Theorising authenticity vis-à-vis, not versus, syntheticity in the age of AI: Introducing a continuum-based framework

Andrew Frank Bradley (Open University of Catalonia) & *María del Mar Palomares Marín* (University of Limerick)

afrank(at)uoc.edu, maria.palomares(at)ul.ie

Abstract

This article addresses contemporary debates surrounding authenticity in the age of Generative Artificial Intelligence (GenAI), where hybrid human-machine productions have given rise to a dual crisis of origin and reception. As GenAI-human co-authored content becomes more ubiquitous, the range of reactions it elicits – from dismissal to acceptance – highlights the need to reassess the relationship between the authentic and the synthetic within modern technologically-mediated and hybrid human-AI ecosystems. To this end, the article introduces the concept of syntheticity as both the degree of technological mediation in production (synthetic origin) and its social perception (synthetic attribution), which may coincide or diverge. Furthermore, the article advances a continuum-based model that reconceptualises authenticity and syntheticity as relational rather than oppositional constructs. This model positions authenticity along two axes: synthetic origin (human to artificial/generated) and reception (authenticated to unauthenticated), yielding four configurations of *Human authenticity*, *Human inauthenticity*, *Synthetic authenticity*, *Synthetic inauthenticity*. Intermediary zones account for hybrid productions (*Authenticated* and *Unauthenticated hybridity*) and contested validation states (*Liminal authentication*). By decoupling origin from reception in the evaluation of authenticity and syntheticity, the model positions hybridity as a legitimate locus of authentic expression. As a conceptual toolkit, the framework offers a new lens for analysing co-authored productions in which human and machine agency coexist and facilitates the critical evaluation of hybrid content.

Keywords

authenticity, syntheticity, artificial intelligence, human-AI hybrid productions, Generative Artificial Intelligence

1 Introduction

Generative Artificial Intelligence (GenAI) involves the use of complex algorithmic computation trained on pre-existing datasets to produce "new" content (Chan and Hu 2023). In practice, this allows users to input a variety of prompts and, in return, generate texts, images, sounds, videos, or a combination of these in mere seconds. Despite its relative infancy, GenAI has already had a profoundly transformative impact on shaping human experiences (Kudina and Verbeek 2019) in and across increasingly diverse contexts.

While Artificial Intelligence (AI) has existed in many forms for decades now (Thakur, Barker and Pathan 2024), the development of GenAI stands out as a revolutionary new phase in its evolution, particularly in terms of its accessibility,

AILing

Bradley, Andrew Frank & Palomares Marín, María del Mar. 2026.
Theorising authenticity vis-à-vis syntheticity.
Special Issue: *The notion of authenticity*. Vol. 4 No.1
DOI: 10.62408/ai-ling.v4i1.39

AI-Linguistica. Linguistic Studies on AI-Generated Texts and Discourses ISSN: 2943-0070

ease of use, and reach among technologically proficient and non-expert users alike. The public release of OpenAI's ChatGPT in November 2022 can be considered to be a pivotal moment that not only accelerated the mainstream adoption of AI, but also democratised GenAI to millions of users. Since then, the GenAI landscape has expanded substantially with the emergence of multiple competing systems and a dramatic increase in global user uptake, with current industry leaders estimating that 1 in 6 people worldwide use GenAI tools (Microsoft 2026).

In the new post-GenAI landscape, where users have the ability to effortlessly generate multimodal outputs on demand, the explosive growth of users is mirrored by a corresponding surge in AI-generated content. For instance, some provisional analyses suggest that at least 30% of text on active websites derives, in some way, from GenAI (Spennemann 2025), while other large-scale studies report that nearly three-quarters of newly created webpages now include AI-generated content (Law, Guan and Soulo 2025). Similarly, in the audio-visual sphere, providers such as YouTube have reported that up to 92% of content creators use GenAI to empower their creative process on its video platform (YouTube 2024). Beyond text and audiovisual media, AI-generated images have also become ubiquitous across a wide range of physical and digital domains (Rapp et al. 2025). Therefore, the exponential growth of AI-generated content, whether fully automated or hybridised through human-machine collaboration, suggests that such productions are likely to become even more pervasive, and, in some contexts, functionally unavoidable.

In terms of verisimilitude and fidelity, early GenAI artefacts exhibited notable limitations, especially in image and video generation, and were often criticised or delegitimised for their conspicuous markers of artificiality (Yang et al. 2022; Borji 2023), or "synthetic" qualities (Ghiurău and Popescu 2024). However, each successive model iteration is able to produce incrementally higher-quality outputs. This poses a significant epistemological challenge when coupled with the sheer volume of generated content now readily available, as contextual cues that once underpinned judgements of origin and credibility are no longer reliable indicators of authorship. To this end, Knott et al. note that "as the differences between AI-generated and human-generated content decrease, it becomes intrinsically harder to adjudicate individual cases" (2024: 62). Floridi (2024) similarly highlights how GenAI's impact on content creation calls into question conventional boundaries surrounding authorship, challenging established distinctions between human and machine contributions, as well as how such productions are understood and categorised.

In this context, authenticity emerges as a central, if contested, concept through which users or observers attempt to negotiate validity and legitimacy amid proliferating hybrid forms of co-production. These dynamics invite renewed reflection on what it means for something to be "authentic"; how "authorship" or attribution is determined; the authentication processes through which "authenticity" is assessed, and even what it means to be "human" itself (Beerends and Aydin

2025). Traditionally, authenticity has been theorised as a multidimensional and complex process of verification (Newman and Smith 2016) that is tied to humans and their creations (Dutton 2003). Although the expansion of digital and online communicative practices has complicated these associations (Mishan 2017), authenticity remains pervasive in contemporary culture (Martínez 2020) and is a central concern in daily life (Guignon 2004). Moreover, it continues to feature prominently in academic discussions, particularly in relation to digitalisation and the Internet (Mercier and Abidin 2025).

However, little attention has been paid to how AI-generated and hybrid productions intersect with non-naturalist and post-GenAI formulations of authenticity, particularly at the level of conceptual modelling. In fact, there is a tendency in both popular and academic discourses for the "authentic" to be coupled with human-origin productions versus the "inauthentic" with AI-generated or hybrid creations (cf. discussions in Bui, Filimonau and Sezerel 2024; Beerends and Aydin 2025; Erdocia, Schneider and Migge 2025; Wu and Lingel 2025). In this paper, we argue that such binary categorisations fail to capture the complexity of AI-mediated content creation. In response, we propose a conceptual, continuum-based model for theorising authenticity in the era of human-machine collaboration, with a specific interest in GenAI co-productions and ecologies. The framework advances two central claims: first, that the authentic should be understood in relation to the synthetic, as opposed to its inherent counterpart; and by extension, that hybrid human-machine (including GenAI) productions can, under certain conditions, be legitimately experienced and evaluated as authentic. By doing so, this paper aims to contribute to ongoing debates concerning authenticity across different disciplines (e.g., Newman and Smith 2016; Pinner 2016; Kalpokas 2024; Lavazza 2025; Lee 2025, among many others) and serve as a tool for assisting in the navigation of current and future AI-human experiences and landscapes.

This article functions as a conceptual and theoretical synthesis that draws on illustrative empirical examples to develop a theory-building framework. The paper is structured as follows. Section 2 offers a conceptual synthesis of existing frameworks of authenticity across different fields, drawing in particular on Newman and Smith's (2016) seminal review of the literature, and extends their analysis to highlight definitional ambiguities and theoretical tensions that are sharpened in the age of GenAI and hybrid human-machine production. The empirical studies cited throughout are drawn selectively from existing scholarship and function as illustrative cases, rather than as an exhaustive review of specific domains. Section 3 examines how authenticity is renegotiated in contemporary mediated communication and interaction, introducing two interrelated crises of post-GenAI authenticity: a crisis of origin, in which provenance becomes increasingly indeterminate, and a crisis of perception, in which authenticity judgements are shaped by attribution and reception rather than by content alone. This section situates these crises within communication studies, AI-mediated interaction, and human-AI collaboration research. Section 4 introduces the

proposed theory-building framework and advances a structured conceptual model for analysing how authenticity and syntheticity are validated, negotiated, and contested across contexts of human-machine and GenAI collaboration. Section 5 concludes the article by discussing the framework's applications, limitations, and implications for future research.

2 Authenticity: a brief overview

Authenticity is a multidimensional concept with no static or singular meaning, whose interpretation(s) evolve over time, vary across disciplines, and circulate in public discourses to serve different social and evaluative aims. Traditionally, authenticity has been associated with qualities such as truthfulness (Williams 2002), originality (Bialystok 2014), sincerity (Trilling 1972), nativeness (Woolard 2016), and (social) virtue (Guignon 2008), among many others. While authenticity has frequently been theorised in relation to human subjects and lived experiences, it has also been extended to non-human referents such as institutions, objects and actions (Bialystok 2014), and, in more recent decades, to digital artefacts and environments (Mishan 2017; Mercier and Abidin 2025). This breadth of application positions authenticity as a polysemous concept that resists a universally applicable framework. At the same time, however, such conceptual flexibility does not imply complete indeterminacy but instead draws attention to whether authenticity has any stable points of reference.

Addressing this, in their seminal examination of existing research, Newman and Smith (2016) successfully outline four areas of convergence across multiple different definitional typologies of authenticity. The authors find that discussions about or theorisations of authenticity tend to gravitate towards one of four distinct categories: historical authenticity, categorical authenticity, value authenticity, and self-authenticity. Rather than proposing a universal definition, their framework offers an analytic synthesis that identifies recurring themes while documenting how authenticity is evaluated across different disciplines. In what follows, we discuss and expand upon these in turn.

2.1 Historical authenticity

In relation to historical authenticity, Newman and Smith describe this as involving "the evaluation of an object's unique spatiotemporal history" (2016: 612). Through an external process of validation (i.e. by an expert), an original item (e.g. the Mona Lisa) is deemed either legitimate or not, and thus historically (in)authentic. Historical authenticity appears under different labels and closely corresponds to several related constructs, such as indexical authenticity in marketing (Grayson and Martinec 2004), nominal authenticity in aesthetics (Dutton 2003), objective authenticity in tourism (Wang 1999) and pure authenticity in advertising (Beverland, Lindgreen and Vink 2008). Despite their disciplinary specificity, these

constructs converge around a shared emphasis on provenance and historical linkage. For instance, Grayson and Martinec (2004) invoke indexicality to describe the traceable link between an object and a specific time and place in the past, through which evaluations of authenticity are formed in relation to consumer products. In the context of artistic production, Dutton (2003) similarly refers to an item's nominal authenticity as the genuine identity of an item, distinguishing it from plagiarism and forgery. Likewise, Wang (1999) emphasises objective authenticity through the uniqueness of objects with a verifiable history, or those considered one of a kind. Finally, pure authenticity is proposed by Beverland, Lindgreen and Vink (2008) to explain the verifiable link between traditional practices and historical cues that confirm if an item has remained unchanged over time and remains true to its origin. Across fields, and within this conceptual lens, historical cues constitute an integral part of the authentication process, which is typically both binary and rooted in origin.

Moreover, Groth, Block and Newman (2023) affirm that authenticity judgements about objects are contingent upon their perceived connection to a person, time, or place. This, in turn, highlights three crucial aspects of an item's origin: essence transfer, spatiotemporal connection, and belief. First, the connection between essence and authenticity stems from the perception that an object can "absorb" an immaterial essence through contact with people, especially the original author or through processes of contagion (Newman and Bloom 2012: 558), thereby enhancing its perceived authenticity. Second, the authenticity of an item is anchored in a specific time and/or place, such that a "real" spatiotemporal connection is understood to exist between an object, its author, and its time or place of creation (Lehman et al. 2019). Third, authenticity judgements may also involve beliefs about whether, and to what degree, a work embodies the essence of its author and its historical context. This is particularly widespread in literary domains where authenticity is typically grounded in the conviction that a work is genuinely (i.e. authentically) the product of a specific author (Martínez 2020).

Historical authenticity can also be subject to degrees of evaluation, particularly in contexts involving mass-produced goods, where distinctions between first editions or signed copies acquire heightened value through their perceived proximity to an author or moment in time (Newman 2019). As Grayson and Martinec (2004: 299) argue, perceptions of indexicality (i.e. what distinguishes the real and valued object from its copies) may vary according to individual interpretations of essence. For example, a book signed by an author may be viewed as less indexical compared to a book personally handed by the writer, which in turn may be experienced as carrying a stronger trace of the author's essence. In contrast, in cases involving unique artworks such as the Mona Lisa, judgements of authenticity and value extend beyond individual evaluation or historical cues, and instead rely on accepted forms of nominal authentication.

2.2 Categorical authenticity

While historical authenticity is frequently conceptualised through provenance, scholars have also highlighted the role of interpretative variation in shaping (historical) authenticity evaluations. In such cases, subjectivity arises from individually formed perceptions of an object's connection to a specific person, time, or place (cf. the notion of *essence transfer* discussed above). This form of subjectivity, however, continues to rest on the assumption of a verifiable historical link and therefore cannot account for other kinds of authenticity where such a link is fragile, secondary, or nonessential. Under these conditions, authenticity is better understood as categorical (Newman and Smith 2016) and grounded in a subjective, socially constructed epistemological framework (Wang 1999; Carroll and Kovács 2021), where judgements are not necessarily based on an authenticated historical connection but instead can fluctuate across time, space, or community. As Newman and Smith (2016) remark, these various approaches converge in their emphasis on the degree to which an entity is perceived to be true to its associated social type, genre, or category. This category-based authenticity relies on perceived typicality (Busselle and Bilandzic 2008) or the cognitive process of internal validation that occurs between the observer's prior mental expectations and the item or concept under observation itself. As noted by Lee (2020), when there are deviations from this expectancy, claims of inauthenticity often arise. In other words, categorical authenticity concerns the construction of the "authentic" according to the pre-conceived expectations of an observer (or observers) about how something is, or ought to be, and the extent to which something is perceived to be real and true.

Carroll (2015; cf. Carroll and Kovács 2021) similarly refers to this as type authenticity, whereby an attribution of authenticity indicates that a person or other entity fits appropriately into the social category to which it has been assigned or that it has claimed for itself (Lehman, O'Connor and Carroll 2019: 21). In a related formulation, Grayson and Martinec (2004) draw on social semiotics in their exploration of iconic authenticity. This describes evaluations of authenticity that are not inherent to an object (cf. indexical relationships above) but are instead conferred through an evaluator's contextual and personal judgement, even when an entity is not itself materially "real". Other theorists employ different nomenclature to reflect similarly category-based interpretations, such as *expressive authenticity* (Dutton 2003), *constructed authenticity* (Wang 1999), and *approximate authenticity* (Beverland, Lindgreen and Vink 2008).

Categorical authenticity is therefore central to many aspects of contemporary human socialisation, such as how cultural products are consumed, how identities are performed, as well as how languages are evaluated and used. As Carroll and Kovács state, the external negotiation of socially constructed authenticity relies significantly on moments when "someone verbalizes something that invokes the concept of authenticity or perhaps uses the word 'authentic' or any host of closely associated words" (2021: 2). In this sense, categorical authenticity

becomes observable not only through implicit judgements and evaluation, but also through explicit (metalinguistic) labelling and everyday interactions.

In the case of languages, this is evident with the kinds of categories language users often ascribe to certain linguistic varieties or speakers. One ubiquitous language ideological belief documented by sociolinguists, language educators, and linguistic anthropologists concerns the construction of the "native speaker". This category of speaker is often idealised as a model language user on account of having acquired a language "naturally" through intergenerational transmission, instead of learning it deliberately at school or later in life (Pinner 2014; Llurda and Mocanu 2024). This analytic conflation of authenticity and naturalism (Woolard 2016) extends to the construction of "authentic speakers" (cf. Eckert 2003; Coupland 2003), who are categorised as local(ised) actors rooted in a particular place, who produce language(s) that belong to that specific location or culture. Underpinning such categories of both "native" and "authentic" speakers is a naturalist assumption that the natural has a positive valence and is therefore valued or valuable, while that which is inauthentic (i.e. the mechanical and artificial) is spurious and suspect (Coupland 2014; Woolard 2016: 32). Within this ideological framing of linguistic authenticity, so-called "authentic language" cannot easily be appropriated by outsiders as it is seen to belong to an exclusive in-group. Delegitimising social labels such as "non-native speaker" or "new speaker" (O'Rourke and Ramallo 2013) are typically mobilised in instances of perceived "inauthentic" language or other illegitimate linguistic practices such as having an L2 accent. Since these judgements are contextual interpretations of the world (Mishan 2017), categorical authenticity is necessarily multiple, negotiable, and layered.

2.3 Value and Self authenticity

The next two categories that Newman and Smith (2016) identify in their convergence analysis involve forms of authenticity that are highly agentive and rooted in deeply personal, subjective assessments, making them less amenable to fixed or externally verifiable criteria. The first is value authenticity, for which "observers can place varying degrees of importance on the extent to which one ought to conform to norms or moral values" (*ibid*: 613), and the second is self-authenticity, which relates to one's own individual sense of self. The question of whether attending Midnight Mass on Christmas Eve constitutes an authentic expression of (certain) Christian faith(s) would be an example of the former, while acting in accordance with, or in opposition to, one's true, or authentic, self (however this may be understood) would exemplify the latter.

The role of observers is key for value authenticity as they proceed to assess the alignment between a person's actions (or an object or experience) and the expression of an "implicitly or explicitly claimed set of values" (Newman 2019: 10). On the other hand, self-authenticity emphasises a complex, and difficult to define ideation of the self. Popular discourse on authenticity often entails "being

true to what someone (or something) truly is" (Guignon 2008: 277). As Taylor (1991) observes, within the Western tradition, the Romantic era introduced an expressivist understanding of the self that paved the way for the "age of authenticity" of the 1960s. Since then, being true to oneself has revolved around the principle of originality, that is, the idea of having something inner, true, and unique, such as an inner voice or a particular way of being human. This understanding is commonly positioned in contrast to imitation or conformity with models imposed by religious, political, societal, or intergenerational pressures (McEvoy 2009).

Other theorists, including Taylor (1991) himself, have challenged this understanding of the self, particularly the notion of a stable inner-self or core. Instead, Taylor (*ibid*) suggests that authenticity is dialogic, while for Guignon (2004) it constitutes a social virtue. More recently, discourses of self-realisation and personal fulfilment have revolved around the idea of expressing one's own inner (true) self (Guignon 2004), especially in online settings (Mercier and Abidin 2025), where both value and self-authenticity have become important for mediated authenticity. In this context, Bailey et al. (2020) posit that authenticity is the unobstructed expression of one's self, which is measured by the degree of congruence between an individual's sense of self and their observable expressions of self, and which also must remain consistent across online-offline spaces and different social media or internet platforms.

2.4 Definitional challenges

As Newman and Smith's (2016) convergence analysis demonstrates, authenticity operates through multiple and context-dependent evaluative processes. Therefore, outlining the nuances of this typological framework is challenging, and as the authors remark, "there are many instances that blur the lines between these distinctions" (2016: 614). This sentiment is echoed by others, such as Dammann et al. (2021) who acknowledge the disadvantageous semantic heterogeneity of the term, or Bialystok who doubts "the possibility of generating a complete philosophical account of authenticity" (2014: 1). However, convergent analyses of authenticity research indicate that authenticity functions primarily as a process of verification across its different manifestations (Newman and Smith 2016: 613-614, cf. Lee 2020; Beerends and Aydin 2021). Newman and Smith (*ibid*) find that the authentication necessary for authenticity judgements is organised along two dimensions: the type of entity that is evaluated (i.e. object vs. agent) and the source of information that is consulted (i.e. external vs. internal). Assessments of authenticity may therefore rely on either objective or subjective criteria, or on combinations of both, at varying degrees in different contexts. Authenticity is thus not fixed nor absolute, but fluid and multifaceted. Furthermore, authenticity is ubiquitous and highly influential, as it informs not only how people interpret the world but also how they behave in different social or cultural settings.

Notwithstanding the plurality of ways in which authenticity can be expressed, interpreted, or reified, we agree with Carroll and Kovács (2021) and others (Pinner 2014; Lee 2020) that authenticity is a positive attribution at its core. For instance, consumers prefer authentic products (Lehman, O'Connor and Carroll 2019), while language learners and teachers find that authentic materials and assessments are better for enhancing critical thinking, problem-solving, and collaboration skills (Vlachopoulos and Makri 2024). However, if authenticity in its various conceptualisations is understood as a markedly positive experience or interpretation of the world, then its opposite is, explicitly or by implication, subtractive, inferior, less genuine, untrue, or false in some way (Wilt, Thomas and McAdams 2019; see Michael and Shuttleworth 2021 for an example in the context of ethics).

This evaluative asymmetry becomes particularly salient in cases where authenticity is associated with behaviours that challenge moral or societal ideas and standards. Several authors have expressed concerns about the different interpretations of what it means to be authentic. Taylor (1991) describes the "culture of authenticity" as an expression of soft relativism and moral individualism, driven by the goal of self-fulfilment reflecting a self-absorbed and narcissist view of personal realisation. Something that Taylor denounced as part of the "most degraded, absurd or trivialized forms" [of the culture of authenticity] (1991: 29). While authenticity is critiqued here in relation to how humans try to be authentic, this interpretation closely aligns with Newman and Smith's (2016) concept of self-authenticity. For Guignon this ideal "makes a very heavy demand on you" (2004: 76), criticising the common assumption that self-fulfilment is achieved only through the expression of one's true self, while also challenging the widely held belief that one's goal in life is to become authentic without further discussion or questioning.

In digital and social media contexts, these pressures are further intensified. Haimson et al. describe what they term the *online authenticity paradox*, whereby efforts to achieve online authenticity are "often unreachable or is possible at great personal cost, especially for those with marginalized identities and difficult life experiences" (2021: 1). Feelings of inauthenticity appear when one fails to gain social and virtual acceptance. Users are often penalised for showing excessive authenticity (e.g. "trying too hard") or too little authenticity (i.e. not showing negative aspects of their lives), which can lead to emotional distress, reduced self-esteem, social censure and stigma (Davis 2019). Additionally, Silver, Newman and Small (2020) draw attention to the volatile nature of authenticity, showing how even a small detail may change consumers' perceptions toward authenticity judgements, with significant consequences for individuals and brands alike. In this sense, authenticity is contingent upon achieving specific norms or expectations (cf. categorical and value authenticity), with deviations from these going beyond the "mere removal of the positive characteristics associated with authenticity" (*ibid*: 71). Instead, it may trigger an "inauthenticity aversion", public disapproval and moral condemnation (*ibid*: 71). Although these dynamics are not explicitly framed

as a crisis of authenticity, the simultaneous elevation of authenticity as idealised yet unattainable anticipates a deep tension for the times ahead that merits critical reflection. Significantly, an emerging actor, GenAI, brings in new considerations and dimensions to the authenticity debate.

Against this backdrop, each of the four domains of authenticity in Newman and Smith's (2016) typological framework faces renewed strain and conceptual challenges in the GenAI era. *Historical authenticity*, which traditionally rests on provenance and indexical links to a fixed origin, is destabilised when AI-generated artefacts such as images, music, or text cannot be distinguished from human creations, thus undermining the spatio-temporal verification of authorship. *Categorical authenticity*, contingent upon social constructions of what counts as "real" or "true" by observers, is equally challenged when audiences radically diverge in their responses to AI outputs, some rejecting them outright as inauthentic, while others legitimising them as authentic within new cultural categories (e.g. "AI art"). *Value authenticity*, which relies on the alignment between action and moral norms, is increasingly invoked in debates about whether the use of AI in domains such as education, journalism, or creative industries is an ethically authentic practice or a betrayal of professional standards and social values. Finally, *self-authenticity*, understood as fidelity to one's true self, is problematised in contexts where digital self-presentation is mediated, or even co-authored by AI systems, raising questions about whether hybrid or assisted forms of expression can still be regarded as authentic.

These emerging pressures illustrate how GenAI does not merely complicate existing debates about authenticity but fundamentally influences the conditions under which authenticity has traditionally existed. In particular, they expose the limits of frameworks that treat authenticity as a stable property of origin, authorship, or human embodiment, and point instead to the need for a reconceptualization of how authenticity is negotiated in contemporary human-machine, and specifically human-AI, ecologies. While Newman and Smith's (2016) convergence analysis of authenticity's plural evaluative processes provides an important foundation, a framework capable of integrating technological mediation and synthetic production into this typology is necessary. The following section examines how GenAI has intensified these challenges through dual crises of origin and perception, before advancing such a framework in Section 4.

3 Exploring authenticity in the age of mediated communication, interaction, and synthetic media

Building on these theoretical accounts of authenticity, this section shifts from definitional concerns to the conditions under which authenticity is negotiated and contested in contemporary mediated communication. In particular, it examines how the rise of agentic, generative, and hybrid human-AI systems reshapes assumptions

about agency, authorship, and source, from which two crises of post-GenAI authenticity emerge: one of origin and one of perception.

Further tensions surrounding traditional notions of authenticity emerge within communication and media studies. As interactions between humans and AI systems become more commonplace, long-standing assumptions about communication, agency, and authorship are placed under strain. Contemporary human-AI interaction threatens traditional assumptions or frameworks of authenticity when machines begin to perform in roles traditionally reserved for humans (Guzman, McEwen and Jones 2023). As a result, technology is no longer seen as a medium, but rather as an agentic machine (i.e. source of communication) capable of autonomously constructing meaning rather than just interacting (Sundar and Chen 2023) or mediating communication.

In early theorisations of computer-mediated communication (CMC) when technology served mainly as a medium for transmitting information, authenticity was tied to human senders and receivers, and message integrity itself, that is, whether the message is conveyed as originally intended. At that time, mediated exchanges were perceived as having lower social presence than face-to-face communications (Hohenstein and Jung 2020). Therefore, in early CMC, agency and authenticity remained closely associated with human actors. However, communication with interactive generative-AI chatbots and emerging technologies in the modern era has blurred these assumptions, shifting attention from who is human (or not) to what happens between interlocutors (Dehnert and Mongeau 2022). This highlights the need to rethink communication in the era of AI (Brandtzaeg et al. 2023). To this end, the Human-AI Communication (HA-C) model (Guzman 2018), the Human-AI Interaction Outcomes (HAI-IO) model (Quinlantang 2025), and the HAII-TIME model (Sundar 2020) have been proposed to explain how AI is not merely a communicative mediator or enhancer, but is, or can be, an active partner in communication. These approaches extend agency beyond a human-centric perspective and acknowledge a more symbiotic human-AI relationship, which entails, by extension, a more complex and interactive communicative dynamic (Revolusi and Febriandy 2025).

Alongside these discussions, recent post-humanist and more-than-human perspectives advocate for reframing human-machine relations within broader ecologies of practice, where humans are no longer at the centre but one actor among many (Giaccardi, Redström and Nicenboim 2024). From this standpoint, posthuman subjects would be decentered in a world where humans, human-made and nonhuman conscious machines are interconnected and integrated (Nicenboim et al. 2023; Lackey and Papacharissi 2024). However, emerging concerns appear when AI becomes a source in the communication process with the possibility of machine agency undermining human autonomy or agency (Sundar and Chen 2023). As the media-equation literature notes, humans tend to interact with media and technology in the same way that they do with humans (Reeves and Nass 1996). Similarly, frameworks such as *Computers-Are-Social-Actors* (CASA) (Nass,

Steuer and Tauber 1994; Nass and Moon 2000) defend the position that humans mindlessly treat machines as if they were social entities with gender, personalities, feelings or intentions (Lee 2024) even when users were consciously aware that machines are different from humans. The *Social Presence Theory* (Short, Williams and Christie 1976) reinforces this account by suggesting that multimodal media can convey feelings of proximity for their use of visual and verbal cues (Calefato and Lanubile 2010). Furthermore, AI-based media often rely on anthropomorphic cues, thereby asserting agency alongside humans (Sundar and Chen 2023). These may also foster feelings of homophily among users and increase perceived socialness in interactions (Sundar 2020). Along these lines, recent research suggests that GenAI chatbots produce human-like conversational cues (e.g. mimicking politeness), thus enhancing the realism of interaction and users' satisfaction (Kim and Lee 2025). They can also foster conditions in which users develop para-social relationships (Marriott and Pitardi 2023) and evoke rich and deep emotional experiences (Li and Zhang 2024) with a high social presence.

In this regard, phenomena such as the *uncanny valley effect* – a concept originally proposed in the 1970s by Mori to describe a shift from empathy toward increasingly humanlike machines to feelings of unease or revulsion as they approach, but fail to attain, lifelike appearance (Mori, MacDorman and Nageki 2012) – raise further concerns about how human-likeness in new AI systems may disrupt social perception and interaction. Studies on technological affordances suggest that the novel affordances of AI can trigger cognitive heuristics by shaping users' psychological responses, and foster actions that shape how they process and respond to media content (Sundar and Chen 2023). A more critical approach comes from Lee (2024), who argues that it is not yet fully understood how prior beliefs about machines may bias users' positive or negative reactions to AI systems. As these become more agentic, generative, and context-aware, these traditional frameworks may require refinement or at least invite further questioning in the exploration of authenticity, reinforcing what Kalpokas et al. describe as an anthropocentric logic of automated policing that positions GenAI as a problem to be governed rather than a mode of legitimate production (Kalpokas, Kalpokienė and Šalaševičiūtė 2024).

Taken together, these dynamics suggest that authenticity in AI-mediated interactions emerges from complex negotiations among agency, source awareness, perception and belief pointing to a broader crisis in how authenticity is attributed and perceived. Across these developments, one recurring feature becomes especially salient: the increasing prevalence of hybrid configurations of agency, in which human and machine contributions are co-constituted rather than separable within the same communicative act. It is this condition of hybridity, and not GenAI per se, that problematises traditional assumptions about authorship and authenticity, and which therefore warrants focused attention.

3.1 Hybridity: human-AI collaboration

While authenticity in AI-mediated interaction has been shown to depend on negotiations among agency, source awareness, perception, and belief, the growing role of GenAI in content production also shifts these negotiations toward the generated output itself. Hybridity, agency, and authorship become central in this discussion.

The notion that AI can augment, amplify or automate human intelligence dates to AI's early developments (Sundar 2020), yet contemporary systems, including GenAI, enable forms of human-AI synergy that were previously (technically) unattainable. In this context, tensions between machine agency and human agency emerge as users are hesitant to cede control over decision-making processes, system behaviour (Mieczkowski 2022) or authorship. Boundaries between human and machine agency often blur in practices of co-creation and collaboration resulting in hybridity understood as distributed and co-creative agency, whereby humans (consciously or not) accept the transfer of varying levels of agency to AI systems, and both participate in the creative process (Rammert 2008; Sundar 2020).

GenAI has disrupted not only the long-standing naturalist assumption that there is an exclusively human creative process (Rodríguez Valdés 2025), but also the binary opposition between augmentation and automation. It does so by exceeding the distinction between AI as human extension and AI as machine substitution, instead operating through hybrid configurations of distributed creative agency (Zafar, Ali and Yasin 2025). An example of such hybridity is synthetic media. More broadly, synthetic media can be understood, as Whittaker et al. (2020) propose, as any media (text, images, audio, video, etc.) that has been artificially generated or manipulated, or as digital content that has been "primarily or exclusively produced using AI-enabled tools" (Kalpokas, Šalaševičiūtė and Lipskė 2024: 3). Although synthetic media is commonly associated with deception, particularly in the context of political disinformation (Hameleers, van der Meer and Dobber 2024), it can also serve harmless creative endeavours, such as artistic experimentation, education, or benign entertainment. Distinct from other forms of media, what is central to synthetic media is the partial or full integration of artificial intelligence technologies specifically. At the same time, it is not solely a technological process or attribute. While some synthetic outputs are produced almost entirely by computational systems, many involve shared human and technological agency (Sundar and Chen 2023). Kalpokas, Šalaševičiūtė and Lipskė (2024) in their exploration of the term note that humans are still involved in synthetic content generation by initiating the creative process, engaging in prompt engineering, as well as evaluating (and modifying) the result. It is for this reason that Tron Gianet, Di Caro and Rapp (2024) contend that views of GenAI are shifting, with the tool-collaborator divide becoming less clear as users are increasingly shifting their roles toward curation, selection, and co-production.

Hybridity in human-AI systems exposes the distributed nature of agency, however this redistribution across human and non-human actors does not automatically redefine authenticity nor move synthetic creations away from inauthenticity. Instead, it crystallises into a more fundamental problem: the increasing difficulty of determining where a work comes from at all, inaugurating what many scholars have identified as a crisis of origin.

3.2 The first crisis of post-GenAI authenticity: origin

Amid this AI driven ecosystem, what is novel about GenAI outputs in the present is their scale, accessibility, cost (Rae 2024) and creative plausibility. On a fundamental level, it puts into question what it means for something to be "authentic" in the age of GenAI. On this point, scholars have declared a new crisis of authenticity (Jacobs 2024; Lee 2025) upon reflecting on the destabilising impact caused by different social actors not being able to determine and discern the original source of a message (whether human or machine), as GenAI creations can be comparable to and even surpass human content (Pedersen and Ritter 2024; Baptista et al. 2025; Gherheș et al. 2025) and are often indistinguishable, for example, in the context of images (Giridhar et al. 2025). Another significant area of concern regarding synthetic media over the past decade has been the ability to create deepfakes, which are synthetically generated or manipulated images, audio, or videos that depict people saying or doing things that they never did (Lee 2025: 86). This has prompted concerns about its misuse and harm to society (Masood et al. 2022) in that some deepfakes can pass as "authentic" under user scrutiny (Köbis, Doležalová and Soraperra 2021). Furthermore, the verisimilitude of deepfake outputs is only likely to get more sophisticated (and therefore harder to authenticate as human-created) with time. AI-generated voice cloning undergoes a similar path of rapid development and technical sophistication. Barrington, Cooper and Farid (2025) studied whether 600 participants could distinguish the identity (human or AI-generated) of different audio clips. They concluded that audiences can be easily tricked into thinking that the AI-generated clones were indeed real human speakers, since some could not reliably distinguish them from natural human speech.

Against the backdrop of this mediatic context, we agree with Lee in that "instances of authenticity scandals, authenticity illusions, and authenticity puzzles are endless and continue to grow" (2025: 26). This highlights how the assumption that authenticity rests on human origin is becoming increasingly unstable, as synthetic/artificial outputs can no longer be easily dismissed as inauthentic based on provenance alone. As a result, the crisis of origin gives way to a separate yet concurrent crisis - one of perception - where questions of authenticity rely less on source, but on how audiences evaluate and respond to different configurations of human-machine hybrid productions.

3.3 The second crisis of post-GenAI authenticity: perception

A parallel crisis concerns the shifting perceptions of authenticity in response to perceived origin, as human-AI collaborations and synthetic media continue to polarise different audiences. Lee (2024) argues that it is not yet fully understood how prior beliefs about machines may bias users' positive or negative reactions to AI systems. Once users identify the source, their beliefs about machines can shape how that source influences their subsequent reactions, challenging the CASA assumption that human-machine interaction is largely mindless.

At one end of this extreme, advocates of GenAI (e.g. AI Maximalists) champion the inclusion of generative and algorithmic technologies in all facets of life. This view is underpinned, in part, by what Kim and Koo define as machine heuristics, or "viewing AI as a neutral and unbiased entity that enhances its credibility and contributes to the belief that AI-generated content offers a more objective basis for evaluating materials" (2024: 3). Similarly, Beerends and Aydin (2025) explore the concept of inauthenticity-as-ideal for which the value of (Gen)AI lies precisely in its difference from human production, and for which artificiality is valued as its own distinct aesthetic or ethical value in its own right. On the other extreme, skeptics or critics express negative AI bias (Wasdahl 2024) or AI aversion (Qin et al. 2025). Such responses have produced derogatory or deauthenticating labels for both tools and outputs, such as "clanker" for AI systems (Romo 2025) or "AI slop" for AI-generated artefacts (Roddy and Bridges 2025), and in some cases have led to institutional prohibitions on GenAI technologies, for instance in schools and universities seeking to safeguard academic integrity (Fine Licht 2025).

Amid this polarisation, recent empirical studies have shown a proliferation of divergent perspectives in which authenticity judgements vary according to perceived source rather than the output or quality thereof. Kirk and Givi (2024) note that consumers reduced positive word of mouth and customer loyalty when they believed emotional marketing communications were written by AI versus a human, an effect that was mediated by moral disgust prompted by perceived inauthenticity. Similarly, other studies have shown how people's behaviour and attitudes shift, sometimes drastically, according to the degree of transparency of GenAI disclosure in digital content creation (Brüns and Meißner 2024; Schilke and Reimann 2025). For example, Altay and Gilardi's (2024) experimental study of AI-generated news headlines found that labelling content as AI-generated reduced perceived accuracy and willingness to share, even when the headlines were factually correct. Comparable sentiments are shared by Lermann Henestrosa and Kimmerle (2024), who discovered that labelling content as AI-written lowered source and message credibility. Strikingly, the negative perception diminished when AI involvement was described as partial (e.g. "drafted" or "edited" by AI) or AI was adopted to assist rather than replace humans (Brüns and Meißner 2024), implying that audiences intuitively applied a flexible model of automation, even if authenticity was not explicitly under discussion. Further evidence comes from education and

media contexts where GenAI is used in the creative process. In these cases, a shift in perception towards writers, journalists, or content creators was found. Frederick et al. (2025) discovered in their study that adult learners preferred human or a combination of human-AI teaching materials over exclusively AI-generated ones, with full AI automation harming instructors' perceived credibility and competence. Similarly, Rae (2024) reported that when content creators were discovered to have used GenAI, they suffered reputational consequences, even though the quality and reception of their written content remained unaffected. These examples highlight the crisis of perception at the heart of post-GenAI authenticity, where evaluations shift not with content quality, but source disclosure and attribution.

In this context, while earlier signs in the technological landscape pointed to a gradual erosion of authenticity in its naturalist assumption, the modern landscape introduces new dynamics. First, synthetic media have become increasingly sophisticated and human-like, and users are changing the way they interact with and perceive such content. Second, the very meaning of authenticity in relation to emergent synthetic and artificial products remains underexplored. Recent studies, such as those above, show that perceptions of authenticity fluctuate in response to perceived source, level of AI automation, and framing of AI involvement rather than to output quality or origin alone. Faced with this instability of provenance, many audiences respond not by rethinking authenticity, but by reasserting a simplified opposition between human authenticity and synthetic inauthenticity, particularly in relation to human-AI co-productions.

3.4 Human authenticity versus GenAI inauthenticity

In contemporary discussions of synthetic content, as well as hybrid human-AI authorship, two persistent patterns emerge in the literature. First, authenticity is often asserted but rarely defined (Huang et al. 2025; Lee 2025). Second, there is a tendency to treat origin and reception of authenticity as tightly coupled, consequently framing the authentic as human-origin, and the inauthentic as synthetically or artificially generated in a dichotomous relationship.

For instance, in early adoptions of GenAI (model GPT-2), co-creative experiences, such as the human-AI co-writing tool studied by Yang et al. (2022), revealed a clear divide between human authenticity and AI inauthenticity. The low-quality outputs of GenAI at that time, described as "incoherent and tenuous text" (*ibid*: 7), encouraged early perceptions of such outputs as pejoratively "synthetic" rather than a constructive product of human-AI hybrid collaboration, as they were seen to be useful for no more than inspiration. More recent research by Ju, Kim and Im (2024) on virtual influencers shows that this dynamic persists beyond writing. In their study, users engaged more with hyper-realistic, non-brand operated virtual influencers precisely because they exhibited greater perceived humanlike authenticity. Similarly, Jia et al. (2024) find that in AI-human collaborative news

production, the disclosure of hybrid roles such as AI assistant or AI collaborator reduced perceived humanness and favourability.

More explicit examples can be found in the domains of sociolinguistics and language education. Traditional language authorities, such as the Spanish Real Academia Española (RAE), claim that language generated by machine-learning technologies is inauthentic and destroys the authentic Spanish of human Spanish speakers (Erdocia, Schneider and Migge 2025), while participants in Werdiningsih, Marzuki and Rusdin's (2024) study of using ChatGPT in writing-focused EFL classes reported that they needed their own human input for their writing to be considered authentic or to preserve an authentic voice in the writing process. Moreover, Allaithy and Zaki (2025) in the context of Arabic language teaching suggest that human insight is necessary before AI-generated content reaches students, and that teachers should act as "linguistic gatekeepers" to ensure quality control. Elsewhere, the divide becomes more absolute. For instance, Draxler et al. (2024) describe cases where AI tools are treated as purely passive instruments rather than co-authors, with full authorial attribution and ownership granted to the human user. In some legal contexts, such as U.S copyright law, AI-generated work cannot (at the time of writing) be copyrighted since in that legal framework authorship necessitates human origin and creation (Caldwell 2023).

The above examples reflect a consistent reinforcement of a normalised division between AI and human authorship, where technological or algorithmic mediation is assumed to compromise authenticity rather than contribute to it. As discussed in Section 2, this assumption is underpinned by a naturalist and hierarchical anthropocentric belief (Kalpokas 2023; Kalpokienė and Kalpokas 2023; Kalpokas 2024) that legitimises human productions as superior or primary while positioning non-human or technologically-mediated creations as, in some way, secondary or inferior. This dichotomisation leaves little conceptual room for categories such as human inauthenticity or synthetic authenticity, nor for the many hybrid configurations of human-machine collaboration that may arise and which may or may not ultimately be validated as authentic. To address this gap, in the subsequent section we reconceptualise the concept of authenticity and propose a continuum-based framework.

4 Rethinking authenticity's relationship with syntheticity: introducing a continuum-based framework

By reframing authenticity's conceptual relationship with technologically-mediated productions and human-machine collaborations, we align with recent work by Kalpokienė and Kalpokas (2023) and others (e.g. Köbis, Doležalová and Soraperra 2021; Lavazza 2025), who argue that synthetic outputs can, under certain conditions, be experienced and evaluated as authentic, and their call for embracing a post-humanist model wherein creativity can be shared between humans and machines instead of being limited to human-centred binaries. This resonates with

Beerends and Aydin's (2021; 2025) interactionist authenticity approach for which an AI artefact does not need to display human-like characteristics to be experienced as authentic, as it may be authenticated or deauthenticated through negotiating processes of language, socialisation, and cognition (cf. categorical authenticity in Section 2). These state-of-the-art contributions underscore the need for fluid and future-proof conceptual frameworks that move beyond binary divisions and instead account for the complex and shifting relationships between origin, mediation, and audience reception regarding authenticity. Responding to this need, in the next sections we propose a continuum-based framework that reconceptualises authenticity as vis-à-vis, not versus, syntheticity. Here, the "authentic" and the "synthetic" are not positioned as mutually exclusive categories, but as poles on a spectrum with the wide range of hybrid human-AI co-productions and shifting audience evaluations occupying the many gradations in between.

4.1 Defining syntheticity in relation to authenticity

Compared to authenticity, the *synthetic* has generally posed fewer definitional challenges across disciplinary contexts. While its specific applications vary, this concept has generally been anchored in questions of construction and origin, rather than evaluation or moral judgements. For instance, Kantian epistemology distinguishes between analytic and synthetic propositions (Anderson 2015). Here, the synthetic denotes that which is produced through the synthesis of experience rather than conceptual analysis. A related usage features in the scholarship of Applied Linguistics and Second Language Acquisition, where pedagogical approaches are often categorised as either *analytic* or *synthetic* (Richards and Rodgers 2001). In this context, synthetic curricula rely on instructional models in which language is taught as discrete or isolated components that learners are expected to synthesise progressively into meaningful and communicative language use.

Within media studies and (digital) technology research – and most pertinently for the present discussion – the "synthetic" has become closely associated with content constructed through human-directed synthesis, artificiality, or increasingly through computational or algorithmic generation (Kalpokas 2021). This functions, therefore, as an ontological descriptor about how an artefact comes into being and the degree of technological mediation involved. However, this understanding of the synthetic is insufficient for capturing the complexity and diversity of how artefacts are encountered and evaluated in modern practices in the current GenAI landscape. It is at this juncture that we introduce *syntheticity* as a distinct analytical concept. Syntheticity is the condition – whether actual and/or attributed – of a work, utterance, or artefact as technologically generated or mediated, together with the social evaluations that mark it as distinct from exclusively human-origin production.

In this dual conceptualisation, syntheticity is not exclusively reducible to machine-made provenance or non-human artifice. Rather, it functions as a relational construct that operates along two analytically distinct but interrelated dimensions. The first, *synthetic origin*, refers to the degree to which an item is produced or shaped through computational, algorithmic, or technologically-mediated processes. The second, *synthetic attribution* (or *perceived syntheticity*), refers to the extent to which an artefact is individually or socially constructed, categorised, or labelled as synthetic by audiences regardless of, or due to, its mode of production. Crucially, these two dimensions do not necessarily coincide. An artefact may exhibit a high degree of synthetic origin while being perceived as human-authored, or conversely, may involve minimal technological mediation yet be experienced as synthetic all the same. Therefore, this proposed formulation of syntheticity denotes both the state of being synthetic (that is, of mechanical, computational or algorithmic origin) and the state of being perceived as synthetic. This distinction aligns with empirical findings by Clerwall (2014), Graefe et al. (2018), Fischer and Läubli (2020), Jacobs (2024), Jones et al. (2025), and Lee (2025), who demonstrate that modern audiences increasingly struggle to reliably distinguish synthetic from human output (cf. Section 3.2), and that assumptions about origin can override or displace the evidential authority of provenance itself (cf. Section 3.3). In such cases, beliefs about origin becomes as consequential as provenance itself, such that works of synthetic origin may be valued as legitimate, credible, or experientially authentic, just as human-origin works may be dismissed as contrived, manipulative, or false.

Moreover, human-AI hybridity can be understood as a specific configuration within syntheticity. As mentioned before, many synthetic outputs involve shared human and technological agency (Sundar and Chen 2023). By hybridity, we refer to this subset: a particular configuration of synthetic origin in which human and machine contributions are integrated in co-constitutive ways, ranging from minimal mediation (e.g. grammar checking with a word processor) to extensive co-authorial creation (e.g. AI-assisted art or writing). In such cases, syntheticity does not erase human authorship but reconfigures it, producing degrees of hybridity that challenge simple binaries of authentic versus inauthentic or human versus machine. Significantly, hybrid artefacts often involve a misalignment between synthetic origin and synthetic attribution, such that substantial technological mediation may coexist with perceptions of human authorship, or minimal mediation may nonetheless trigger synthetic labelling. Hybrid productions can be, and often are, authenticated when they resonate with audiences or align with evaluative norms. Empirical studies on human-AI collaboration support this claim: Jia et al. (2024) describe shifting perceptions of AI assistants in journalism, while Moorhouse and Wong (2025) argue that GenAI can enhance authenticity when framed within the author's communicative intent. As such, hybridity operates as a diagnostic case within this model, highlighting how authenticity judgements emerge from the dynamic interaction between production and reception, rather than from provenance alone.

4.2 Conceptualising authenticity vis-à-vis syntheticity: a continuum-based model

As explored earlier, much of the public discourse surrounding authenticity in AI-human cultural production has tended to default to a binary opposition between authenticity and the synthetic, most often understood in terms of synthetic origin. Within such framings, the "authentic" is imagined as distinctly humanly authored, while the "inauthentic" is relegated to the realm of computational or algorithmic (co-)production. The following framework (see Figure 1) challenges this essentialised binary by introducing a two-axis continuum that disaggregates origin from reception, thereby accounting for more dynamic theorisations of authenticity in relation to syntheticity in the age of generative technologies. This framework thus facilitates the conceptualisation of authenticity, not as a fixed attribute of human or technologically-mediated origin, but as a concept that emerges from the alignment, or misalignment, between its origin and reception.

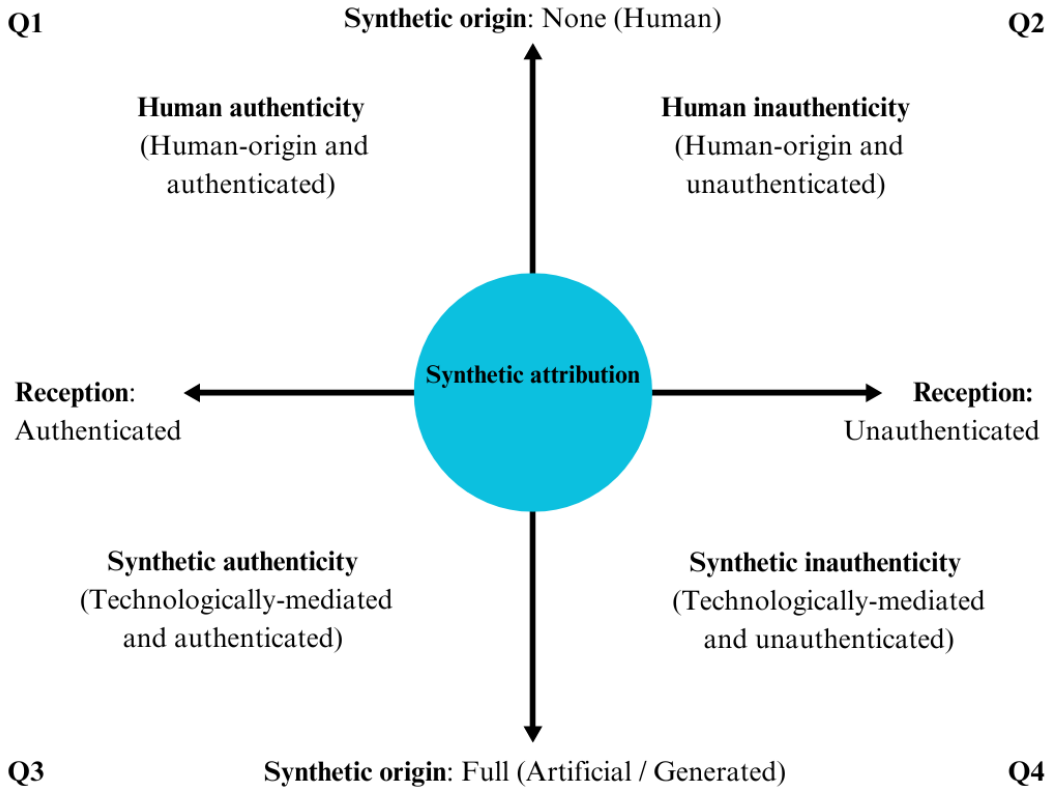

Figure 1: Continuum-based four-quadrant model of authenticity vis-à-vis syntheticity. Created by the authors.

In this model, the vertical axis, Synthetic origin, ranges from *None* (Human) at the top to *Full* (Artificial/Generated) at the bottom. This axis captures whether a work is produced through human agency, hybrid and technologically-mediated

configurations, or through primarily, if not fully, automated computational or algorithmic processes. The horizontal axis (Reception) traces the evaluative trajectory from *Authenticated* (left) to *Unauthenticated* (right),[1] accounting for the range of social, cultural, and institutional judgements through which creations may be experienced as genuine and meaningful (i.e. authentic), or conversely dismissed as contrived or illegitimate (i.e. inauthentic).

The intersection of these two axes yields four analytically distinct but interconnected quadrants: *Human authenticity* (Q1), *Human inauthenticity* (Q2), *Synthetic authenticity* (Q3), and *Synthetic inauthenticity* (Q4). Each quadrant therefore represents a specific configuration of origin and reception, while remaining permeable to movements as evaluations shift over time or across contexts. The model also foregrounds synthetic attribution, that is, the perception, labelling, or attribution of synthetic involvement, as a relational process that may operate independently of actual origin. As represented by the central circle in Figure 1, synthetic attribution does not constitute a third axis; rather, it may traverse and modulate all quadrants by representing how works are interpreted, authenticated, or rejected, regardless of their procedural origin. In this sense, syntheticity is not confined exclusively to origin or reception alone (i.e. synthetic origin or synthetic attribution), but may operate at either level independently, or emerge through their interaction, depending on how production conditions and reception dynamics align. Therefore, syntheticity is best understood not as a bounded category, but as a quality that may traverse origin and reception alike: it may inhere in artefacts as a feature of their production, or circulate as a discursive label applied through processes of authentication or unauthentication.

In relation to the quadrants, the first, *Human authenticity* (Q1, top left) pertains to works of human origin that are also authenticated by audiences as "authentic" in some meaningful way. At the extremes, this quadrant reflects the conventional conflation of human provenance with the authentic. A diary written from personal experience, a signed manuscript, ritual or religious performances, or first-hand reporting in investigative journalism therefore all exemplify this alignment. Conceptually, this quadrant aligns closest with traditional historical authenticity (Newman and Smith 2016), Dutton's (2003) nominal authenticity, and Grayson and Martinec's (2004) indexical authenticity, all of which underscore the verifiable link between (human) artefact and its (human) origin. It also frequently, but not exclusively, overlaps with categorical authenticity (Newman and Smith 2016), insofar as human provenance often functions as a default criterion for what counts as "real" or legitimate (i.e. authenticated) within established categories.

[1] The term *unauthentication* is employed as opposed to other related concepts such as *inauthentication* or *deauthentication* (cf. Beerends and Aydin 2025) as these imply discursive or ideological processes of rejection and therefore presuppose an agentive stance. By contrast, *unauthentication* is intended to be more semantically broad so as to encompass not only explicit disavowal (or related affective positions) but, crucially, also the other non-positional stances toward hybrid or synthetic forms, such as indifference, ignorance, or obliviousness.

Moving downward along the vertical axis, *Synthetic authenticity* (Q3, bottom left) represents works of synthetic or artificial origin that are nonetheless authenticated as legitimate or experientially valuable. This quadrant problematises the aforementioned assumption that authenticity must be tethered to human provenance, demonstrating instead that technological mediation does not preclude, but can in fact constitute its own authenticity. This configuration builds on long-standing research in communication and interaction studies showing that users routinely engage with AI systems, avatars, and computational agents as social actors (Reeves and Nass 1996; Sundar and Chen 2023; Lee 2024). AI-driven conversational assistants adopted for customer service, algorithmically composed music performed in concert, or virtual influencers embraced as authentic by their communities all exemplify this configuration. What these cases share is that authenticity is attributed not through provenance alone but through validation (i.e. authentication) in some way. Related research includes the study of Ramadan and Ramadan (2025), who note that virtual avatars can enhance the authenticity of virtual social connections when they incorporate highly realistic features, such as facial animations, reinforcing anthropomorphic cues (Sundar and Chen 2023; Sundar 2020) and social presence (Short, Williams and Christie 1976). In this example, the virtual avatar embodies an online persona and reflects a user's personal preferences (i.e. hair styles, choice of apparel, etc.), thereby fulfilling a sense of self-authenticity by aligning digital representations with one's perceived true authentic self. As mentioned above, avatars can also be considered as social actors (Nass, Steuer and Tauber 1994; Nass and Moon 2000) without needing to be perceived as human in order to remain authentic. As CASA and subsequent communication media research demonstrate, artificial agents can be experienced as authentic if they behave as "good" machines, that is, by acting in ways that are predictable and appropriate to their assigned social roles. Huang and Jung (2022) demonstrate the continued relevance of the framework for contemporary forms of *Synthetic authenticity*. They show that different kinds of virtual characters (e.g. smart speakers, avatars, 3D projections) retain authenticity as long as they meet user's expectations for machine behaviour. New AI companions can therefore feel authentic for users when they act in ways consistent with their assigned roles, such as offering friendship, deep emotional experiences (Li and Zhang 2024) or engaging in playful interactions.

Synthetic authenticity emerges not because audiences forget that an artefact is artificial or generated, but because technological mediation itself becomes integrated into the criteria by which authenticity is evaluated. Syntheticity, in other words, can itself become the grounds of authenticity when audiences accept technological mediation as part of a work's legitimacy (cf. Beerends and Aydin 2025; Kalpokas, Šalaševičiūtė and Lipskė 2024).

If synthetic origin can be authenticated, then the converse is also possible: *Human inauthenticity* (Q2, top right) represents productions that, despite being of human provenance, are unauthenticated by their audiences. This quadrant reflects

that authenticity is an ongoing evaluative process of validation, contingent on judgements of legitimacy, sincerity, and alignment with cultural expectations (cf. categorical and self-authenticity in Section 2). Examples that fall into the remit of this quadrant would be the production of counterfeit artwork or jewellery (i.e. a fake Rolex watch) that is dismissed on account of its status as a replica or copy despite being human-made, or instances of cultural appropriation such as New Age shamans performing Native American rituals contested as inauthentic imitations (Aldred 2000).

The final quadrant, *Synthetic inauthenticity* (Q4, bottom right), completes the model by representing works of synthetic and technologically-mediated origin that are also perceived to be inauthentic. Here, unauthentication emerges through two primary psychologically distinct pathways: epistemic unauthentication, where rejection stems from perceived deception, falsity, or discrepancy between advertised and actual origin; and affective unauthentication, where rejection arises from perceptual discomfort, uncanny valley effects, or unease triggered by near-human mimicry.

Epistemic unauthentication manifests in multiple forms. At the extreme end of the reception continuum, this is exemplified by categorical or integrity-based rejections in which synthetic origin itself is incompatible with authenticity judgements. Discourses surrounding "AI slop" (Roddy and Bridges 2025; Romo 2025) exemplify this logic. Here, items are invalidated as inauthentic due to the presumption that algorithmic generation inextricably undermines their creativity, effort, or quality. In other words, synthetic attribution functions as a pejorative and essentialist evaluative label that marks generated or human-AI mediated productions as inauthentic by default.

Synthetic inauthenticity may also be driven by deception-based unauthentication processes, where audiences experience or construe a discrepancy between advertised and factual origin. In such cases, synthetic origin and negative reception reinforce one another. This aligns with the abovementioned research on deepfakes (Masood et al. 2022; Lee 2025), or inauthenticity aversion (Silver, Newman and Small 2020), which posits that when synthetic artefacts are presented as human-made, audiences can respond with moral condemnation, distrust, or rejection by interpreting the misalignment as a form of deception.

Affective unauthentication, by contrast, can occur even in the absence of explicit deception. Drawing on uncanny valley research, studies demonstrate that artefacts which approach, but fail to fully achieve, human likeness often provoke feelings of eeriness, coldness, or soullessness (Mori, MacDorman and Nageki 2012; Kätsyri et al. 2015; MacDorman and Chattopadhyay 2016). This reflects often unconscious responses to subtle inconsistencies in digital representations of human appearance, behaviour, or expressivity, rather than rationalised judgements about origin. For example, this mechanism is evidenced in empirical work on human-AI interaction in which highly anthropomorphised chatbots or virtual AI agents can trigger discomfort and negative emotional reactions when perceived as being "too

human" yet insufficiently authentic (Lee 2024; Li and Zhang 2024; Giri et al. 2026). In such cases, S*ynthetic inauthenticity* emerges not from obvious artificiality or poor quality, but from affective dissonance produced by near-human mimicry. As Mieczkowski (2022) argues, perceptions of increasing machine agency are often accompanied by a corresponding reduction in perceived human agency, producing feelings of vulnerability or confusion. When GenAI systems appear overly agentic or self-aware, this imbalance can trigger a crisis of authenticity, leading users to reject specific manifestations of human-like synthetic interaction despite their technical sophistication.

Crucially, recognising these multiple pathways to unauthentication reinforces the central argument of this article: syntheticity is not inherently inauthentic. Rather, S*ynthetic inauthenticity* can manifest through reception, whether via perceived deception, perceptual unease, or both. As with authenticity more broadly, rejection in Q4 is not an intrinsic property of artefacts or technologies, but the outcome of complex socially and psychologically mediated processes of evaluation, negotiation, and verification (Newman and Smith 2016; Beerends and Aydin 2025).

In practical terms, the framework works as follows: the vertical axis tracks whether an item or idea is made by humans, machines, or a combination of the two. The horizontal axis tracks whether audiences accept this as authentic or reject it as inauthentic. This, in turn, creates four broad possibilities: authentically human-made (Q1), inauthentically human-made (Q2), authentically synthetic (Q3), and inauthentically synthetic (Q4). Between these extremes lie various hybrid collaborations that may be validated or contested to different degrees, as well as works whose authenticity status remains unsettled or disputed (see Figure 2 in Section 4.3). In summary, the four quadrants provide a more flexible and relational account of authenticity and syntheticity in the age of generative technologies. They show that these are not fixed properties of origin alone, but the emergent outcome of an interplay between both provenance and observer reception, mediated by processes of synthetic attribution where relevant. In addition, the proposed framework also creates space to analyse the intermediary zones between these quadrants, where works do not fall neatly into the extremes of fully authentic or inauthentic, human or synthetic.

4.3 Between the extremes: (Un)Authenticated hybridity and Liminal authentication

While the four quadrants illustrate the extremes of origin and reception, much contemporary production resides in the spaces between them. To capture this complexity, the framework introduces two intermediary categories: *Authenticated* and *Unauthenticated hybridity* (situated vertically between Q1-Q3 and Q2-Q4 respectively) and *Liminal authentication* (situated between the extremes of the central horizontal axis). These zones illuminate the ways in which authenticity and

syntheticity are in flux in contexts where provenance and reception are neither straightforward nor stable. These are illustrated in the following Figure 2:

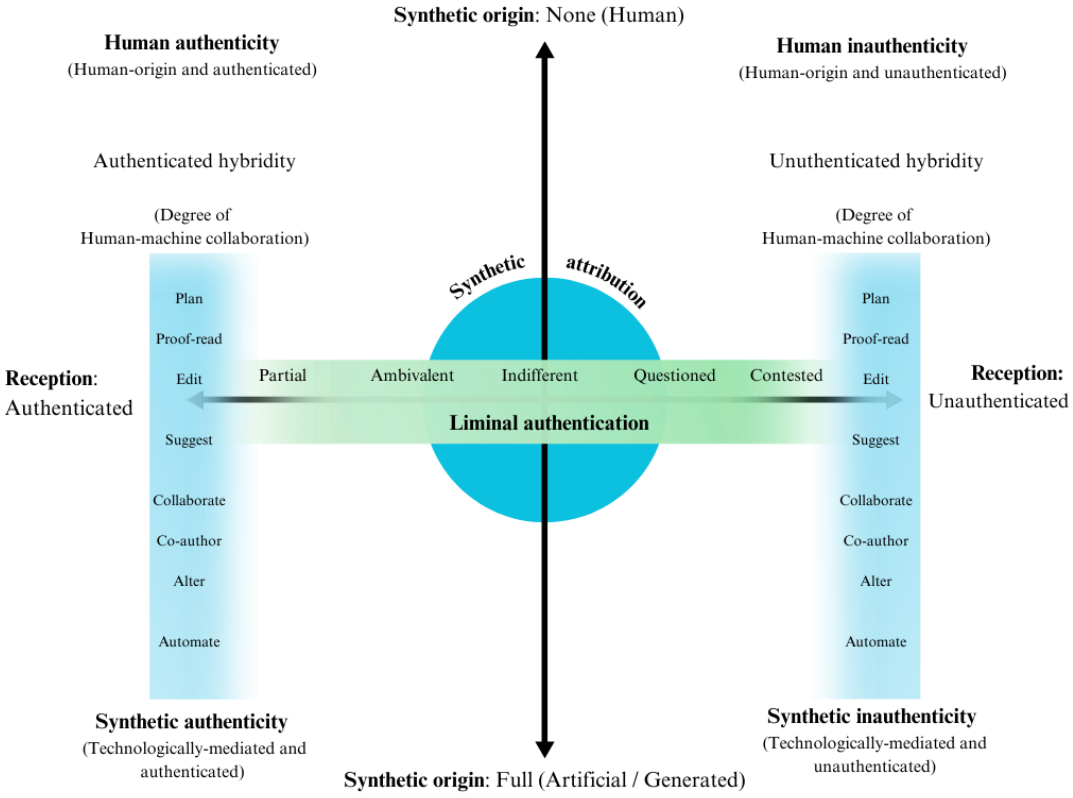

Figure 2: Complete continuum-based and four quadrant model for authenticity in relation to origin and reception including intermediary zones across the horizontal axis (*Liminal authentication*) and vertical axes (*Authenticated* and *Unauthenticated hybridity*). Created by the authors.

Authenticated and *Unauthenticated hybridity* represent divergent receptional outcomes of synthetic attribution. In this sense, while hybridity describes the structural configuration of human-machine collaboration, it is the evaluation of syntheticity, specifically synthetic attribution, that determines whether such configurations are authenticated or rejected, and to which degree. In this framework, the concept of *Authenticated hybridity* is proposed to explain the range of works of mixed human and computational origin that are nonetheless authenticated as legitimate or experientially authentic. Rather than treating technological mediation as a disqualifying factor, this category highlights the rich diversity of co-authored productions in which human and machine agency co-exist (Zafar, Ali and Yasin 2025). As discussed in Section 3, hybridity can take many forms, spanning from minimal intervention, such as grammar correction or stylistic refinement, to more substantive forms, such as deepfake generation or iterative image co-creation. At the far end of this continuum are co-authorial practices in

which human and (Gen)AI contributions are equally visible (or acknowledged) in the final product. Across these multiple variations, it is imperative to note that hybridity does not dissolve authenticity but instead can contribute to it. This reflects recent empirical findings that partial AI involvement can often be evaluated more positively than full automation (Brüns and Meißner 2024; Zhang et al. 2025), and that co-creative configurations can sustain perceptions of human authorship and intentionality (Jia et al. 2024; Moorhouse and Wong 2025).

Conversely, the opposing vertical continuum of *Unauthenticated hybridity* represents a gradual spectrum where different degrees of technological mediation coincide with audience suspicion or even full rejection. Plagiarised essays (Huang et al. 2025), heavily automated or generated cinematographical productions criticised as being "soulless" (Halperin and Rosner 2025), and politically deceptive deepfake or cheapfakes (Hameleers 2024) all exemplify this trajectory, with minor forms of hybridity typically being situated closer to the authenticated end (i.e. planning, proofreading, or minor editing), while major collaborations or substitution of human multimodal creativity are often positioned further toward inauthenticity. In such cases, negative synthetic attribution attached to hybrid productions amplifies rather than mitigates perceptions of inauthenticity. Therefore, in cases of *Authenticated hybridity*, audiences validate hybrid human-AI productions as authentic because authenticity is experienced not through provenance alone, but through resonance, credibility, and perceived validity; in unauthenticated cases, the same configurations are rejected when synthetic attribution reframes these qualities and leads audiences to interpret technological or generative intervention as a marker of inauthenticity.

The horizontal band of *Liminal authentication* is positioned between these poles, cutting across all four quadrants centrally to capture works whose reception is not yet decided and is, in some way, partial, contested, or reversible. *Liminal authentication* thus marks the indeterminate middle ground along the reception axis where creations are neither fully validated (i.e. authenticated) nor wholly rejected (i.e. unauthenticated) as they remain subject to negotiation, re-evaluation, or even uncertainty. For example, this is reflected in scenarios where a human-authored text is under investigation due to suspicion of AI authorship in an institutional setting, or in instances in which students are directly falsely accused (Stone 2024). Indeed, AI detection systems have been shown to habitually misclassify the writing of students as machine-generated despite student work being wholly human-authored, especially in the case of L2 language learners or users with non-standard linguistic styles (Liang et al. 2023; Wu et al. 2025). *Liminal authentication* is relevant here because the student work occupies a suspended state of validation. Initially, their work is accused (by machine or another social actor) of having a synthetic origin via synthetic attribution. Next, it undergoes a process of investigation and negotiation (i.e. interpersonally or institutionally), before those involved reach a final (but not necessarily permanent) decision: authentication or unauthentication. Significantly, this liminal condition does not only operate at the point of evaluation,

but prospectively also. Authors may engage in pre-emptive self-censorship or extensive revision in an effort to avoid accusations of AI usage (Perkins et al. 2024; Bannister 2025). Such cases illustrate how *Liminal authentication* in its different manifestations highlights the instability of authenticity itself (Lee 2020) and reveals how outputs can shift across evaluative boundaries over time in response to changing cultural, institutional, or technological framings.

5 Applications and limitations

By decoupling ontological origin from phenomenological reception, the proposed framework avoids essentialising *authentic* as "human" and *synthetic* as "inauthentic". In so doing, it marks a conceptual shift away from the binary logic that has long cast *authenticity* as natural, embodied, and humanly authored, while relegating *syntheticity* to the realm of the artificial, the deceptive, or the illegitimate. Instead, the model reconceptualises authenticity while also introducing syntheticity as coexisting within a common interpretative space, especially in the context of the emerging era of GenAI.

The practical value of this shift lies in the framework's two-axis structure that provides scholars and practitioners alike with a scalable tool for classifying and empirically testing works across different domains and cultural contexts. The framework is sufficiently abstract and descriptive to account for the multitude of ever-changing scenarios in which authenticity and syntheticity discussions unfold, while still offering pertinent distinctions that can guide case-specific analyses. Moreover, the framework refrains from imposing normative or prescriptive claims about what ought to be "authentic", "synthetic" or the in-between. In other words, the purpose of this article has not been to champion a particular stance or to comment on the influence of AI in the current landscape. Indeed, productive critical conversations are already taking place in areas such as post-humanism (cf. Kalpokiene and Kalpokas 2023; Nicenboim et al. 2023; Giaccardi, Redström and Nicenboim 2024; Lackey and Papacharissi 2024), AI-mediated or human-AI relationships (cf. Lee 2024; Battisti 2025), the existential crisis of what it means to be human (cf. Lackey and Papacharissi 2024; Beerends and Aydin 2025), or the consequences of anthropomorphising machines or algorithms (cf. Placani 2024; Simas and Ulbricht 2024). What this framework offers is a structured conceptual framework for analysing how authenticity and syntheticity are validated, negotiated, or contested across shifting technological and cultural settings.

5.1 Applications

The framework makes four key contributions to current debates on authenticity and human-machine communication in the age of GenAI. First, it introduces the concept of *syntheticity*, understood in terms of origin and/or attribution, as a construct that operates alongside, and in relation to, authenticity in modern communicative and

mediated landscapes. The model illustrates how syntheticity and authenticity are not inherently mutually exclusive but may coexist in a shared evaluative field. In so doing, the second area of contribution highlights the need to decouple origin from reception in the theorisation of post-GenAI authenticity. As discussed in Section 3, contemporary discussions often conflate human provenance with authenticity and synthetic provenance with inauthenticity. By separating these dimensions, the proposed model both accommodates and accounts for additional configurations such as *Human inauthenticity* and *Synthetic authenticity* within the same theoretical framework. Third, the framework aligns with research that positions hybridised human-machine co-productions as a legitimate locus of authenticity (Rammert 2008; Sundar 2020; Sundar and Chen 2023; Kalpokas, Šalaševičiūtė and Lipskė 2024). The categories of *(Un)Authenticated hybridity* recognise that co-authored or mediated productions can be legitimised and accepted when they resonate with audiences and their validation processes (or not). Finally, the model accounts for perceptual and temporal dynamism in authenticity and syntheticity judgements. Through the concept of *Liminal authentication*, it captures ambiguity, misclassification, and shifts over time, while also recognising that authenticity evaluations are non-linear and socially mediated. Furthermore, a high degree of hybridity (or even full synthetic automation) does not determine whether an output or creation will be authenticated or not, as the scales of *Authenticated* and *Unauthenticated hybridity* are not predictive of reception outcome. In other words, the model reflects non-linear mapping in which greater syntheticity of origin does not necessarily entail diminished authentication (though of course it still may do this), and by extension minimal technological mediation does not guarantee audience acceptance. This responds to recent calls to move beyond static notions of authenticity in technologically-mediated communication (Lee 2020; Huang et al. 2025) and complements current discussions that post-digital authenticity is shaped by contextual, temporal, and communal factors (Köbis, Doležalová and Soraperra 2021; Kalpokas, Šalaševičiūtė and Lipskė 2024).

While elements of the proposed framework resonate with established insights from human-machine communication (e.g. distributed agency, hybridity, co-production) and communication and media research, the present model does not seek to redescribe these phenomena under new terminology. Rather, its contribution lies in integrating these strands within a single, explicitly relational theoretical framework that jointly accounts for origin, attribution, and reception in authenticity judgements. Unlike existing models that foreground either production dynamics (e.g. agency, co-creation) or audience perception (e.g. disclosure, credibility effects) in relative isolation, the continuum-based framework proposed in this article illustrates how provenance (i.e. synthetic origin) and validation (i.e. authentication) do not always align, while introducing syntheticity as the construct through which divergences between these dimensions become visible. This enables observers to classify cases in which human-authored work is rejected, synthetic or technologically-mediated work is authenticated, and hybrid productions oscillate

between authenticated and unauthenticated across contexts or over time. In doing so, the model provides a scalable structure for comparing authenticity judgements across domains (such as education, journalism, and creative industries), for empirically testing the effects of attribution and disclosure, and for tracking how works move between authenticated, liminal, and unauthenticated states.

In terms of practical applications, particularly within digital media literacy, the framework enables researchers and practitioners to identify and analyse systematic patterns of synthetic attribution that reveal underlying biases, such as in cases where audiences classify human works as synthetic or vice versa. Repeated patterns of misclassification, whether based on stylistic cues, cultural stereotypes, or assumptions about what "authentic" expression should look like, may serve as diagnostic indicators of certain ideological predispositions in audience evaluation. This application is supported by recent empirical findings that demonstrate that authenticity judgements are often shaped less by provenance itself than by disclosure or labelling practice (Clerwall 2014; Fischer and Läubli 2020; Altay and Gilardi 2024; Lermann Henestrosa and Kimmerle 2024). By mapping such cases onto the framework's quadrants and intermediary zones, future research can systematically disentangle perceptual and attributional bias from evaluative assessment, thereby tracing where rejection reflects assumptions about synthetic origin rather than deficiencies in content quality or communicative effectiveness.

Another immediate application lies in higher education, where human-AI authorship is increasingly present (Kasneci et al. 2023; Bannister 2025). In curriculum design and AI literacy initiatives, the framework may enable educators to help students critically examine not only the degree of synthetic origin involved in a task but also how synthetic attribution and reception shape the evaluation of their work or that of others. Rather than treating AI use as a binary presence or absence, the two-axis continuum-based framework can encourage students to reflect on how varying forms of technological mediation can be interpreted, legitimised, or problematised. Similarly, in language teacher education and assessment research, the model offers an analytic tool for examining how educators respond to AI-assisted student work: whether partial human authorship mitigates concerns about academic integrity, or whether any perceived synthetic involvement prompts categorical rejection. Mapping such evaluations onto the continuum makes visible the often implicit assumptions that govern grading decisions and integrity judgements. In this way, the framework can inform the development of assessment policies and pedagogical guidelines that move beyond rigid binary notions of authenticity toward more transparent and context-sensitive approaches. Future research might also explore whether explicit engagement with the continuum model (or adapted versions of it) supports students' digital literacy, metacognitive awareness, and evaluative practices in hybrid learning environments.

5.2 Limitations

Despite these applications and the theoretical contributions to the discussion of authenticity and syntheticity in the age of GenAI, the framework has several limitations that warrant acknowledgment. To begin, the model is not holistic and applies most directly to content, output, and authorship in the context of human creations, human-machine communication, and/or synthetic media (Kalpokas, Šalaševičiūtė and Lipskė 2024). As discussed in Section 2, *authenticity* is a polysemous construct and extending the model beyond the area of cultural production risks stretching its analytic utility beyond its intended scope. In addition, the empirical research and normative assumptions mobilised throughout the article are largely grounded in Western institutions and domains, such as higher education and media production. Therefore, the framework's application in other contexts (i.e. the Global South) cannot be assumed.

At present, the framework remains theoretically rather than empirically grounded. While scholarship is drawn upon to illustrate the framework's components through selected examples, no further studies have yet applied the model systematically across domains or audiences. Future empirical research is required to test the model's reliability and validity in practice and across diverse settings.

As for the visual representation of the model in Figures 1 and 2, they necessarily simplify complex socio-technological relationships and outcomes. The horizontal and vertical axes of Synthetic origin and Reception are conceptualised as singular scalar dimensions, whereas real-world authentication processes are often plural and overlapping. For example, AI-generated work might be embraced as authentic within one evaluative frame (e.g. for its aesthetic innovation) while simultaneously rejected within another (e.g. for its lack of originality). Such cases of simultaneous (un)authentication are difficult to place on a single continuum. In addition, the intermediary continua introduced in Section 4.3 are not exhaustive. Due to constraints of space and scope, it has not been possible to visualise all potential gradients or transitional pathways between quadrants, for instance, a diagonal continuum between Q1 (*Human authenticity*) and Q4 (*Synthetic inauthenticity*) which may capture scenarios in which both origin and reception shift simultaneously over time. Finally, while the framework accommodates a wide range of human-machine collaborative configurations in principle, it does not yet differentiate between specific modes of collaboration. Emerging work has begun to quantify and typologise human-AI co-creative practices in domains such as music production (Davis et al. 2025) and information science (Salma, Hijón-Neira and Pizarro 2025), though future refinements of the model could build on this emerging research to clarify how distinct collaborative roles, such as editing, curating, co-authoring, or automating, shape audience evaluations and mediate the relationship between authenticity and syntheticity.

6 Conclusion

This article has examined how *authenticity* has been theorised across different disciplines such as philosophy, sociology, linguistics, and communication studies, highlighting its multidimensional character and persistent definitional challenges. It has further shown how the rise of GenAI technologies has intensified these challenges by producing two interlinked crises: a crisis of origin, in which provenance is destabilised by the growing difficulty of distinguishing between human and synthetic outputs, and a related crisis of perception, in which authenticity judgements increasingly depend on disclosure, framing, and social reception rather than on content quality alone.

Against this backdrop, the central contribution of this article has been to introduce a continuum-based framework that reconceptualises authenticity vis-à-vis syntheticity not as mutually exclusive categories but as mutually shaping constructs. By decoupling origin from reception, the model makes analytically visible forms of misalignment that are otherwise obscured in binary formulations. In other words, the framework demonstrates how human work can be unauthenticated, synthetic work authenticated, and hybrid human-machine productions (including GenAI) can serve as legitimate sites of authentic expression. The framework further foregrounds intermediary zones of technologically-mediated hybridity and *Liminal authentication*, thereby accounting for the dynamic and negotiable nature of authenticity judgements in the rapidly evolving age of generative technologies.

While the framework is presently theoretical and domain-specific, its applications are wide-ranging. It offers a shared conceptual and terminological tool for analysing how provenance and reception interact in authenticity judgements across technosocial domains. The framework also provides a basis for empirical extension, inviting future research to examine how authenticity judgements vary across different forms of human-machine collaboration and social contexts. As GenAI technologies advance, alongside others yet to be developed, we anticipate that the framework will remain a useful lens for mapping the shifting relationships between human creativity, machine mediation, and audience validation.

Conflicts of interest

The authors declare no conflicts of interest regarding the publication of this contribution.

References

Anderson, R. Lanier. 2015. *The Poverty of Conceptual Truth: Kant's Analytic/Synthetic Distinction and the Limits of Metaphysics*. Oxford: Oxford University Press.

Aldred, Lisa. 2000. Plastic shamans and astroturf sun dances: new age commercialization of Native American spirituality. *American Indian Quarterly* 24(3). 329–352. http://www.jstor.org/stable/1185908 (last accessed on 26/02/2026).

Allaithy, Ahmed & Zaki, Mai. 2025. Evaluation of AI-generated reading comprehension materials for Arabic language teaching. *Computer Assisted Language Learning.* 1–33. https://doi.org/10.1080/09588221.2025.2474037

Altay, Sacha & Gilardi, Fabrizio. 2024. People are skeptical of headlines labeled as AI-generated, even if true or human-made, because they assume full AI automation. *PNAS Nexus* 3(10). 1–11. https://doi.org/10.1093/pnasnexus/pgae403

Bailey, Erica R. & Matz, Sandra C. & Youyou, Wu & Iyengar, Sheena S. 2020. Authentic self-expression on social media is associated with greater subjective well-being. *Nature Communications* 11. 1–9. https://doi.org/10.1038/s41467-020-18539-w

Bannister, Peter. 2025. ParadAIse L0st? *Higher Education Research & Development*. Comment. 1–10. https://doi.org/10.1080/07294360.2025.2586653

Baptista, João Pedro & Rivas-de-Roca, Rubén & Gradim, Anabela & Pérez-Curiel, Concha. 2025. Human-made news vs AI-generated news: a comparison of Portuguese and Spanish journalism students' evaluation. *Humanities and Social Science Communications* 12. 1–9. https://doi.org/10.1057/s41599-025-04872-2

Barrington, Sarah & Cooper, Emily & Farid, Hany. 2025. People are poorly equipped to detect AI-powered voice clones. *Scientific Reports* 15. 1–9. https://doi.org/10.1038/s41598-025-94170-3

Battisti, Davide. 2025. Second-person authenticity and the mediating role of AI: a moral challenge for human-to-human relationships? *Philosophy & Technology* 38(1). 1–19. https://doi.org/10.1007/s13347-025-00857-w

Beerends, Siri & Aydin, Ciano. 2021. Negotiating authenticity in technological environments. *Philosophy & Technology* 34. 1665–1685. https://doi.org/10.1007/s13347-021-00480-5

Beerends, Siri & Aydin, Ciano. 2025. Negotiating the authenticity of AI: how the discourse on AI rejects human indeterminacy. *AI & Society* 40. 263–276. https://doi.org/10.1007/s00146-024-01884-5

Beverland, Michael B. & Lindgreen, Adam & Vink, Michiel W. 2008. Projecting authenticity through advertising: consumer judgments of advertisers' claims. *Journal of Advertising* 37(1). 5–15. https://doi.org/10.2753/JOA0091-3367370101

Bialystok, Lauren. 2014. Authenticity and the limits of philosophy. *Dialogue* 53. 271–298. https://doi.org/10.1017/S001221731300111X

Borji, Ali. 2023. Qualitative failures of image generation models and their application in detecting deepfakes. *Image and Vision Computing* 137. 1–21. https://doi.org/10.1016/j.imavis.2023.104771

Brandtzaeg, Petter Bae & You, Yukun & Wang, Xi & Lao, Yucong. 2023. "Good" and "bad" machine agency in the context of human-AI communication: The case of ChatGPT. In Degen, Helmut & Ntoa, Stavroula & Moallem, Abbas (eds), *Proceedings of the HCI International 2023 - Late Breaking Papers* (Copenhagen, July 23-28, 2023), 3–23. https://doi.org/10.1007/978-3-031-48057-7_1

Brüns, Jasper David & Meißner, Martin. 2024. Do you create your content yourself? Using generative artificial intelligence for social media content creation diminishes perceived brand authenticity. *Journal of Retailing and Consumer Services* 79. 1–15. https://doi.org/10.1016/j.jretconser.2024.103790

Bui, Hien Thu & Filimonau, Viachaslau & Sezerel, Hakan. 2024. AI-thenticity: exploring the effect of perceived authenticity of AI-generated visual content on tourist patronage intentions. *Journal of Destination Marketing and Management* 34. 1–11. https://doi.org/10.1016/j.jdmm.2024.100956

Busselle, Rick W. & Bilandzic, Helena. 2008. Fictionality and perceived realism in experiencing stories: a model of narrative comprehension and engagement. *Communication Theory* 18(2). 255–280. https://doi.org/10.1111/j.1468-2885.2008.00322.x

Caldwell, Mackenzie. 2023. What is an 'author'? Copyright authorship of AI art through a philosophical lens. *Houston Law Review* 61(2). 411–442. https://houstonlawreview.org/article/92132-what-is-an-author-copyright-authorship-of-ai-art-through-a-philosophical-lens

Calefato, Fabio & Lanubile, Filippo. 2010. Chapter 6 - Communication media selection for remote interaction of Ad Hoc groups. In Zelkowitz, Marvin V. (ed.), *Advances in Computers* 78, 271–313. https://doi.org/10.1016/S0065-2458(10)78006-2

Carroll, Glenn R. & Kovács, Balázs. 2021. Authenticity: meanings, targets, audiences and third parties. *Research in Organizational Behavior* 41. 1–13. https://doi.org/10.1016/j.riob.2021.100149

Carroll, Glenn R. 2015. Authenticity: attribution, value, and meaning. In Scott, Robert A. & Kosslyn, Stephen M. (eds), *Emerging Trends in the Social and*

Behavioral Sciences, 1–13. Hoboken: Wiley Online Library. https://doi.org/10.1002/9781118900772.etrds0020

Chan, Cecilia Ka Yuk & Hu, Wenjie. 2023. Students' voices on generative AI: perceptions, benefits and challenges in higher education. *International Journal of Educational Technology in Higher Education* 20. 1–18. https://doi.org/10.1186/s41239-023-00411-8

Clerwall, Christer. 2014. Enter the robot journalist: users' perceptions of automated content. *Journalism Practice* 8(5). 519–531. https://doi.org/10.1080/17512786.2014.883116

Coupland, Nikolas. 2003. Sociolinguistic authenticities. *Journal of Sociolinguistics* 7(3). 417–431. https://doi.org/10.1111/1467-9481.00233

Coupland, Nikolas. 2014. Language, society and authenticity: themes and perspectives. In Lacoste, Véronique & Leimgruber, Jakob & Breyer, Thiemo (eds), *Indexing Authenticity: Sociolinguistic Perspectives*, 14–39. Berlin: Mouton de Gruyter.

Dammann, Olaf & Friederichs, Katja M. & Lebedinski, Sabine & Liesenfeld, Kerstin M. 2021. The essence of authenticity. *Frontiers in Psychology* 11. 1–6. https://doi.org/10.3389/fpsyg.2020.629654

Davis, Jenny L. 2019. Authenticity, digital media, and person identity verification. In Stets, Jan E. & Serpe, Richard T. (eds), *Identities in Everyday Life*, 93–111. New York: Oxford University Press.

Davis, Nicholas & Clemens, Michael & Browne, Eric & Rezwana, Jeba. 2025. Unlocking the black box of Artificial Media with quantified and explainable co-creative AI systems. In Zagalo, Nelson & Keller, Damian (eds), *Artificial Media. Emerging Trends in Narratives, Education and Creative Practice*, 21–48. Cham: Springer.

Dehnert, Marco & Mongeau, Paul A. 2022. Persuasion in the age of artificial intelligence (AI): theories and complications of AI-based persuasion. *Human Communication Research* 48(3). 386–403. https://doi.org/10.1093/hcr/hqac006

Draxler, Fiona & Werner, Anna & Lehman, Florian & Hoppe, Matthias & Schmidt, Albrecht & Buschek, Daniel & Welsch, Robin. 2024. The AI ghostwriter effect: When users do not perceive ownership of AI-generated text but self-declare as author. *ACM Transactions on Computer-Human Interaction* 31(2). 1–40. https://doi.org/10.1145/3637875

Dutton, Denis. 2003. Authenticity in arts. In Levinson, Jerrold (ed.), *The Oxford Handbook of Aesthetics*, 258–274. New York: Oxford University Press.

Eckert, Penelope. 2003. Elephants in the room. *Journal of Sociolinguistics* 7(3). 392–397. https://doi.org/10.1111/1467-9481.00231

Erdocia, Iker & Schneider, Britta & Migge, Bettina. 2025. Language in the age of AI technology: from human to non-human authenticity, from public governance to privatised assemblages. *Language in Society* 54. 1–21. https://doi.org/10.1017/S004740452500017X

Fine Licht, Karl. 2025. Rethinking the ethics of GenAI in higher education: a critique of moral arguments and policy implications. *Journal of Applied Philosophy* 42(4). 1317–1337. https://doi.org/10.1111/japp.70026

Fischer, Lucas & Läubli, Samuel. 2020. What's the difference between professional human and machine translation? A blind multi-language study on domain-specific MT. In Way, Andy & Forcada, Mikel L. & Galuščáková, Petra & Popel, Martin (eds), *Proceedings of the 22nd Annual Conference of the European Association for Machine Translation* (Lisbon, November 3-5, 2020), 215–224. https://aclanthology.org/2020.eamt-1.23.pdf

Floridi, Luciano. 2024. On the future of content in the age of artificial intelligence: Some implications and directions. Editor letter. *Philosophy & technology* 37(112). 1–11. https://doi.org/10.1007/s13347-024-00806-z

Frederick, Austin & Lin, Xialing & Edwards, Chad & Edwards, Autumn. 2025. AI and human generation of classroom content: adult learners' perceptions. *Communication Education* 74(2). 168–181. https://doi.org/10.1080/03634523.2025.2466502

Gherheș, Vasile & Fărcașiu, Marcela Alina & Cernicova-Buca, Mariana & Coman, Claudiu. 2025. AI vs. human-authored headlines: evaluating the effectiveness, trust, and linguistic features of ChatGPT-generated clickbait and informative headlines in digital news. *Information* 16(2). 1–21. https://doi.org/10.3390/info16020150

Ghiurău, David & Popescu, Daniela Elena. 2024. Distinguishing reality from AI: approaches for detecting synthetic content. *Computers* 14(1). 1–33. https://doi.org/10.3390/computers14010001

Giaccardi, Elisa & Redström, Johan & Nicenboim, Iohanna. 2024. The making(s) of more-than-human design: introduction to the special issue on more-than-human design and HCI. *Human Computer Interaction* 40(1-4). 1–16. https://doi.org/10.1080/07370024.2024.2353357

Giri, Arunangshu & Chakrabarty, Dipanwita & Lim, Weng Mark & Mangla, Sachin Kumar & Hollebeek, Linda D. 2026. Can generative artificial intelligence enhance brand performance in tourism? A mixed-method study integrating service-dominant logic, social exchange theory, and uncanny valley theory. *Tourism Management Perspectives* 60. 1–18. https://doi.org/10.1016/j.tmp.2025.101442

Giridhar, Akash & Ravi, Anirudh & Kumar, Anjar & Charan B. & Vanamala, Honganur Rajanna. 2025. Detecting AI-generated artwork: an ensemble model for enhancing visual authenticity in the Digital Age. In *Proceedings 2025*

International Conference on Inventive Computation Technologies ICICT (Kirtipur, April 23-25, 2025), 475–481. Piscataway: IEEE. http://doi.org/10.1109/ICICT64420.2025.11004722

Graefe, Andreas & Haim, Mario & Haarmann, Barbara & Brosius, Hans-Bernd. 2018. Readers' perception of computer-generated news: credibility, expertise, and readability. *Journalism* 19(5). 595–610. https://doi.org/10.1177/1464884916641269

Grayson, Kent & Martinec, Radan. 2004. Consumer perceptions of iconicity and indexicality and their influence on assessments of authentic market offerings. *Journal of Consumer Research* 31(2). 296–312. https://doi.org/10.1086/422109

Groth, Oden & Block, Lauren & Newman, George E. 2023. Lost in digitalization: the physical format of creative work affects authenticity perceptions. *Psychology of Aesthetics, Creative, and the Arts* 19(6). 1578–1592. https://dx.doi.org/10.1037/aca0000628

Guignon, Charles. 2004. *On Being Authentic*. Oxon: Routledge.

Guignon, Charles. 2008. Authenticity. *Philosophy Compass* 3(2). 277–290. https://doi.org/10.1111/j.1747-9991.2008.00131.x

Guzman, Andrea L. 2018. *Human-machine communication. Rethinking communication, technology, and ourselves*. New York: Peter Lang.

Guzman, Andrea L. & McEwen, Rhonda & Jones, Steve. 2023. Introduction to the handbook. In Guzman, Andrea L. & McEwen, Rhonda & Jones, Steve (eds), *The SAGE Handbook of Human-Machine Communication*, xli-xlvi. London: Sage.

Haimson, Oliver L. & Liu, Tianxiao & Zhang, Ben Zefeng & Corvite, Shanley. 2021. The online authenticity paradox: what being "authentic" on social media means, and barriers to achieving it. In Nichols, Jeff (ed.), *Proceedings of the ACM on Human-Computer Interaction* 5 CSCW2, 1–18. https://doi.org/10.1145/3479567

Halperin, Brett A. & Rosner, Daniela K. 2025. "AI is soulless": Hollywood film workers' strike and emerging perceptions of generative cinema. *ACM Transactions on Computer-Human Interaction* 32(2). 1–27. https://doi.org/10.1145/3716135

Hameleers, Michael & van der Meer, Tony G & Dobber, Tom. 2024. They would never say anything like this! Reasons to doubt political deepfakes. *European Journal of Communication* 39(1). 56–70. https://doi.org/10.1177/02673231231184703

Hameleers, Michael. 2024. Cheap versus deep manipulation: the effects of cheapfakes versus deepfakes in a political setting. *International Journal of Public Opinion Research* 36(1). 1–19. https://doi.org/10.1093/ijpor/edae004

Hohenstein, Jess & Jung, Malte. 2020. AI as a moral crumple zone: the effects of AI-mediated communication on attribution and trust. *Computers in Human Behavior* 106. 1–13. https://doi.org/10.1016/j.chb.2019.106190

Huang, Dongpeng & Hash, Nicole & Cummings, James J. & Prena, Kelsey. 2025. Academic cheating with generative AI: exploring a moral extension of the theory of planned behaviour. *Computers and Education: Artificial Intelligence* 8. 1–16. https://doi.org/10.1016/j.caeai.2025.100424

Huang, Junru & Jung, Younbo. 2022. Perceived authenticity of virtual characters makes the difference. *Frontiers in Virtual Reality* 3. 1–15. https://doi.org/10.3389/frvir.2022.1033709

Jacobs, Bart. 2024. The authenticity crisis. *Computer Law & Security Review: The International Journal of Technology Law and Practice* 53. 1–5. https://doi.org/10.1016/j.clsr.2024.105962

Jia, Huiyan & Appelman, Alyssa & Wu, Mu & Bien-Aimé, Steve. 2024. News bylines and perceived AI authorship: effects on source and message credibility. *Computers in Human Behavior: Artificial Humans* 2(2). 1–9. https://doi.org/10.1016/j.chbah.2024.100093

Jones, Cameron R. & Rathi, Ishika & Taylor, Sydney & Bergen, Benjamin K. 2025. People cannot distinguish GPT-4 from a human in a Turing test. In *FAccT'25: Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency* (Athens, June 23-26, 2025), 1615–1639. https://doi.org/10.1145/3715275.3732108

Ju, Naan & Kim, Terry & Im, Hyunjoo. 2024. Fake human but real influencer: the interplay of authenticity and humanlikeness in virtual influencer communication? *Fashion and Textiles* 11. 1–27. https://doi.org/10.1186/s40691-024-00380-0

Kalpokas, Ignas. 2021. Problematising reality: the promises and perils of synthetic media. *SN Social Sciences* 1(1). 1–11. https://doi.org/10.1007/s43545-020-00010-8

Kalpokas, Ignas. 2023. Work of art in the age of its AI reproduction. *Philosophy & Social Criticism* 51(8). 1268–1286. https://doi.org/10.1177/01914537231184490

Kalpokas, Ignas. 2024. *Technological governance and escapism in times of accelerated change*. Cham: Palgrave Macmillan.

Kalpokas, Ignas & Kalpokienė, Julija & Šalaševičiūtė, Vaiva. 2024. Rage against the machine: automated policing of generative AI. In Kuldova, Teresa Østbø & Gundhus, Helene Oppen Ingebrigtsen & Wathne, Christin T. (eds), *Policing and Intelligence in the Global Big Data Era, Volume II. Palgrave's Critical Policing Studies*, 259–286. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-031-68298-8_10

Kalpokas, Ignas & Šalaševičiūtė, Vaiva & Lipskė, Monika. 2024. Technology as a threat or a solution? The challenges of responding to synthetic Media. *Baltic Journal of Law & Politics* 16(2). 1–22. https://doi.org/10.2478/bjlp-2023-0010

Kalpokienė, Julija & Kalpokas, Ignas. 2023. Creative encounters of a posthuman kind – anthropocentric law, artificial intelligence, and art. *Technology in Society* 72. 1– 8. https://doi.org/10.1016/j.techsoc.2023.102197

Kasneci, Enkelejda & Sessler, Katharina & Küchemann, Stefan & Bannert, Maria & Dementieva, Daria & Fischer, Frank & Gasser, Urs & Groh, Georg & Poquet, Oleksandra & Sailer, Michael et al. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences* 103. 1–9. https://doi.org/10.1016/j.lindif.2023.102274

Kätsyri, Jari & Förger, Klaus & Mäkäräinen, Markus & Takala, Tapio. 2015. A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology* 6. 1–16. https://doi.org/10.3389/fpsyg.2015.00390

Keller, Damián & Zagalo, Nelson. 2025. Exploring the frontiers of artificial media. In Zagalo, Nelson & Keller, Damián (eds), *Artificial Media. Emerging Trends in Narratives, Education and Creative Practice*, 1–20. Cham: Springer.

Kim, Hyoseok & Koo, Thomas K. B. 2024. The impact of generative AI on syllabus design and learning. *Journal of Marketing Education* 48(1). 20–41. https://doi.org/10.1177/02734753241299024

Kim, Hayeon & Lee, Sang Wo. 2025. Sorry, it's my fault: politeness, attribution, and anthropomorphism in managing generative AI hallucinations. *International Journal of Information Managemen*t 86. 1–15. https://doi.org/10.1016/j.ijinfomgt.2025.102996

Kirk, Colleen P. & Givi, Julian. 2024. The AI-authorship effect: understanding authenticity, moral disgust, and consumer responses to AI-generated marketing communications. *Journal of Business Research* 186. 1–14. https://doi.org/10.1016/j.jbusres.2024.114984

Knott, Alistair & Pedreschi, Dino & Jitsuzumi, Toshiya & Leavy, Susan & Eyers, David & Chakraborti, Tapabrata & Trotman, Andrew & Sundareswaran, Sundar & Baeza-Yates, Ricardo & Biecek, Przemyslaw et al. 2024. AI content detection in the emerging information ecosystem: new obligations for media and tech companies. *Ethics and Information Technology* 26. 1–14. https://doi.org/10.1007/s10676-024-09795-1

Köbis, Nils C. & Doležalová, Barbora & Soraperra, Ivan. 2021. Fooled twice: people cannot detect deepfakes but think they can. *iScience* 24(11). 1–17. https://doi.org/10.1016/j.isci.2021.103364

Kudina, Olya & Verbeek, Peter-Paul. 2019. Ethics from within: Google Glass, the Collingridge Dilemma, and the mediated value of privacy. *Science, Technology & Human Values* 44(2). 291–314. https://doi.org/10.1177/0162243918793711

Lackey, Cait & Papacharissi, Zizi. 2024. Machine ex machina: a framework decentering the human in AI design praxis. *Human-Machine Communication* 8. 7–25. https://doi.org/10.30658/hmc.8.1

Lavazza, Andrea. 2025. AI, authenticity, and the chatbot between us: commentary on Battisti. Commentary. *Philosophy & Technology* 38(49). 1–5. https://doi.org/10.1007/s13347-025-00880-x

Law, Ryan & Guan, Xibeijia & Soulo, Tim. 2025. "74% of new webpages include AI content (study of 900k pages)". *Ahrefs blog*. https://www.ahrefs.com/blog/what-percentage-of-new-content-is-ai-generated/ (last accessed on 26/02/2026).

Lee, Eun-Ju. 2020. Authenticity model of (mass-oriented) computer-mediated communication: Conceptual explorations and testable propositions. *Journal of Computer-Mediated Communication* 25(1). 60–73. https://doi.org/10.1093/jcmc/zmz025

Lee, Eun-Ju. 2024. Minding the source: toward an integrative theory of human-machine communication. *Human Communication Research* 50. 184–193. https://doi.org/10.1093/hcr/hqad034

Lee, Eun-Ju. 2025. Authenticity at the heart of communication. *Journal of Communication* 75(2). 85–89. https://doi.org/10.1093/joc/jqaf005

Lehman, David W. & O'Connor, Kieran & Carroll, Glenn R. 2019. Acting on authenticity: individual interpretations and behavioral responses. *Review of General Psychology* 23(1). 19–31. https://doi.org/10.1177/1089268019829470

Lehman, David W. & O'Connor, Kieran & Kovács, Balázs & Newman, George E. 2019. Authenticity. *Academy of Management Annals* 13(1). 1–42. https://doi.org/10.5465/annals.2017.0047

Lermann Henestrosa, Angelica & Kimmerle, Joachim. 2024. The effects of assumed AI vs. human authorship on the perception of a GPT-Generated text. *Journalism and Media* 5(3). 1085–1097. https://doi.org/10.3390/journalmedia5030069

Li, Hang & Zhang, Renwen. 2024. Finding love in algorithms: deciphering the emotional contexts of close encounters with AI chatbots. *Journal of Computer-Mediated Communication* 29(5). 1–13. https://doi.org/10.1093/jcmc/zmae015

Liang, Weixin & Yuksekgonul, Mert & Mao, Yining & Wu, Eric & Zou, James. 2023. GPT detectors are biased against non-native English writers. *Patterns* 4(7). 100779. 1–4. https://doi.org/10.1016/j.patter.2023.100779

Llurda, Enric & Mocanu, Vasilica. 2024. English in Spain: education, attitudes and native-speakerism. *World Englishes* 43(2). 315–331. https://doi.org/10.1111/weng.12651

Marriott, Hanna R. & Pitardi, Valentina. 2023. One of the loneliest number… two can be as bad as one. The influence of AI friendship apps on users, wellbeing and addiction. *Psychology and Marketing* 41(1). 86–101. https://doi.org/10.1002/mar.21899

Martínez, Matías. 2020. Authenticity in narratology and in literary studies. In Fludernik, Monica & Ryan, Marie-Laure (eds), *Narrative Factuality: A Handbook*, 521–532. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110486278-036

Masood, Momina & Nawaz, Mariam & Malik, Khalid Mahmood & Javed, Ali & Irtaza, Aun & Malik, Hafiz. 2022. Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward. *Applied Intelligence* 53. 3974–4026. https://doi.org/10.1007/s10489-022-03766-z

MacDorman, Karl F. & Chattopadhyay, Debaleena. 2016. Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition* 146. 190–205. https://doi.org/10.1016/j.cognition.2015.09.019

McEvoy, James. 2009. Living in an age of authenticity: Charles Taylor on identity today. *Australasian Catholic Record* 86(2). 161–172.

Mercier, Faye & Abidin, Crystal. 2025. Authenticity as discursive gaze: a critical review of social media research on authenticity. *Journal of Communication* 75(5). 398–404. https://doi.org/10.1093/joc/jqaf038

Michael, Kyle & Shuttleworth, James. 2021. *The History and Ethics of Authenticity. Meaning, Freedom and Modernity*. London: Bloomsbury Academic.

Microsoft. 2026. "Global AI adoption in 2025. A widening digital divide". Microsoft. https://www.microsoft.com/en-us/research/wp-content/uploads/2026/01/Microsoft-AI-Diffusion-Report-2025-H2.pdf (last accessed on 26/02/2026).

Mieczkowski, Hanna Nicole. 2022. *AI-mediated communication: Examining agency, ownership, expertise, and roles of AI systems*. Stanford: Stanford University. Ph.D. Dissertation. https://purl.stanford.edu/nz776rw2839

Mishan, Freda. 2017. 'Authenticity' in the digital era. In Maley, Alan & Tomlinson, Brian (eds), *Authenticity in Material Development for Language Learning*, 10–24. Cambridge: Cambridge Scholars Publishing.

Mori, Masahiro & MacDorman, Karl F. & Nageki, Norri. 2012. The uncanny valley [From the field]. *IEEE Robotics & Automation Magazine* 19(2). 98–100.

Moorhouse, Benjamin Luke & Wong, Kevin M. 2025. *Generative Artificial Intelligence and Language Teaching*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009618823

Nass, Clifford & Moon, Youngme. 2000. Machines and mindlessness: social responses to computers. *Journal of Social Issues* 56(1). 81–103. https://doi.org/10.1111/0022-4537.00153

Nass, Clifford & Steuer, Jonathan & Tauber, Ellen R. 1994. Computers as social actors. In Adelson, Beth & Dumais, Susan & Olson, Judith (eds), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, April 24-28, 1994), 72–78. New York: Association for Computing Machinery. https://doi.org/10.1145/191666.191703

Newman, George E. & Bloom, Paul. 2012. Art and authenticity: the importance of originals in judgments of value. *Journal of Experimental Psychology* 141(3). 558–569. https://doi.org/10.1037/a0026035

Newman, George E. & Smith, Rosanna K. 2016. Kinds of authenticity. *Philosophy Compass* 11(10). 609–618. https://doi.org/10.1111/phc3.12343

Newman, George E. 2019. The psychology of authenticity. *Review of General Psychology* 23(1). 8–18. https://doi.org/10.1037/gpr0000158

Nicenboim, Iohanna & Oogjes, Doenja & Biggs, Heidi & Nam, Seowoo. 2023. Decentering through design: bridging posthuman theory with more-than-human design practices. *Human Computer Interaction* 40(1-4). 195–220. https://doi.org/10.1080/07370024.2023.2283535

O'Rourke, Bernadette & Ramallo, Fernando. 2013. Competing ideologies of linguistic authority amongst new speakers in contemporary Galicia. *Language in Society* 42(3). 287–305. https://doi.org/10.1017/S0047404513000249

Pedersen, Carsten L. & Ritter, Thomas. 2024. Digital authenticity: towards a research agenda for the AI-driven fifth phase of digitalization in business-to-business marketing. *Industrial Marketing Management* 123. 162–172. https://doi.org/10.1016/j.indmarman.2024.10.005

Perkins, Mike & Roe, John & Vu, Binh H. & Postma, Dirk & Hickerson, David & McGaughran, James & Khuat, Huyen Q. 2024. Simple techniques to bypass GenAI text detectors: implications for inclusive education. *International Journal of Educational Technology in Higher Education* 21(1). 1–25. https://doi.org/10.1186/s41239-024-00487-w

Peterson, Richard A. 2005. In search of authenticity. *Journal of Management Studies* 42(5). 1083–1098. https://doi.org/10.1111/j.1467-6486.2005.00533.x

Pinner, Richard S. 2014. The authenticity continuum: toward a definition incorporating international voices. *English Today* 30(4). 22–27. https://doi.org/10.1017/S0266078414000364

Pinner, Richard S. 2016. *Reconceptualising Authenticity for English as a Global Language*. Bristol: Multilingual Matters.

Placani, Adriana. 2024. Anthropomorphism in AI: hype and fallacy. *AI and Ethics* 4. 691–698. https://doi.org/10.1007/s43681-024-00419-4

Qin, Xin & Zhou, Xiang & Chen, Chen & Wu, Dongyuan & Zhou, Hansen & Dong, Xiaowei & Cao, Limei & Lu, Jackson G. 2025. AI aversion or appreciation? A capability–personalization framework and a meta-analytic review. *Psychological Bulletin* 151(5). 580–599. https://doi.org/10.1037/bul0000477

Quinlantang, Rae Francis. 2025. The HAI-IO model: a framework for understanding the human-AI communication process. *Human Machine Communication* 10. 207–221. https://doi.org/10.30658/hmc.10.9

Rae, Irene. 2024. The effects of perceived AI use on content perceptions. In Floyd Mueller, Florian (ed.), *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, May 11-16, 2024), 1–14. New York: Association for Computing Machinery. https://doi.org/10.1145/3613904.3642076

Ramadan, Zahi & Ramadan, Jad. 2025. AI avatars and co-creation in the metaverse. *Consumer Behavior in Tourism and Hospitality* 20(1). 131–147. https://doi.org/10.1108/CBTH-07-2024-0246

Rammert, Werner. 2008. Where the action is: distributed agency between humans, machines, and programs. TUTS – Working Papers, 4-2008. Berlin. https://www.ssoar.info/ssoar/handle/document/1233 (last accessed on 26/02/2026).

Rapp, Amon & Di Lodovico, Chiara & Torrielli, Federico & Di Caro, Luigi. 2025. How do people experience the images created by generative artificial intelligence? An exploration of people's perceptions, appraisals, and emotions related to Gen-AI text-to-image model and its creations. *International Journal of Human-Computer Studies* 193. 1–16. https://doi.org/10.1016/j.ijhcs.2024.103375

Reeves, Byron & Nass, Clifford. 1996. *The Media Equation. How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge: Cambridge University Press.

Revolusi, Prabu & Febriandy, Radians Frisna. 2025. Human AI communication (HA-C): transforming the role of technology in human interaction. *Multidisciplinary Science Journal* 8(3). 1–14. https://doi.org/10.31893/multiscience.2026201

Richards, Jack C. & Rodgers, Theodore S. 2001. *Approaches and Methods in Language Teaching* (2nd ed.). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511667305

Roddy, Stephen & Bridges, Brian. 2025. Cybernetic resurgences: machine music beyond AI slop. In Zagalo, Nelson & Keller, Damián (eds), *Artificial Media.*

Emerging Trends in Narratives, Education and Creative Practice, 95–116. Cham: Springer.

Rodríguez Valdés, María. 2025. The use of artificial media as a tool for the creation of artistic images through collaborative methodologies. In Zagalo, Nelson & Keller, Damián (eds), *Artificial Media. Emerging Trends in Narratives, Education and Creative Practice*, 159–174. Cham: Springer

Romo, Vanessa. 2025. "It's 2025, the year we decided we need a widespread slur for robots". *NPR News*, August 6, 2025. https://www.vpm.org/npr-news/npr-news/2025-08-06/its-2025-the-year-we-decided-we-need-a-widespread-slur-for-robots (last accessed on 26/02/2026).

Salma, Zainab & Hijón-Neira, Raquel & Pizarro, Celeste. 2025. Designing co-creative systems: five paradoxes in human-AI collaboration. *Information* 16(10). 1–19. https://doi.org/10.3390/info16100909

Schilke, Oliver & Reimann, Martin. 2025. The transparency dilemma: how AI disclosure erodes trust. *Organizational Behavior and Human Decision Processes* 188. 1–16. https://doi.org/10.1016/j.obhdp.2025.104405

Short, John & Williams, Ederyn & Christie, Bruce. 1976. *The Social Psychology of Telecommunications*. London: John Wiley and Sons.

Silver, Ike & Newman, George & Small, Deborah A. 2020. Inauthentic aversion: moral reactance toward tainted actors, actions, and objects. *Consumer Psychology review* 4(1). 70–82. https://doi.org/10.1002/arcp.1064

Simas, Gustavo & Ulbricht, Vania. 2024. Human-AI interaction: an analysis of anthropomorphization and user engagement in conversational agents with a focus on ChatGPT. In *Intelligent Human Systems Integration IHSI 2024*, 119. 454–464. http://doi.org/10.54941/ahfe1004510

Spennemann, Dirk H. R. 2025. Generative Artificial Intelligence and the future of public knowledge. *Knowledge* 5(3). 1–19. https://doi.org/10.3390/knowledge5030020

Stone, Brian W. 2024. Generative AI in higher education: uncertain students, ambiguous use cases, and mercenary perspectives. *Teaching of Psychology* 52(3). 347–356. https://doi.org/10.1177/00986283241305398

Sundar, S. Shyam. 2020. Rise of machine agency: a framework for studying the psychology of human-AI interaction (HAII). *Journal of Computer-Mediated Communication* 25(1). 74–88. https://doi.org/10.1093/jcmc/zmz026

Sundar, S. Shyam & Chen, Jin. 2023. From CASA to TIME: machine as a source of media effects. In Guzman, Andrew L. & McEwen, Rhonda & Jones, Steve (eds), *The SAGE Handbook of Human-Machine Communication*, 63–70. London: Sage. https://doi.org/10.4135/9781529782783.n9

Taylor, Charles. 1991. *The Ethics of Authenticity*. Cambridge: Harvard University Press.

Thakur, Kutub & Barker, Helen G. & Pathan, Al-Sakib Khan. 2024. *Artificial Intelligence and Large Language Models*. Oxon: CRC Press. https://doi.org/10.1201/9781003474173

Trilling, Lionel. 1972. *Sincerity and Authenticity*. Cambridge: Harvard University Press.

Tron Gianet, Eric & Di Caro, Luigi & Rapp, Amon. 2024. Redefining the user in human-generative AI collaboration: insights from music composition. In *Proceedings of the 8th International Workshop on Cultures of Participation in the Digital Age CoPDA 2024* (Arenzano, June 3, 2024). Aachen: CEUR Workshop Proceedings. https://www.ceur-ws.org/Vol-3701/paper6.pdf (last accessed on 26/02/2026).

Vlachopoulos, Dimitrios & Makri, Agoritsa. 2024. A systematic literature review on authentic assessment in higher education: best practices for the development of 21st century skills, and policy considerations. *Studies in Educational Evaluation* 83. 1–13. https://doi.org/10.1016/j.stueduc.2024.101425

Wang, Ning. 1999. Rethinking authenticity in tourism experience. *Annals of Tourism Research* 26(2). 349–370. https://doi.org/10.1016/S0160-7383(98)00103-0

Wasdahl, Alex. 2024. Machine credibility: how news readers evaluate AI-generated content. *InterActions: UCLA Journal of Education and Information Studies* 19(1). 1–28. http://dx.doi.org/10.5070/D41.7196

Werdiningsih, Indah & Marzuki & Rusdin, Diyenti. 2024. Balancing AI and authenticity: EFL students' experiences with ChatGPT in academic writing. *Cogent Arts & Humanities* 11(1). 1–15. http://doi.org/10.1080/23311983.2024.2392388

Whittaker, Lucas & Kietzmann, Tim C. & Kietzmann, Jan & Dabirian, Amir. 2020. 'All around me are synthetic faces': the mad world of AI-generated media. *IT Professional* 22(5). 90–99. https://doi.org/10.1109/MITP.2020.2985492

Williams, Bernard. 2002. *Truth and Truthfulness: an Essay in Genealogy*. Princeton: Princeton University Press.

Wilt, Joshua A. & Thomas, Sarah & McAdams Dan P. 2019. Authenticity and inauthenticity in narrative identity. *Heliyon* 5(7). 1–13. https://doi.org/10.1016/j.heliyon.2019.e02178

Woolard, Kathryn A. 2016. *Singular and Plural: Ideologies of Linguistic Authority in 21st century Catalonia*. Oxford: Oxford University Press.

Wu, Junchao & Yang, Shu & Zhan, Runzhe & Yuan, Yulin & Chao, Lidia & Wong, Derek. 2025. A survey on LLM-generated text detection: necessity, methods, and future directions. *Computational Linguistics* 51(1). 275–338. https://doi.org/10.1162/coli_a_00549

Wu, Wanyan & Lingel, Jessa. 2025. "I am Neuro, who are you?": Performances of authenticity in an experimental AI livestream. *New Media and Society*. 1–21. https://doi.org/10.1177/14614448251406904

Yang, Daijin & Zhou, Yanpeng & Zhang, Zhiyuan & Li, Toby Jia-Jun & LC, Ray. 2022. AI as an active writer: interaction strategies with generated text in human-AI collaborative fiction writing. In Smith-Renner, Alison & Amir, Ofra (eds), *Joint Proceedings of the IUI 2022 Workshops: APEx-UI, HAI-GEN, HEALTHI, HUMANIZE, TExSS, SOCIALIZE co-located with the ACM International Conference on Intelligent User Interfaces* (Helsinki, March 21-22), 56–65. Helsinki: CEUR Workshop Proceedings. https://ceur-ws.org/Vol-3124/#paper6 (last accessed on 26/02/2026).

YouTube. 2024. "A.I. and creators. The future of tech and creativity." YouTube. https://www.youtube.com/howyoutubeworks/ai-and-creators/?utm_source=linkedin&utm_medium=social&utm_campaign=foc_launch_kc&utm_content=post_9.24.24 (last accessed on 23/08/25).

Zafar, Muhammad Bilal & Ali, Hassnian & Yasin, Talha. 2025. Reimagining human creativity and learning in the age of generative AI: a multi-method meta-thematic synthesis. *Next Research* 2(4). 1–18. https://doi.org/10.1016/j.nexres.2025.100802

Zhang, Longyu & Fang, Cong & Wang, Stephen Jia & Wang, Yao & Luo, Shijian. 2025. Students' attitudes and sentiments toward AI-generated images: deep learning-based social media text mining. *Interactive Learning Environments*. 1–26. https://doi.org/10.1080/10494820.2025.2545964