Persona Marketing on Social Media: The Evolution from Constructed Authenticity to Consumer Trust

Runqi Wang[a]

[a]School of Business and Management,Queen Mary University of London, UK

Abstract

Purpose – The purpose of this paper is to investigate the evolution of persona marketing on social media, specifically how brands and influencers move from staging constructed authenticity to building long-term consumer trust. The study aims to provide a conceptual framework that clarifies this transition and explores the impact of emerging technologies such as generative AI and virtual influencers on trust dynamics.

Design/Methodology/Approach – Based on a literature synthesis approach, this research draws from 58 peer-reviewed sources across marketing, media studies, psychology, and digital communication. The paper analyzes theoretical trends and case studies to identify three interlinked phases in persona marketing: constructed authenticity, performative persona management, and trust-based engagement.

Findings – This paper studies how persona marketing evolves from emotionally staged performances to sustained trust relationships shaped by algorithmic visibility, parasocial interaction, and ethical branding. It highlights that while authenticity opens the door to engagement, trust becomes the primary asset in long-term consumer-brand relationships.

Research Implications – In the management of digital branding, the findings suggest a strategic shift: marketers must move beyond performative authenticity and focus on designing trustworthy, emotionally consistent personas. In an era of AI-driven content, transparency, ethical standards, and platform-aware design are essential for maintaining consumer confidence.

[a] First Author, E-mail: bzxywyn@163.com

I. Introduction

Digital technology now permeates almost every aspect of daily life, and this shift has fundamentally altered consumer culture, changing the ways brands and individuals converse, form relationships, and attach meaning to products and services. In this fresh landscape, persona marketing has moved to the front of social media, allowing companies and influencers to create and nurture polished, lifelike identities that engage audiences on personal, emotional, and psychological grounds. In contrast to older campaigns that spotlighted product specifications and pushed messages in a one-way flow, today's strategy leans on story, interaction, and an aura of genuineness to create emotional ties and foster lasting trust (Marwick & boyd, 2011).

This rise of persona marketing tracks closely with the growth of popular social-media platforms like Instagram, TikTok, YouTube, and Douyin. Each site invites users not only to showcase products but also to perform identities. Selfies, vlogs, story updates, livestreams, and algorithmic boosts turn both independent creators and corporate accounts into narrative curators, sharing snippets, values, and feelings that feed into a broader branding plan (Abidin, 2016; Duffy, 2017). The logo stops being a static mark and starts acting like a living character. Such a change signals a fundamental move away from brand-controlled communication toward engagement that centres on the persona itself.

A central characteristic of persona marketing is its reliance on what scholars call constructed authenticity. In an effort to look trustworthy and approachable, brands and social-media influencers routinely highlight small flaws, candid moments, and behind-the-scenes clips that imitate the shape of everyday life. Yet Gaden and Dumitrica (2015) point out that this mask of honesty is usually skillfully arranged, aimed first and foremost at boosting relatability, driving engagement, and opening fresh revenue streams. The contradiction follows: the tighter the strategy, the more the audience expects the result to feel loose, off-the-cuff, and unfiltered in order to work. Under these conditions, authenticity shifts emphasis from factual accuracy to the far slipperier ground of emotional credibility.

Performative authenticity thus becomes an entry point to the more lasting asset of consumer trust. In today's crowded online market, where advertisements blend with user-generated posts, skepticism runs high and genuine trust feels rare yet extraordinarily valuable. The 2023 Edelman Trust Barometer shows nearly seventy percent of shoppers say they are likelier to purchase from brands they see as honest and clear about their practices. What counts as honesty, however, has shifted; it hinges less on corporate authority and more on consistent stories, heartfelt delivery, and emotional steadiness kept over months and years (Labrecque, 2014; Urban et al., 2000). Seen this way, persona marketing stops being just a visibility hack and turns into a psychological tool for building trust in an economy that trades principally in attention.

At the same time, the way these digital personae are shaped and sustained depends heavily on what the platform makes possible-its design features, community norms, and hidden ranking signals that nudge visibility up or down. On TikTok, for example, the algorithm favours clips with quick, strong emotional hits, while Instagram Stories reward regular, private-seeming check-ins. The persona therefore arises not just from an authors creative wish but from a messy interplay of audience reactions, platform logic, and the ever-shifting

rhythm of algorithmic visibility (Bucher, 2012; Marwick, 2015).

Another key force reshaping persona marketing today is the rise of parasocial relationships- one-sided emotional bonds audiences form with media figures they perceive as approachable, familiar, and trustworthy. First described by Horton and Wohl in the 1950s, this phenomenon has intensified in the social-media era, where platforms imitate the give-and-take of conversation. Users can comment, react, and slide into DMs with influencers or brand pages, creating the illusion of closeness and real investment even when no genuine two-way exchange occurs (Chung & Cho, 2017). These quasi-friendship ties matter; when nurtured with care, they can spark stronger loyalty, drive purchases, and turn followers into passionate brand advocates (Jin & Ryu, 2020).

At the same time, the lines around trust and authenticity are being further obscured by artificial personas-AI-powered influencers and digitally stitched brand avatars that appear human but are pure code. Virtual stars such as Lil Miquela and Imma now boast millions of fans despite never stepping off a studio floor. While some shoppers greet these characters as playful novelties or a peek into tomorrows marketing, others worry about the ethics and emotional fallout of believing in something that doesnt breathe (Moustakas et al., 2020). As generative AI gets sharper and cheaper, it forces brands and scholars alike to revisit what authenticity means and what, in turn, future consumers will consider worthy of their trust.

Scholarship on influencer culture, digital branding, and AI-mediated communication has grown rapidly, yet the insights it produces still sit in separate disciplinary silos. Marketing journals track how influencers drive consumer engagement and lift brand equity; media studies analyse the way online celebrities shape personal and collective identities, and psychologists investigate the trust that develops through one-way, or parasocial, interactions. What these conversations lack is a unifying theory that maps the journey from staged authenticity to trust-centered influence.

To fill that void, the present paper pulls together evidence from marketing, communication, psychology, and digital sociology. The review shows that persona marketing moves through three overlapping phases: (1) constructed authenticity, (2) performative persona management, and (3) trust-based engagement. Although researchers often treat these phases as distinct, the evidence reveals them instead as different lenses on the same changing process. The study finally discusses how platform features, audience demands, and shifting technologies continuously reconfigure the stages and the meanings attached to digitally curated selves.

By reframing authenticity as a tool rather than a final goal, this study treats trust as the real prize in effective persona marketing. It shows that modern consumers value transparent, steady, and emotionally thoughtful voices far more than faultless polish, provided those voices match their own beliefs and daily realities. With this insight, the research offers fresh theoretical ground for scholars while also giving todays marketers clear, usable strategies for the AI-driven, post-authentic world of online communication.

II. Literature Review

Persona marketing, a distinctive strand of digital branding, draws on insights from marketing theory, media studies, social psychology, and digital ethnography. As social media platforms have reshaped everyday communication, the way brands and influencers build and showcase their identities has grown more visible-and more consequential for audiences. This section surveys core scholarship along three lines: (1) the theoretical roots of persona work, (2) the deliberate performance of authenticity, and (3) the shift toward trust-based engagement.

2.1 Theoretical Origins: Persona as Performance

The conceptual backbone of persona marketing rests on Erving Goffmans (1959) classic dramaturgical lens. In The Presentation of Self in Everyday Life, Goffman likens social life to a stage where setting, audience, and assigned roles shape each actors display. People polish a front-stage image to meet expectations while tucking away candid moments into the backstage. In social media s logic, however, those curtains fray, blending public and private and complicating even the most casual self-representation.

Platforms such as Instagram, TikTok, and YouTube push users-brands and everyday posters alike-to keep up an ongoing show for a mixed crowd. Hogan (2010) builds on Goffman to say that online presenters do two jobs at once: they perform and they curate. Unlike face-to-face talk, posting online means picking, cropping, and scheduling material that stays up and can be watched again and again. Because each piece is polished and permanent, the carefully built self starts to act like a kind of digital currency, or what van Dijck (2013) calls "digital capital."

Things get trickier when what Marwick and boyd (2011) call context collapse kicks in. On any given feed, friends, ex-classmates, co-workers, fans, and strangers all show up together, swirling the usual boundaries. For influencers and brands, that mix forces them to draft a single, consistent image that keeps every group happy yet still reads as real. The result is less casual, more planned: every like, share, and hashtag are weighed in advance to squeeze out maximum attention and cultural buzz.

From a branding viewpoint, recent developments echo the rise of the "branded self"-a phrase Hearn (2008) used to show how people turn their character, daily habits, and beliefs into a revenue stream. For influencers, the cultivated persona acts as both merchandise and pledge. Every post, reply, and paid partnership either reinforces or undermines that promise. Keeping the image steady and attractive adds to what Duffy (2017) terms "aspirational labor"-the unseen, often unpaid effort needed to keep a sellable self in view.

Brands have borrowed these tactics for their own aims. To seem more human, companies now imitate influencer habits, using casual language, sharing behind-the-scenes glimpses, and even giving their accounts catchable traits (Khamis, Ang, & Welling, 2017). Whether it is Wendy's cheeky tweets or Duolingos lively TikTok skits, businesses craft showy identities meant to make customers feel closer and more relatable.

Yet platforms do more than simply mediate human expression; their design norms actively sculpt what

counts as performance online. Features like image filters, text captions, threaded replies, and real-time livestreams configure discrete channels through which acts of publicity unfold. Bucher and Helmond (2018) remind us that recommendation algorithms then sift these channels, elevating content that crosses preset engagement thresholds and, in turn, granting greater visibility to the personas behind that content. Visibility emerges as a collaborative effect, the product of technical affordances, audience responses, and shifting platform etiquette.

The performative logics of persona marketing also resonate with symbolic interactionism, which views identity as a constantly negotiated outcome of social encounters. Within this framework, a circulating economy of likes, comments, and shares delivers instant forms of validation, quietly rewarding certain aesthetics while discouraging others. Gradually, this rhythm hardens into a recognisable personal brand, a construct that moulds user behaviour even as it mirrors audience expectations (Kozinets, de Valck, Wojnicki, & Wilner, 2010).

Although carefully devised and managed, digital personas often feel genuinely real to audiences chiefly because they use familiar social formats-conversational language, steady storytelling, and visible emotion. Persuasiveness increases when these performances match typical cultural roles for example, the friendly millennial, the down-to-earth neighbor, or the ambitious founder. In line with Thompson and Haytko, a brands cultural standing rests less on verifiable fact than on how coherently its symbols and stories work together.

Taken together, the academic roots of persona marketing lie in performance theory, identity studies, and branding discourse. Whether created by individuals or by firms, the digital persona functions as a strategic tool-always performed, constantly polished, and judged inside the native logic of each platform. This ongoing act of presentation and adjustment sets the stage for building what later sections will call authenticity.

2.2 Constructed Authenticity: Simulated Realness in the Age of Influence

Today, marketers consider authenticity the beating heart of social-media outreach, because users increasingly value exchanges that feel emotionally honest. In reality, however, persona marketing seldom springs from genuine spontaneity; it emerges instead from painstaking design-a staged act engineered to look effortless and real. Termed constructed authenticity, this uneasy truth reveals a central irony of online life: acceptance as \"authentic\" requires a performance so convincing that viewers never suspect artifice is at work.

Gaden and Dumitrica (2015) argue that such authenticity on platforms like Instagram or YouTube behaves less like a fixed trait and more like a repeatable performance. Influencers braid story arcs, behind-the-curtain shots, raw confessions, and relaxed slang to weave an aura of closeness and approachability. Together these techniques \"humanize the brand\" and create the illusion of moral warmth-even when viewers know the post was paid and planned.

Researchers argue that the authenticity people see online is not a flat-out lie; it simply tracks how audience taste changes over time (Abidin, 2016; Audrezet, de Kerviler, & Moulard, 2018). After interviewing dozens of fashion influencers, Audrezet and her co-authors map out three clear ways creators stake their credibility: first, moral authenticity, in which they highlight personal values and causes; second, passionate authenticity, where

raw enthusiasm-about, say, vintage jackets-shines through; third, transparent authenticity, in which sponsored posts and partnerships are flagged up-front. Though each leans on a different feeling trust, loyalty, attachment they all aim to craft a believable online self.

That balancing act between performance and honesty shows up loud and clear in the way content looks. For years platforms such as Instagram have nudged users toward perfectly staged, high-gloss photos, creating what Abidin (2016) calls aspirational relatability-a style that makes an influencer feel familiar while still suggesting wealth, travel, and know-how. Yet even within that polished feed creators sneak in "imperfections" messy hair, haste-filter selfies, or quick notes about anxiety-or burnout in order to keep inspiration from sliding into fantasy.

TikTok and similar short-form video apps now celebrate a rough, handheld look, where messy authenticity-wobbly edits, jokes that fall flat, and barely scripted clips-feels more genuine than the slick production once favored (Keller, 2022). Many younger viewers have started to see high-gloss branding as a sign of manipulation, so persona marketers must tune in to these new platform rules and age-specific moods if they want to succeed. Because audience tastes change so fast online, cultural attunement is no longer a bonus; in short-form video, it is the price of admission.

The economic part of manufactured authenticity adds another layer of complexity to the model. Duffy (2017) shows that many female digital workers still juggle the need to appear genuine while also making money, a balancing act she illustrates in her interviews with todays influencers. Their audience wants spontaneity, yet the same audience also expects posts to comply with brand guidelines, hit platform algorithms, and generate revenue for the creator. Pulling this off demands not only emotional control but also hours of editing, careful self-revelation, and a persistent awareness of emerging trends, all of which push the line between private moments and a polished public image ever closer together.

Put more simply, manufactured authenticity is less a paradox than a tool that lets online figures build credibility and draw in followers on an emotional level. Creators stage this impression through the words they write, the photos they share, the comments they answer, and the small glimpses of honesty they offer, yet every element must be adjusted as audience moods and site cultures shift. Even with all that work, honesty counts for only part of the story-trust is the bigger prize that persona-based marketing really aims to win and keep over years. The following section looks at the steps by which these crafted performances turn into lasting trust and explains why, in todays digital marketplace, trust has become the most valuable form of currency.

2.3 Trust-Based Engagement: From Performance to Psychological Commitment

Within todays fast-moving social-media marketplace, trust has become the rarest and most important asset a brand tones calories willing to invest in online persona and image-building. Authenticity may kick-start conversation by showing a hint of humanity, yet it is durable, credible trust that keeps viewers emotionally involved over months and years. Different from conventional advertising, where trust flowed from corporate pedigree or product pedigree, trust in persona-marketing rests on steady emotional tone, shared values, and believable sincerity (Urban, Sultan, & Qualls, 2000).

2.3.1 Parasocial Trust: Intimacy Without Reciprocity

Even online, trust remains relational, shaped by what Horton and Wohl (1956) once called parasocial interaction-a one-way attachment viewers form toward someone who never meets them face to face. Today that phenomenon stretches far beyond Hollywood stars. Influencers and branded avatars use stories, livestreams, and instant-message buttons to craft ersatz familiarity that feels almost reciprocal. For many users, these features create the inviting illusion of a friendly dialogue, convincing them the persona is both emotionally open and morally answerable to the community (Labrecque, 2014).

Research in social psychology consistently demonstrates that consumers are more likely to trust another person-including a digital persona-when they believe that individual genuinely understands and cares about them (Sweeney & Swait, 2008). Influencers who maintain an emotionally stable presence across tone, story arc, and guiding principles create a baseline of cognitive reliability. After repeated interactions, followers start to regard these creators as near-friends who can weigh right from wrong and to whom they owe ongoing loyalty (Chung & Cho, 2017).

2.3.2 Components of Trust in Persona Marketing

Trust, however, does not exist as a single, simple idea. In their study of online celebrities, Freberg and colleagues (2011) break it into three distinct yet overlapping parts:

Expertise-what the influencer knows or can do in a given field; attractive-qualities such as warmth and charm that create immediate emotional pull; trustworthiness-the impression that the person is honest, open, and ethically grounded.

Of the three, trustworthiness usually casts the longest shadow over how people behave online. Jin and Ryu (2020) found that followers engage more deeply with creators seen as morally sincere, even when those creators offer only modest expertise or average looks. That finding echoes Edelman's 2023 Trust Barometer, which shows almost 70 percent of shoppers worldwide prefer brands they believe share their core values, valuing alignment over price or convenience.

A critical element in trust formation is emotional authenticity-the extent to which the personas affective expressions appear sincere and situationally appropriate. As Marwick 2015 notes, digital audiences are increasingly media-literate. They recognize that social media content is performative, but still demand emotional transparency. Performances of vulnerability, humor, or moral reflection-when perceived as genuine-deepen trust, even in highly commercial contexts.

2.4 The Erosion and Reconstruction of Trust in AI-Mediated Persona Environments

Technological advances in social media have normalised AI-created avatars and synthetic influencers, revealing both exciting opportunities and deep uncertainty about how trust is constructed and eroded online.

Traditional, human-led marketing shapes identity through visible emotion and everyday fluctuation; in contrast, an AI persona results from deliberate design choices, data patterns, and continual adjustments to performance. For that reason, researchers are beginning to ask whether trust can survive when seemingly spontaneous behaviour is actually machine-coded, and what new markers of credibility emerge when the familiar signals fall silent.

2.4.1 From Human Empathy to Algorithmic Consistency

In campaigns led by flesh-and-blood people, observers typically measure trust by spotting genuine emotions, willing self-disclosure, and a track record of taking responsibility (Audrezet et al., 2018). Those cues, rooted in lived experience, paint a portrait of humanness and invite audiences to let their guard down. Inside AI-centered spaces, however, those markers vanish or show up only as carefully timed simulations. Virtual spokesfigures such as Lil Miquela or Imma can "feel," endorse global causes, and banter with followers-yet every response is prewritten, overseen by human teams, and finally delivered by lines of code (Moustakas et al., 2020).

People recognise that many online characters are computer-generated or staged, yet a surprising number still invest real feelings in them, forming what scholars call parasocial bonds with these ghosts of data (Berryman & Kavka, 2018). Such attachments reveal that trust can grow in the absence of flesh and blood, as long as a story clicks together, the jokes land, and the interface feels familiar. If a digital face keeps delivering sentiment, cultural relevance, and behaviour that rarely surprises, audiences may crown it "trustworthy." 2.4.2 The Algorithm as Co-Curator of Persona Inside an algorithmic playground, decision-making code takes centre stage in telling us who these identities will be, and, eventually, whether we will trust them at all. Recommendation engines do much more than feed funny memes; they steer conversation patterns, chart visibility curves, and colour an accounts emotional tone (Bucher & Helmond, 2018). By that logic, a feed does not merely distribute avatars-it volunteers as a ghostwriter, blessing consistency, cheerfulness, and predictability while muting ambiguity, dissent, or messy human contradiction. As a result, digital characters now march toward engagement rather than complexity,forcing a sort of flattening in their crafted personalities. Behind the glossy surface, emotional realism often settles for readymade sympathy, stock cause-line associations, and a statistical remix of past likes. Such polish may win trust quickly; over time, however, it quietly strips away the nervous, unpredictable, even vulnerable edge that human authenticity used to give us (Marwick, 2015).

### 2.4.3 Trust Transfer: Human-AI Hybrid Personas

Scholars are now examining hybrid online personas in which artificial intelligence aids influencers by drafting posts, managing comments, or even mimicking their tone. An influencer can quietly ask ChatGPT to write a catchy caption, then publish the result as if it originated entirely from them. This practice forces us to rethink trust: does an audience still accept the persona as genuine when most of the voice comes from code rather than a human mind?

Two recent studies in human-computer interaction suggest that trust shifts from people to machines as long as someone we recognize still supervises the system (Glikson & Woolley, 2020). If followers believe a human keeps moral authority over the AI input, they often treat the blend as trustworthy. Yet the bond fractures once the public thinks responsibility has been severed, especially when a tool spreads false claims or disguises its own involvement (Floridi & Cowls, 2019).

In online spaces increasingly shaped by artificial intelligence, our old ideas of trust-woven from personal honesty, moods, and face-to-face signs-no longer hold up. Researchers now speak of algorithmic trust, suggesting that faith in a system should rest instead on how clear, predictable, and accountable that system is, rather than on any designer's reassuring smile (Wagner, 2019). To build such trust, platforms must adopt solid guardrails: they should openly declare when AI is at work, explain how a choice was reached, and impose clear moral rules on any automated content creation.

Marketers therefore need to move past mere showmanship and focus on trust design, checking every piece of a brands voice, style, and emotion not only for click-through rates but for ethical soundness. The change is all the more pressing because deepfake videos, synthetic voices, and affordable tampering tools make believable falsehoods disturbingly easy to produce.
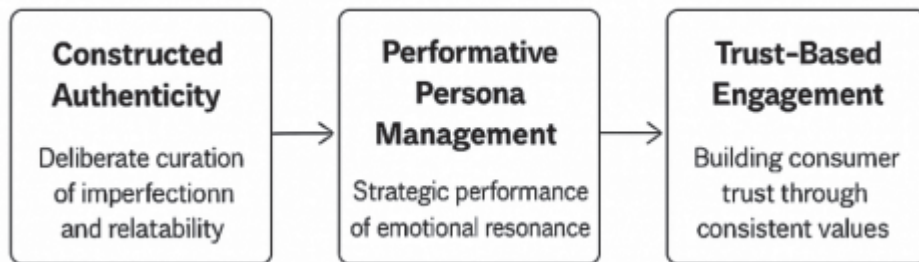
Fig1. The Evolution of Persona Marketing: From Constructed Authenticity to Consumer Trust
Source: Author's own elaboration

The conceptual model shows a gradual journey through three interlinked phases of persona marketing: constructed authenticity, performative persona management, and trust-based engagement. Moving from one phase to the next, emotional connection grows, narratives become richer, and platform-driven interactions multiply, paving the way for long-lasting consumer trust.

III. Research Method

To study how persona marketing on social media evolved-from crafted authenticity to consumer trust-the

project uses a literature-synthesis approach. By pulling together insights from across disciplines, this method forms a unified theoretical map that mirrors the changing ways people build digital identities and interact with audiences.

3.1 Methodological Approach: Literature Synthesis

A literature synthesis differs from a meta-analysis or systematic review in that it values conceptual blending over the pooling of raw data. It lets scholars draw on ideas from sociology, communication, psychology and beyond, especially when a topic lacks a single theory or agreed-upon measures (Tranfield, Denyer, & Smart, 2003).

The goal of the current synthesis is not merely to chart what scholars have said, but to trace the intellectual path of persona marketing-from authenticity-centered strategies to those grounded in trust. The exercise also opens space to critique how new players-like AI-generated influencers and synthetic personas-may unsettle the ways audiences decide who or what deserves their trust.

3.2 Data Collection and Source Selection

The investigation draws on peer-reviewed journal articles, academic books, and a handful of industry white papers published between 2000 and 2024. Collection proceeded in line with a formal search protocol applied to key databases, namely Scopus, Web of Science, JSTOR, and Google Scholar.

Searching employed phrases like persona marketing, authenticity and social media, parasocial trust, influencer credibility, virtual influencers, and algorithmic visibility. Boolean logic refined and linked the terms into coherent queries. In addition, backward and forward citation tracing pinpointed both classic and newer relevant studies.

From an initial pool of around 150 documents, 58 items were chosen for detailed coding. Selected works satisfied three criteria: (1) they were peer-reviewed or respected trade publications; (2) they provided theoretical insight into digital persona, authenticity, or trust; and (3) they addressed current practices on major online platforms. Highly technical AI papers or research lacking behavioral or branding focus were set aside.

3.3 Analytical Strategy and Thematic Integration

After gathering the archival materials, a three-part interpretive framework guided their examination. In the first stage each item received margin annotations that recorded defining arguments, core terms, and cited models. From those annotations a set of codes emerged-emphasis on performance, emotional sincerity, parasocial ties, the rise of digital trust, influencer ethics, and governance of synthetic selves.

The second stage grouped the discrete codes into larger clusters that matched the studys overarching questions. Doing so revealed three overlapping stages in the trajectory of persona marketing.

First, constructed authenticity functions as an intentional performance; second, performative personas evolve into continuous emotional identities; and third, trust-driven engagement becomes the aspirational bond between brands and consumers.

The final step placed the clustered themes on a timeline, showing how changing audience expectations, shifting platform rules, and advancing technology enabled each new layer. That layered history permitted a synthesis that clarifies how both human and algorithmic personas adaptively build trust over time.

3.4 Scope and Limitations

Though a literature synthesis provides useful overview, it also carries clear drawbacks. Most notably, the project avoids first-hand data collection, so it cannot track real user behaviour in live settings. Instead, findings lean on broad patterns and scholarly argument, not on direct observation.

Second, the review sample drew almost entirely from English journals, a choice that skews discussion toward North American and European views of authenticity, branding, and trust. Efforts to include work on Douyin, WeChat and East Asian virtual stars helped, yet still leave unexplored country-specific norms and platform habits for future authors to address.

Finally, time lags in academic publishing mean some rapid topics, for example AI avatars, deep fakes and generative branding tools, are only treated as yesterday's news. Because technology moves quickly, its real-world effects may already outlive the theories offered here.

IV. Discussion

A review of the current scholarship shows that digital branding has moved away from well-rehearsed personas and toward engagement grounded in organizational trust. In what follows, I examine the theoretical and practical consequences of this trend, situating it within the wider media environment and offering critical reflections on its ethical and technological implications.

4.1 From Constructed Authenticity to Emotional Commitment

Early influencer campaigns portrayed authenticity through staged imperfections, backstage glimpses, and heartfelt confessions. Such performances produced the "realness" many viewers now expect, because they provided a softer counterpoint to formal corporate rhetoric. Over time, however, audiences have grown savvier, and a polished illusion of genuineness no longer converts attention. Authenticity has become a threshold rather than an end; it opens the door, but sustaining trust demands emotionally steady, values-driven exchanges.

Research increasingly positions emotional labor at the core of contemporary persona marketing. Digital creators must weave a consistent story across platforms while adjusting to volatile audience moods (Duffy,

2017). Although this strategic affective work strengthens parasocial ties, it simultaneously risks creator and brand burnout, ethical strain, and damage to public reputation.

Firms have mirrored this model by launching emotionally charged stories-from mission-centred ads to CEO-led, human-facing posts-in hopes of close-knit audience rapport. The danger, however, lies in a shallow, scripted sheen. When such displays lack genuine organisational backing, they invite charges of hypocrisy, so-called woke-washing, or blatant opportunism (Vredenburg et al., 2020).

4.2 The Psychological Mechanics of Digital Trust

A central conclusion from the review is that trust online springs less from formal institutional stature than from relational experience. It builds gradually, through emotionally steady actions-consistent messaging, fair disclosure, and prompt engagement-that pile up day after day. Parasocial Interaction theory (Horton & Wohl, 1956) still clarifies this process, revealing how audiences establish robust, if one-sided, attachments to digital figures that feel remarkably like everyday social bonds.

This evidence points to the idea that trust is built less on objective fact and more on how believable a presentation feels on an emotional level. As Marwick (2015) observes, most users of social media recognise that online profiles are carefully staged. What they really look for is affective sincerity, a sense that the persona shares their moral code, and steady behaviour over time. People apply these same tests to human influencers and to algorithm-made characters, although the specific yardsticks may vary.

In everyday terms, trust shows up as likes, repeat purchases, and unsolicited word-of-mouth. At the same time, it remains extremely delicate. One undisclosed sponsorship, an obvious algorithm tweak, or a seemingly hollow post can tip a community from acceptance to quiet unfollowing or even open boycott. Repairing that bond, once damaged, usually takes far longer than building it in the first place.

4.3 Trust in the Age of AI Personas: A Shifting Paradigm

The most pressing takeaway from this research is the accelerating appearance of AI-driven personae. Virtual stars such as Lil Miquela, along with branded chatbots, are pushing the limits of what we classify as a "person." These digital figures promise unbroken availability, uniform tone, and reach that never sleeps, yet they also force us to rethink old ideas about authenticity, responsibility, and who is ultimately answerable when things go wrong (Moustakas et al., 2020).

Our examination indicates that people are becoming more willing to overlook the artificiality of digital characters, especially when these characters offer either entertainment or practical help. Yet, confidence in the experience rests less on how lifelike the persona appears and more on clear design, understandable algorithms, and visible human supervision (Glikson & Woolley, 2020). Trust thus shifts from the face offering the service to the systems and teams that built and maintain that face.

This change alters how authenticity and trust function in online marketing. The field is moving from seeing

authenticity purely as raw emotion to viewing trust instead as the product of careful design-a trend that carries both strategic openings and ethical pitfalls. When the role of AI is hidden, consumers are likely to feel misled, prompting serious questions about consent, manipulation, and who is accountable for harm (Floridi & Cowls, 2019).

V. Conclusion

This article charts the evolving landscape of persona marketing on social media, focusing on the way deliberately constructed online selves move from seeming genuine to actually attracting consumer trust. By synthesizing previous scholarship, it outlines the evolution of persona marketing in three phases: staged authenticity, the performative persona, and trust-based engagement. Each stage corresponds to larger shifts in the emotional, technological, and ethical context of digital branding.

In its earliest form, persona marketing relied on artfully copied signs of authenticity, as influencers and brands introduced minor flaws or fleeting moments of vulnerability to appear relatable. Audiences, however, quickly honed their skepticism; the mere display of authenticity proved insufficient. What originated as a brief performance evolved into a longer choreography-emotional steadiness, shared values, and-visible transparent cues emerging across multiple touchpoints. Trust, therefore, assumed primary status as the psychological strand linking consumers to brands over time. It is not simply a sequel to authenticity; rather, it grows from steady, emotionally clear exchanges between the constructed self and its followers.

Research indicates that contemporary online trust is no longer anchored in traditional signposts-such as institutional status or recognised expertise-rather, it develops through routine, interpersonal exchanges. Users begin to feel confident in a persona when it consistently appears genuine, replies swiftly, and demonstrates a stable set of ethical principles. That impression emerges not only from the content of the message but also from the technologies that deliver it-algorithms, community guidelines, and the viewers own interpretation. Understood this way, persona marketing becomes more than a tactical move; it operates as an intertwined social, technical, and emotional ecosystem.

A significant theoretical gain lies in the studys ability to combine Goffmans stage imagery, parasocial bonds, influencer authenticity, and machine mediation into one usable map. In doing so, it breathes new life into classic notions of self-presentation and demonstrates that algorithm-driven identities and even AI-generated acts now reshape the performance itself. The resulting framework signals how online personas shift across platforms, weighing emotional labour, audience expectations, and the incessant lure of visibility.

These findings go straight to the desks of digital marketers and creative units that shape online messaging. Now that persona marketing has moved past its pilot phase and is being rolled out at scale, the challenge is no longer to appear authentic but to keep winning trust in an era that prizes transparency yet turns dubious in seconds. Practitioners therefore have to ask how every persona is tuned for emotional resonance, kept within humane and ethical bounds, and communicated clearly enough that the audience knows who-and what-its really

meeting. Credibility comes now, less from a show-stopping headline, and more from constant value alignment, prompt reply to doubts, and a strategy that does not keep switching the rules.

At the same time, AI-generated personas introduce a new layer of complexity to trust management. Consumers routinely encounter synthetic influencers whose smile is scripted by code yet strategically guided by human choices, and that fusion reshapes their yardstick for credibility. The evidence suggests that approval in these domains rests not only on emotion that looks genuine but also on candid disclosure of the design steps, crystal-clear attribution of authorship, and guardrails that prevent harm while still letting creators innovate.

Despite its grounding in extensive theory, the review leans almost exclusively on English-language publications, leaving a wealth of non-Western scholarship largely ignored. Future studies should place the proposed model in the wild-watching diverse users interact with AI avatars across languages, cultures, and operating systems. Specific questions to pursue include whether trust registers more in a click than in a comment, what design features trigger genuine belief in a persona, and under what circumstances platform rules extend or trim an avatars sail.

When considered as a whole, the evidence paints persona marketing as a fragile, trust-based contract rather than a crude copycat game, one constantly shaped by message quality, the immediate environment, and shifting crowd moods. In a crowded digital bazaar where authentic and fabricated selves rub elbows, trust emerges as the most valuable coin-and also the one that breaks or tarnishes with startling ease. The brand takeaway is straightforward: map a avatars life cycle, defend its credibility with informed design, and intervene swiftly whenever confidence begins to lose its shine.

References

Abidin, C. (2016). 'Aren't These Just Young, Rich Women Doing Vain Things Online?': Influencer Selfies as Subversive Frivolity. Social Media + Society, 2(2), pp.1–17.

Audrezet, A., de Kerviler, G. and Moulard, J.G. (2018). Authenticity under threat: When social media influencers need to go beyond self-presentation. Journal of Business Research, 117, pp.557–569.

Berryman, R. and Kavka, M. (2018). Crying on YouTube: Vlogs, self-exposure and the productivity of negative affect. Convergence, 24(1), pp.85–98.

Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. New Media & Society, 14(7), pp.1164–1180.

Bucher, T. and Helmond, A. (2018). The affordances of social media platforms. The SAGE Handbook of Social Media, pp.233–253.

Campbell, M.C. and Grimm, P.E. (2019). The challenges of marketing brands in a politically polarized marketplace. Journal of Marketing, 83(1), pp.79–89.

Chung, S. and Cho, H. (2017). Fostering parasocial relationships with celebrities on social media: Implications for celebrity endorsement. Psychology & Marketing, 34(4), pp.481–495.

Duffy, B.E. (2017). (Not) Getting Paid to Do What You Love: Gender, Social Media, and Aspirational Work. Yale University Press.

Edelman. (2023). Edelman Trust Barometer Global Report 2023. [online] Available at: https://www.edelman.com/trust-barometer

Floridi, L. and Cowls, J. (2019). A unified framework of five principles for AI in society. Harvard Data Science Review, 1(1).

Freberg, K., Graham, K., McGaughey, K. and Freberg, L.A. (2011). Who are the social media influencers? A study of public perceptions of personality. Public Relations Review, 37(1), pp.90–92.

Gaden, G. and Dumitrica, D. (2015). The 'real deal': Strategic authenticity, politics and social media. First Monday, 20(1).

Glikson, E. and Woolley, A.W. (2020). Human trust in artificial intelligence: Review of empirical research. Academy of Management Annals, 14(2), pp.627–660.

Goffman, E. (1959). The Presentation of Self in Everyday Life. Anchor Books.

Hearn, A. (2008). "Meat, mask, burden": Probing the contours of the branded self. Journal of Consumer Culture, 8(2), pp.197–217.

Hogan, B. (2010). The presentation of self in the age of social media: Distinguishing performances and exhibitions online. Bulletin of Science, Technology & Society, 30(6), pp.377–386.

Horton, D. and Wohl, R.R. (1956). Mass communication and para-social interaction: Observations on intimacy at a distance. Psychiatry, 19(3), pp.215–229.

Jin, S.A.A. and Ryu, E. (2020). "I'll buy what she's #wearing": The roles of envy toward and parasocial interaction with influencers in Instagram celebrity-based brand endorsement and social commerce. Journal of Retailing and Consumer Services, 55, p.102121.

Keller, J. (2022). Influencer Culture and the Authenticity Illusion: Anxieties and Aspirations in the Social Media Age. Routledge.

Khamis, S., Ang, L. and Welling, R. (2017). Self-branding, 'micro-celebrity' and the rise of Social Media Influencers. Celebrity Studies, 8(2), pp.191–208.

Kozinets, R.V., de Valck, K., Wojnicki, A.C. and Wilner, S.J.S. (2010). Networked narratives: Understanding word-of-mouth marketing in online communities. Journal of Marketing, 74(2), pp.71–89.

Labrecque, L.I. (2014). Fostering consumer–brand relationships in social media environments: The role of parasocial interaction. Journal of Interactive Marketing, 28(2), pp.134–148.

Marwick, A. (2015). Instafame: Luxury selfies in the attention economy. Public Culture, 27(1), pp.137–160.

Marwick, A. and boyd, d. (2011). To see and be seen: Celebrity practice on Twitter. Convergence, 17(2), pp.139–158.

Morhart, F.M., Malär, L., Guèvremont, A., Girardin, F. and Grohmann, B. (2015). Brand authenticity: An integrative framework and measurement scale. Journal of Consumer Psychology, 25(2), pp.200–218.

Moustakas, E., Lamba, R., Mahmoud, D. and Zheng, J. (2020). Virtual influencers: When AI meets marketing. Journal of Strategic Marketing, 31(2), pp.245–262.

Sirgy, M.J. (1982). Self-concept in consumer behavior: A critical review. Journal of Consumer Research, 9(3), pp.287–300.

Sweeney, J.C. and Swait, J. (2008). The effects of brand credibility on customer loyalty. Journal of Retailing and Consumer Services, 15(3), pp.179–193.

Tranfield, D., Denyer, D. and Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. British Journal of Management, 14(3), pp.207–222.

Urban, G.L., Sultan, F. and Qualls, W.J. (2000). Placing trust at the center of your Internet strategy. MIT Sloan Management Review, 42(1), pp.39–48.

van Dijck, J. (2013). The Culture of Connectivity: A Critical History of Social Media. Oxford University Press.

Vredenburg, J., Kapitan, S., Spry, A. and Kemper, J.A. (2020). Woke branding: Activism or exploitation? Journal of Public Policy & Marketing, 39(4), pp.444–460.

Wagner, B. (2019). Algorithmic accountability: A primer. Internet Policy Review, 8(2), pp.1–15.