Journal of Advertising Research

ISSN: 0021-8499 (Print) 1740-1909 (Online) Journal homepage: www.tandfonline.com/journals/ujar20

Forgiving the Algorithm? Consumer Responses to AI- vs. Human-Created Misrepresentative Ads

Kshitij Bhoumik

To cite this article: Kshitij Bhoumik (13 Jul 2026): Forgiving the Algorithm? Consumer Responses to AI- vs. Human-Created Misrepresentative Ads, Journal of Advertising Research, DOI: 10.1080/00218499.2026.2677339

To link to this article: https://doi.org/10.1080/00218499.2026.2677339

Forgiving the Algorithm? Consumer Responses to AI- vs. Human-Created Misrepresentative Ads

KSHITIJ BHOUMIK
Leeds University Business School, University of Leeds
k.bhoumik@leeds.ac.uk

As brands increasingly use generative AI for advertising, ethical challenges arise when such ads misrepresent marginalized communities. Across three studies, we demonstrate that consumers are less forgiving of AI-created misrepresentative ads than human-created ones because AI ads are perceived as involving less effort, reducing ad attitudes and forgiveness. Extending attribution theory, we show that perceived effort becomes the primary responsibility cue when the proximate creator lacks moral agency, yet institutional accountability persists. This mechanism operates even when consumers' personal identities are implicated (LGBTQIA+ participants evaluating Pride ads). However, strong brand reputation buffers these effects, informing responsible AI integration in advertising.

Advertising has long been more than a tool for selling products—it functions as a cultural institution that reflects and shapes social norms, identities, and values (McFall 2004; Schroeder and Borgerson 2005). Because of this cultural role, misrepresentation in advertising—through stereotyping, exclusion, or tokenistic inclusion—has consistently provoked criticism and consumer backlash. From the hypersexualization of women in earlier decades (Browne 1998; Eisend 2019; Kang 1997) to recent campaigns criticized for superficial portrayals of ethnic and LGBTQIA+ identities (Grau and Zotos 2018; Vredenburg et al. 2020), such missteps reveal how advertising can perpetuate inequities. For instance, Dove's 2017 body wash ad, which appeared to depict a black woman transforming into a white woman, sparked outrage for racial insensitivity. Despite calls for inclusivity and authenticity, these failures remain common (Burgess, Wilkie, and Dolan 2023; Campbell et al. 2025).

Keywords: Artificial intelligence; advertising ethics; consumer forgiveness; perceived effort; brand reputation

Submitted June 22, 2025;
revised May 4, 2026;
accepted May 12, 2026.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/00218499.2026.2677339.

Management Slant

- Consumers react more negatively to misrepresentative ads created by AI than by humans because AI signals lower care and oversight.
- Perceived effort is the key lever shaping responses to AI-generated advertising. When campaigns reference marginalized or identity-based communities, consumers scrutinize whether meaningful human care and diligence were involved.
- In identity-relevant contexts (e.g., LGBTQIA+ advertising), simply labeling content as "AI-generated" can backfire by triggering low-effort inferences. Brands should position AI as a supportive tool embedded within visible human oversight and cultural review.
- Strong brand reputations buffer firms from backlash to AI missteps, whereas weaker brands face heightened scrutiny and must rely more on human refinement and ethical review.

Misrepresentation is not a trivial error but a cultural and commercial risk: when brands reproduce stereotypes or rely on superficial symbolism, they alienate the very communities they seek to engage (Eisend, Muldrow, and Rosengren 2023). Thus, advertising misrepresentation exposes a tension between commercial motives and cultural responsibility (Schroeder and Zwick 2004). Consumers are acutely attuned to these dynamics and often respond to exploitative or insincere representation with skepticism, anger, or boycott behavior (Wilkie et al. 2023).

Into this sensitive environment, generative artificial intelligence (AI) has rapidly entered advertising workflows. Brands increasingly rely on AI to generate text, imagery, and video at scale, offering efficiency gains but also raising new representational risks. AI systems trained on biased datasets can reproduce or amplify stereotypes (Benjamin 2019; Zou and Schiebinger 2018), and their outputs often lack the nuance and contextual sensitivity consumers expect in identity-relevant domains (Sundar 2020). Recent controversies underscore these risks. In 2023, Levi's announced plans to use AI-generated models to "increase diversity," but critics argued that algorithmic substitutes undermine genuine inclusion (Allyn 2023). In 2024, Skechers' AI-generated Vogue campaign featuring synthetic models was condemned as unrealistic and inauthentic (Surnit 2024). When representation is delegated to machines, questions of accountability and social responsibility become newly salient.

Despite growing awareness of representational failures in advertising, little empirical research has examined how consumers respond when such failures are created by AI versus humans. Although consumers generally react negatively to misrepresentative ads (Åkestam 2017; Johnson and Grier 2012), AI involvement may trigger distinct inferences about oversight and care. What remains unclear is whether AI-created errors prompt different reactions—and why? Do consumers view AI-generated missteps as more forgivable because machines lack intent, or less forgivable because automation signals insufficient care?

Existing research on misrepresentation has primarily examined human-created campaigns, emphasizing stereotyping and tokenistic diversity portrayals (Campbell et al. 2025; Eisend 2010). In parallel, work on AI in advertising has largely focused on its operational advantages, including personalization and efficiency (Kumar et al. 2019; Loureiro, Guerreiro, and Tussyadiah 2021; Vakratsas and Wang 2021), while emerging research highlights concerns surrounding algorithmic bias and representational fairness (Campbell 2023; Hoffmann 2019). However, AI's increasing use introduces a distinct attributional challenge: consumers must assign responsibility when the proximate creator is nonhuman yet institutional accountability persists. This represents a shift from traditional attribution frameworks, which assume morally agentic actors. When AI creates misrepresentation, consumers cannot meaningfully assess intent as they would for human creators. Instead, they rely on perceived effort as a signal of whether brands exercised adequate oversight.

Consumers rely on perceived effort as the key explanatory mechanism: AI-authored misrepresentations evoke lower perceived effort, leading to less favorable attitudes and reduced forgiveness.

To address this gap, the present research draws on Attribution Theory (Weiner 2012) to examine how ad creator identity (AI vs. human) shapes consumer forgiveness following representational failures. We propose that consumers rely on perceived effort as the key explanatory mechanism: AI- authored misrepresentations evoke lower perceived effort, leading to less favorable attitudes and reduced forgiveness. Importantly, this attributional process persists in identity-relevant contexts where consumers' personal identities are directly implicated. Across three experimental studies, we test this framework while ruling out perceived intent as alternative explanation. Study 1 demonstrates that AI-authored ads generate lower attitudes and forgiveness than human-created ads, an effect driven by perceived effort rather than perceived intent. Study 2 replicates these findings in LGBTQIA+-themed advertising, testing the mechanism among LGBTQIA+ participants to demonstrate robustness when moral stakes are elevated. Study 3 tests boundary conditions by examining brand reputation: while AI-generated misrepresentation typically elicits harsher judgments, strong reputations buffer these effects through pathways independent of the effort-attribution mechanism.

By integrating advertising ethics, attribution theory, and consumer trust in AI, this research makes three key contributions (Castleberry, French, and Carlin 1993; Zinkhan 1994). First, it extends scholarship on advertising misrepresentation by shifting attention from representational accuracy to the perceived process of ad creation (Campbell et al. 2025; Gardete 2013). Forgiveness hinges not only on what an ad depicts but also on who is perceived to have created it and how much effort is inferred. Second, it advances attribution theory by identifying perceived effort—not intent—as the primary attribution guiding forgiveness when the proximate creator is nonhuman yet institutional accountability

persists. Third, it identifies brand reputation as a key boundary condition (De Chernatony 1999), showing reputational capital can mitigate ethical risks in AI advertising.

THEORETICAL FRAMEWORK

Misrepresentation in Advertising

Advertising not only connects brands with consumers but also shapes cultural values and social norms (McFall 2004). Yet research shows that advertising frequently misrepresents marginalized groups through exclusion or reductive stereotypes (Eisend, Muldrow, and Rosengren 2023; Zotos and Grau 2016). Such failures are especially damaging in identity-relevant contexts where consumers expect authenticity and inclusivity.

Misrepresentation can also take the form of symbolic rather than substantive inclusivity. Pride-themed campaigns without genuine organizational commitment, for example, are often labeled "rainbow-washing"—signaling allyship for commercial gain rather than meaningful advocacy (Gurrieri, Brace-Govan, and Cherrier 2016). Similar backlash arises when brands co-opt movements such as Black Lives Matter or LGBTQIA+ activism without demonstrating credible support (Vredenburg et al. 2020). Consumers are not passive recipients; they infer motives, evaluate responsibility, and respond accordingly. Perceived inauthenticity and lack of inclusivity diminish consumers' positive attitudes toward advertising (Eisend 2010; Grau and Zotos 2018; Wilkie et al. 2023).

Attribution Theory (Weiner 2012) helps explain how consumers form these moral evaluations. Judgments depend not only on the nature of the misstep but also on characteristics of the ad creator. When consumers encounter misrepresentation, they treat it as a negative event and make causal inferences about responsibility. Responsibility judgments hinge on causal dimensions such as controllability and intentionality (Fehr, Gelfand, and Nag 2010; Weiner 2012). Controllability is often inferred from effort—whether the creator invested sufficient care to prevent misrepresentation—whereas intentionality captures whether the misrepresentation reflects deliberate purpose. These dimensions raise an important question: how do consumers apply such attributional cues when ads are produced by fundamentally different types of agents—humans versus AI systems?

Ad Creator Identity, Perceived Effort, and Consumer Forgiveness

Generative AI is increasingly integrated into advertising workflows, enabling rapid, low-cost creation of visual and textual content (Huh, Nelson, and Russell 2023). Yet consumers do not interpret AI-generated content in the same way they interpret human-created content, particularly when representational failures occur. A central challenge in judging misrepresentation is identifying who is responsible. Because assigning responsibility presupposes an agent, the characteristics of the ad creator—human or AI—shape how consumers interpret the lapse and form moral evaluations.

Human creators are typically viewed as capable of reflection, contextual understanding, and creative judgment—capacities associated with moral agency and care (Coeckelbergh 2020; Gray and Wegner 2012). AI systems, in contrast, are seen as efficient but detached tools that operate through statistical patterning rather than social understanding. When misrepresentation occurs, people expect that human creators could have exercised oversight, whereas AI-generated content is associated with minimal sensitivity to cultural nuance. These initial assumptions form the basis for attributional differences in how consumers evaluate identical misrepresentations depending on the creator's identity.

Importantly, AI-generated advertising represents a distinct attributional case rather than merely a context of ambiguity. Classic attribution theory was developed to explain how observers infer responsibility for the actions of morally agentic actors assumed to possess intentions and moral awareness (Kelley and Michela 1980; Weiner 2012). AI systems differ fundamentally: they produce outputs that resemble human decisions but lack mental states and moral agency. Simultaneously, they operate within institutional arrangements in which brands deploy these systems, set their parameters, and retain ultimate responsibility. Consumers must therefore evaluate outcomes within a hybrid structure where agency is partially automated yet institutionally anchored. This configuration shifts the cues used to assess responsibility. Because AI lacks mental states, consumers cannot evaluate the proximate creator's intent in the same way they would for human creators. As a result, responsibility judgments shift toward assessments of effort. Prior attribution models centered on human actors do not explicitly account for this redistribution of agency. When the proximate actor lacks moral agency yet institutional accountability persists, effort becomes a central diagnostic cue for judging blameworthiness.

Our framework thus extends attribution theory by demonstrating how responsibility judgments adapt when the creator is nonhuman operating within institutional accountability. Building on this insight, we propose that in representational contexts—where harm is often ambiguous rather than overtly malicious—consumers rely on perceived effort as the primary diagnostic cue. Intent is notoriously difficult to infer when advertising missteps reflect oversight rather than purposeful discrimination (Malle and Knobe 1997). Perceived effort signals controllability and responsibility (Weiner 2012). High effort implies attempts to consider cultural nuance and avoid stereotypes, whereas low effort signals

negligence. These attributional tendencies become particularly pronounced in AI-driven advertising. Because AI tools are perceived as automated, impersonal, and lacking contextual awareness, consumers infer that little effort was invested in producing culturally sensitive content (Longoni, Bonezzi, and Morewedge 2019; Sundar 2020). AI-generated missteps therefore appear more preventable and less excusable. Human creators, by contrast, are afforded greater moral latitude: even when mistakes occur, consumers assume some creative engagement and effortful judgment.

These attributional differences have direct implications for forgiveness—a key outcome in the aftermath of advertising failures. Forgiveness reflects willingness to reduce blame and extend understanding toward the responsible agent (Fehr, Gelfand, and Nag 2010; McCullough, Fincham, and Tsang 2003; McCullough et al. 1998; Xie and Peng 2009). Unlike immediate attitudinal responses, forgiveness captures consumers' willingness to grant moral leniency and preserve the brand relationship despite the misstep (Tsarenko and Tojib 2015). In advertising contexts, consumer forgiveness is critical for brand recovery: it predicts whether consumers will continue engaging with the brand, recommend it to others, or give it a second chance following representational failures (Chung and Beverland 2006). When misrepresentation is attributed to AI—and thus perceived as reflecting low effort and institutional negligence—consumers are less willing to extend forgiveness. We therefore predict:

> **H1:** Consumers are less forgiving of misrepresentative ads when they are created by AI (vs. by a human) (Figure 1).

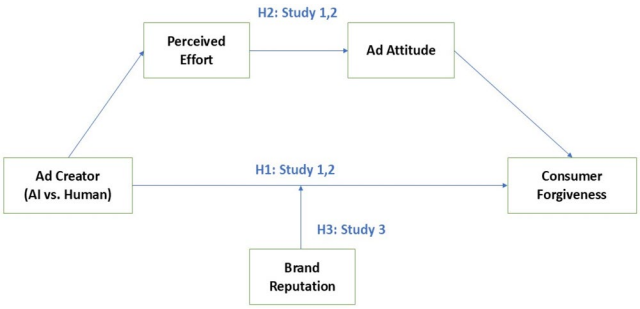

Figure 1 Conceptual Model

Perceived Effort and Ad Attitude as Sequential Mediators

The effect of ad creator identity on forgiveness operates through a sequential process. Perceived effort influences not only forgiveness but also more proximal evaluations of the advertisement. Advertising research demonstrates that perceptions of thoughtfulness and authenticity enhance ad attitudes, whereas ads perceived as careless elicit skepticism and negative evaluations (Aaker, Garbinsky, and Vohs 2012; Campbell and Kirmani 2000). In representation-sensitive contexts, effort cues become especially consequential. When consumers believe a brand invested care in depicting diverse identities, they evaluate the ad more positively, even if imperfections remain (Gerrath et al. 2026). Conversely, when ads are attributed to AI—a production process associated with efficiency rather than contextual sensitivity—perceived effort declines. When misrepresentation is attributed to low effort, the lapse appears negligent and less excusable, reducing both ad attitudes and forgiveness.

Integrating attributional reasoning with evaluative responses, we propose that ad creator identity first influences perceived effort, which then shapes ad attitudes, and ultimately consumer forgiveness. This serial mediation reflects how attributional inferences (effort) translate into evaluative judgments (ad attitudes) before manifesting in moral outcomes (forgiveness).

> **H2(a)**: The negative effect of AI (vs. human) ad creator on consumer forgiveness is sequentially mediated by lower perceived effort and less favorable attitudes toward the ad.

Testing the Mechanism in Identity-Relevant Advertising Contexts

We extend this framework to identity-relevant advertising, testing the effort-based mechanism in LGBTQIA+-themed advertising for several reasons. First, LGBTQIA+ consumers have been systematically subjected to tokenistic representation, rainbow-washing (superficial Pride Month gestures without substantive support), and commodification of their identities for commercial gain (Cheng, Zhou, and Yao 2023; Vredenburg et al. 2020). These experiences create heightened expectations around authentic representation: LGBTQIA+ consumers evaluate whether brands demonstrate genuine care and cultural competence rather than performative allyship.

Second, when misrepresentation occurs in LGBTQIA+ advertising, the implications extend beyond reputational concerns to encompass social and identity-related harm (Montecchi et al. 2024). Misrepresentation can perpetuate stereotypes and violate trust built through community engagement (Eisend, Muldrow, and Rosengren 2023), making LGBTQIA+ consumers particularly attentive to signals that brands exercised adequate cultural sensitivity.

Third, Pride Month advertising represents a highly visible domain where brands face scrutiny over representational ethics (Champlin and Li 2020), making it a theoretically informative context for testing whether effort-based attributions shape forgiveness when moral stakes are elevated and consumers' personal identities are directly implicated.

Importantly, heightened scrutiny does not imply a different attributional process but rather intensifies the salience of effort cues when consumers evaluate whether a brand exercised adequate care (Johnson and Grier 2012). If the effort-based mechanism reflects a general attributional process—as attribution theory suggests (Weiner 2012)—it should operate even when consumers' own identities are directly implicated, and moral stakes are elevated. Accordingly, we do not expect a different process in identity-relevant contexts, but rather the persistence of the same effort-based pathway.

> **H2(b):** Among LGBTQIA+ participants exposed to LGBTQIA+-themed advertising, the negative effect of AI (vs. human) ad creator on consumer forgiveness is sequentially mediated by lower perceived effort and less favorable attitudes toward the ad.

Moderating Role of Brand Reputation

While AI-generated ads are often evaluated less favorably due to lower perceived effort, brand reputation operates as a boundary condition that shapes how consumers interpret such cues. Brand reputation reflects accumulated judgments of a firm's trustworthiness, credibility, and consistency over time (Fombrun and Shanley 1990; Walsh et al. 2009). Strong reputations function as "trust reserves," enabling consumers to extend goodwill, discount negative information, and grant brands the benefit of the doubt when missteps occur (Aaker, Garbinsky, and Vohs 2012; Dawar and Lei 2009). By contrast, weak-reputation brands lack this buffer.

From an attribution-theory perspective, brand reputation functions as a higher-order cue that shapes how consumers interpret responsibility judgments (Kelley 1973; Weiner 2012). As established, AI involvement shifts responsibility assessments toward evaluations of institutional effort and oversight. However, consumers must still determine whether perceived low effort reflects stable brand characteristics (systemic carelessness) or a situational lapse. Brand reputation provides prior information about the brand's typical standards of diligence and cultural sensitivity, guiding this interpretation. For strong-reputation brands, consumers externalize causality, attributing missteps to isolated mistakes or atypical processes rather than stable brand characteristics, thereby dampening blame-oriented attributions and the negative effect of AI involvement. For weak-reputation brands, consumers make internal, stable attributions, seeing AI-driven misrepresentation as consistent with carelessness or lack of cultural awareness. Weak reputation amplifies the diagnostic value of AI involvement, intensifying negative reactions. This logic aligns with research showing that strong reputations buffer firms by directing attributions toward temporary or situational causes (Dawar and Lei 2009; Sohn and Lariscy 2015), whereas weak reputations lead consumers to rely on heuristic cues—such as the belief that "AI = low effort"—to form harsher judgments. Related theories of moral capital and uncertainty reduction similarly suggest that strong reputations reduce interpretive uncertainty and afford brands "moral credit" during failures (Blanken, De Ven, and Zeelenberg 2015; Tsarenko and Tojib 2015). Conversely, when reputation is weak, AI-based misrepresentation reinforces expectations of indifference or corner-cutting (Kaplan et al. 2023).

In sum, brand reputation functions as a higher-level attributional cue: strong reputations attenuate, and weak reputations amplify the negative effect of AI (vs. human) ad creators on forgiveness.

> **H3:** The negative effect of AI (vs. human) ad creator on consumer forgiveness is moderated by brand reputation, such that the difference in forgiveness between AI- and human-created ads will be significant when brand reputation is low but attenuated when brand reputation is high.

OVERVIEW OF STUDIES

We conducted three studies to test our framework across progressively refined contexts. Study 1 examines the main effect of ad creator identity (AI vs. human) on consumer forgiveness of a misrepresentative advertisement and tests the serial mediation of perceived effort and attitude toward the ad while ruling out perceived intent. Study 2 tests the robustness of this attributional mechanism in a socially sensitive, identity-relevant context by using a misrepresentative LGBTQIA+-themed advertisement and a quota sample of LGBTQIA+ and non-LGBTQIA+ participants, while ruling out perceived intent and ad relevance as alternative explanations. Study 3 investigates brand reputation as a boundary condition, testing whether strong brand reputation buffers the negative effects of AI involvement.

To ensure generalizability across English-speaking Western markets, we collected data from two independent participant platforms—one primarily U.S.-based (CloudResearch) and one primarily U.K.-based (Prolific). Using both platforms broadened regional representation, enabled targeted LGBTQIA+ recruitment for Study 2, and minimized potential sampling bias, strengthening confidence that the observed effects are not platform-specific (Peer et al. 2017). Both platforms independently verify participant location and language fluency, supporting data quality and cultural relevance. Participants were required to be at least 18, fluent in English, and located in either the UK or the US; no other inclusion criteria were imposed. Education level was not recorded to reduce

survey length, though both platforms draw from demographically diverse adult panels.

We report gender distribution and mean age in the main text and provide age range and ethnic composition in Appendix–G. Sample sizes exceeded recommended thresholds for detecting medium effects (Cohen 1992; Kang 2021). Attention checks were included (Appendix–H), although all results hold without exclusions.

STUDY 1: AD CREATOR IDENTITY AND CONSUMER FORGIVENESS: THE MEDIATING ROLES OF PERCEIVED EFFORT AND AD ATTITUDE

Overview and Purpose

Study 1 investigates how human- versus AI-created misrepresentative ads shape consumer forgiveness through perceived effort and ad attitude. We pretested the ad stimulus for misrepresentation, validity, and scale reliability (Appendix A and E).

Method

Participation and Design. A total of 161 participants were recruited from the Cloud Research platform (66.4% female, 32.9% male, 0.7% non-binary; M_{age} = 40.76 years). Participants were randomly assigned to one of the two conditions, ad creator: AI vs. human.

Stimulus and Procedure. Participants viewed a digital ad for *New Barbie Spa and Salon*, a fictitious local salon. The ad featured the tagline *"Fair Skin is Beautiful Skin!"* and served as a misrepresentation stimulus by equating beauty with lighter skin tones (see Appendix A for full ad). In the AI condition, the caption read: *"This ad was designed using cutting-edge artificial intelligence that analyzes patterns to create highly targeted content."* In the human condition, the caption stated: *"This ad was crafted by professional human designers utilizing knowledge of consumer behavior and creative principles."*

Measures. After viewing the ad, participants completed a series of measures. Attitude toward the ad was assessed using a four-item, seven-point bipolar scale (α = .94; e.g., "This ad was good/bad"), adapted from MacKenzie, Lutz, and Belch (1986). Consumer forgiveness was measured with three seven-point Likert items (α = .93; e.g., "I will forgive the creator of this ad for making a mistake"), adapted from Xie and Peng (2009). Perceived effort was captured through a four-item, seven-point Likert scale (α = .97; e.g., "The ad creator has put a lot of effort into the ad"), adapted from Mohr and Bitner (1995). Perceived intentionality was captured through a four-item seven-point Likert scale (α = .77; e.g., "The advertisement was created with a clear purpose in mind"; adapted from Campbell and Kirmani 2000). To confirm the effectiveness of the manipulation, participants completed a two-item check (α = .89; e.g., "Who created this ad?" 1 = AI/7 = Human). Finally, participants provided basic demographic information, including age, gender, and ethnicity.

Results

Manipulation Check. Participants in AI condition believed that ad is created by AI and participants in human condition believed that ad is created by human (M_{AI}=1.61, SD = 1.23 vs. M_{human}=5.80, SD = 1.37; $F(1, 160)$= 415.94, p< .001, η^2 =0.772), indicating that our manipulation worked successfully.

Dependent Variable(s). A one-way analysis of variance reveal that participants were less forgiving when the misrepresented ad was created by AI than when it was created by human (M_{AI} =3.62, SD = 1.55 vs. M_{human}= 4.17, SD = 1.62; $F(1, 160)$=4.80, p = 0.03, η^2=0.097).

Sequential Mediation Analysis. We used Hayes PROCESS model-6 (Preacher and Hayes 2008) to conduct sequential mediation with perceived effort as M1, attitude toward ad as M2, and consumer forgiveness as the dependent variable. We found an indirect effect of ad creator type on forgiveness through perceived effort and ad attitude (β=-0.41; SE = 0.1507; 95% CI: −0.7423, −0.1527), making the direct effect insignificant (β = 0.15; SE = 0.2163; 95% CI: −0.2787, 0.5758) and suggesting full sequential mediation. We further examined whether including attitude as a second mediator adds explanatory value beyond perceived effort alone. Both the effort-only path (Exp → Effort → Forgiveness; Ind1 = 0.2691, 95% CI [0.0547, 0.5316]) and the serial path through effort and attitude (Exp → Effort → Attitude → Forgiveness; Ind3 = 0.2689, 95% CI [0.1089, 0.4780]) were significant. Importantly, the attitude-only path (Ind2) was not significant (–0.1378, 95% CI [–0.3215, 0.0100]), suggesting that attitude contributes meaningfully only when preceded by effort. A contrast test confirmed no significant difference between the effort-only and serial paths ($C2$ = 0.0003, 95% CI [–0.3387, 0.3084]). These findings indicate that while perceived effort is a primary mechanism, including attitude helps capture a more complete psychological process that leads to forgiveness.

Ruling Out Alternate Mechanism. We also examined whether perceived intentionality could account for the observed effect. A serial mediation model replacing perceived effort with perceived intentionality yielded a non-significant indirect effect (β = 0.0193; SE = 0.0962; 95% CI: −0.1709, 0.2263), indicating that perceived intentionality does not explain the relationship between ad creator and forgiveness. This supports our theorizing that perceived effort, rather than perceived intent, drives consumer forgiveness.

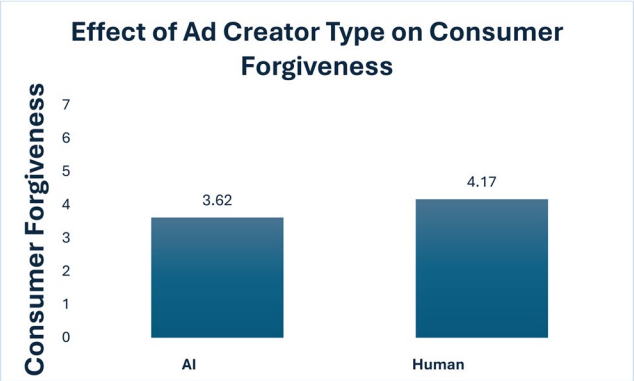

Figure 2 Effect of Ad Creator on Consumer Forgiveness (Study 1)

Control Variables. Participants found both ad equally misrepresentative (M_{AI} =2.55, SD = 1.74 vs. M_{human} = 2.97, SD = 1.78; F(1, 160)=2.28, p>.10, η^2=0.07). Next, we tested whether consumer forgiveness differed by ethnicity, specifically white vs. non-white. Results showed no significant interaction between ad creator and ethnicity (p=.714), and estimated marginal means of forgiveness were comparable across groups (M_{white} = 3.84, SD = 1.63 vs. $M_{non\text{-}white}$ = 4.02, SD = 1.53; F (1, 160)= 0.398, p = 0.529, η^2=0.04).

Discussion of Findings

Study 1 provides initial support for our proposed mechanism. Consistent with H1, participants were less forgiving of misrepresentative ads created by AI than by humans (Figure 2). Supporting H2, this effect was serially mediated by perceived effort and ad attitude: AI-created ads were seen as involving lower effort, which reduced ad attitudes and, in turn, forgiveness. Including perceived intentionality as an additional mediator did not account for the relationship, underscoring the unique explanatory role of perceived effort.

Although attribution theory highlights perceived intent as central to forgiveness (Fehr, Gelfand, and Nag 2010; Weiner 2012), our findings indicate that intent is less diagnostic in AI-driven contexts. When causes of misrepresentation are unclear, consumers instead evaluate how much care and oversight the ad reflects, rendering perceived effort a stronger signal of responsibility and moral concern.

STUDY 2: AD CREATOR IDENTITY AND CONSUMER RESPONSE IN LGBTQIA+ CONTEXTS: EVIDENCE FROM QUOTA-BASED SAMPLING

Overview and Purpose

Study 2 builds directly on Study 1 with two objectives. First, it tests whether the previously identified attributional mechanism replicates in an identity-relevant advertising context (LGBTQIA+-themed ads, H2a). Second, it examines whether this mechanism operates specifically among consumers for whom the ad content is identity-relevant (LGBTQIA+ participants, H2b). Together, these objectives assess the robustness and generalizability of the proposed serial mediation when misrepresentation concerns are tied to personal social identity and moral stakes are elevated. Additionally, Study 2 rules out perceived intent and perceived relevance as alternative mechanisms, demonstrating that effort-based attributions remain the primary driver of forgiveness responses.

Method

Participation and Design. We recruited 340 participants from Prolific (59.4% female, 38.0% male, 2.3% non-binary; 50% self-identified as LGBTQIA+; Mage = 39.29 years) in exchange for monetary compensation. Recruitment followed a quota-based approach to ensure adequate representation of LGBTQIA+ participants. This design tests whether the serial mediation replicates in the full sample (H2a) and whether it operates among LGBTQIA+ participants specifically (H2b). Analyzing LGBTQIA+ participants only for H2b ensures that identity relevance is unambiguous while avoiding potential confounding from heterogeneous patterns of cause involvement among non-LGBTQIA+ consumers. Participants were randomly assigned to one of two conditions (ad creator: AI vs. human) and read a short caption identifying the ad creator, identical to that used in Study 1.

Stimulus and Procedure. Following the same procedure and measures as in Study 1, we created (see Appendix B) and pretested an advertisement for misrepresentation (see Appendix F) adapted to an identity-relevant context. The ad targeted the LGBTQIA+ community and featured the fictional brand *True Glow* with the tagline "True Pride, True Glow." While the ad visually referenced Pride Month, it relied on limited-edition messaging (e.g., "available only this Pride Month—don't miss a chance to shine") without articulating substantive support for the LGBTQIA+ community, reflecting a form of symbolic or commodified inclusion. Along with the ad, participants read the same caption identifying the ad as AI- or human-created as in Study 1 (full stimulus in Appendix B).

Measures. Participants completed the same items on perceived effort, perceived intent, attitude toward the ad, forgiveness and manipulation check as in Study 1. Participants also completed items on perceived ad relevance (four-item 7-point Likert, e.g.,

"How much do you feel this ad was designed to be relevant to you"), an additional construct relevant in identity sensitive context. Finally, participants completed a three-item measure of perceived representation accuracy (e.g., "To what extent does this ad accurately represent the LGBTQIA+ community") and demographics (age, gender, ethnicity).

Results

Manipulation Check. Participants correctly identified the ad creator (M_{AI} = 1.58, SD = 1.07 vs. M_{human}=5.32, SD = 1.34; F (1, 338)=806.00, $p < .001$, η^2 =0.705), confirming that the manipulation worked as intended.

Dependent Variable(s). A one-way ANOVA revealed that participants exposed to an AI-generated misrepresentative ad were less forgiving than those who viewed a human-generated ad (M_{AI} = 3.72, SD = 1.36 vs. M_{human}=4.32, SD = 1.27; F(1, 338)=17.52, $p < .001$; η^2 =0.50).

Sequential Mediation Analysis(H2a). Using Hayes PROCESS Model 6 (Preacher and Hayes 2008), we tested perceived effort (M1) and ad attitude (M2) as sequential mediators. Results revealed an indirect effect of ad creator type on forgiveness through perceived effort and ad attitude (β=–0.18, SE = 0.0716, 95% CI [–0.35, –0.07]). This finding replicates the attributional mechanism identified in Study 1.

Identity-Relevant Replication (H2b). To test whether the serial mediation operates among consumers for whom the ad content is identity-relevant, we estimated the same mediation model among LGBTQIA+ participants only (N = 169). This analytical approach addresses the concern that non-LGBTQIA+ participants may differ in cause involvement, ensuring a clean test when identity relevance is unambiguous. Results revealed a significant indirect effect of ad creator on forgiveness through perceived effort and ad attitudes (indirect effect = 0.15, 95% CI [0.03, 0.33]), supporting H2b. This finding demonstrates that the effort-based attributional mechanism operates when consumers' own social identities are directly implicated by the misrepresentation.

Ruling Out Alternate Explanations. To examine the mechanisms through which ad creator (AI vs. human) influences consumer forgiveness, we conducted a parallel mediation analysis including perceived effort, perceived intentionality, and perceived relevance as mediators. Among the three proposed mediators, only perceived effort significantly mediated the relationship between ad creator and forgiveness (*indirect effect* =.23, 95% *CI* [.13, .35]). Neither perceived intentionality nor perceived relevance emerged as significant mediators, as their indirect effects were non-significant (intentionality: *indirect effect* =.002, 95% CI [–.01, .03]; relevance: *indirect effect* = .02, 95% CI [–.03, .09]). Contrast analyses further confirmed that the indirect effect via effort was significantly stronger than those via intentionality or relevance.

Discussion of Findings

Study 2 reinforces perceived effort as the central mechanism through which consumers differentiate between AI- and human-created misrepresentative ads. Consistent with Study 1, ads attributed to human creators were perceived as involving greater effort, leading to more favorable ad attitudes and greater forgiveness, supporting H2a.

Importantly, Study 2 rules out alternative explanations salient in identity- based advertising contexts. Neither perceived intentionality nor perceived ad relevance mediated the effect of ad creator on forgiveness. Although perceived relevance was positively associated with forgiveness overall, it did not account for differential responses to AI- versus human-created ads, suggesting that relevance shapes baseline evaluations but does not substitute for effort-based inferences.

Finally, the serial mediation operated among LGBTQIA+ participants for whom the ad content was identity-relevant (H2b), demonstrating that the mechanism extends to contexts where consumers' personal identities are directly implicated by the misrepresentation. Together, these findings show that consumer forgiveness depends primarily on whether the ad reflects a meaningful effort rather than on whether it appears targeted or relevant to one's identity.

STUDY 3: ROLE OF BRAND REPUTATION

Overview and Purpose

Study 3 tests whether brand reputation alters consumer responses to AI- versus human-created misrepresentative ads. While Studies 1 and 2 demonstrated that consumers are less forgiving of misrepresentative ads created by AI (vs. humans) because they attribute lower effort to AI, Study 3 examines whether a strong brand reputation can buffer this effect.

Pilot Study

A pretest was conducted to validate the brand reputation manipulation. We recruited 101 participants via CloudResearch (64% female, 36% male, M_{age} = 41.84 years). Participants viewed the *New Barbie Spa and Salon* ad (as in Study 1), followed by a description of the brand's reputation. In the high-reputation condition, the brand was described as *"widely recognized for its high-quality services, ethical*

business practices, and dedication to customers." In the low-reputation condition, the brand was described as *"frequently criticized for inconsistent quality, outdated facilities, and unprofessional staff."* Participants then rated the brand's reputation on a three-item, seven-point scale (e.g., "This brand has a strong reputation for delivering high-quality products and services," 1 = strongly disagree, 7 = strongly agree). As expected, participants rated the brand as significantly more reputable in the high-reputation condition than in the low-reputation condition ($M_{highrep}$ = 5.36, SD = 1.51 vs. M_{lowrep} = 1.90, SD = 1.29; F (1, 100) = 150.47, $p < 0.001$) confirming successful manipulation.

Main Study

Participation and Design. A total of 260 participants were recruited via Prolific (56.1% female, 43.5% male; 0.3% binary; M_{age} = 43.57 years). Participants were randomly assigned to 2 (Ad creator: AI vs. Human) *2 (Brand Reputation: High vs. Low) between-subjects design.

Stimulus and Procedures. Participants first came across an ad for the brand *New Barbie Spa and Salon* with a caption on AI/Human ad developer like in previous studies, followed by an editorial article discussing the reputation of the brand (using validated stimuli from the pilot).

Measures. Participants then completed items on forgiveness (α = 0.87), attitude (α = 0.96), and manipulation checks for ad creator (α = 0.88) and brand reputation(α = 0.97), and questions on demographics.

Results

Manipulation Check. Participants correctly identified the ad creator (M_{AI} =1.48, SD = 0.90 vs. M_{human}=5.14, SD = 1.87; F(1, 258)=399.34, p< .001, η^2=0.60). Further, participants in the higher reputation condition indeed rated the brand as highly reputed than participants in the lower reputation condition (M_{hbr} =5.62, SD = 1.20 vs. M_{lbr}=1.85, SD = 1.24; F (1, 258) = 614.06, $p < .001$, η^2 =0.70).

Dependent Variables and Conditional Effects. There was the main effect of both ad creator (M_{AI} =3.22, SD = 1.47 vs. M_{human}=3.70, SD = 1.37; F (1, 258) =7.54, $p < .01$, η^2 =0.03) and brand reputation (M_{hbr}=3.72, SD = 1.39 vs. M_{lbr}= 3.20, SD = 1.44; F (1, 258) =8.70, $p < .01$, η^2 =0.03) on consumer forgiveness, although the interaction was not significant ($p > 0.10$). Conditional effects analyses showed that under low brand reputation, participants were less forgiving of AI-created ads compared to human-created ads (M_{AI} =2.93, SD = 1.41 vs. M_{human} = 3.47, SD = 1.43; F (1, 128) = 4.68, p =.03). Similarly, ad attitudes were lower for AI (vs. human) creators in this condition (M_{AI} = 3.11, SD = 1.56 vs. M_{human} = 3.66, SD = 1.42; F (1, 128) = 4.44, p =.038). However, these differences become insignificant in high reputation condition for both forgiveness (M_{AI} =3.51, SD = 1.48 vs. M_{human}=3.84, SD = 1.28; F (1, 128) = 3.15, $p > .10$) and attitudes (M_{AI} = 4.27, SD = 1.69 vs. M_{human}= 4.61, SD = 1.62; F (1, 128) = 1.30, p = .255).

Discussions of Findings

Study 3 demonstrates that brand reputation functions as a boundary condition shaping how consumers interpret ad creator identity. When brand reputation is low, AI-created ads elicit less favorable attitudes and reduced forgiveness than human-created ads, replicating the effects observed in Studies 1 and 2. Consistent with H3, these differences disappear under high reputation, indicating that brand reputation overrides the negative inferences consumers typically draw about AI involvement.

GENERAL DISCUSSION

This research examines how consumers respond to AI- versus human-generated misrepresentative advertising. Across three studies, we demonstrate that AI-generated misrepresentations elicit less forgiveness than human-created ones, with this effect mediated by perceived effort rather than intent. By identifying effort as the central mechanism, we extend attribution theory to address how consumers assign responsibility when the proximate creator is nonhuman yet institutional accountability persists.

Critically, Study 2 demonstrates this effort-based mechanism operates in identity-relevant contexts where moral stakes are elevated. Testing among LGBTQIA+ consumers evaluating Pride-themed advertising, we show the serial mediation remains robust when consumers' personal identities are directly implicated, addressing concerns about AI's cultural sensitivity limitations (Huang and Rust 2021).

We also identify brand reputation as a moderator: high-reputation brands are buffered from the AI penalty, though our moderated serial mediation analysis reveals this buffering operates outside the effort-attribution pathway—we speculate this occurs through accumulated brand equity. For low-reputation brands lacking such goodwill, the AI penalty persists. These findings advance AI ethics debates in marketing (Milano, Taddeo, and Floridi 2020) by demonstrating that moral judgments hinge on inferred care and oversight rather than intent, highlighting ethical challenges when deploying AI in representationally sensitive domains.

THEORETICAL CONTRIBUTIONS

This research contributes to advertising, consumer psychology, and human–AI interaction literature by examining how ad creator identity (AI vs. human) shapes consumer responses to representational missteps.

First, we extend attribution theory (Kelley 1973; Weiner 1985, 2012) to contexts where the proximate creator is nonhuman, yet institutional accountability persists. Classic attribution theory explains responsibility judgments about morally agentic actors assumed to possess intentions and moral awareness. AI systems differ fundamentally: they produce outputs resembling human decisions but lack mental states and moral agency while operating within institutional arrangements in which brands retain responsibility. We show that under these conditions, responsibility judgments shift from intent-based to effort-based evaluations. Because consumers cannot meaningfully infer intent from nonhuman creators, perceived effort becomes the primary diagnostic cue for whether brands exercised adequate care and oversight when deploying AI technology. This extends attribution theory by demonstrating how responsibility judgments adapt when nonhuman production is embedded within institutional accountability.

Second, we advance understanding of consumer responses to AI-generated advertising beyond content-level features. Prior work has emphasized message attributes such as argument strength or executional quality (Shavitt, Lowrey, and Haefner 1998; Smith, Chen, and Yang 2008). We show that creator identity constitutes a distinct evaluative dimension. When ads misrepresent marginalized groups, AI-generated versions elicit more negative reactions through reduced perceptions of effort. This extends research on algorithmic agents in morally complex domains (Castelo, Bos, and Lehmann 2019; Longoni, Bonezzi, and Morewedge 2019), suggesting that consumers evaluate AI-generated content more critically in contexts requiring cultural sensitivity.

We further identify perceived effort as the mechanism linking ad creator identity to consumer forgiveness. Prior research on forgiveness has emphasized the role of transgression severity and relationship strength (Tsarenko and Tojib 2015; Wei, Liu, and Keh 2020). Our findings indicate that forgiveness also depends on perceived production effort: AI-generated misrepresentations signal insufficient care and oversight, reducing ad attitudes and forgiveness. Replication with LGBTQIA+ participants demonstrates that this mechanism persists when consumers' identities are directly implicated, extending forgiveness research by showing that creation-process attributions operate alongside transgression characteristics and relationship factors.

Third, we extend research on brand reputation and consumer responses to corporate missteps (Aaker, Vohs, and Mogilner 2010; Greyser 2009) by demonstrating that reputation moderates reactions to AI-created misrepresentations. While a strong reputation buffers the negative effect of AI involvement on forgiveness, this buffering operates independently of the effort-attribution pathway. We speculate that high-reputation brands draw on accumulated relational capital that provides direct routes to forgiveness, functioning as a parallel mechanism rather than altering attributional processing. This finding enriches theories of brand equity and crisis management by showing that reputational assets and attributional processes operate as complementary pathways shaping responses to brand failures.

Finally, this research contributes to debates on AI transparency and disclosure in ethically sensitive domains (Milano, Taddeo, and Floridi 2020). Although disclosure is often assumed to enhance fairness and trust, our findings suggest that revealing AI authorship can backfire in representationally sensitive contexts. When misrepresentation occurs, AI disclosure signals reduced oversight and care, lowering perceived effort and forgiveness. Transparency policies may therefore require contextual calibration, as disclosure in domains demanding cultural sensitivity and ethical judgment may undermine rather than enhance consumer confidence.

MANAGERIAL IMPLICATIONS

This research offers actionable guidance for managers integrating AI into advertising workflows. While AI provides efficiency, consumer responses depend on ad content and perceptions of creator identity and effort. Managers must therefore consider how audiences interpret these cues while deploying AI.

First, human oversight is essential, especially for identity or cultural campaigns. Even when AI generates draft content, human creators should refine outputs to ensure cultural sensitivity and authenticity. Consumers more readily forgive human errors, which they interpret as isolated lapses, whereas AI errors are viewed as signs of systemic negligence or insufficient care. Assigning AI to socially sensitive messaging amplifies reputational risk—one that

Consumers more readily forgive human errors, which they interpret as isolated lapses, whereas AI errors are viewed as signs of systemic negligence or insufficient care.

"AI-generated" labels may backfire in representational contexts by highlighting automation and reinforcing low-effort inferences.

can be mitigated through clear oversight and rigorous quality assurance.

Second, disclosure strategies require nuance. Our findings indicate that simple "AI-generated" labels may backfire in representational contexts by highlighting automation and reinforcing low-effort inferences. Brands need not conceal AI use, but selective disclosure, for instance, positioning AI as a supporting tool ("AI-assisted, reviewed by our creative team") signals efficiency while preserving human care. Before launching AI-generated campaigns in sensitive domains, brands should conduct A/B testing with diverse consumer panels to assess whether AI disclosure undermines perceived effort and ad attitudes. Emphasizing human involvement is particularly important in campaigns related to diversity, equity, or cultural identity.

Third, brand reputation strongly shapes consumer reactions to AI involvement. Established brands with "trust reserves" benefit from reputational buffering as consumers attribute missteps to situational factors rather than indifference or lack of care. Weaker brands lack this protection; their AI use may instead signal cost-cutting and heighten scrutiny. For lower-reputation brands, we recommend explicitly communicating human oversight in campaign messaging (e.g., "Reviewed by our diversity and inclusion team") and prioritizing transparency about creative processes to build trust before deploying AI at scale.

LIMITATIONS AND FUTURE RESEARCH

This research advances understanding of responses to AI- versus human-generated misrepresentative advertising, yet several limitations open avenues for future work.

First, although we experimentally manipulated ad creator identity, consumers in practice infer creator identity from stylistic cues such as overly polished visuals or generic copy. Future work should examine perceived rather than disclosed creator identity to test whether misclassification produces similar attributional outcomes.

Second, while perceived effort emerged as the dominant mechanism, it may not capture all psychological bases of forgiveness. Additional mediators such as perceived warmth or creativity (Sundar 2020) may also shape consumer judgments. Future research could differentiate among AI roles (e.g., generative vs. assistive) or varying levels of human–AI collaboration to reveal how these configurations alter attributional reasoning.

Third, contextual moderators beyond brand reputation warrant attention. Industry domain, ad placement, and cultural norms may shape tolerance for AI-driven misrepresentation. Consumers may accept algorithmic creativity in technology or entertainment but react negatively in identity-sensitive domains. Cross-cultural work could illuminate how national differences in automation trust or privacy orientation shape responses to AI-authored ads.

Fourth, while restricting H2b to LGBTQIA+ participants addresses cause involvement heterogeneity for that hypothesis, we did not measure cause involvement among non-LGBTQIA+ participants in the full sample analysis (H2a). As a result, involvement may have varied across participants (e.g., allies vs. uninvolved individuals), potentially influencing how the ad was processed. Although the observed effects are consistent with our theorized mechanism, future research should explicitly measure and control for cause involvement to rule out this alternative explanation and examine whether the effort-based attribution process varies across levels of involvement in identity-relevant contexts.

Fifth, while brand reputation buffers high-reputation brands from the AI penalty, our moderated mediation analysis (Model 59; Web Appendix I) reveals this buffering operates independently of the effort-attribution pathway. The underlying mechanism is not directly tested in our data. We therefore speculate that this pattern occurs because high-reputation brands appear to benefit from accumulated relational capital providing direct routes to forgiveness (Ahluwalia, Burnkrant, and Unnava 2000; Cheng, White, and Chaplin 2012). Future research should identify specific mechanisms of this parallel effect (e.g., brand trust, relationship strength) and assess whether it extends to marginalized communities, who may evaluate brands through histories of representation and community trust.

Established brands with "trust reserves" benefit from reputational buffering as consumers attribute missteps to situational factors rather than indifference or lack of care.

Finally, our samples were predominantly White, limiting demographic diversity. Future research should examine whether reactions to AI-created misrepresentations are stronger among racially minoritized consumers, who may be more sensitive to representation and brand missteps.

DISCLOSURE STATEMENT

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

FUNDING

No funding was received.

ABOUT THE AUTHOR

Kshitij Bhoumik is a Lecturer in Marketing at Leeds University Business School, University of Leeds. His research examines how AI mediates consumer-brand relationships, affects consumer wellbeing, and shapes advertising ethics for marginalized communities. His work has been published in the *Journal of Advertising Research*, *Journal of Advertising*, *Journal of Interactive Marketing*, *Psychology & Marketing*, and *Journal of Business Research*. He is also the author of the forthcoming textbook AI in Marketing (Kogan Page/Hachette, 2027).

ORCID

Kshitij Bhoumik http://orcid.org/0000-0001-7566-271X

REFERENCES

Aaker, J., E. N. Garbinsky, and K. D. Vohs. 2012. "Cultivating Admiration in Brands: Warmth, Competence, and Landing in the "Golden Quadrant." *Journal of Consumer Psychology* 22 (2): 191–194. https://doi.org/10.1016/j.jcps.2011.11.012.

Aaker, J., K. D. Vohs, and C. Mogilner. 2010. "Nonprofits Are Seen as Warm and for-Profits as Competent: Firm Stereotypes Matter." *Journal of Consumer Research* 37 (2): 224–237. https://doi.org/10.1086/651566.

Ahluwalia, R., R. E. Burnkrant, and H. R. Unnava. 2000. "Consumer Response to Negative Publicity: The Moderating Role of Commitment." *Journal of Marketing Research* 37 (2): 203–214. https://doi.org/10.1509/jmkr.37.2.203.18734.

Åkestam, N. 2017. *Understanding Advertising Stereotypes: Social and Brand-Related Effects of Stereotyped versus Non-Stereotyped Portrayals in Advertising*. Stockholm, Sweden: Stockholm School of Economics.

Allyn, B. 2023. "Levi's announced it will use AI-generated models to increase diversity, and people aren't buying it." *Business Insider* https://www.businessinsider.com/levis-ai-generated-company-models-diversity-backlash-2023-3

Benjamin, R. 2019. "Assessing Risk, Automating Racism." *Science (New York, N.Y.)* 366 (6464): 421–422. https://doi.org/10.1126/science.aaz3873.

Blanken, I., N. Van De Ven, and M. Zeelenberg. 2015. "A Meta-Analytic Review of Moral Licensing." *Personality & Social Psychology Bulletin* 41 (4): 540–558. https://doi.org/10.1177/0146167215572134.

Browne, B. A. 1998. "Gender Stereotypes in Advertising on Children's Television in the 1990s: A Cross-National Analysis." *Journal of Advertising* 27 (1): 83–96. https://doi.org/10.1080/00913367.1998.10673544.

Burgess, A., D. C. H. Wilkie, and R. Dolan. 2023. "Brand Approaches to Diversity: A Typology and Research Agenda." *European Journal of Marketing* 57 (1): 60–88. https://doi.org/10.1108/EJM-09-2021-0696.

Campbell, M. C., and A. Kirmani. 2000. "Consumers' Use of Persuasion Knowledge: The Effects of Accessibility and Cognitive Capacity on Perceptions of an Influence Agent." *Journal of Consumer Research* 27 (1): 69–83. https://doi.org/10.1086/314309.

Campbell, C. 2023. "Ready or Not, Generative AI is Here to Stay: Advertisers Need More Research to Harness the Benefits of AI Technologies." *Journal of Advertising Research* 63 (3): 202–204. https://doi.org/10.2501/JAR-2023-019.

Campbell, C., S. Sands, B. McFerran, and A. Mavrommatis. 2025. "Diversity Representation in Advertising." *Journal of the Academy of Marketing Science* 53 (2): 588–616. https://doi.org/10.1007/s11747-023-00994-8.

Castleberry, S. B., W. French, and B. A. Carlin. 1993. "The Ethical Framework of Advertising and Marketing Research Practitioners: A Moral Development Perspective." *Journal of Advertising* 22 (2): 39–46. https://doi.org/10.1080/00913367.1993.10673402.

Castelo, N., M. W. Bos, and D. R. Lehmann. 2019. "Task-Dependent Algorithm Aversion." *Journal of Marketing Research* 56 (5): 809–825. https://doi.org/10.1177/0022243719851788.

Champlin, S., and M. Li. 2020. "Communicating Support in Pride Collection Advertising: The Impact of Gender Expression and Contribution Amount." *International Journal of Strategic Communication* 14 (3): 160–178. https://doi.org/10.1080/1553118X.2020.1750017.

Cheng, Y., X. Zhou, and K. Yao. 2023. "LGBT-Inclusive Representation in Entertainment Products and Its Market Response: evidence from Field and Lab." *Journal of Business Ethics* 183 (4): 1189–1209. https://doi.org/10.1007/s10551-022-05075-4.

Cheng, S. Y., T. B. White, and L. N. Chaplin. 2012. "The Effects of Self-Brand Connections on Responses to Brand Failure: A New Look at the Consumer–Brand Relationship." *Journal of Consumer Psychology* 22 (2): 280–288. https://doi.org/10.1016/j.jcps.2011.05.005.

Chung, E., and M. Beverland. 2006. "An Exploration of Consumer Forgiveness Following Marketer Transgressions." *Advances in Consumer Research* 33 (1).

Coeckelbergh, M. 2020. *AI Ethics*. Cambridge, MA: MIT Press.

Cohen, J. 1992. "A Power Primer." *Psychological Bulletin* 112 (1): 155–159. https://doi.org/10.1037/0033-2909.112.1.155.

Dawar, N., and J. Lei. 2009. "Brand Crises: The Roles of Brand Familiarity and Crisis Relevance in Determining the Impact on Brand Evaluations." *Journal of Business Research* 62 (4): 509–516. https://doi.org/10.1016/j.jbusres.2008.02.001.

De Chernatony, L. 1999. "Brand Management through Narrowing the Gap between Brand Identity and Brand Reputation." *Journal of Marketing Management* 15 (1-3): 157–179. https://doi.org/10.1362/026725799784870432.

Eisend, M. 2010. "A Meta-Analysis of Gender Roles in Advertising." *Journal of the Academy of Marketing Science* 38 (4): 418–440. https://doi.org/10.1007/s11747-009-0181-x.

Eisend, M. 2019. "Gender Roles." *Journal of Advertising* 48 (1): 72–80. https://doi.org/10.1080/00913367.2019.1566103.

Eisend, M., A. F. Muldrow, and S. Rosengren. 2023. "Diversity and Inclusion in Advertising Research." *International Journal of Advertising* 42 (1): 52–59. https://doi.org/10.1080/02650487.2022.2122252.

Fehr, R., M. J. Gelfand, and M. Nag. 2010. "The Road to Forgiveness: A Meta-Analytic Synthesis of Its Situational and Dispositional Correlates." *Psychological Bulletin* 136 (5): 894–914. https://doi.org/10.1037/a0019993.

Fombrun, C., and M. Shanley. 1990. "What's in a Name? Reputation Building and Corporate Strategy." *Academy of Management Journal* 33 (2): 233–258. https://doi.org/10.2307/256324.

Gardete, P. M. 2013. "Cheap-Talk Advertising and Misrepresentation in Vertically Differentiated Markets." *Marketing Science* 32 (4): 609–621. https://doi.org/10.1287/mksc.2013.0772.

Gerrath, M. H., K. Bhoumik, A. Biraglia, A. Ulqinaku, and G. Viglia. 2026. "How the Timing of LGBT+ Brand Activism Affects Consumer Responses: Better Early Than Late?" *Journal of Advertising Research* 66 (1): 47–62. https://doi.org/10.1080/00218499.2025.2485414.

Grau, S. L., and Y. C. Zotos. 2018. "Gender Stereotypes in Advertising: A Review of Current Research." *Current Research on Gender Issues in Advertising:* 3–12.

Gray, K., and D. M. Wegner. 2012. "Feeling Robots and Human Zombies: Mind Perception and the Uncanny Valley." *Cognition* 125 (1): 125–130. https://doi.org/10.1016/j.cognition.2012.06.007.

Greyser, S. A. 2009. "Corporate Brand Reputation and Brand Crisis Management." *Management Decision* 47 (4): 590–602. https://doi.org/10.1108/00251740910959431

Gurrieri, L., J. Brace-Govan, and H. Cherrier. 2016. "Controversial Advertising: transgressing the Taboo of Gender-Based Violence." *European Journal of Marketing* 50 (7/8): 1448–1469. https://doi.org/10.1108/EJM-09-2014-0597.

Hoffmann, A. L. 2019. "Where Fairness Fails: data, Algorithms, and the Limits of Antidiscrimination Discourse." *Information, Communication & Society* 22 (7): 900–915. https://doi.org/10.1080/1369118X.2019.1573912.

Huang, M. H., and R. T. Rust. 2021. "A Strategic Framework for Artificial Intelligence in Marketing." *Journal of the Academy of Marketing Science* 49 (1): 30–50. https://doi.org/10.1007/s11747-020-00749-9.

Huh, J., M. R. Nelson, and C. A. Russell. 2023. "ChatGPT, AI Advertising, and Advertising Research and Education." *Journal of Advertising* 52 (4): 477–482. https://doi.org/10.1080/00913367.2023.2227013.

Johnson, G. D., and S. A. Grier. 2012. "What about the Intended Consequences?" *Journal of Advertising* 41 (3): 91–106. https://doi.org/10.2753/JOA0091-3367410306

Kang, M. E. 1997. "The Portrayal of Women's Images in Magazine Advertisements: Goffman's Gender Analysis Revisited." *Sex Roles* 37 (11-12): 979–996. https://doi.org/10.1007/BF02936350.

Kang, H. 2021. "Sample Size Determination and Power Analysis Using the G* Power Software." *Journal of Educational Evaluation for Health Professions* 18: 17. https://doi.org/10.3352/jeehp.2021.18.17.

Kaplan, A. D., T. T. Kessler, J. C. Brill, and P. A. Hancock. 2023. "Trust in Artificial Intelligence: Meta-Analytic Findings." *Human Factors* 65 (2): 337–359. https://doi.org/10.1177/00187208211013988.

Kelley, H. H. 1973. "The Processes of Causal Attribution." *American Psychologist* 28 (2): 107–128. https://doi.org/10.1037/h0034225.

Kelley, H. H., and J. L. Michela. 1980. "Attribution Theory and Research." *Annual Review of Psychology* 31 (1): 457–501. https://doi.org/10.1146/annurev.ps.31.020180.002325.

Kumar, V., B. Rajan, R. Venkatesan, and J. Lecinski. 2019. "Understanding the Role of Artificial Intelligence in Personalized Engagement Marketing." *California Management Review* 61 (4): 135–155. https://doi.org/10.1177/0008125619859317.

Longoni, C., A. Bonezzi, and C. K. Morewedge. 2019. "Resistance to Medical Artificial Intelligence." *Journal of Consumer Research* 46 (4): 629–650. https://doi.org/10.1093/jcr/ucz013.

Loureiro, S. M. C., J. Guerreiro, and I. Tussyadiah. 2021. "Artificial Intelligence in Business: State of the Art and Future Research Agenda." *Journal of Business Research* 129: 911–926. https://doi.org/10.1016/j.jbusres.2020.11.001.

McCullough, M. E., F. D. Fincham, and J. A. Tsang. 2003. "Forgiveness, Forbearance, and Time: The Temporal Unfolding of Transgression-Related Interpersonal Motivations." *Journal of Personality and Social Psychology* 84 (3): 540–557. https://doi.org/10.1037/0022-3514.84.3.540.

McCullough, M. E., K. C. Rachal, S. J. Sandage, E. L. Worthington, Jr, S. W. Brown, and T. L. Hight. 1998. "Interpersonal Forgiving in Close Relationships: II. Theoretical Elaboration and Measurement." *Journal of Personality and Social Psychology* 75 (6): 1586–1603. https://doi.org/10.1037/0022-3514.75.6.1586.

McFall, L. 2004. "*Advertising: A Cultural Economy*."

MacKenzie, S. B., R. J. Lutz, and G. E. Belch. 1986. "The Role of Attitude toward the Ad as a Mediator of Advertising Effectiveness: A Test of Competing Explanations." *Journal of Marketing Research* 23 (2): 130–143. https://doi.org/10.2307/3151660.

Malle, B. F., and J. Knobe. 1997. "The Folk Concept of Intentionality." *Journal of Experimental Social Psychology* 33 (2): 101–121. https://doi.org/10.1006/jesp.1996.1314.

Milano, S., M. Taddeo, and L. Floridi. 2020. "Recommender Systems and Their Ethical Challenges." *AI & SOCIETY* 35 (4): 957–967. https://doi.org/10.1007/s00146-020-00950-y.

Mohr, L. A., and M. J. Bitner. 1995. "The Role of Employee Effort in Satisfaction with Service Transactions." *Journal of Business Research* 32 (3): 239–252. https://doi.org/10.1016/0148-2963(94)00049-K.

Montecchi, M., M. R. Micheli, M. Campana, and H. J. Schau. 2024. "From Crisis to Advocacy: tracing the Emergence and Evolution of the LGBTQIA+ Consumer Market." *Journal of Public Policy & Marketing* 43 (1): 10–30. https://doi.org/10.1177/07439156231183645.

Peer, E., L. Brandimarte, S. Samat, and A. Acquisti. 2017. "Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research." *Journal of Experimental Social Psychology* 70: 153–163. https://doi.org/10.1016/j.jesp.2017.01.006.

Preacher, K. J., and A. F. Hayes. 2008. "Asymptotic and Resampling Strategies for Assessing and Comparing Indirect Effects in Multiple Mediator Models." *Behavior Research Methods* 40 (3): 879–891. https://doi.org/10.3758/brm.40.3.879.

Schroeder, J. E., and J. L. Borgerson. 2005. "An Ethics of Representation for International Marketing Communication." *International Marketing Review* 22 (5): 578–600. https://doi.org/10.1108/02651330510624408

Schroeder, J. E., and D. Zwick. 2004. "Mirrors of Masculinity: Representation and Identity in Advertising Images." *Consumption Markets & Culture* 7 (1): 21–52. https://doi.org/10.1080/1025386042000212383.

Shavitt, S., P. Lowrey, and J. Haefner. 1998. "Public Attitudes toward Advertising: More Favorable than You Might Think." *Journal of Advertising Research* 38 (4): 7–22. https://doi.org/10.1080/00218499.1998.12466562.

Smith, R. E., J. Chen, and X. Yang. 2008. "The Impact of Advertising Creativity on the Hierarchy of Effects." *Journal of Advertising* 37 (4): 47–62. https://doi.org/10.2753/JOA0091-3367370404.

Sohn, Y. J., and R. W. Lariscy. 2015. "A "Buffer" or "Boomerang?"—The Role of Corporate Reputation in Bad Times." *Communication Research* 42 (2): 237–259. https://doi.org/10.1177/0093650212466891.

Surnit. 2024. "Skechers Faces Backlash Over AI-Generated Ad in Vogue's December Issue." https://www.designrush.com/news/skechers-faces-backlash-over-ai-generated-ad-in-vogue-december-issue?

Sundar, S. S. 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.

Tsarenko, Y., and D. Tojib. 2015. "Consumers' Forgiveness after Brand Transgression: The Effect of the Firm's Corporate Social Responsibility and Response." *Journal of Marketing Management* 31 (17-18): 1851–1877. https://doi.org/10.1080/0267257X.2015.1069373.

Vakratsas, D., and X. Wang. 2021. "Artificial Intelligence in Advertising Creativity." *Journal of Advertising* 50 (1): 39–51. https://doi.org/10.1080/00913367.2020.1843090.

Vredenburg, J., S. Kapitan, A. Spry, and J. A. Kemper. 2020. "Brands Taking a Stand: Authentic Brand Activism or Woke Washing?" *Journal of Public Policy & Marketing* 39 (4): 444–460. https://doi.org/10.1177/0743915620947359.

Walsh, G., V. W. Mitchell, P. R. Jackson, and S. E. Beatty. 2009. "Examining the Antecedents and Consequences of Corporate Reputation: A Customer

Perspective." *British Journal of Management* 20 (2): 187–203. https://doi.org/10.1111/j.1467-8551.2007.00557.x.

Wilkie, D. C. H., A. J. Burgess, A. Mirzaei, and R. M. Dolan. 2023. "Inclusivity in Advertising: A Typology Framework for Understanding Consumer Reactions." *Journal of Advertising* 52 (5): 721–738. https://doi.org/10.1080/00913367.2023.2255252.

Wei, C., M. W. Liu, and H. T. Keh. 2020. "The Road to Consumer Forgiveness is Paved with Money or Apology? The Roles of Empathy and Power in Service Recovery." *Journal of Business Research* 118: 321–334. https://doi.org/10.1016/j.jbusres.2020.06.061.

Weiner, B. 1985. "Attribution Theory." In *Human Motivation*, 275–326. New York, NY: Springer.

Weiner, B. 2012. *An Attributional Theory of Motivation and Emotion*. New York, NY: Springer Science & Business Media.

Xie, Y., and S. Peng. 2009. "How to Repair Customer Trust after Negative Publicity: The Roles of Competence, Integrity, Benevolence, and Forgiveness." *Psychology & Marketing* 26 (7): 572–589. https://doi.org/10.1002/mar.20289.

Zinkhan, G. M. 1994. "International Advertising: A Research Agenda." *Journal of Advertising* 23 (1): 11–15. https://doi.org/10.1080/00913367.1994.10673427.

Zotos, Y. C., and S. L. Grau. 2016. "Gender Stereotypes in Advertising: exploring New Directions." *International Journal of Advertising* 35 (5): 759–760. https://doi.org/10.1080/02650487.2016.1203555.

Zou, J., and L. Schiebinger. 2018. "AI Can be Sexist and Racist—It's Time to Make It Fair." https://doi.org/10.1038/d41586-018-05707-8