketing, Nazara, Nugrahaningsih, and Abdillah (2025) report that trust is built through perceived expertise, consistency, and ethical transparency, and that authenticity itself is constructed through personal storytelling and a selective choice of brand endorsements.

These findings sit comfortably with the broader literature. Wiratha and Kemalasari (2025) find that de-influencing shapes Gen Z purchase intention, with source credibility working as the mechanism in between. Earlier on, Audrezet, de Kerviler, and Guidry Moulard (2020) had already shown that influencer authenticity rests on pulling off two things at once, passionate self-expression and transparent disclosure, both of which get harder to hold onto as commercial pressures climb. Authenticity, then, is not merely a matter of style. It is a functional condition for persuasion to work at all.

5.2 From authenticity to staged authenticity

There is a paradox at the heart of influencer marketing, since its very success laid the ground for its own delegitimation. As brand partnerships piled up, as disclosure rules sharpened scrutiny, and as the emotional registers of influencer content slid into formula, consumers began to read authenticity as something performed. We call this perceived staged authenticity, by which we mean the consumer's recognition that the relational, emotional, and confessional modes of influencer content have been turned to commercial ends.

Perceived staged authenticity is not the same as plain dishonesty. An influencer may well use and enjoy a product. The difficulty is that the consumer can no longer be sure of it, because the platform's structural incentives, algorithmic rewards for engagement and financial rewards for sponsorship, turn out content that mimics genuine recommendation no matter how sincere it really is underneath. Writing in the *Asian Journal of Communication*, Kim, Lau, and Lee (2025) show that in low influencer-product fit scenarios, non-narrative content styles sharply increase consumer skepticism. Putting it another way, the formal features of influencer content move trust on their own, apart from whatever the message itself happens to be saying.

What has really shifted is the consumer's reading practice. Content that five years ago would have passed as a personal recommendation gets read today as a commercial performance, even where the sincerity behind it has not changed at all. And once perceived staged authenticity becomes the default lens, influencers find it structurally hard to win back credibility by piling on more authenticity, since the audience now treats every extra gesture of sincerity as one more piece of evidence of staging. This is the ground on which de-influencing emerges. It is the audience's reply to a credibility crisis that simply cannot be settled from inside the original frame of persuasion.

6. Audience Correction Labor and De-influencing as Unpaid Evaluative Work

6.1 Defining de-influencing

De-influencing reached public discourse mostly through social media, TikTok above all, as a practice in which creators actively talk people out of buying, expose sponsored content as misleading, or offer critical counter-evaluations of products that have been pushed heavily. In both popular and academic discussion, early framings tended to file de-influencing under anti-consumption behavior, reading it as a rejection of the platform's commercial logic.

That framing does not take us far enough analytically. Penttinen, Ciuchita, and Pedersen (2026) offer a sharper conceptualization. They define de-influencing content as organic social media posts in which influencers discourage purchases, and they show that this content works to repair the consumer trust that the flood of sponsored endorsements has weakened. Tellingly, their

findings show that de-influencing raises an influencer's trustworthiness by easing suspicion about commercial motives. The effect falls away, however, once de-influencing shows up next to a sponsored post for the same brand. The reframing matters. Rather than rejecting consumption, de-influencing recalibrates the way credibility gets read in algorithmic commerce.

It is worth being precise about how this influencer-led de-influencing relates to the audience correction labor that is the focus of this article, since the two are not the same. Influencer-led de-influencing is content a creator produces, whereas audience correction labor is work the audience performs. The two can stand in three relations. They complement each other when a creator's honest discouragement is taken up and circulated by audiences as a credibility signal, so that each reinforces the other. They amplify each other when audience scrutiny pushes creators toward pre-emptive candor, which in turn invites further checking. And they cancel each other when influencer-led de-influencing is read as merely another sponsored move, as happens when it appears beside a paid post for the same brand (Penttinen, Ciuchita, & Pedersen, 2026); in that case the creator's gesture collapses and audience correction does the trust work on its own. The framework that follows is concerned chiefly with this second, audience-performed activity.

Even this sharper framing is still incomplete, because it casts de-influencing as something certain influencers do on behalf of other consumers. In reality, the consumers themselves are doing most of the work. The comments, stitches, screenshots, peer warnings, comparison threads, and "don't buy this" videos that gather around promotional content do not come from influencers. They come from audiences. And they are produced for free.

6.2 Audience correction labor as the central construct

De-influencing, we argue, is best read as unpaid evaluative work carried out by consumers. We give it the name audience correction labor, the process by which consumers spot misleading claims, expose inconsistencies, put together counter-recommendations, and generate credibility signals that help other consumers find their way through algorithmic e-commerce.

Box 1. Definition of audience correction labor.

Nominal definition. The unpaid evaluative work through which consumers test, correct, and re-signal the credibility of influencer content, and in doing so help other consumers navigate algorithmic e-commerce.

Dimensions. (i) Consumer due diligence, the individual-level activity of checking claims, weighing source credibility, and verifying evidence before acting; (ii) reverse influence, the market-level process by which aggregated audience reactions become public credibility signals.

Proximal determinants. Perceived staged authenticity, sponsorship saturation, algorithmic opacity, accumulated persuasion knowledge, and digital literacy.

Conceptual boundaries. Correction labor is corrective rather than promotional, unpaid and unprotected by disclosure rules, and distinct both from skepticism, which is an attitude, and from anti-consumption, which rejects the purchase. It is an activity that produces shareable information.

Framing it as labor sets de-influencing inside a larger literature on digital labor on platforms. Terranova (2000) was among the first to observe that the digital economy runs on activity that is given freely yet stays unwaged, enjoyed and exploited at one and the same time. Duffy (2017) carried this further with her account of aspirational labor, the unpaid work content creators take on in the hope of monetizing it later. Abidin (2016) looked at visibility labor, the work followers do to help influencers reach and hold their place on the platform. Audience correction labor sits apart from all three. It does not chase monetization, it does not boost an influencer's visibility, and it does not feed the platform's engagement metrics in the way those earlier concepts describe. Where they are promotional, it is reactive; where they are productive in the platform's own terms, it is regulatory. Its face is turned toward other consumers, not toward the platform or the influencer.

Correction labor qualifies as labor in a real sense, since it costs time, creates value for others, and goes uncompensated. Take the consumer who writes a careful comparison comment under a sponsored post, who screenshots and annotates a misleading claim, or who pieces together a thread documenting how a product fails. That person produces genuine informational value. Other consumers draw on that value. Platforms gain from the engagement it stirs up. Brands feel its effects on their credibility. The corrector, meanwhile, gets no payment, no formal recognition, and none of the regulatory protection that the disclosure rules give influencers. This is the underlying condition we are putting a name to.

Unpaid does not have to mean unrewarded. Consumers may take on correction work because it earns them social capital, reputational value, moral satisfaction, or a sense of belonging in a digital community. A careful warning comment, a well-supported counter-recommendation, or a documented exposure of misleading claims can mark the corrector out as a knowledgeable and trustworthy member of the audience. So correction labor creates value for other consumers and for platforms, and on top of that it can hand the corrector symbolic rewards. The distinction is worth keeping in view. The labor is unwaged, but it can still be socially incentivized, and that is part of why it keeps going at scale even though no formal compensation is on offer.

6.3 Consumer due diligence as the individual-level form

At the individual level, correction labor takes the shape of consumer due diligence, namely checking claims, weighing source credibility, looking up other evaluations, and adjusting a purchase decision whenever something fails to add up. The term is chosen on purpose. Due diligence points to active investigation rather than passive reception, since the consumer has stopped treating influencer content as advice and has begun treating it as evidence, something to be examined before it is acted on.

Moving from skepticism to due diligence carries three analytical consequences. It leaves room for influence to keep working, since due diligence does not kill the purchase, it sets conditions on it. It also ties correction labor to a well-developed line of research on consumer information processing, verification behavior, and trust management in digital settings (McKnight & Chervany, 2001; Rousseau et al., 1998), as well as to the persuasion knowledge model (Friestad & Wright, 1994) that sits underneath most current skepticism research. And it lets us separate consumers who simply turn away from influencer content from those who put it through structured verification, a distinction with bite in both theory and practice.

In AI-mediated commerce, consumer due diligence covers a number of observable behaviors, including scanning comment sections for counter-evidence, hunting down independent reviews, asking around in peer networks (WhatsApp groups, Reddit threads, platform comment aggregators), comparing options across platforms, and picking out sponsorship signals or undisclosed conflicts of interest. None of this is mere skepticism. Taken together, these practices amount to a verification protocol that runs in parallel with the platform's own persuasion architecture.

It helps to be clear about how this departs from older accounts of consumer skepticism. Classical models treat skepticism as a fixed disposition that filters incoming persuasive messages. Consumer due diligence, by contrast, is an activity. It generates new information. It eats up time. And it leaves digital traces behind, such as saved comparison links, screenshots, comments dropped into threads, questions posted to peer groups. Each consumer who checks something in private also produces traces that the next consumer will lean on. That is how individual verification scales up into a market effect.

Table 1. How audience correction labor works.

Dimension	Consumer due diligence	Reverse influence
Level	Individual	Collective / market
Main action	Checking, comparing, verifying	Public correction and counter-signals
Trigger	Suspicion, staged authenticity, unclear sponsorship	Sponsored-content saturation and peer dissatisfaction
Space	Searches, comments, peer groups	Comments, stitches, Reddit, WhatsApp, Facebook groups
Effect	Delays, conditions, or prevents purchase	Reshapes credibility of brands/influencers
Why it is labor	Takes time and produces useful information	Creates unpaid market-level quality control
Theoretical anchor	Persuasion knowledge and consumer skepticism (Friestad & Wright, 1994; Obermiller & Spangenberg, 1998)	Electronic word-of-mouth (Bickart & Schindler, 2001; Hennig-Thurau et al., 2004; Cheung & Thadani, 2012)

Source: Authors.

7. Reverse Influence and the Market Effect of Correction Labor

Individual verification practices do not remain private for long. They build up in comment sections, group chats, stitched videos, and product threads where other consumers can come across them. Once they are visible to others, what began as private due diligence becomes something else, a credibility signal with effects at the level of the market. We call this reverse influence, the collective process by which audiences turn their reactions to influencer content into public credibility signals that reshape how legitimate influencers, brands, products, and platform recommendations appear to be.

A sponsored post does not stop moving the moment it is posted. It enters a second life, one in which comments, stitches, screenshots, peer warnings, and counter-reviews gather around it and rework what it means. By this stage the audience is doing more than reacting. It is rewriting the terms on which influence can succeed at all.

7.1 From individual verification to collective correction

Consumer due diligence captures what individuals do. It does not yet capture what happens when all those individual acts of verification pile up in public or semi-public spaces. The comments under a sponsored video, the stitched response, the screenshot passed around a WhatsApp group, the Reddit thread that gathers up bad experiences are not just records of private skepticism. They are inputs into a shared evaluative environment, and the next viewer runs into that environment before, during, or instead of the original promotional content.

Reverse influence marks a change in the direction the influence flows. In the classical model, influence runs from influencer to audience, with the audience as the endpoint. In AI-mediated commerce, that direction no longer holds steady. When audiences comment, expose contradictions, pass around bad experiences, or push back on exaggerated claims, they create consumer evaluations that travel alongside the promotional content, and at times ahead of it. The credibility, the visibility, and the commercial value of the original content now hang partly on what the audience does with it once it has been posted.

Reverse influence is neither the same as individual verification nor a form of anti-consumption. It is the point at which consumers stop being the endpoint of influencer persuasion and turn into active agents who redirect, contest, or cancel out influence inside AI-mediated commerce. The very comment section the algorithm reads as an engagement signal becomes a space where the audience pushes back against the content the algorithm is busy amplifying.

7.2 Spaces of reverse influence

Reverse influence clusters in spaces where audience reactions stay visible, searchable, or socially trusted. Among these are comment sections under sponsored posts, stitched and duet

videos that the algorithm distributes next to the original, product-specific Reddit threads, WhatsApp and Facebook groups, and dedicated de-influencing accounts. They vary in reach and in how long they last, yet they share three features. They are open to other consumers. They build up evidence in the form of screenshots and dated experiences. And they let later viewers read promotional content through the correction work that came before. This is what marks reverse influence off from ordinary eWOM. It is not just one more information channel; it is a corrective layer fastened to the promotional content itself.

7.3 Mechanisms of reverse influence

Three mechanisms run in parallel. The first is public correction, where a consumer flags an undisclosed sponsorship, an inaccurate claim, or a product that failed to perform as promised. The second is counter-recommendation, where a consumer puts forward an alternative product, often cheaper or longer-lasting, and explains why. The third is credibility challenge, where a consumer calls into question the influencer's expertise, the brand's track record, or the platform's enforcement of disclosure rules, frequently by pointing to a pattern that runs across several posts or several creators.

None of these mechanisms needs coordination. A single visible comment, when upvoted or pinned, can shift the interpretive frame through which later viewers read the promotional content. One stitched video with high engagement can go head to head with the content it responds to. One negative thread, once search surfaces it, can set a consumer's first impression of a product. And the asymmetry matters here, since producing reverse influence costs far less than producing the original promotional content, so that a fifteen-second comment can undo a sponsored campaign that took weeks to plan.

7.4 Reverse influence and the influencer's response

Reverse influence also changes where the influencer stands. Influencers more and more make their content knowing that the audience may respond publicly, and that this response will shape how credible the next sponsored post turns out to be. Out of this comes what we call anticipatory authenticity, namely pre-emptive disclosure, the explicit naming of products the influencer does not recommend, and visible engagement with negative comments. Whether or not such gestures actually restore credibility, they confirm one thing, that the influence relation is no longer one-directional. The audience has been built into the design of the content before any reaction even arrives. We offer anticipatory authenticity here as an emergent construct rather than a tested one. It is not formalized in the propositions below, since its antecedents and its actual effect on credibility remain open questions, and we flag it instead as a priority for future conceptual and empirical work.

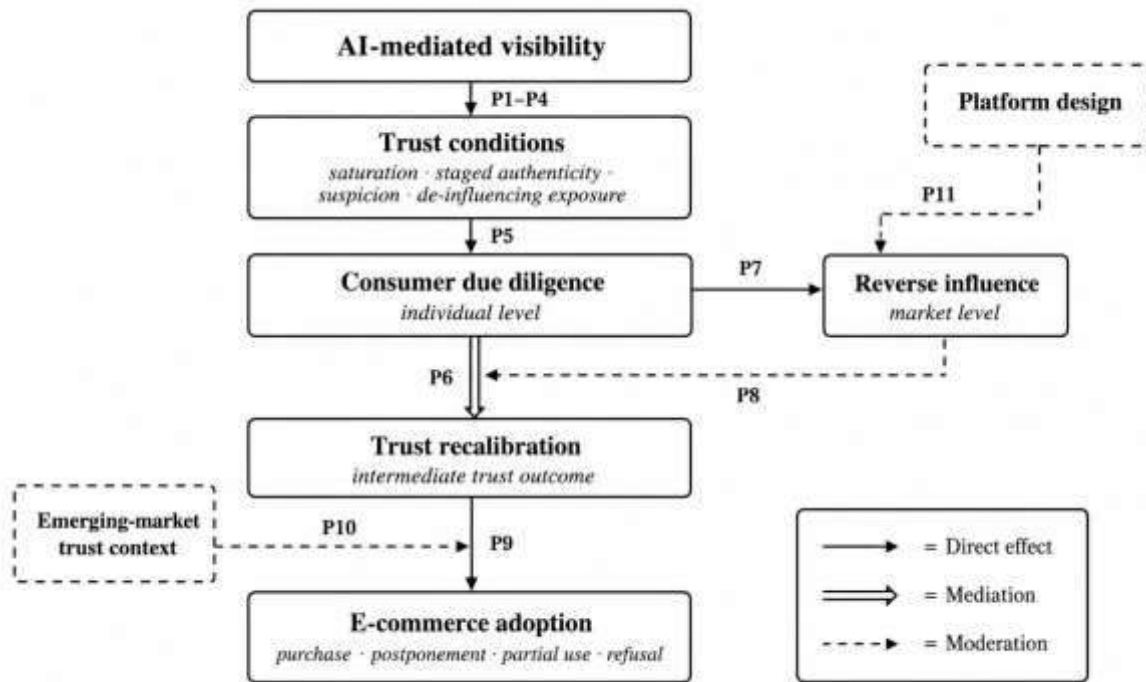
8. Conceptual Framework and Propositions

8.1 Framework overview

The framework connects the conditions of AI-mediated commerce to e-commerce adoption outcomes by way of four linked layers, namely AI-mediated visibility, trust conditions, audience correction labor (in both its individual and its market-level faces), and trust recalibration. Each variable is grounded in the literature reviewed above, and the relationships among them are set out as propositions in Section 8.2. Figure 1 lays out the full model.

The framework keeps the individual form of correction labor (consumer due diligence) apart from its market-level effect (reverse influence). Trust recalibration is the intermediate outcome that carries correction labor through to e-commerce adoption, tying private verification together with collective credibility signals.

Figure 1. Conceptual model linking correction labor, trust recalibration, and e-commerce outcomes.



Source: Authors.

8.2 Propositions

The propositions below are drawn from the framework and from the literature reviewed in the preceding sections.

The propositions can be read in four groups. P1 through P5, together with P7 and P9, state direct effects along the causal chain that runs from algorithmic visibility to adoption. P6 is a mediation. P8 and the platform-design proposition P11 are moderations. P10 is a contextual proposition specific to emerging markets; because it identifies the conditions under which off-platform correction labor most strongly shapes adoption, it functions as a contextual moderator, and Figure 1 accordingly groups it with the moderation relationships. Each proposition is labeled accordingly below, and every relationship in the figure is annotated with its proposition.

P1. AI-mediated algorithmic visibility increases consumer exposure to influencer-based product recommendations beyond the consumer's intentional information-seeking behavior.

P2. Influencer saturation, produced by algorithmic amplification, increases the consumer's perception of staged authenticity in promotional content.

P3. Perceived staged authenticity increases consumer suspicion toward promoted products and toward the credibility of endorsing influencers.

P4. Consumer suspicion increases the probability of de-influencing exposure, either through algorithmic recommendation of counter-content or through deliberate consumer search.

P5. De-influencing exposure is likely to activate audience correction labor in its individual form, consumer due diligence, including verification behavior, comparative evaluation, and peer consultation prior to purchase.

P6. Consumer due diligence mediates the relationship between consumer suspicion and trust recalibration, such that structured verification reduces the negative effect of suspicion on trust. (Mediation.)

P7. The accumulation of individual due diligence practices in public and semi-public spaces produces reverse influence, the market-level effect of correction labor, as audience reactions reshape the credibility of influencers, brands, products, and platform recommendations.

P8. Reverse influence is expected to moderate the relationship between de-influencing exposure and trust recalibration, such that strong reverse influence within a peer network strengthens the effect of de-influencing exposure on trust adjustment. (Moderation.)

P9. Trust recalibration influences e-commerce adoption outcomes, namely purchase, postponement, partial use, or refusal, in proportion to the degree to which audience correction labor produces positive or negative evaluative outcomes.

P10. In emerging-market contexts where formal institutional trust is weak, off-platform correction work, through private messaging groups, kinship networks, and informal peer channels, is expected to exert a stronger effect on e-commerce adoption than equivalent correction activity performed on-platform. (Contextual proposition.)

P11. Platform design moderates the market-level effect of correction labor, such that design choices that amplify corrective content, such as visible comment ranking and the surfacing of critical threads, strengthen the effect of reverse influence on trust recalibration, while choices that suppress it, such as comment filtering and the down-ranking of counter-content, weaken that effect. (Moderation.)

Three of the propositions call for comment. At the individual level, P6 is the most consequential of all. Its mediation claim leans on McKnight and Chervany's (2001) distinction between trusting intentions and trusting beliefs, and on the weight Rousseau et al. (1998) put on experience-based trust calibration. The market-level effect comes in with P7 and P8. P7 picks up the move from individual verification to public credibility signals, since protective activities also leave traces that other consumers then use as inputs to their own credibility judgments. P8 casts reverse influence as a moderator, such that wherever peer networks already host strong reverse influence, through active comment sections, scam-watch threads, and established critical accounts, the effect of de-influencing exposure on trust recalibration should be larger. P10 is the proposition that the Moroccan discussion in Section 9.4 speaks to most directly. Where institutional trust mechanisms are weak, the informal correction work moving through kinship-anchored peer channels carries an outsized weight in shaping adoption decisions, and that is a claim future research can put to the test.

These propositions also have limits, since correction labor does not always build trust. First, the trust-repairing effect assumed in P5 and P6 depends on the correction being accurate; erroneous or malicious correction, including coordinated review bombing, can damage credible sellers and degrade rather than improve the information environment. Second, the effects are weaker where consumers lack the persuasion knowledge and digital literacy needed to tell sound correction from noise (Obermiller & Spangenberg, 1998), so the framework presumes a baseline of evaluative competence. Third, the moderating role of reverse influence in P8 and P11 can reverse sign when correction is dismissed as untrustworthy in its own right, which is more likely where dispositional or institutional distrust is high (McKnight & Chervany, 2001). These conditions mark the limits within which the propositions are expected to behave as stated, and each is itself a candidate for empirical test.

9. Discussion

9.1 Implications for brands and influencer marketing practice

For brand managers and influencer practitioners, the practical implication hinges on how de-influencing gets read. Read only as anti-consumption behavior, it invites suppression, that is, holding back counter-content, pushing harder on sponsored messaging, or walking away from the channel. Read instead as audience correction work, the unpaid evaluative activity that now shapes how trust is built in AI-mediated commerce, it calls for something else. Brands and influencers have to make sure their partners' claims are accurate, that sponsorships are disclosed, that product quality can hold up under scrutiny, and that the influencer's credibility

is solid enough to weather both individual investigation and market-level challenge. We call this orientation verification-readiness. This reading follows directly from P5 through P9, where due diligence and reverse influence are what convert consumer suspicion into trust recalibration and, in turn, into adoption.

Penttinen, Ciuchita, and Pedersen (2026) find that de-influencing raises an influencer's trustworthiness, at least where it is not contaminated by an adjacent sponsored post, and this suggests that brands willing to allow and even amplify honest critique may gain from the credibility it lends to the influencers who take it on. The finding lines up with the broader literature on the paradox of transparency, where disclosing conflicts of interest, counter to intuition, tends to raise trust rather than lower it (Evans et al., 2017; De Veirman, Cauberghe & Hudders, 2017; Khalil et al., 2025). What shifts under algorithmic commerce is that disclosure is no longer the only signal the audience reads. They also read the correction work gathered around the influencer, the influencer's handling of negative comments, and how consistent the recommendations stay over time.

At the collective level, reverse influence concentrates in identifiable spaces, comment sections, response videos, scam-watch threads, and peer groups, which leaves brands and influencers with both the opportunity and the obligation to engage with those spaces in the open. Engaging here is not the same as suppressing comments. It means responding where people can see it, acknowledging concerns that are legitimate, and correcting the product when the criticism is fair. The cost of ignoring reverse influence is not one lost negative comment. It is the slow build-up of an evidentiary record that the next consumer will run into before ever running into the brand.

9.2 Implications for e-commerce adoption

For e-commerce adoption, one implication is hard to miss. Conventional adoption models lean on platform usability, price, convenience, and perceived ease of use. What the framework adds is harder to see, namely the credibility conditions that audience correction labor brings about. When consumers in digital commerce decide whether to buy online, they go not only on platform features but also on whether the trust cues around a promoted product hold up once they look closely.

Two consequences follow from this. The first is that strategies pushing e-commerce adoption through better platform design alone may miss a large part of the trust problem. If consumers are filtering platform recommendations through correction labor, whether their own or that of others, then improving only the promotional layer means improving just one side of a two-sided decision. The second is that adoption outcomes are not binary. The framework anticipates delayed adoption, where consumers buy only after long verification; selective adoption, where they buy in some product categories but not others, depending on how much correction labor is around; and resistance, where they refuse to buy online in categories whose correction signals stay consistently negative. These graded outcomes are what P9 formalizes, and P10 conditions them on context.

For empirical research, these distinctions count. Adoption studies that record only purchase versus non-purchase pass over the more common middle-ground outcomes, the ones where correction labor leads to conditional or partial adoption. Research that takes up the framework ought to spell out which adoption outcome it is measuring, and ought to treat the strength and availability of correction labor in the relevant product category as an explanatory variable, set beside the more familiar antecedents of adoption.

9.3 Implications for platform design

Platform design is caught up in this as well. Algorithmic visibility is not neutral; it shapes trust. Platforms that optimize for engagement alone, with no thought for what saturation and staged authenticity do to trust, risk building the very conditions for their own commercial decline.

Examining how consumers respond to AI-generated advertising content, Jiang, Zheng, and Kong (2025) find that transparency about algorithmic provenance is a precondition for any lasting consumer trust in AI-mediated commerce. De-influencing is what fills the gap when that transparency is missing.

Reverse influence makes the platform design problem thornier still. The very engagement signals the algorithm rewards, comments, replies, stitches, are now what audiences use to push back against the content being amplified. A platform that reads all engagement as positive risks amplifying critical content right alongside the promotional content it attacks, and over time that may erode the platform's commercial value for its brand partners. A platform that suppresses critical engagement, on the other hand, risks damaging the very credibility infrastructure on which consumer trust rests. Neither path settles the tension cleanly. What a platform can do is stop reading engagement signals as uniformly positive, and that distinction is where any serious design response has to start. From a governance perspective, audience correction labor functions as decentralized verification, in which consumers informally perform, at scale, an audit function that platforms and regulators do not reliably provide. Read this way, reverse influence is an emergent informal-governance mechanism, and its effects on market efficiency and trust merit attention even in places where formal disclosure regimes are already in force. P8 and P11 state these platform-level moderations formally, the first for the strength of reverse influence and the second for the design choices that amplify or suppress it.

9.4 Emerging markets and the Moroccan context

The framework set out here is global in its theoretical reach. Its implications, though, show up most clearly in places where digital commerce is still earning its consumers' trust. Morocco makes a useful illustrative case. E-commerce adoption in Morocco has grown a great deal over the past decade, yet it stays held back by trust deficits, by worries about product quality, delivery reliability, and the authenticity of online recommendations. Secondary indicators bear this out. Morocco counted 35.3 million internet users in January 2025, corresponding to an internet penetration rate of 92.2% (DataReportal, 2025). At the same time, e-commerce adoption remains more limited than general internet access. Online shoppers accounted for around 21.6% of Morocco's population in 2025, while cash-on-delivery remained the preferred payment method for online shopping in 2023, accounting for 58% of e-commerce purchases (Ou, 2025). These figures suggest that the expansion of digital access has not fully removed trust and payment-related concerns in online commerce. Persistent reliance on payment at, or close to, delivery can be read as an expression of limited trust in remote online exchange, the same trust deficit that off-platform correction work addresses on the credibility side. Against this backdrop, the correction work the framework describes is hardly a new behavior at all. It is older habits of social verification carried over into digital channels.

In Morocco, as in comparable emerging markets, audience correction labor may matter all the more because so much of the credibility work around online buying takes place off-platform, through WhatsApp groups, Instagram comments, Facebook groups, family recommendations, and warnings between peers. These informal trust channels can move e-commerce adoption even while they stay invisible to platform analytics. The visible engagement economy does not always pick up the real channels through which consumers settle the question of whether to buy online. P10, introduced back in Section 8.2, puts this hypothesis into a testable form.

Read in this light, audience correction labor can be seen as the digital counterpart of the cash-on-delivery practices that are still widespread in Moroccan e-commerce. Cash-on-delivery lets a consumer cut risk by holding off payment until the product is physically in hand; off-platform correction labor lets a consumer hold off trust until the claims have been checked through peers, comments, screenshots, or earlier experiences. Both practices answer the same underlying need, the consumer's wish to reduce uncertainty before fully committing to an online transaction. So

the staying power of cash-on-delivery as a payment method in Morocco can be read as a symptom of the same trust-deficit problem that correction work goes some way toward addressing on the credibility side.

What this means in practice is that the right dependent variable in this setting is not influencer purchase intention but e-commerce adoption itself, that is, whether the consumer accepts, postpones, partly uses, or refuses online purchasing in a given category. On the framework's logic, correction labor in kinship-anchored peer channels looks like an especially decisive trust signal in categories where the perceived risk of buying is high. A cultural dimension is worth flagging too. In markets where interpersonal trust has long counted for more than institutional trust, the move from one-way persuasion to two-way correction may feel less like a disruption and more like a continuation of habits that already exist. The framework's two-level structure, with correction labor as due diligence at the individual level and as reverse influence at the market level, maps onto the way digital consumers in emerging markets already pair private investigation with peer consultation.

None of this is empirical yet. Work that draws on Moroccan samples, survey-based testing of the framework's propositions, qualitative investigation of correction practices inside WhatsApp and Facebook groups, or comparative studies across emerging-market settings would do a great deal to advance both the testing of the framework and the wider literature on e-commerce adoption in trust-constrained environments.

9.5 Limitations and directions for future research

The limitations fall into three kinds. The first are limitations of design. The framework is conceptual, so its propositions await empirical testing across consumer segments, cultural contexts, and product categories. The second concern theoretical scope. The framework treats reverse influence as a collective phenomenon without modeling the network structures along which it travels, sets the role of regulation largely to one side, and does not yet specify when accurate correction gives way to mistaken or strategic correction. The third concern contextual generalization. The framework is illustrated mainly through an emerging-market lens, and its weighting of interpersonal over institutional trust may not transfer cleanly to mature markets, so comparative testing is needed to mark its boundary conditions. The paragraphs that follow develop each of these in turn.

One limitation in particular calls for attention. Audience correction labor should not be romanticized. The same mechanisms that let consumers expose misleading claims can just as easily produce inaccurate correction, review bombing, reputational harm, or call-outs driven by misinformation. Unpaid auditors do not always work from complete information, and their corrective activity is not held to the verification standards that govern formal regulation or professional auditing. Correction work can repair trust in algorithmic commerce, but it can also unsettle trust when the correction itself proves wrong. Future research should keep accurate correction, mistaken correction, and strategic negative influence apart from one another, and should look into the conditions under which each one comes to dominate.

Correction work raises questions about platform design as well. Platforms gain from the engagement that audience correction stirs up, yet critical engagement may eat into the commercial value of sponsored content. Future research should look at when platforms amplify, tolerate, or cut back the visibility of such correction work. This work also creates hidden economic value, since it reduces information asymmetry and throws off informal quality-control signals, even if its exact size stays hard to pin down.

Mixed-method designs, ones that pair survey-based quantitative testing with qualitative investigation of verification practices, would be well suited to capturing how trust recalibration unfolds in algorithmic environments. Comparative studies across mature and emerging digital commerce markets would help mark out the boundary conditions of the framework.

Longitudinal work that tracks how reverse influence develops within particular product categories would help separate stable correction practices from episodic backlash. And platform-level studies could examine how recommendation systems respond to, and in turn reshape, the reverse influence dynamics described here.

10. Conclusion

De-influencing is not well captured by the label of anti-consumption behavior. It is closer to a form of unpaid evaluative work, audience correction labor, that consumers carry out under the conditions of AI-mediated commerce. This work shows up on two levels. As an individual practice, it runs as consumer due diligence, the verification, comparison, and peer consultation that come before a purchase. As a market effect, it runs as reverse influence, where audience reactions gather in public and semi-public spaces and feed back into the market, reshaping how credible influencers, brands, products, and platform recommendations appear.

The article makes three contributions. The first is conceptual. It introduces audience correction labor as a construct that names the unpaid evaluative work consumers now perform around influencer content, gives it an operational definition and two dimensions, and situates it within the digital-labor tradition while marking what sets it apart, namely its corrective, regulatory, and unpaid character.

The second is integrative. By treating consumer due diligence at the individual level and reverse influence at the market level as two faces of one activity, the article bridges the persuasion-knowledge and skepticism literatures, on one side, and the electronic word-of-mouth tradition, on the other, which have largely developed apart.

The third is the framework itself. It links AI-mediated visibility, trust conditions, correction labor, and trust recalibration to e-commerce adoption, and expresses these links as eleven testable propositions, so that de-influencing can be studied as trust-building work rather than as anti-consumption.

For practice and policy, the framework carries a clear message. Brands and influencers gain more from verification-readiness, from accurate claims, visible disclosure, and open engagement with criticism, than from suppressing corrective content. Platforms face a sharper design choice, since reading all engagement as positive amplifies critique alongside promotion, while filtering critique erodes the credibility infrastructure consumers rely on; surfacing accurate correction while curbing manipulation is the more sustainable course. Regulators, for their part, may come to treat informal audience correction as a complement to formal disclosure regimes rather than as noise to be managed.

The propositions now need to be tested across consumer segments, cultural contexts, and product categories. Morocco, alongside comparable emerging-market settings, offers a particularly useful testing ground, since off-platform correction labor plays a central role in shaping e-commerce adoption precisely where formal institutional trust is at its weakest. Studying how that correction work operates in such settings can sharpen our grasp of how trust is built, adjusted, and at times broken in AI-mediated commerce, and of how the unpaid evaluative work of ordinary consumers has become part of the infrastructure through which digital markets function.

Four research directions follow directly from this framework. First, a survey-based test of the propositions using validated scales for perceived staged authenticity, consumer due diligence, reverse influence, and trust recalibration, with adoption measured as a graded rather than binary outcome. Second, a network study that traces how corrective content travels across comment sections, stitched videos, and peer groups, so that the diffusion assumed in P7 can be observed directly. Third, a design experiment that varies whether a platform amplifies or suppresses corrective content, providing a direct test of P11. Fourth, comparative fieldwork in emerging

markets, Morocco among them, examining off-platform correction in WhatsApp and Facebook groups in order to test P10 against on-platform activity.

References

- (1). Abidin, C. (2016). Visibility labour: Engaging with influencers' fashion brands and #OOTD advertorial campaigns on Instagram. *Media International Australia*, 161(1), 86–100. <https://doi.org/10.1177/1329878X16665177>
- (2). Audrezet, A., de Kerviler, G., & Guidry Moulard, J. (2020). Authenticity under threat: When social media influencers need to go beyond self-presentation. *Journal of Business Research*, 117, 557–569. <https://doi.org/10.1016/j.jbusres.2018.07.008>
- (3). Bickart, B., & Schindler, R. M. (2001). Internet forums as influential sources of consumer information. *Journal of Interactive Marketing*, 15(3), 31–40. <https://doi.org/10.1002/dir.1014>
- (4). Boerman, S. C., Willemsen, L. M., & Van Der Aa, E. P. (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, 82–92. <https://doi.org/10.1016/j.intmar.2016.12.002>
- (5). Cheung, C. M. K., & Thadani, D. R. (2012). The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decision Support Systems*, 54(1), 461–470. <https://doi.org/10.1016/j.dss.2012.06.008>
- (6). DataReportal. (2025). Digital 2025: Morocco. Kepios. <https://datareportal.com/reports/digital-2025-morocco>
- (7). De Veirman, M., Cauberghe, V., & Hudders, L. (2017). Marketing through Instagram influencers: The impact of number of followers and product divergence on brand attitude. *International Journal of Advertising*, 36(5), 798–828. <https://doi.org/10.1080/02650487.2017.1348035>
- (8). Duffy, B. E. (2017). (Not) getting paid to do what you love: Gender and aspirational labor in the social media economy. Yale University Press.
- (9). Eisend, M., & Tarrahi, F. (2022). Persuasion knowledge in the marketplace: A meta-analysis. *Journal of Consumer Psychology*, 32(1), 3–22. <https://doi.org/10.1002/jcpy.1258>
- (10). Evans, N. J., Phua, J., Lim, J., & Jun, H. (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>
- (11). Friestad, M., & Wright, P. (1994). The Persuasion Knowledge Model: How people cope with persuasion attempts. *Journal of Consumer Research*, 21(1), 1–31. <https://doi.org/10.1086/209380>
- (12). Gefen, D., & Straub, D. W. (2000). The relative importance of perceived ease of use in IS adoption: A study of e-commerce adoption. *Journal of the Association for Information Systems*, 1(1), Article 8. <https://doi.org/10.17705/1jais.00008>
- (13). George, A., Shibu, M., Joseph, E. T., & Sunny, P. (2025). Impact of social media influencer marketing on customer purchase intention in the fashion industry: A systematic literature review. *Frontiers in Communication*, 10, 1676901. <https://doi.org/10.3389/fcomm.2025.1676901>
- (14). Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, 18(1), 38–52. <https://doi.org/10.1002/dir.10073>

- (15). Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, 15(4), 635–650. <https://doi.org/10.1086/266350>
- (16). Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, 10(1–2), 18–26. <https://doi.org/10.1007/s13162-020-00161-0>
- (17). Jiang, S., Zheng, W., & Kong, H. (2025). Trust or skepticism? Unraveling the communication mechanisms of AIGC advertisements on consumer responses. *Journal of Theoretical and Applied Electronic Commerce Research*, 20(4), 339. <https://doi.org/10.3390/jtaer20040339>
- (18). Khalil, T., Mahmood, C. K., Singh, H., Aftab, F., Khan, H., Manzoor, I., & Abdulhayee, K. A. (2025). The digital de-influencing wave: Redefining trust in online communities. *Frontiers in Communication*, 10, 1600657. <https://doi.org/10.3389/fcomm.2025.1600657>
- (19). Kim, D. H., Lau, S. Y., & Lee, S. (2025). The skepticism equation in influencer marketing: The interplay of influencer–product fit and content style in shaping ad skepticism and purchase intention. *Asian Journal of Communication*, 35(4), 286–305. <https://doi.org/10.1080/01292986.2025.2497504>
- (20). 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>
- (21). MacInnis, D. J. (2011). A framework for conceptual contributions in marketing. *Journal of Marketing*, 75(4), 136–154. <https://doi.org/10.1509/jmkg.75.4.136>
- (22). Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business Horizons*, 52(4), 357–365. <https://doi.org/10.1016/j.bushor.2009.03.002>
- (23). McKnight, D. H., & Chervany, N. L. (2001). What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology. *International Journal of Electronic Commerce*, 6(2), 35–59. <https://doi.org/10.1080/10864415.2001.11044235>
- (24). Nazara, D. S., Nugrahaningsih, H., & Abdillah, F. (2025). The future of influencer marketing: Trust, authenticity, and consumer behavior in the social media age. *Jurnal Informatika Ekonomi Bisnis*, 7(2), 345–349. <https://doi.org/10.37034/infec.v7i2.1138>
- (25). Obermiller, C., & Spangenberg, E. R. (1998). Development of a scale to measure consumer skepticism toward advertising. *Journal of Consumer Psychology*, 7(2), 159–186. https://doi.org/10.1207/s15327663jcp0702_03
- (26). Ou, X. (2025, December 17). E-commerce in Morocco – statistics & facts. Statista. <https://www.statista.com/topics/9376/e-commerce-in-morocco/>
- (27). Penttinen, V., Ciuchita, R., & Pedersen, S. (2026). Let me deinfluence you! The impacts of deinfluencing content. *Journal of the Academy of Marketing Science*. Advance online publication. <https://doi.org/10.1007/s11747-026-01141-9>
- (28). Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review*, 23(3), 393–404. <https://doi.org/10.5465/amr.1998.926617>
- (29). Sardar, S., & Vijay, T. S. (2025). Social media influencers and purchase intention: A review and research agenda. *International Journal of Consumer Studies*, 49(3), e70046. <https://doi.org/10.1111/ijcs.70046>
- (30). Terranova, T. (2000). Free labor: Producing culture for the digital economy. *Social Text*, 18(2/63), 33–58. https://doi.org/10.1215/01642472-18-2_63-33
- (31). Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>

- (32). Wiratha, A., & Kemalasari, N. (2025). The impact of de-influencing on social media on Gen Z's purchase intention: The mediating role of source credibility. *Jurnal Informatika Ekonomi Bisnis*, 7(4), 768–773. <https://doi.org/10.37034/infeb.v7i4.1292>
- (33). Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.