Journal of Marketing Analytics
https://doi.org/10.1057/s41270-026-00534-7

ORIGINAL ARTICLE

"Made with AI" but why? How consumers interpret beneficiary-framed AI disclosures in advertising

Timo Schreiner[1] · Yakov Bart[2] · Koen Pauwels[2] · Cosima Schütze[1] · Fabian Buder[1,2]

Received: 4 May 2026 / Accepted: 13 August 2026

Abstract
As generative AI becomes more prevalent in advertising, firms increasingly face requirements to disclose AI involvement. Although prior research shows that such disclosures may generate negative consumer responses, it remains unclear whether explanatory disclosures can mitigate these effects. This research examines beneficiary-framed AI disclosures, explanations communicating why AI was used and who benefits, across three studies. In a controlled text-based experiment (Study 1), more specific explanations improve evaluations relative to minimal and less specific AI labels. However, beneficiary framing effects are not systematic, and practically negligible. Moreover, these benefits do not generalize to more realistic advertising contexts. Across two Instagram-style ad studies (Studies 2A and 2B), explanatory disclosures fail to improve consumer responses and, in some cases, lead to more negative evaluations with effects that are either statistically equivalent to zero or significantly negative. Across studies, AI aversion emerges as a robust predictor of negative responses, suggesting that disclosure effects are driven more by consumers' prior beliefs than by the specific framing of explanations. The findings suggest caution in adding explanatory disclosures, as default inferences of firm-serving motives are difficult to override.

Keywords Generative AI · AI disclosure · Advertising effectiveness · Beneficiary framing · AI aversion · Null effects

Introduction

Since ChatGPT's public launch in late 2022, generative AI continues transforming marketing practices by promising efficiency and creativity gains. However, its use in advertising is highly controversial. A striking example is Coca-Cola's repeated use of AI-generated Christmas campaigns, which triggered substantial public backlash over two consecutive years, with critics pointing to uncanny visuals, perceived low quality, and concerns that AI was replacing human creativity (Di Placido 2025). Such reactions highlight a broader tension: while firms embrace generative AI as a powerful creative tool, consumers often respond with skepticism or even backlash. At the same time, regulatory pressure is intensifying. Regulators such as the European Union, alongside platforms including TikTok, Meta, and YouTube, increasingly require AI-generated advertising content to be explicitly labeled (Bickert 2024; European Union, 2024; TikTok, 2025; Youtube team, 2024). As a result, transparency is no longer optional but increasingly mandatory, shifting the central managerial question from whether to disclose AI use to how to do so effectively.

Empirical evidence consistently shows that merely disclosing AI involvement reduces ad and brand attitudes, purchase intentions, engagement, and behavioral metrics such as donations to charities or click-through-rates (Arango et al. 2023; Brüns and Meißner 2024; De Cicco et al. 2025; Narayanan 2025; Qiu et al. 2025; To et al. 2025). Thus, regulatory compliance with AI disclosure may come at the cost of marketing effectiveness. While emerging research suggests explanatory disclosures may help reducing such negative responses (Kirk and Givi 2025; Narayanan 2025; Zhang and Hur 2025), much less is known about whether consumers' interpretations of why firms use AI can be actively shaped. Recent work has explicitly called for research examining

✉ Timo Schreiner
timo.schreiner@nim.org

[1] Nuremberg Institute for Market Decisions, Steinstraße 21, 90419 Nuremberg, Germany

[2] D'Amore-McKim School of Business, Northeastern University, 202 Hayden Hall, 360 Huntington Avenue, MA 02115 Boston, USA

Published online: 27 August 2026

whether disclosures providing a reason for AI usage can shape consumer inferences and mitigate disclosure effects (Brüns and Meißner 2024; Carney et al. 2026). To address this question, we propose *beneficiary framing* as a structured explanation approach that explicitly communicates why AI was used and who benefits. Drawing on attribution theory (Kelley 1973; Weiner 1985), we argue that consumers use AI disclosures to infer firms' underlying motives for employing AI. Beneficiary framing seeks to redirect these inferences away from default assumptions of efficiency- or cost-driven motives toward more legitimate or prosocial justifications (Grewal et al. 2025; Lowe et al., 2025; To et al. 2025). However, even if disclosures provide alternative explanations, consumers may not necessarily accept them. Persuasion knowledge research suggests that providing justifications can increase scrutiny and trigger resistance, particularly when disclosures are perceived as attempts to legitimize self-serving firm behavior (Koning and Voorveld 2025; Qiu et al. 2025; Wortel et al. 2024). Whether explanatory disclosures can overcome these persistent default inferences therefore remains an open question. If such default assumptions dominate consumers' interpretations, variations in explanation specificity or beneficiary framing may exert little influence on consumer responses. Accordingly, the present research not only examines whether explanatory disclosures improve consumer responses but also explicitly considers the possibility that they produce informative null effects. Responding to recent calls for more rigorous treatment of null findings in marketing research, we therefore complement conventional hypothesis testing with equivalence testing to distinguish practically negligible from inconclusive effects (Harms and Lakens 2018; Lakens, 2017; Petrescu and Krishen 2025; Sarstedt 2026).

A second factor likely to shape the effectiveness of explanatory disclosures is the extent to which disclosure information is attended to and processed (Boerman et al. 2015; Wojdynski and Evans 2016). Consequently, attributional processing may occur only when consumers actually notice and recognize the disclosure, potentially limiting the effectiveness of explanatory disclosure strategies in realistic advertising environments. To address this, we compare beneficiary-framed AI disclosures in settings that vary in the degree to which disclosure information is likely to be attended to and processed. Specifically, we contrast a controlled, text-based setting (Study 1), in which disclosure information is prominently presented without further visual context, with more realistic Instagram-style advertising environments (Studies 2A and 2B), where attention to disclosures is neither guaranteed nor required.

Accordingly, this research contributes to different streams of literature. First, it contributes to the literature on AI disclosures in advertising by examining whether explanatory disclosures can reshape consumers' inferences regarding firms' motives for using AI and identifying the limits of beneficiary framing as an attributional intervention. Second, it extends prior disclosure research and contributes to the marketing analytics literature by evaluating whether variations in AI disclosure design meaningfully influence advertising effectiveness across controlled and realistic advertising environments. Third, it contributes to the growing literature on informative null effects by demonstrating how equivalence testing can distinguish informative from inconclusive null findings.

Conceptual background and hypotheses

AI disclosure as an attribution cue

Although covert AI-generated stimuli can outperform human-made content (Hartmann et al. 2025; Heitmann et al. 2025), consumer reactions shift markedly once AI involvement is disclosed. Prior research consistently shows that AI disclosure can reduce perceived authenticity, sincerity, and human effort, while increasing skepticism, persuasion knowledge, and perceptions of manipulation (Brüns and Meißner 2024; Kirk and Givi 2025; Koning and Voorveld 2025; To et al. 2025; Wortel et al. 2024). Importantly, these negative effects appear to be driven not by disclosure per se, but by the specific attribution of AI involvement. In contrast to AI disclosure, human authorship disclosure generally does not produce comparable negative effects (Qiu et al. 2025; Schilke and Reimann 2025), suggesting that consumers do not simply react to source transparency, but to what AI signals about the production process and the firm's underlying intentions.

These responses can be understood through the lens of the persuasion knowledge model, which proposes that consumers develop knowledge about marketers' motives and persuasion tactics and use this knowledge to recognize, interpret, evaluate, and cope with persuasion attempts (Friestad and Wright 1994). Disclosure cues can activate conceptual persuasion knowledge by helping consumers recognize the persuasive or commercially motivated nature of a message. Such recognition can subsequently activate more critical attitudinal persuasion knowledge, including skepticism, distrust, perceived bias, and disliking of the persuasion attempt (Boerman et al. 2015; Campbell et al. 2013; Van Reijmersdal et al. 2023). In this process, disclosure recognition can produce a "change of meaning", whereby consumers reinterpret communication that initially appears noncommercial or authentic as a strategically motivated persuasion attempt (Friestad and Wright 1994; Kim et al., 2021; Van Reijmersdal et al. 2023).

One particularly relevant motive-related judgment arising from such processing is perceived manipulative intent. Rather than reflecting only a general negative evaluation, perceived manipulative intent captures an attribution that the advertiser has deliberately employed an inappropriate, unfair, deceptive, or unduly controlling persuasion tactic (Campbell 1995; Kim et al., 2021; Pierre et al., 2026). Prior disclosure research, including recent work on AI disclosures, identifies perceived manipulative intent as a central mechanism through which disclosure cues reduce consumer responses by activating persuasion knowledge and suspicion about firms' strategic motives (Kim et al., 2021; Pierre et al., 2026).

Building on this perspective, attribution theory provides a useful complementary perspective to explain this pattern (Kelley 1973; Weiner 1985). Whereas persuasion knowledge explains why consumers begin to scrutinize a message and question its persuasive intent, attribution theory explains the specific causes, motives, and beneficiaries they infer from the firm's decision to use AI. Prior research shows that individuals routinely interpret others' actions in terms of underlying motives and goals, even when behavior is shaped by situational factors, and that such motive inferences play a central role in shaping subsequent evaluations (Reeder et al. 2004; Verlegh et al. 2013). First evidence in the context of AI usage in food production shows that consumers may infer cost-cutting motives from the deployment of AI, suggesting that firms should avoid emphasizing efficiency gains and instead communicate value-added intentions such as product quality or long-term benefits (Ruan et al. 2026). By extension, in advertising contexts where AI is commonly used to automate tasks and increase efficiency in advertising, often reducing production costs (Grewal et al. 2025; Lowe et al., 2025; Reisenbichler et al. 2022), and is associated with reduced human effort in content creation (To et al. 2025), a minimal disclosure such as "AI-generated" may activate the inference that the firm used AI primarily for its own operational benefit. Such inferences can lower the perceived legitimacy and persuasive effectiveness of the message.

At the same time, disclosures may also generate a competing positive pathway by increasing perceived transparency. Disclosure research shows that recognized disclosures can simultaneously activate conceptual persuasion knowledge and increase perceptions that the message openly communicates its paid nature and sponsor identity (Van Reijmersdal et al. 2023). This distinction may be especially relevant in AI-assisted advertising. More elaborated disclosures that clarify AI authorship levels, specify the role of AI in the creation process, or provide reasons for AI use can attenuate negative responses (Brüns and Meißner 2024; Carney et al. 2026; Kirk and Givi 2025; Zhang and Hur 2025). However, such disclosures vary systematically in the extent to which they provide meaningful interpretive guidance, leaving open when and why they are effective.

Building on this logic, we argue that structured AI disclosures that explain *why* and *for whose benefit* AI was used should be evaluated more favorably than minimal labels that merely state AI involvement without interpretive guidance.

H1 *AI usage disclosures with explanations lead to more favorable consumer responses than minimal AI usage disclosures.*

Explanation specificity

Research on recommender systems shows that specific explanations enhance transparency and trust more than vague statements by reducing "black box" perceptions (Schreiner et al. 2019; Tintarev and Masthoff 2012). Similarly, generic explanations (e.g., "AI was used for efficiency") may be too ambiguous to alter attributional inferences meaningfully, whereas specific explanations clearly articulate rationale and expected outcomes.

Much of the existing literature on AI disclosure in advertising relies on minimal labels such as "AI-generated" or "created by AI," which merely indicate the presence of AI without clarifying its precise role, degree of involvement, or purpose. Even more elaborate disclosures differ substantially in how much information they provide. Some disclosures specify whether AI fully created the content or merely assisted or edited it (Brüns and Meißner 2024; De Cicco et al. 2025; Kirk and Givi 2025), whereas others clarify the effort involved in using AI tools or provide reasons for AI use (Carney et al. 2026; Zhang and Hur 2025). These studies suggest that more informative disclosures can attenuate negative reactions by making the production process easier to interpret.

From an attributional perspective, specificity should matter because it reduces interpretive uncertainty. The more clearly a disclosure communicates why AI was used, the less consumers need to rely on default assumptions about automation, cost-cutting, or reduced human effort. Specific explanations should therefore be more effective than vague ones in redirecting attributions toward more legitimate interpretations of AI use and reducing reliance on default negative inferences.

H2 *More specific explanations for AI use lead to more favorable consumer responses than low-specificity explanations.*

Beneficiary framing

We conceptualize beneficiary framing as a form of emphasis framing (Druckman 2001) that shapes judgments by making certain considerations more salient. In this sense, beneficiary framing moves beyond descriptive transparency by offering a motivational justification of AI use, helping consumers interpret why AI was employed and for whose benefit.

Existing studies have operationalized such explanations using concrete benefit-oriented statements, for example emphasizing improved convenience or content quality for consumers (e.g., "to keep you updated faster"; Ali et al. 2025; "to enhance visual appeal"; Zhang and Hur 2025), firm-related gains (e.g., "saving valuable resources"; Arango et al. 2023; "for cost efficiency"; Zhang and Hur 2025), or broader societal and ethical motives (e.g., "to protect real children's privacy"; Arango et al. 2023; "benefit the welfare and wellbeing of the wider community"; Sands et al. 2025). Zhang and Hur (2025) show that privacy-protection framing significantly improves evaluations compared to cost-efficiency framing. Similarly, Arango et al. (2023) demonstrate that ethical motives mitigate negative effects, whereas instrumental motives (e.g., cost savings) exacerbate them. Complementing these findings, Narayanan (2025) shows that providing a plain reason for AI use (e.g., referring to deforestation) without a clear justification does not necessarily improve evaluations relative to generic AI disclosures, suggesting that the effectiveness of explanations depends on how the underlying rationale is framed. These examples illustrate that AI disclosures already vary systematically in the beneficiaries they highlight, even though prior research has not examined these variations in a unified framework.

This distinction closely aligns with prior research on emphasis framing in marketing, which differentiates between self-benefit and other-benefit frames. Such frames activate distinct motivational processes, including self-interest versus prosocial responses such as empathy and moral evaluation, and have been shown to systematically shape consumer judgments and behavior (White and Peloza 2009; Yang et al. 2015). Prior research comparing self-benefit and other-benefit appeals further suggests that their effectiveness is context-dependent rather than uniform, with both types of framing generating favorable responses depending on situational factors, product characteristics, and consumer traits (Ryoo et al. 2025; White and Peloza 2009; Yucel-Aybat and Hsieh 2021). These different framings are unlikely to be evaluated equally, because they imply different underlying motives for AI adoption. Previous research shows that highlighting a concrete beneficiary helps consumers interpret and evaluate firm actions by linking them to meaningful motives, moral justifications, or mental accounts, thereby shaping perceived appropriateness and response tendencies (Fisher et al. 2023; Moes et al. 2022; Yucel-Aybat and Hsieh 2021).

Research on social distance further suggests that actions benefiting socially proximate targets are evaluated more favorably than those benefiting more distant entities (Schreiner and Baier 2021; Wiebe et al. 2017). Socially proximate beneficiaries are typically processed more concretely and perceived as more relevant, whereas distant beneficiaries are construed more abstractly, which can influence how persuasive such appeals are (Saeed et al. 2024). In AI disclosure contexts, consumer benefits are self-relevant, whereas company benefits may appear self-serving, and societal benefits gain normative legitimacy. Thus, both consumer-benefit and other-benefit framings should outperform firm-benefit framings:

H3a *AI usage explanations emphasizing consumer benefits lead to more favorable consumer responses than those emphasizing company benefits.*

H3b *AI usage explanations emphasizing benefits for others (society, environment) lead to more favorable consumer responses than those emphasizing company benefits.*

Null effects and theoretical uncertainty

A central premise of this research is that the effectiveness of explanatory AI disclosures is theoretically plausible but empirically uncertain. Although explanations may influence attributional inferences, they may equally increase scrutiny or highlight persuasive intent, such that their incremental impact relative to a baseline disclosure may be limited or even negative.

Therefore, the present research explicitly treats null results as theoretically and practically meaningful outcomes, rather than as a lack of evidence. An excessive focus on statistical significance can create the misleading impression that only significant findings are informative, even though statistically significant results may be trivial and null results may be highly informative (Petrescu and Krishen 2025; Rigdon 2026). This perspective is especially relevant given the long-standing marginalization of null findings in marketing and business research, reflected in phenomena such as the file-drawer problem (Hubbard and Armstrong 1992; Landis et al. 2014; Petrescu and Krishen 2025). Recent methodological work therefore recommends focusing on effect estimates, interval estimates, and the remaining uncertainty surrounding those estimates rather than relying on dichotomous significance decisions (Bultez and Herrmann 2025; Sarstedt et al. 2024). Accordingly, the

key question is not simply whether an estimate differs from zero, but whether the available evidence is sufficiently precise to rule out substantively meaningful effects (Sarstedt 2026). This distinction gives rise to two fundamentally different types of null findings: informative nulls, in which estimates are sufficiently precise to rule out effects of practical relevance, and inconclusive nulls, in which the available evidence remains too imprecise to exclude meaningful effects (Ehrlich et al. 2026; Lakens, 2017; Sarstedt 2026).

In the present context, informative null effects would suggest that explanatory AI disclosures do not meaningfully improve consumer responses relative to simpler formats. Such a result would be theoretically relevant and managerially useful, because it would indicate that more elaborate disclosure strategies add complexity without generating meaningful benefits. By contrast, inconclusive null effects would imply that the evidence is not yet precise enough to determine whether meaningful effects exist.

Beyond that, null findings must be interpreted cautiously, as they may also arise from factors such as limited power, weak manipulations, or measurement issues (Favero et al. 2025; Kane 2025). To distinguish between informative and inconclusive null effects, the present research complements conventional significance testing with equivalence testing (Harms and Lakens 2018; Lakens, 2017).

Null effects at the aggregate level do not necessarily imply uniform responses but may reflect heterogeneity in how consumers interpret AI disclosures. Research on cause-related marketing shows that the same prosocial cue can evoke different self- and other-oriented attributions across consumers, leading to heterogeneous effects on choice and evaluation (Arora and Henderson 2007; Schreiner and Baier 2021). Similarly, explanatory AI disclosures may trigger different interpretations depending on whether consumers perceive them as consumer-benefiting, socially beneficial, firm-serving, or manipulative. Aggregate null effects may therefore arise not only because explanations are ineffective, but also because divergent attributional responses offset one another.

AI aversion as boundary condition

One important driver of such divergent interpretations may be individual's AI aversion, defined as the tendency to distrust or avoid algorithmic systems even when they perform as well as or better than humans (Dietvorst et al. 2015). Prior research on AI disclosure suggests that consumers with more negative predispositions toward AI react more critically to disclosed AI use, showing stronger skepticism, less favorable attitudes, and lower behavioral intentions (Qiu et al. 2025; Wortel et al. 2024). Conceptually, AI aversion can be understood as a prior belief structure that shapes how disclosure information is interpreted. Consumers high in AI aversion are more likely to associate AI with inauthenticity, low human effort, impersonality, or manipulative intent. As a result, they should evaluate AI-related disclosures less favorably overall (Qiu et al. 2025). Moreover, because they enter the persuasion situation with stronger negative priors, they may be less receptive to explanatory attempts designed to legitimize AI use. Even when firms provide structured explanations or emphasize favorable beneficiaries, highly AI-averse consumers may ignore these disclosures or interpret them with suspicion.

Accordingly, we expect AI aversion to exert both a direct negative effect on evaluations and a moderating effect on the effectiveness of explanatory AI disclosures.

H4a *AI disclosures are evaluated less favorably by individuals with higher AI aversion.*

H4b *AI aversion negatively moderates the effectiveness of explanatory AI disclosures (vs. minimal disclosure), such that AI use explanations are less effective among consumers with higher AI aversion.*

Attention to AI disclosures

The effectiveness of explanatory AI disclosures may further depend on whether disclosure information is actually noticed and processed. Much of the existing literature relies on controlled or hypothetical exposure scenarios in which disclosure information is prominently presented. However, real-world advertising environments, such as social media, are characterized by constraint attention, fast-paced content consumption and dense information (Boyd 2010).

Evidence suggesting that explanatory disclosures can mitigate negative responses largely stems from settings in which disclosure information is highly salient or even enforced. For instance, some studies highlighting attenuating effects of beneficiary framing or explanatory content present disclosures directly within the stimulus in a visually prominent manner or require participants to attend to them for a minimum duration (Zhang and Hur 2025). Similarly, some studies suggest that varying the degree of AI authorship (compared to labeling content as fully AI-generated) can mitigate negative effects, although these findings often rely on explicit primers or introductory statements that ensure participants process the disclosure before evaluating the stimulus (Brüns and Meißner 2024; De Cicco et al. 2025). Other work employs hypothetical or text-based scenarios in which disclosure information is centrally embedded in the narrative and cannot be easily overlooked (Kirk and Givi 2025).

More naturalistic evidence remains limited. Carney et al. (2026) examine AI disclosure effects in a social media-like environment and find that disclosure reduces engagement. Importantly, only around half of participants correctly recognized the AI disclosure in their experiments, illustrating how easily such cues can be overlooked in realistic social media environments. In exploratory analyses, the negative effects of AI disclosures were substantially stronger among participants who noticed the disclosure, suggesting that consumers' recognition and processing of the disclosure cue may play an important role in determining disclosure effectiveness.

This observation is theoretically important because prior disclosure research suggests that disclosure effects unfold through multiple information-processing stages. Disclosure cues must first attract sufficient visual attention to be recognized (Boerman et al. 2015; Wojdynski and Evans 2016). Once recognized, disclosures can activate persuasion knowledge and trigger inferential processes regarding the persuasive intent underlying the message, ultimately influencing subsequent consumer evaluations (Boerman et al. 2015; Campbell et al. 2013; Eisend et al. 2020; Wojdynski and Evans 2016). Consequently, disclosure effects should depend not merely on the presence of a disclosure cue, but on consumers' actual recognition and processing of that cue. This distinction is particularly relevant in social media environments, where disclosure information competes with visually rich advertising content for limited attentional resources (Breves et al. 2024; Pittman and Haley 2023).

Applying this framework to AI disclosures suggests recognition of the disclosure constitutes a necessary precondition for attributional processing. Once noticed, AI disclosures can function as salient source cues that trigger attributional and inferential processes associated with AI-generated content, including reduced perceived authenticity and human effort, as well as increased skepticism and persuasion knowledge (Brüns and Meißner 2024; Kirk and Givi 2025; Wortel et al. 2024). Consequently, consumers who notice the AI disclosure should respond less favorably than those who do not. Moreover, explanatory disclosures should outperform minimal AI labels only when the disclosure is actually processed.

H5a *Participants who notice the AI disclosure will report less favorable advertising outcomes than participants who do not notice the disclosure.*

H5b *The positive effect of explanatory AI disclosures (vs. minimal disclosure) will be stronger among participants who notice the AI disclosure.*

While hypotheses H1–H4b are examined across all three studies, the hypotheses H5a and H5b concerning disclosure recognition and processing are additionally tested in the more realistic advertising environments of Studies 2A and 2B.

Study 1

Design and procedure

In a within-subjects design (pre-registered: https://aspredicted.org/m29gt8.pdf), we vary disclosure statements along two dimensions: For each beneficiary type, we implement two concrete benefit themes (personalization vs. product quality for consumer benefits; efficiency vs. speed-to-market for company benefits; employee workload reduction vs. environmental impact for other benefit framing), yielding 12 beneficiary-framing statements plus a baseline disclosure ("Created with AI"). Importantly, Study 1 employs a highly controlled, text-based setting in which disclosure statements are presented in isolation, without any accompanying visual or contextual advertising elements. This design ensures that the disclosure itself receives focal attention and allows us to isolate consumers' immediate responses to different disclosure formulations while minimizing the influence of advertisement-specific visual and contextual factors. Similar text- and scenario-based approaches have been used in recent research on AI-generated marketing communications to isolate consumers' responses to AI-disclosure statements before examining more realistic advertising settings (Brüns and Meißner 2024; Kirk and Givi 2025; Sands et al. 2025). This trade-off is addressed in Studies 2A and 2B, which embed the disclosures within complete Instagram advertisements to examine whether the observed effects generalize to more realistic advertising environments.

Between subjects, participants were randomly assigned to one of four advertising contexts (automotive, soft drink, corporate social responsibility campaign, local business). These contexts were selected to capture variation in product involvement and situational framing. The automotive context represents a high-involvement category, whereas the soft drink context reflects a low-involvement, habitual consumption setting. The CSR context introduces a prosocial framing, which has been frequently examined in prior research on AI disclosures in advertising and where firm motives might be evaluated differently compared to product-related settings (Arango et al. 2023; Baek et al. 2024; De Cicco et al. 2025; Ilg and Stadtelmann 2025; Lee et al. 2025; Narayanan 2025). The local business context (framed as "a local bike shop in your area") reflects a setting in which the use of AI in advertising may be perceived as more

acceptable due to limited marketing resources. Because prior research does not provide clear predictions regarding how advertising context interacts with explanatory AI disclosures, advertising context was treated as an exploratory between-subject factor.

Participants first evaluated the baseline disclosure, then all 12 statements in randomized order, each presented individually to capture immediate reactions. The exact statement wording and theoretical grounding are provided in Appendix A.

Participants and measures

Participants (N=339) were recruited via Prolific in the U.S. (M_{age} = 43.6 years; 51.9% male) in March 2026. No participants were excluded, as all preregistered data quality criteria (i.e., attention checks, survey completion, and minimum completion time) were met. The main dependent variable was overall evaluation of the disclosure statement, measured using a single-item scale ("How do you evaluate this statement overall?"; 1=very negative, 7=very positive), capturing participants' holistic judgment of each disclosure statement. To assess the perceived credibility of the disclosure statements, participants rated each statement on a five-item 7-point semantic differential scale adapted from established source credibility measures (MacKenzie and Lutz 1989; Newell and Goldsmith 2001; Ohanian 1990) and recent work on post credibility in social media advertising (Boerman et al. 2012, 2017; Brüns and Meißner 2024). To verify whether the beneficiary framing manipulation was perceived as intended, participants indicated who they believed primarily benefits from the use of AI ("me personally", "the company", "other people/society", "don't know/prefer not to answer"). Individual differences in AI aversion were measured using the four-item Attitudes toward AI Scale (AIAS-4; Grassini 2023) and experience with generative AI was measured using four items assessing familiarity with AI-generated text and images as well as the use of generative AI tools on a five-point scale, similar to recent measures (Koning and Voorveld 2025; Wu et al. 2025). The data structure reflects repeated evaluations of disclosure statements within participants, which is accounted for in the analysis using linear mixed-effects models. Detailed item wording, scale properties, and reliability statistics are reported in Appendix B.

A sensitivity analysis for the key within-subject contrast indicated that the sample (N=339) provided 80% power to detect small effects (d≈0.15), corresponding to about 0.16 scale points on the 7-point measure. To complement null-hypothesis significance testing, we conducted equivalence tests using the two one-sided tests (TOST) procedure. A non-significant difference alone does not establish the absence of a substantively meaningful effect, because the corresponding confidence interval may remain compatible with effects of theoretical or practical relevance. Following recent recommendations, inference should therefore be based on the remaining uncertainty surrounding the estimated effect and its relationship to a predefined region of practical equivalence (ROPE) rather than on statistical significance alone (Anvari and Lakens 2021; Sarstedt 2026). Equivalence testing provides exactly this assessment and allows us to assess whether the data are sufficiently precise to rule out differences exceeding a prespecified smallest effect size of interest (SESOI; Harms and Lakens 2018; Lakens, 2017). We specified equivalence bounds of ±0.20 scale points on the seven-point response scale. This threshold reflects our substantive goal of identifying whether alternative disclosure formulations produce differences large enough to meaningfully distinguish them in consumers' evaluations. Differences smaller than one-fifth of a scale point were considered unlikely to provide a practically relevant basis for preferring one disclosure formulation over another. Given the observed variability of the focal outcome, overall statement evaluation (SDs ranging from 1.60 to 2.14), these bounds correspond to standardized effects of approximately d=0.09 to d=0.13, well below Cohen's (1988) benchmark of d=0.20 for a small effect. Differences whose confidence intervals fell entirely within these bounds were therefore interpreted as statistically equivalent and too small to substantively distinguish the disclosure statements, representing informative null findings. By contrast, results were considered inconclusive when the confidence interval remained compatible with both negligible and substantively meaningful differences.

Results

Explanatory disclosures overall did not outperform the minimal AI label (Δ=0.07, p = .69; see Table 1), providing no support for H1. While seven out of twelve individual statements received numerically higher evaluations than the baseline, these differences were not statistically significant at the aggregate level, suggesting that simply adding explanations does not systematically improve responses. The 90% confidence interval for the contrast fell entirely within the predefined equivalence bounds (90% CI [− 0.02, 0.16]; see Table 2), indicating that explanatory disclosures are statistically equivalent to the baseline.

However, explanation specificity mattered: specific explanations exceeded the baseline (Δ=0.20, p = .07), whereas unspecific explanations did not (Δ = − 0.06, p = .57). Linear mixed-effects analyses further confirm a significant positive effect of specificity (b=0.27, p < .01: Table 1), supporting H2. Consistent with this result, the corresponding contrast between high and low specificity exceeded the

Table 1 Study 1 mixed-effects regression model predicting overall evaluation (main effects model)

Predictor	Fixed effects
AI aversion (centered)	− 0.65***
AI experience (centered)	− 0.10 (p = .16)
Statement specificity: specific (vs. unspecific)	0.27***
Beneficiary: consumer (vs. company)	0.06 (p = .41)
Beneficiary: other (vs. company)	− 0.08 (p = .23)
Ad context: soft drinks (vs. cars)	− 0.25 (p = .15)
Ad context: CSR (vs. cars)	− 0.19 (p = .28)
Ad context: local business (vs. cars)	0.13 (p = .45)

Unstandardized coefficients. Reference categories: ad context = cars; beneficiary = company; information specificity = unspecific

Models include random intercepts for respondents and disclosure statements

*** $p < .001$, ** $p < .01$, * $p < .05$, $^{\dagger}p < .10$

N = 4068 observations from 339 respondents excluding the baseline evaluation. Marginal $R^2 = 0.33$; Conditional $R^2 = 0.69$

Table 2 Summary of hypothesis tests and equivalence results (Study 1)

Hypothesis	Contrast	Δ (Estimate)	90% CI	SESOI	Equivalence conclusion
H1	Explanations vs. baseline	0.07	[− 0.02, 0.16]	±0.20	Equivalent
H2	High vs. low specificity	0.27	[0.17, 0.37]	±0.20	Not equivalent
H3a	Consumer vs. company	0.06	[− 0.05, 0.16]	±0.20	Equivalent
H3b	Other vs. company	− 0.08	[− 0.19, 0.02]	±0.20	Equivalent

predefined equivalence bounds (Δ = 0.27, 90% CI [0.17, 0.37]; Table 2).

Beneficiary framing had no significant effect, providing no support for H3a or H3b (Table 1). These contrasts met the prespecified equivalence criterion. The comparison between consumer- and company-benefit framing yielded a 90% confidence interval of [− 0.05, 0.16], and the comparison between other- and company-benefit framing yielded a confidence interval of [− 0.19, 0.02]. These results indicate that beneficiary framing does not produce practically meaningful differences in this setting.

To better understand how participants interpret explanatory disclosures, we examined participants' attributions of the primary beneficiary of AI use. Overall, respondents predominantly attributed the benefits of AI use to the company (58–96% across statements), suggesting a strong default inference that AI in advertising primarily serves firm interests. Nevertheless, attribution patterns varied systematically by framing. Consumer- and environmental-benefit explanations were most likely to be interpreted as intended, with up to one-third of respondents identifying the focal beneficiary. In contrast, employee-benefit explanations were rarely interpreted as benefiting others and were instead predominantly reattributed to the company.

AI aversion negatively predicted evaluations (b = − 0.65, $p < .01$), supporting H4a. To test H4b, we included an interaction term between AI aversion and disclosure type (explanatory disclosure vs. minimal AI label) in a mixed-effects model. This interaction was not significant (b = − 0.04, $p = .31$), providing no support for H4b. AI experience did not exert a significant effect on evaluations.

Advertising context did not significantly affect evaluations (all p's > 0.10), indicating that the effects of explanation specificity and AI aversion are largely consistent across contexts.

Notably, credibility ratings were consistently lower for all explanatory disclosures compared to the baseline AI label. This likely reflects a conceptual difference between the measures: whereas the baseline disclosure does not advance a concrete claim and is therefore difficult to judge as implausible, explanatory statements introduce specific justifications that invite greater scrutiny.

Discussion

The results provide a nuanced picture of how consumers respond to explanatory AI disclosures. Explanation specificity showed a positive effect in this controlled setting, suggesting that more detailed justifications can improve evaluations. In contrast, beneficiary framing does not systematically influence responses, suggesting that highlighting consumer or societal benefits does not outweigh clearly stated firm-related motives. Importantly, attribution data provide insight into this null effect: although explanations shifted perceived beneficiaries to some extent, consumers predominantly inferred that AI use serves company interests. As a result, beneficiary framing was insufficient to meaningfully alter these default attributions, limiting its impact on overall evaluations. Consistent with this pattern, equivalence testing indicates that these null effects reflect the absence of practically meaningful differences rather than inconclusive results. At the same time, AI aversion remains the dominant predictor of evaluations, underscoring the importance of pre-existing attitudes toward AI. However, these findings are derived from a highly controlled setting in which disclosure information is necessarily processed. This raises the question of whether similar effects would emerge under more realistic exposure conditions, which we address in the subsequent Studies 2A and 2B.

Study 2A

Design and procedure

Study 2 A (pre-registered: https://aspredicted.org/qx2d84.pdf) examines whether the disclosure effects observed in Study 1 generalize to a more realistic advertising context and whether they depend on consumers' attention to the disclosure. To this end, we employed a mixed within–between-subjects design. Participants were randomly assigned one of the three beneficiary framings (consumer, company, or other) and exposed to three Instagram-style advertisements for a car brand. The order of ad presentation was randomized to avoid sequence effects. The stimuli were based on an AI-generated social media post from Hyundai, which served as a reference and was subsequently adapted and re-generated using a generative AI model (Sora) to create standardized experimental stimuli. Hyundai was selected as a globally established, mid-market automobile brand with high familiarity yet relatively neutral brand associations.

The ads were designed to closely resemble real social media content, including visual imagery, caption text, and an AI disclosure embedded within the post. The three ads differed only in the type of disclosure: a minimal disclosure ("Created with AI"), an unspecific explanatory disclosure, and a specific explanatory disclosure adapted from promising explanations identified in Study 1 (see Fig. 1). This within-subjects manipulation allows for direct comparison of disclosure formats under naturalistic viewing conditions in which disclosure cues compete with other elements of the ad for attention.

Participants and measures

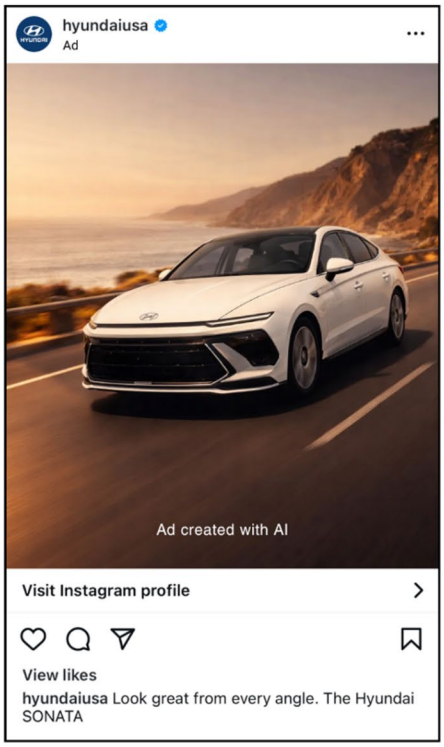
Participants ($N = 359$) were recruited via Prolific in the U.S. (M_{age} = 40.0 years; 47.4% male) in March 2026. Following preregistered exclusion criteria, no participants were excluded from the analysis.

After viewing each advertisement, participants reported attitude toward the ad, attitude toward the brand, and engagement intention using established multi-item scales. Ad attitude was measured with three semantic differential items ("bad/good", "unfavorable/favorable", "unpleasant/pleasant") adopted from MacKenzie and Lutz (1989). Brand attitude was measured with five items ("bad/good", "unfavorable/favorable", "unpleasant/pleasant", "unappealing/appealing", "not likeable/likeable") based on MacKenzie and Lutz (1989) and Spears and Singh (2004). Engagement intention was assessed with four items capturing typical social media behaviors (e.g., liking, commenting, sharing), adapted from Aljarah et al. (2025) and Giakoumaki and Krepapa (2020). Accordingly, hypotheses H1-H4b are tested using these outcome measures in the more realistic advertising context.

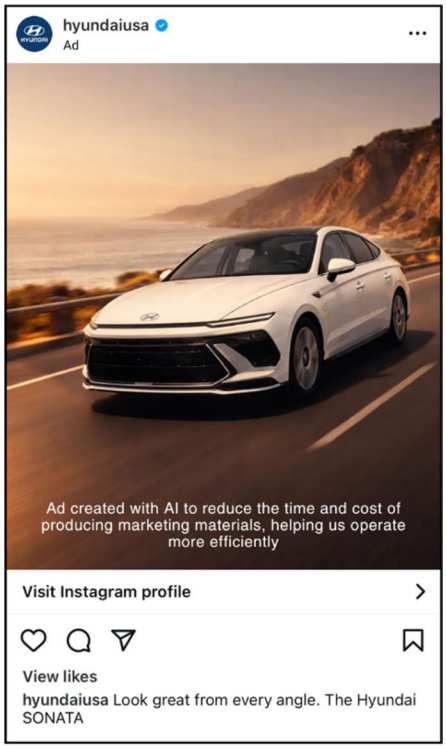
To capture underlying psychological processes, participants also evaluated each ad in terms of perceived effort and perceived manipulative intent, which were treated as exploratory mediators. Perceived effort was measured using four semantic differential items (e.g. "Creating this post seems (not) time-intensive"), adapted from De Kerviler et al. (2022) and To et al. (2025). Perceived manipulative intent was assessed using three items capturing perceived persuasion attempts and fairness of the ad (e.g., "The advertiser tried to manipulate the audience in ways I don't like"), based on established persuasion knowledge measures (Koning and Voorveld 2025; Wortel et al. 2024).

To assess attention to the disclosure, participants indicated after viewing all ads whether each advertisement explicitly mentioned the creator of the content. Based on these responses, a binary indicator of AI disclosure noticing was constructed for each ad, capturing whether participants recognized the presence of an AI disclosure. This variable serves as a key moderator in the analysis. In addition, participants reported the perceived reason for AI use (quality improvement, company efficiency, environmental impact) and the perceived primary beneficiary of AI use as in Study 1. These follow-up questions were only administered for respondents who indicated that they had noticed the AI disclosure in at least one ad. These measures serve as manipulation checks for the disclosure content (see Appendix C for full list of items). Finally, individual differences in AI aversion were measured as in Study 1. Detailed item wording, scale properties, and reliability statistics are reported in Appendix B.

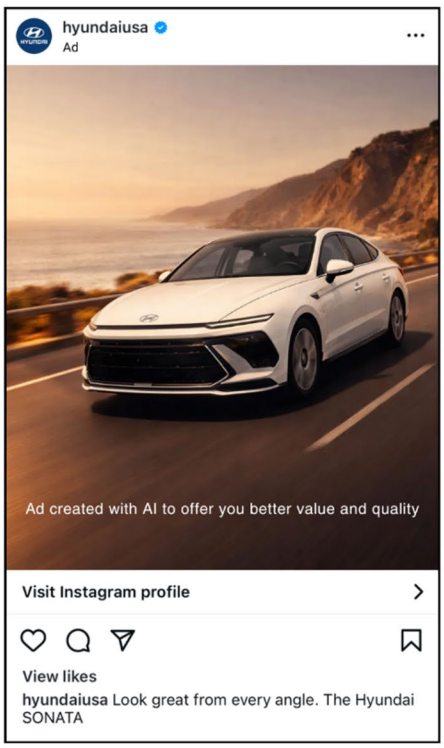
A sensitivity analysis indicated that the within-subject design ($N = 359$) provided 80% power to detect small effects ($d \approx 0.15$), corresponding to raw-score differences of approximately 0.09 to 0.17 scale points across the focal outcomes. The between-subject beneficiary framing contrasts were powered to detect effects of approximately $d \approx 0.36$, corresponding to raw-score differences of approximately 0.63 to 0.70 points. Consequently, these between-subject comparisons were considerably less precise than the within-subject disclosure contrasts. As in Study 1, we complemented null-hypothesis significance testing with equivalence tests, using bounds of ± 0.20 scale points. Because the standardized equivalent of a raw-score bound depends on the variability of the respective outcome and contrast, the ± 0.20-point bounds represented standardized difference-score effects of approximately $d = 0.17$ to 0.33 for the within-subject comparisons and standardized mean differences of approximately $d = 0.10$ to 0.12 for the between-subject comparisons. Confidence intervals falling entirely within these bounds were interpreted as evidence that the corresponding



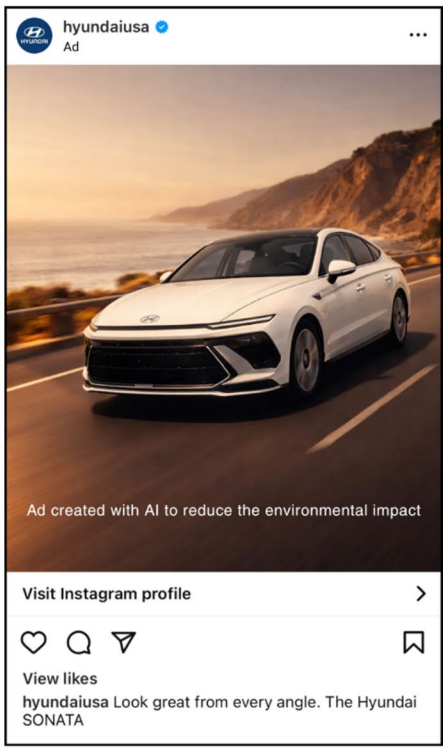
(a) Minimal AI disclosure



(b) Specific explanatory disclosure (company-benefit)



(c) Unspecific explanatory disclosure (consumer-benefit)



(d) Unspecific explanatory disclosure (other-benefit)

Fig. 1 Example stimuli used in Study 2A. Notes. Example stimuli showing (**a**) minimal AI disclosure, (**b**) specific explanatory company-benefit disclosure, (**c**)unspecific explanatory consumer-benefit disclosure, and (**d**) unspecific explanatory other-benefit disclosure. Stimuli were generated with GenAI (Sora) based on a real Hyundai Instagram post

Table 3 Study 2A planned contrasts testing disclosure explanations and beneficiary framing

Contrast	Ad attitude	Brand attitude	Engagement intention
Explanation vs. baseline (H1)	− 0.14*	-0.08 (p = .13)	− 0.07*
High vs. low specificity (H2)	0.09 (p = .15)	0.09† (p = .10)	− 0.01 (p = .80)
Consumer vs. company (H3a)	− 0.13 (p = .54)	0.07 (p = .73)	− 0.15 (p = .44)
Other vs. company (H3b)	0.00 (p = .99)	0.09 (p = .63)	0.01 (p = .97)
Explanation x AI aversion (H4b)	0.01 (p = .89)	-0.01 (p = .67)	− 0.03 (p = .15)
Explanation (vs. baseline) x disclosure noticing (H5b, difference-in-difference)	− 0.24 (p = .12)	-0.24† (p = .08)	− 0.20*

Unstandardized coefficients from linear mixed-effects models. Planned contrasts estimated using estimated marginal means. H5b is tested using a difference-in-differences contrast comparing the effect of explanatory disclosures (vs. baseline) across disclosure noticing conditions. Reference categories: beneficiary = company; disclosure type = minimal disclosure / baseline. Models include random intercepts for respondents

*** $p < .001$, ** $p < .01$, * $p < .05$, $^{\dagger}p < .10$

N = 1077 observations from 359 respondents

Table 4 Summary of hypothesis tests and equivalence results (study 2A)

Hypothesis	Outcome	Δ (Estimate)	90% CI	SESOI	Equivalence conclusion
H1	AA	− 0.14	[− 0.23, − 0.04]	±0.20	Not equivalent
	BA	− 0.08	[− 0.16, 0.01]		Equivalent
	EI	− 0.07	[− 0.12, − 0.02]		Equivalent
H2	AA	0.09	[− 0.01, 0.2019]	±0.20	Not equivalent
	BA	0.09	[0.00, 0.19]		Equivalent
	EI	− 0.01	[− 0.07, 0.05]		Equivalent
H3a	AA	− 0.13	[− 0.48, 0.22]	±0.20	Inconclusive
	BA	0.07	[− 0.26, 0.40]		Inconclusive
	EI	− 0.15	[− 0.48, 0.17]		Inconclusive
H3b	AA	0.00	[− 0.33, 0.34]	±0.20	Inconclusive
	BA	0.09	[− 0.22, 0.41]		Inconclusive
	EI	0.01	[− 0.31, 0.32]		Inconclusive

AA = Attitude toward the ad; BA = Attitude toward the brand; EI = Engagement intention. "Inconclusive" indicates that equivalence could not be established because the confidence interval extended beyond the equivalence bounds

differences were too small to substantively distinguish the disclosure conditions.

Results

Consistent with Study 1, explanatory disclosures did not improve advertising outcomes, again providing no support for H1. Instead, explanatory statements slightly reduced ad attitude and engagement intention, indicating that adding explanations can backfire in realistic settings, while brand attitude was not significantly affected (Table 3). Equivalence tests provide some perspective on these results: for brand attitude and engagement intention, the observed differences were statistically equivalent within the prespecified ± 0.20-point equivalence bounds (Table 4), indicating no practically meaningful effect, whereas the negative effect on ad attitude exceeded the predefined equivalence bounds, suggesting a small but meaningful negative impact.

More specific explanations did not outperform unspecific ones (H2 not supported; Table 3). Equivalence tests further indicate that these differences are generally negligibly small. For brand attitude and engagement intention, the 90% confidence interval fell entirely within the prespecified equivalence bounds (brand attitude: Δ = 0.09, 90% CI [0.00, 0.19]; engagement intention: Δ = − 0.01, 90% CI [− 0.07, 0.05]), while for ad attitude, the confidence interval marginally exceeded the predefined equivalence bounds (Δ = 0.09, 90% CI [− 0.01, 0.2019]), meaning the result is inconclusive with respect to equivalence.

Beneficiary framing again had no significant effect (H3a-b not supported; Table 3). For beneficiary framing (H3a-b), none of the contrasts met the equivalence criterion. Confidence intervals were wide and extended beyond the predefined bounds across all outcomes (e.g., ad attitude: consumer vs. company Δ = − 0.13, 90% CI [− 0.48, 0.22]; other vs. company Δ = 0.00, 90% CI [− 0.33, 0.34]). These results are therefore inconclusive rather than informative null findings, meaning that the data cannot distinguish between negligible beneficiary-framing effects and effects that would exceed our threshold of practical relevance.

As a manipulation check, we examined perceived beneficiary attributions for non-baseline disclosure exposures. In line with the findings from Study 1, respondents again predominantly attributed the use of AI to firm-serving motives. Across the three beneficiary contexts, between 82% and 92% of respondents inferred that the primary beneficiary was the company, regardless of the communicated explanation. Even when alternative beneficiaries were explicitly framed, attributions towards the intended beneficiaries remained rare: in the consumer condition, only 3% of respondents identified the consumer as the primary beneficiary, and in the other-beneficiary condition, 13% did so.

These results suggest that beneficiary framing had even less impact on perceived attributions in the realistic advertising setting compared to Study 1, which helps explain the absence of significant beneficiary-framing effects.

Replicating Study 1, AI aversion strongly predicted negative responses across outcomes (ad attitudes: b = − 0.60, brand attitudes: b = − 0.53, engagement intentions: b = − 0.49, all p's < 0.01), supporting H4a, while no moderation effects emerged (H4b; Table 3).

Consistent with H5a, participants who noticed the AI disclosure reported significantly less favorable evaluations across outcomes (ad attitude: b = − 1.04, brand attitude: b = − 0.86, engagement intentions: b = − 0.43, all p's < 0.01). Participants correctly identified the AI creator in 83% of ad exposures. However, recalling the disclosure content was substantially lower: among ads for which the AI disclosure was noticed, the communicated benefit was correctly identified in only 59% of cases. This discrepancy suggests that while disclosure cues are often noticed, their substantive content is not always fully processed.

Contrary to H5b, explanatory disclosures were not more effective when the disclosure was noticed. A difference-in-differences test indicated no significant moderation for ad attitudes (Δ = − 0.24, p = .12) and only marginal effects for brand attitudes (Δ = − 0.24, p = .08). For engagement intentions, a significant difference emerged (Δ = − 0.20, p = .03). Examining the simple effects underlying these interactions, explanatory disclosures had little to no effect when AI was recalled as the creator of the ad (ad attitudes: Δ = -0.03, p = .62; brand attitudes: Δ = 0.03, p = .54; engagement intentions: Δ = − 0.01, p = .86) but tended to reduce evaluations when the creator was not noticed. Thus, the interaction reflects an attenuation of the negative effect rather than a stronger positive effect of explanatory disclosures, providing no support for H5b.

As an additional robustness check, we used correct identification of the communicated benefit as a stricter indicator of disclosure processing. Among explanatory disclosures, correctly identifying the communicated benefit was associated with significantly less favorable evaluations across all outcome variables (ad attitude: b = − 0.40, p < .001; brand attitude: b = − 0.33, p < .001; engagement intention: b = − 0.14, p = .006), suggesting that more elaborate processing of explanatory content not only fails to mitigate negative responses but may even reinforce them.

Exploratory mediation analyses provide further insight into the underlying mechanisms. Perceived effort did not mediate the effect of explanatory disclosures on any of the outcome variables, as the indirect effects were small and not statistically significant across ad attitude, brand attitude, and engagement intention (all confidence intervals included zero). In contrast, perceived manipulative intent emerged as a consistent mediator. Explanatory disclosures increased perceptions of manipulative intent, which in turn negatively affected advertising outcomes. The indirect effects via manipulative intent were significant for ad attitude (b = − 0.19, 95% CI [− 0.24, − 0.14]), brand attitude (b = − 0.16, 95% CI [− 0.20, − 0.11]), and engagement intention (b = -0.06, 95% CI [− 0.08, − 0.04]).

Compared to the controlled setting of Study 1, the findings from Study 2 A suggest that the positive effects of explanation specificity do not generalize to a realistic social media environment. Moreover, AI disclosure effects appear to depend less on the explanatory content itself than on consumers' broader interpretations of AI-generated advertising. To assess the robustness and generalizability of these findings beyond a single product category, Study 2B replicates the design in a low-involvement advertising context.

Study 2B

Design and procedure

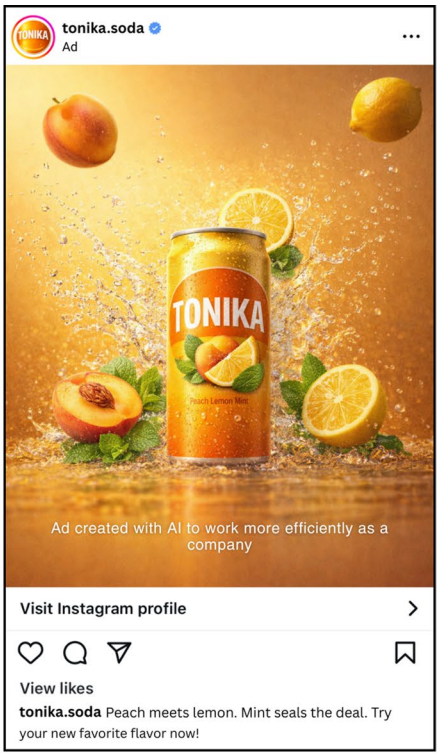
Study 2B (pre-registered: https://aspredicted.org/9fr672.pdf) examines whether the findings from Study 2 A generalize to a lower-involvement product category. To this end, we conducted a direct replication of Study 2 A using Instagram-style advertisements for a fictitious soft drink brand (*Tonika*) instead of an automobile brand (see Fig. 2). The study employed the same mixed within-between-subjects design as Study 2 A. Participants were exposed to three advertisements that varied only in the type of AI disclosure (minimal disclosure, unspecific explanation, specific explanation), while beneficiary framing (consumer, company, other) was manipulated between subjects. The order of ad presentation was randomized.

Participants and measures

Participants (N = 359) were recruited via Prolific in the U.S. (M_{age} = 39.0 years; 43.2% male) in March 2026. In line with the preregistered criteria, all participants were retained in the final sample. Measures were identical to Study 2A. After viewing each advertisement, participants reported attitude toward the ad, attitude toward the brand, and engagement intention. After all ads, participants indicated whether they noticed the AI disclosure, and AI aversion was measured. As in Study 2A, follow-up questions on perceived AI benefits stated in the disclosure and the perceived primary beneficiary were only administered for respondents who indicated that they had noticed the AI disclosure in at least one ad.

A sensitivity analysis indicated that the within-subject design (N = 359) provided 80% power to detect small

(a) Minimal AI disclosure



(b) Unspecific explanatory disclosure (company-benefit)

(c) Specific explanatory disclosure (consumer-benefit)

(d) Specific explanatory disclosure (other-benefit)

Fig. 2 Example stimuli used in Study 2B. Notes. Example stimuli showing (**a**) minimal AI disclosure, (**b**) unspecific explanatory company-benefit disclosure, (**c**) specific explanatory consumer-benefit disclosure, and (**d**) specific explanatory other-benefit disclosure. Stimuli were fully generated using GenAI (Sora)

Table 5 Study 2B planned contrasts testing disclosure explanations and beneficiary framing

Contrast	Ad attitude	Brand attitude	Engagement intention
Explanation vs. baseline (H1)	− 0.27***	-0.19***	− 0.05 (p = .15)
High vs. low specificity (H2)	− 0.09 (p = .18)	-0.01 (p = .91)	0.00 (p = .95)
Consumer vs. company (H3a)	0.02 (p = .93)	0.14 (p = .43)	0.09 (p = .62)
Other vs. company (H3b)	0.06 (p = .76)	0.13 (p = .48)	0.14 (p = .45)
Explanation x AI aversion (H4b)	− 0.09*	-0.04 (p = .21)	− 0.02 (p = .48)
Explanation (vs. baseline) x disclosure noticing (H5b, difference-in-difference)	0.00 (p = .97)	-0.05 (p = .73)	− 0.05 (p = .55)

Unstandardized coefficients from linear mixed-effects models. Planned contrasts estimated using estimated marginal means. H5b is tested using a difference-in-differences contrast comparing the effect of explanatory disclosures (vs. baseline) across disclosure noticing conditions. Reference categories: beneficiary = company; disclosure type = minimal disclosure / baseline. Models include random intercepts for respondents

*** $p < .001$, ** $p < .01$, * $p < .05$, $^{\dagger}p < .10$

N = 1,077 observations from 359 respondents

effects (d ≈ 0.15), corresponding to raw-score differences of approximately 0.10 to 0.17 scale points across the focal outcomes. The between-subject beneficiary framing contrasts were powered to detect effects of approximately d ≈ 0.36, corresponding to raw-score differences of approximately 0.63 to 0.70 points. As in Study 2A, we complemented null-hypothesis significance testing with equivalence tests, using bounds of ± 0.20 scale points. The ± 0.20-point bounds represented standardized difference-score effects of approximately d = 0.17 to 0.29 for the within-subject comparisons and standardized mean differences of approximately d = 0.10 to 0.12) for the between-subject comparisons.

Results

Replicating the findings from the automotive context, explanatory disclosures did not improve advertising outcomes (H1; see Table 5). Instead, they resulted in significantly less favorable evaluations for both ad attitude (Δ = − 0.27, $p < .001$) and brand attitude (Δ = − 0.19, $p < .001$), while engagement intentions were not significantly affected, with the observed difference falling within the predefined equivalence bounds (see Table 6; Δ = − 0.05, 90% CI [− 0.11, 0.01]).

For explanation specificity (H2), the observed differences were small and statistically equivalent within the prespecified ± 0.20-point equivalence bounds across outcomes (ad attitude: Δ = − 0.09, 90% CI [− 0.1996, 0.02]; brand attitude: Δ = − 0.01, 90% CI [− 0.11, 0.01]; engagement intention: Δ = 0.00, 90% CI [− 0.06, 0.07]). This indicates that differences in explanation specificity are negligibly small and can be ruled out as practically meaningful (Table 6).

Table 6 Summary of hypothesis tests and equivalence results (Study 2B)

Hypothesis	Outcome	Δ (Estimate)	90% CI	SESOI	Equivalence conclusion
H1	AA	− 0.27	[− 0.37, − 0.18]	± 0.20	Not equivalent
	BA	− 0.19	[− 0.27, − 0.10]		Not equivalent
	EI	− 0.05	[− 0.11, 0.01]		Equivalent
H2	AA	− 0.09	[− 0.1996, 0.02]	± 0.20	Equivalent
	BA	− 0.01	[− 0.11, 0.01]		Equivalent
	EI	0.00	[− 0.06, 0.07]		Equivalent
H3a	AA	0.02	[− 0.30, 0.33]	± 0.20	Inconclusive
	BA	0.14	[− 0.16, 0.45]		Inconclusive
	EI	0.09	[− 0.22, 0.41]		Inconclusive
H3b	AA	0.06	[− 0.25, 0.37]	± 0.20	Inconclusive
	BA	0.13	[− 0.17, 0.43]		Inconclusive
	EI	0.14	[− 0.17, 0.45]		Inconclusive

AA = Attitude toward the ad; BA = Attitude toward the brand; EI = Engagement intention. "Inconclusive" indicates that equivalence could not be established because the confidence interval extended beyond the equivalence bounds

Beneficiary framing again showed no significant effects (H3a-b not supported). None of the contrasts met the equivalence criterion, with confidence intervals extending beyond the predefined bounds (e.g., ad attitude: consumer vs. company Δ = 0.02, 90% CI [− 0.30, 0.33]; other vs. company Δ = 0.06, 90% CI [− 0.25, 0.37]). Similar patterns emerged across outcomes. Again, these results indicate inconclusive null effects: although point estimates are close to zero, the confidence intervals remain too wide to rule out practically meaningful differences.

As a manipulation check, we examined participants' attributions of the primary beneficiary of AI use for explanatory disclosures. Consistent with Study 1 and Study 2A, respondents predominantly attributed AI use to firm-serving motives. Across conditions, 84% to 92% of participants inferred that the company was the primary beneficiary, irrespective of the communicated explanation. Attribution of the intended alternative beneficiaries remained rare, with only 2% (consumer condition) and 11% (other condition) of respondents identifying the focal beneficiary. These results suggest that beneficiary framing had limited impact on perceived attributions, which again helps explain the absence of significant beneficiary-framing effects (H3a-b).

Consistent with Study 1 and Study 2 A, AI aversion strongly predicted negative responses across all outcome variables (ad attitude: b = − 0.65, brand attitude: b = − 0.66, engagement intention: b = − 0.51, all p's < 0.01), supporting H4a.

In contrast to Study 2A, partial support for H4b was found. A small but significant interaction between explanation and AI aversion emerged for ad attitudes (b = − 0.09, p < .05). As indicated by simple slopes, the negative association between AI aversion and ad attitudes was stronger in the explanation conditions (low specificity: b = − 0.66; high specificity: b = − 0.69) than in the baseline condition (b = − 0.59). This suggests that consumers with higher AI aversion respond more negatively to explanatory disclosures than consumers with lower AI aversion. No significant moderation effects were observed for brand attitudes or engagement intentions.

Consistent with H5a and Study 2A, participants who noticed the AI disclosure (78% of all ad exposures) reported significantly less favorable evaluations across all outcomes (ad attitude: b = -0.84, brand attitude: b = − 0.67, engagement intention: b = − 0.41, all p's < 0.001).

Contrary to H5b, no evidence was found that explanatory disclosures were more effective when noticed. A difference-in-differences test of the explanation effect across disclosure noticing conditions indicated that the effect of explanatory disclosures did not differ between noticing conditions across ad attitudes (Δ = 0.00, p = .97), brand attitudes (Δ = − 0.05, p = .73), and engagement intentions (Δ = -0.05, p = .55).

The robustness check, again using benefit noticing as a stricter measure of disclosure processing, showed a similar pattern to Study 2A. Among ad exposures for which the benefit question was asked, the communicated benefit was correctly identified in approximately 60% of cases. Consistent with the disclosure-noticing analyses, correctly identifying the benefit was associated with less favorable evaluations, although the effects were again smaller in magnitude (ad attitude: b = − 0.47, p < .001; brand attitude: b = − 0.38, p < .001; engagement intention: b = − 0.11, p = .06). This pattern again suggests that more elaborate processing of explanatory content does not mitigate negative responses.

Both perceived effort and perceived manipulative intent emerged as significant mediators. Explanatory disclosures reduced perceptions of effort, which in turn negatively affected advertising outcomes. The indirect effects via perceived effort were significant for ad attitude (b = − 0.08, 95% CI [− 0.12, -0.03]), brand attitude (b = − 0.07, 95% CI [− 0.11, − 0.03]), and engagement intention (b = − 0.05, 95% CI [− 0.07, − 0.02]). Consistent with Study 2 A, perceived manipulative intent remained the dominant mechanism. Explanatory disclosures increased perceptions of manipulative intent, which in turn negatively affected advertising outcomes. The indirect effects via manipulative intent were substantial and significant for ad attitude (b = − 0.30, 95% CI [− 0.36, − 0.24]), brand attitude (b = − 0.26, 95% CI [− 0.32, − 0.21]), and engagement intention (b = − 0.13, 95% CI [− 0.16, − 0.10]).

General discussion

Across the three studies, the results provide a differentiated picture across hypotheses.

For H1, explanatory disclosures did not systematically improve consumer responses. In Study 1, the aggregate effect of explanations relative to a minimal AI label was small and statistically equivalent within the prespecified ± 0.20-point equivalence bounds, indicating the absence of practically meaningful differences. While in the more realistic settings of Studies 2A and 2B, explanations also did not improve evaluations, they in some cases even led to significantly more negative responses compared to the minimal baseline. Overall, across all three studies, the findings suggest that simply providing any explanation for why AI is used is not sufficient to improve consumer responses.

A more nuanced pattern emerges for H2. In the controlled setting of Study 1, more specific explanations outperformed both vague explanations and the minimal baseline, suggesting that additional detail can improve evaluations. However, this effect did not generalize to more realistic contexts. In Studies 2A and 2B, explanation specificity had no meaningful impact, with effects that did not exceed the prespecified SESOI, indicating that more detailed explanations do not improve consumer responses under typical exposure conditions.

For H3, beneficiary framing produced no significant effects across all three studies. However, the interpretation of these null effects differs by context. In Study 1, beneficiary framing effects were statistically equivalent within the prespecified ± 0.20-point equivalence bounds, indicating the absence of meaningful differences under controlled

conditions. In contrast, the findings from Studies 2A and 2B remain inconclusive because the between-subject estimates were not sufficiently precise to rule out small but potentially meaningful effects. Overall, these findings provide no evidence that beneficiary framing robustly improves consumer responses. This pattern is further supported by respondents' attribution of the perceived beneficiaries of AI use. In Study 1, although participants overwhelmingly inferred that AI primarily serves firm interests (58–96% across statements), attribution patterns still varied systematically with framing. Consumer- and environmental-benefit explanations were relatively more likely to be interpreted as intended, with up to one-third of respondents identifying the focal beneficiary. In contrast, in the more realistic settings of Studies 2A and 2B, this variation largely disappeared. Across conditions, between 82% and 93% (Study 2A) and 84% and 92% (Study 2B) of respondents attributed the primary benefit to the company, regardless of the communicated explanation. Correct attribution of alternative beneficiaries was rare, with at most 13% of respondents identifying the intended beneficiary. This suggests that in realistic advertising environments, consumers rely strongly on a default inference that AI use serves firm interests, which is difficult to override through explanatory framing.

Another key insight across Studies 2A and 2B is that AI disclosures operate under conditions of limited attention and competing cues. Although most participants noticed the presence of an AI disclosure, substantially fewer correctly recalled the communicated benefit. Importantly, noticing the disclosure was associated with more negative evaluations (in line with H5a), whereas there was no consistent evidence that processing the explanatory content improved consumer responses (providing no support for H5b). This interpretation is further supported by the finding that participants who correctly identified the communicated benefit did not respond more favorably. Instead, more accurate recall of the explanation was associated with more negative evaluations. Because disclosure recognition was assessed after stimulus exposure rather than experimentally manipulated, these findings should be interpreted cautiously and should not be taken as causal evidence that disclosure recognition itself caused the observed effects. Despite this methodological limitation, these findings suggest that explanatory AI disclosures lose much of their potential once they are embedded in realistic advertising stimuli. In such environments, the AI cue itself appears to function as a dominant signal that outweighs the intended persuasive value of the explanation. This interpretation is consistent with recent evidence by Berner and Dickel (2026), who found that AI labels increased reactance and reduced credibility despite showing no measurable effect on visual attention in an eye-tracking experiment, suggesting that AI labels may primarily function as salient signals rather than as detailed informational messages. Hence, explanatory disclosures may not only fail to reduce skepticism but, once processed more deeply, may even amplify scrutiny of firms' motives by activating persuasion knowledge and attributional reasoning.

Across studies and in line with H4a, AI aversion emerged as a robust and strong predictor of negative responses. In contrast, there was limited and inconsistent support for H4b: no moderation effects were observed in Studies 1 and 2A, and only a small interaction emerged in Study 2B for attitudes toward the ad. This pattern suggests that disclosure effects are shaped less by the precise wording of explanations than by consumers' preexisting beliefs about AI. In addition, manipulative intent consistently emerged as a key mechanism underlying negative responses in more realistic settings, and perceived effort played a complementary role in the soft-drink replication. Together with the disclosure-processing findings, this pattern supports our theoretical argument that explanatory AI disclosures can increase consumers' scrutiny of firms' motives, thereby offsetting the intended benefits of greater transparency.

Theoretical contributions

The present findings highlight an important boundary condition for prior research on AI disclosures. Much of the existing literature relies on controlled settings in which disclosure information is prominently presented. Under such conditions, explanatory framing can shape attributional inferences and, in some cases, mitigate negative responses (Arango et al. 2023; Zhang and Hur 2025). By embedding disclosure cues in more naturalistic social media advertisements, the present research responds to recent calls for marketing scholarship to remain closer to the "real world" of marketing and marketplace decision making (Van Heerde et al. 2021). In doing so, the results demonstrate that effects observed under controlled exposure conditions do not necessarily generalize to realistic advertising environments. This pattern helps explain why explanation specificity produced some positive effects in the controlled setting of Study 1 but failed to improve consumer responses once disclosures were embedded in realistic social media advertisements in Studies 2A and 2B.

A useful way to interpret these findings is to conceptualize AI disclosure as a heuristic source cue whose effects depend on the processing environment. While prior research shows that explanatory framing can shape attributional inferences under controlled conditions, it can also activate unfavorable inferences such as reduced authenticity, lower human effort, diminished credibility, and increased persuasion resistance (Brüns and Meißner 2024; Bui et al. 2024; Kirk and Givi

2025; Koning and Voorveld 2025; To et al. 2025; Wortel et al. 2024). The present findings suggest that in realistic settings these latter processes tend to dominate. Across all three studies, consumers primarily attributed AI use to firm-serving motives regardless of the communicated explanation, indicating that brief explanatory cues are unlikely to override these default inferences. Moreover, in Study 1, explanatory disclosures were perceived as less credible than minimal labels, suggesting that introducing specific justifications invites greater scrutiny than simply disclosing AI use. As a result, explanatory cues may generate competing processes: on the one hand providing additional interpretive information, but on the other hand reinforcing skepticism and resistance. This interpretation is also supported by the mediation results in the more realistic Instagram-style settings (Studies 2A and 2B), where perceived manipulative intent emerged as the primary driver of negative consumer responses. Together with prior evidence that AI-related cues can simultaneously increase positive perceptions (e.g., innovation or appropriateness) while also activating persuasion knowledge, skepticism, and reduced trust (Chen et al. 2025; Koning and Voorveld 2025; Wortel et al. 2024), these findings suggest that AI disclosures may produce attenuated or null overall effects because competing pathways operate simultaneously.

By combining traditional hypothesis testing with equivalence testing, this research distinguishes affirmative evidence of negligible differences from statistically inconclusive null findings. Several contrasts were statistically equivalent to negligible effects, some explanatory disclosures even produced negative effects, whereas the less precise beneficiary-framing comparisons remained inconclusive. More broadly, these findings illustrate how equivalence testing can strengthen theory development by distinguishing the absence of meaningful effects from the absence of sufficient evidence, thereby responding to recent calls for more informative interpretations of null findings in marketing research (Lakens, 2017; Sarstedt 2026).

At the same time, these findings should not be interpreted as evidence that explanatory framing can never matter. Much of the prior evidence suggesting that certain explanations may attenuate disclosure penalties stems from contexts that differ substantially from the present studies. In particular, earlier work has often focused on prosocial or CSR-related settings, such as charity campaigns, environmental communication, or socially responsible brand content (Arango et al. 2023; De Cicco et al. 2025; Narayanan 2025), where specific motives may appear more legitimate or morally congruent with the focal stimulus. Evidence from these contexts suggests that explanatory effectiveness depends strongly on the congruency between the communicated rationale and the advertising context. In such settings, highly context-specific explanations may buffer negative reactions more effectively than in ordinary product advertising. While we do not find meaningful differences between advertising contexts in Study 1, the explanatory AI usage statements were not designed to be context-specific, which may have attenuated potential context-dependent effects. Overall, this raises the possibility that the success of explanatory disclosures depends not only on the framing itself, but also on the congruency between the disclosure text, the underlying AI rationale, and the broader advertising context. In product-related settings such as cars and soft drinks, short explanatory disclosures may be less able to shift inferences about firm self-interest, authenticity, or manipulation.

Finally, in line with several recent studies, AI aversion emerged as a strong and consistent predictor of evaluations of AI disclosure statements across all studies. More broadly, prior research shows that consumers' pre-existing beliefs about AI play a central role in shaping disclosure effects, with negative beliefs such as AI aversion or anxiety amplifying unfavorable responses (Lee et al. 2025; Qiu et al. 2025; Wortel et al. 2024). Related work suggests that these beliefs are rooted in broader perceptions of AI, including its perceived opacity, lack of emotion, rigidity, autonomy, and non-human nature (De Freitas et al. 2023), and in lay beliefs about what AI is competent to do across utilitarian versus hedonic domains (Longoni and Cian 2022).

Managerial implications

Recent work has proposed that firms should enhance AI disclosures by providing richer explanations, highlighting AI's intended benefits, or framing the motives for AI use, and has even introduced a managerial decision framework advocating increasingly informative AI disclosure strategies (Le et al. 2026). Similarly, practitioner guidance increasingly recommends that organizations explain why AI was used, for example by highlighting its intended benefits to contextualize the message and foster trust (Bahr 2024; National Artificial Intelligence Centre 2025; Villegas Bravo 2024; Walther 2026). Current platform-level guidance also assumes that AI disclosures themselves have little impact on user responses. For example, TikTok (2025) states in its community guidelines that labeling AI-generated content does not affect user engagement. Our findings qualify both these academic recommendations and current practitioner guidance. Overall, neither explanation specificity nor beneficiary framing consistently improved consumer responses, suggesting that explanatory disclosure design alone may not meaningfully mitigate disclosure penalties.

From a managerial perspective, the findings suggest that transparency through AI disclosure does not equate to

persuasion effectiveness. While regulatory requirements may necessitate disclosure, firms should not expect that simply adding explanatory or beneficiary-framed justifications will offset potential negative reactions in social media advertising. Instead, practitioners should be aware that AI disclosures may trigger negative perceptions regardless of explanation, that explanatory content is not guaranteed to be processed by consumers, and that efforts to justify AI use may even backfire by increasing perceived manipulative intent.

Accordingly, rather than investing primarily in optimizing disclosure wording, organizations should shift their attention from how AI use is communicated to whether AI should be used in a given advertising context in the first place. The managerial question should not be *"How can we better justify AI use?"* but rather *"Does AI genuinely make this advertisement better?"* Organizations should focus on designing great advertisements and employ AI only when it adds meaningful value to the final advertisement, rather than relying on disclosure strategies to legitimize AI use after the fact. This recommendation becomes particularly important as AI disclosure increasingly shifts from a voluntary communication choice to a legal requirement. This distinction is particularly relevant in light of the EU AI Act, which introduces transparency obligations under Article 50 from August 2026 (EU Artificial Intelligence Act, 2024). Where disclosure is legally required, organizations may often be better served by relying on clear, standardized, and compliant disclosures rather than attempting to justify AI use through increasingly elaborate explanatory statements.

Finally, managers should be careful not to overgeneralize from isolated positive findings in prior research. Explanations that may appear legitimate in prosocial or CSR-related communication do not necessarily transfer to ordinary product advertising. The effectiveness of disclosure strategies may depend heavily on contextual congruency between the explanation, the focal ad, and the broader brand setting. For firms, this means that disclosure design should not be treated as a one-size-fits-all communication task.

Limitations and future research

Several limitations provide avenues for future research. First, while the studies capture more realistic social media environments than much of the prior literature, they remain experimental in nature. Field experiments could further validate the observed effects under real-world conditions. Moreover, disclosure recognition was assessed after stimulus exposure rather than experimentally manipulated. While disclosure recognition represents a theoretically important process through which AI disclosures are expected to influence consumer responses, it may itself have been shaped by other factors such as individual differences in general attention to the stimuli. Therefore, future research could experimentally manipulate disclosure salience or placement or employ eye tracking to directly assess visual attention to AI disclosures.

Second, future research could explore alternative disclosure formats that increase attention and comprehension, such as more prominent visual cues, interactive disclosure elements, multimodal disclosures combining icons with explanatory text, or the AI disclosure icons recently recommended by the European Commission (2026*)* to support compliance with the EU AI Act. Future work should also more systematically examine congruency between the provided explanations and the context by comparing product-related and prosocial settings and by testing whether certain AI rationales are effective only when they align closely with the advertised stimulus and category.

Third, while equivalence testing allows for distinguishing between informative and inconclusive null effects, the results for beneficiary framing in the realistic settings remain inconclusive due to limited precision. Future research should use larger between-subject samples or more statistically efficient designs to obtain more precise estimates of beneficiary-framing effects. Narrower confidence intervals would reduce the remaining uncertainty surrounding these effects and allow researchers to determine more conclusively whether beneficiary framing is genuinely negligible or whether small but practically meaningful effects exist (Sarstedt 2026). Such studies are needed to determine whether these effects are genuinely negligible or whether small but potentially meaningful effects remain undetected.

Next, the present research focuses on AI aversion as a key individual difference but does not account for broader sources of heterogeneity in consumer responses. While AI aversion consistently emerged as a strong predictor of evaluations, little is known about how disclosure effects vary across other individual and contextual factors, such as cultural background, or personality traits. Future research should therefore adopt a more comprehensive perspective on consumer heterogeneity to better understand for whom and under which conditions AI disclosures are more or less effective.

Finally, additional work is needed to examine further potential boundary conditions. For example, responses may vary across product categories, brand positioning, or other contextual characteristics.

Overall, the findings suggest that the central managerial challenge is not how firms disclose AI use, but whether AI genuinely adds value to the advertisement in the first place. Explanatory disclosure strategies appear to have limited ability to overcome consumers' default assumptions about firms' motives for using AI, particularly under realistic advertising conditions.

Appendix

Appendix A: overview of beneficiary framings included in study 1

Beneficiary	Information specificity	Statement	Benefit	Rationale in existing literature
Minimal / baseline	None	"Ad created with AI "	None	
Consumer self-benefit	Specific	"Ad created with AI to tailor this ad more closely to what matters to you, helping us show the product in a way that fits your everyday needs."	Personal relevance / personalization	Moderate personalization increases perceived relevance and utility (Sun et al. 2025). Beneficiary framing increases perceived fit through mental accounting (Fisher et al. 2023).
	Unspecific	"Ad created with AI to tailor this ad to your needs"		
	Specific	"Ad created with AI to offer you better value: less time spent on expensive ad production means more resources can go into product quality."	Higher consumer value / quality focus	Consumer self-benefit framings with financial impact (e.g., price discounts, free shipping) provide high utility to consumers (Schreiner and Baier 2021). Framing savings around a beneficiary increases perceived value (Fisher et al. 2023).
	Unspecific	"Ad created with AI to offer you better value and quality"		
Company self-benefit	Specific	"Ad created with AI to reduce the time and cost of producing marketing materials, helping us operate more efficiently."	Efficiency	GenAI enhances cost efficiency in marketing production tasks (Grewal et al. 2025; Lowe et al., 2025). Automation of routine content generation improves firm productivity (Grewal et al. 2025).
	Unspecific	"Ad created with AI to work more efficiently as a company"		
	Specific	"Ad created with AI to simplify internal workflows and speed up campaign production, allowing us to bring new messages to market faster."	Pace	GenAI increases the speed of marketing content production and deployment (Grewal et al. 2025). Generative systems enable rapid, large-scale content production (Vyas and Vishwakarma 2026).
	Unspecific	"Ad created with AI to speed up how we create our campaigns"		

Beneficiary	Information specificity	Statement	Benefit	Rationale in existing literature
Other benefit	Specific	"Ad created with AI to automate repetitive layout and editing steps, giving our creative team more time to focus on the core idea behind the campaign."	Less routine work	Generative AI reduces routine workload, freeing resources for creative tasks (Hou et al., 2025). AI supports divergent thinking and idea generation (Rogge et al. 2025). Employee-benefit framings reflect socially close beneficiaries (Schreiner and Baier 2021).
	Unspecific	"Ad created with AI to free our creative team from repetitive tasks"		
	Specific	"Ad created with AI to avoid resource-intensive photo and video shoots – helping reduce emissions and waste from our ad production."	Lower environmental footprint	Societal-level benefits require concrete, believable claims (Wiebe et al. 2017).
	Unspecific	"Ad created with AI to reduce the environmental impact"		

Appendix B: constructs and item measures

Construct	Item	Scale	Cronbach's α
Overall evaluation (only study 1)	"How do you evaluate this statement overall?"	7-point scale (1 = very negative, 7 = very positive)	NA
Statement credibility (only study 1) (Boerman et al. 2012; Brüns and Meißner 2024; MacKenzie and Lutz 1989; Newell and Goldsmith 2001; Ohanian 1990)	"How do you evaluate this statement in terms of the following aspects?"	7-point semantic differential	Study 1: 0.99
	Dishonest–Honest		
	Unbelievable–Believable		
	Not trustworthy–Trustworthy		
	Not convincing–Convincing		
	Implausible–Plausible		
AI aversion (REVERSE CODED) (Grassini 2023; Qiu et al. 2025; Wortel et al. 2024)	"To what extent do you agree with the following statements about artificial intelligence (AI)?"	7-point Likert scale (1 = strongly disagree 7 = strongly agree)	Study 1: 0.93 Study 2 A: 0.93 Study 2B: 0.93
	I believe that AI will improve my life.		
	I believe that AI will improve my work.		
	I think I will use AI technology in the future.		
	I think AI technology is positive for humanity.		
Experience with generative AI (only study 1) (Koning and Voorveld 2025; Wu et al. 2025)	"How much experience do you personally have with generative AI in the following areas?"	5-point scale (1 = Not at all experienced 5 = Extremely experienced)	Study 1: 0.86
	Reading AI-generated texts		
	Viewing AI-generated images		
	Using tools to generate text with AI		
	Using tools to generate images with AI		
Attitude toward the ad (only Studies 2A and 2B) (MacKenzie and Lutz 1989)	"What is your overall impression of the ad?"	7-point semantic differential	Study 2 A: 0.99 Study 2B: 0.98
	Bad–Good		
	Unfavorable – Favorable		
	Unpleasant–Pleasant		

Construct	Item	Scale	Cronbach's α
Attitude toward the brand (only Studies 2A and 2B) (MacKenzie and Lutz 1989; Spears and Singh 2004)	"What is your overall impression of the advertised brand?" Bad–Good Unfavorable–Favorable Unpleasant–Pleasant Unappealing–Appealing Not likeable–Likeable	7-point semantic differential	Study 2 A: 0.99 Study 2B: 0.99
Engagement intention (only Studies 2A and 2B) (Aljarah et al. 2025; Giakoumaki and Krepapa 2020)	"Please indicate how much you agree with the following statements about your possible interaction with this post." I would like to see more content from this brand. I would press the "like" button if I saw this post on my social media feed. I would comment on this post on social media. I would share this post on social media with a friend.	7-point Likert scale (1 = strongly disagree 7 = strongly agree)	Study 2 A: 0.93 Study 2B: 0.92
Perceived effort (only Studies 2A and 2B) (De Kerviler et al. 2022; To et al. 2025)	"Please indicate how you perceive the effort that went into creating the post you just saw." Creating this post seems not time-intensive – Creating this post seems very time-intensive Producing this post seems low effort – Producing this post seems high effort This post looks quickly made – This post looks carefully made Overall, this post seems cheaply produced – Overall, this post seems professionally produced	7-point semantic differential	Study 2 A: 0.95 Study 2B: 0.95
Perceived manipulative intent (only Studies 2A and 2B) (Campbell 1995; Koning and Voorveld 2025; Wortel et al. 2024)	"Please indicate how much you agree with the following statements." The advertiser tried to manipulate the audience in ways I don't like. (REVERSE CODED) This post was fair in what was said and shown. The way this ad tries to persuade people seems acceptable to me.	7-point Likert scale (1 = strongly disagree 7 = strongly agree)	Study 2 A: 0.75 Study 2B: 0.80

Items were adopted or adapted from prior literature as indicated by the cited sources. Adaptations include minor wording changes to fit the context of AI-generated advertising and social media environments. Experience with generative AI was adapted and extended to capture both consumption and creation of AI-generated content. The perceived effort scale was adapted and extended from prior research to better capture perceptions of production quality in AI-generated social media content.

Appendix C: items of the manipulation checks used in studies 2A and 2B

Manipulation check	Items	
Notice of AI usage (only Studies 2A and 2B)	Did any of the advertisements you just saw mention who created the content?	Yes, explicitly stating it was created by a human Yes, explicitly stating it was created by AI No, no creator information was mentioned Don't know / prefer not to answer
Stated benefit of AI usage (only Studies 2A and 2B)	According to the advertisement, why was AI used to create this ad?	To offer better quality and value To improve efficiency within the company To reduce the environmental impact Didn't say Don't know / prefer not to answer
Perceived beneficiary	Who do you personally think benefits most from the use of AI in this advertisement?	Me personally Other people / society The company Don't know / prefer not to answer

The items "According to the advertisement, why was AI used to create this ad?" and "Who do you personally think benefits most from the use of AI in this advertisement?" were only shown to participants who indicated that at least one advertisement explicitly stated it was created by AI.

Participants who did not report noticing AI disclosure skipped this question

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Ali, L., F. Ali, M. J. Abdalla, and S. Alotaibi. 2025. Beyond the hype: Evaluating the impact of generative AI on brand authenticity, image, and consumer behavior in the restaurant industry. *International Journal of Hospitality Management* 131:1–12. https://doi.org/10.1016/j.ijhm.2025.104318

Aljarah, A., B. Ibrahim, and M. López. 2025. In AI, we do not trust! the nexus between awareness of falsity in AI-generated CSR ads and online brand engagement. *Internet Research* 35(3):1406–1426. https://doi.org/10.1108/INTR-12-2023-1156

Anvari, F., and D. Lakens. 2021. Using anchor-based methods to determine the smallest effect size of interest. *Journal of Experimental Social Psychology* 96:104159. https://doi.org/10.1016/j.jesp.2021.104159

Arango, L., S. P. Singaraju, and O. Niininen. 2023. Consumer responses to AI-generated charitable giving Ads. *Journal of Advertising* 52(4):486–503. https://doi.org/10.1080/00913367.2023.2183285

Arora, N., and T. Henderson. 2007. Embedded premium promotion: why it works and how to make it more effective. *Marketing Science* 26(4):514–531. https://doi.org/10.1287/mksc.1060.0247

Baek, T. H., J. Kim, and J. H. Kim. 2024. Effect of disclosing AI-generated content on prosocial advertising evaluation. *International Journal of Advertising* 45(1):171–192. https://doi.org/10.1080/02650487.2024.2401319

Bahr, C. 2024. Guide to AI disclaimers: how to create one and why. Usercentrics. URL https://usercentrics.com/guides/website-disclaimers/ai-disclaimer/ (accessed 7.29.26).

Berner, M., and P. Dickel. 2026. It's just a tiny label… or not? How AI labels can backfire in communication. *Journal of Marketing Communications*. https://doi.org/10.1080/13527266.2026.2697265

Bickert, M. 2024. Our Approach to labeling AI-generated content and manipulated media. Meta Newsroom. URL https://about.fb.com/news/2024/04/metas-approach-to-labeling-ai-generated-content-and-manipulated-media/ (accessed 7.29.26).

Boerman, S. C., E. A. Van Reijmersdal, and P. C. Neijens. 2012. Sponsorship disclosure: effects of duration on persuasion knowledge and brand responses: sponsorship disclosure. *Journal of Communication* 62(6):1047–1064. https://doi.org/10.1111/j.1460-2466.2012.01677.x

Boerman, S. C., E. A. Van Reijmersdal, and P. C. Neijens. 2015. Using eye tracking to understand the effects of brand placement disclosure types in television programs. *Journal of Advertising* 44(3):196–207. https://doi.org/10.1080/00913367.2014.967423

Boerman, S. C., L. M. Willemsen, and E. P. Van Der Aa. 2017. This post is sponsored effects of sponsorship disclosure on persuasion knowledge and electronic word of mouth in the context of facebook. *Journal of Interactive Marketing* 38(1):82–92. https://doi.org/10.1016/j.intmar.2016.12.002

Boyd, D. 2010. Streams of content, limited attention: the flow of information through social media. *EDUCAUSE Review* 45(5):26–36.

Breves, P. L., N. Liebers, and Z. M. C. van Berlo. 2024. Followers' cognitive elaboration of sponsored influencer content: the significance of argument quality. *Journal of Interactive Advertising* 24(3):203–214. https://doi.org/10.1080/15252019.2024.2388644

Brüns, J. D., and M. Meißner. 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* 79:103790. https://doi.org/10.1016/j.jretconser.2024.103790

Bui, H. T., V. Filimonau, and H. Sezerel. 2024. AI-thenticity: Exploring the effect of perceived authenticity of AI-generated visual content on tourist patronage intentions. *Journal of Destination Marketing & Management* 34:100956. https://doi.org/10.1016/j.jdmm.2024.100956

Bultez, A., and J.-L. Herrmann. 2025. Value added to marketing research diagnoses by add-ons to p-values. *Journal of Marketing Analytics* 13(2):445–466. https://doi.org/10.1057/s41270-024-00351-w

Campbell, M. C. 1995. When Attention-getting advertising tactics elicit consumer inferences of manipulative intent: the importance of balancing benefits and investments. *Journal of Consumer Psychology* 4(3):225–254. https://doi.org/10.1207/s15327663jcp0403_02

Campbell, M. C., G. S. Mohr, and P. W. J. Verlegh. 2013. Can disclosures lead consumers to resist covert persuasion? The important roles of disclosure timing and type of response. *Journal of Consumer Psychology* 23(4):483–495. https://doi.org/10.1016/j.jcps.2012.10.012

Carney, S., I. Riveros, and S. M. Tully. 2026. Made with AI: consumer engagement with social media containing AI disclosures. *Journal of Consumer Research: ucag013*. https://doi.org/10.1093/jcr/ucag013

Chen, C., D. Zhang, L. Zhu, and F. Zhang. 2025. Exploring consumer responses to AI-generated visual marketing content: a dual-process theory perspective. *International Journal of Consumer Studies* 49(5):e70121. https://doi.org/10.1111/ijcs.70121

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. New York: Routledge. https://doi.org/10.4324/9780203771587

De Cicco, R., B. Francioni, I. Curina, and M. Cioppi. 2025. AI, human or a blend? How the educational content creator influences consumer engagement and brand-related outcomes. *Journal of Services Marketing* 39(10):52–70. https://doi.org/10.1108/JSM-10-2024-0539

De Freitas, J., S. Agarwal, B. Schmitt, and N. Haslam. 2023. Psychological factors underlying attitudes toward AI tools. *Nature Human Behaviour* 7(11):1845–1854. https://doi.org/10.1038/s41562-023-01734-2

De Kerviler, G., N. Heuvinck, and E. Gentina. 2022. Make an effort and show me the love! effects of indexical and iconic authenticity on perceived brand ethicality. *Journal of Business Ethics* 179(1):89–110. https://doi.org/10.1007/s10551-021-04779-3

Di Placido, D. 2025. Coca-Cola sparks backlash with AI-Generated Christmas Ad, Again. Forbes. URL https://www.forbes.com/sites/danidiplacido/2025/11/04/coca-cola-sparks-backlash-with-ai-generated-christmas-ad-again/

Dietvorst, B. J., J. P. Simmons, and C. Massey. 2015. Algorithm aversion: people erroneously avoid algorithms after seeing them err.

Journal of Experimental Psychology: General 144(1):114–126. https://doi.org/10.1037/xge0000033

Druckman, J. N. 2001. The implications of framing effects for citizen competence. *Political Behavior* 23(3):225–256. https://doi.org/10.1023/A:1015006907312

Ehrlich, J. S., C. Kreienbaum, and C. M. Ringle. 2026. Looking at the bigger picture: on the relevance of null effects of model differences across time. *Journal of Marketing Analytics*. https://doi.org/10.1057/s41270-026-00497-9

Eisend, M., E. A. Van Reijmersdal, S. C. Boerman, and F. Tarrahi. 2020. A meta-analysis of the effects of disclosing sponsored content. *Journal of Advertising* 49(3):344–366. https://doi.org/10.1080/00913367.2020.1765909

EU Artificial Intelligence Act. 2024. Implementation Timeline. URL https://artificialintelligenceact.eu/implementation-timeline/ (accessed 7.29.26).

European Union. 2024. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

European Commission. 2026. EU Icons for labelling AI-generated content. URL https://digital-strategy.ec.europa.eu/en/policies/eu-icons-labelling-ai-generated-content (accessed 7.29.26).

Favero, N., P. Belardinelli, L. Zhu, and J. Lahey. 2025. Making null results credible: an overview of design and analytical tools. *Journal of Behavioral Public Administration* 8:1–14. https://doi.org/10.30636/jbpa.81.404

Fisher, G., M. McGranaghan, J. Liaukonyte, and K. C. Wilbur. 2023. Price promotions, beneficiary framing, and mental accounting. *Quantitative Marketing and Economics* 21(2):147–181. https://doi.org/10.1007/s11129-023-09261-0

Friestad, M., and P. Wright. 1994. The persuasion knowledge model: how people cope with persuasion attempts. *Journal of Consumer Research* 21(1):1–31. https://doi.org/10.1086/209380

Giakoumaki, C., and A. Krepapa. 2020. Brand engagement in self-concept and consumer engagement in social media: the role of the source. *Psychology and Marketing* 37(3):457–465. https://doi.org/10.1002/mar.21312

Grassini, S. 2023. Development and validation of the AI attitude scale (AIAS-4): a brief measure of general attitude toward artificial intelligence. *Frontiers in Psychology* 14:1191628. https://doi.org/10.3389/fpsyg.2023.1191628

Grewal, D., C. B. Satornino, T. Davenport, and A. Guha. 2025. How generative AI Is shaping the future of marketing. *Journal of the Academy of Marketing Science* 53(3):702–722. https://doi.org/10.1007/s11747-024-01064-3

Harms, C., and D. Lakens. 2018. Making null effects informative: statistical techniques and inferential frameworks. *Journal of Clinical and Translational Research* 3(2):382–393.

Hartmann, J., Y. Exner, and S. Domdey. 2025. The power of generative marketing: can generative AI create superhuman visual marketing content? *International Journal of Research in Marketing* 42(1):13–31. https://doi.org/10.1016/j.ijresmar.2024.09.002

Heitmann, M., T. P. J. Jansen, M. Reisenbichler, and D. A. Schweidel. 2025. EXPRESS: picture perfect: engaging customers with visual generative AI. *Journal of Marketing* 90(4):74–96. https://doi.org/10.1177/00222429251356993

Hou, J. (Jove), L. Wang, G. Wang, H. J. Wang, and S. Yang. eds. 2025. The double-edged roles of generative AI in the creative process: experiments on design work. *Information Systems Research*. https://doi.org/10.1287/isre.2024.0937

Hubbard, R., and J. S. Armstrong. 1992. Are null results becoming an endangered species in marketing? *Marketing Letters* 3(2):127–136. https://doi.org/10.1007/BF00993992

Ilg, K., and M. Stadtelmann. 2025. Labeling AI-generated content: insights from a field experiment and online survey. *Marketing Review St Gallen* 2025(4):38–46.

Kane, J. V. 2025. More than meets the ITT: A guide for anticipating and investigating nonsignificant results in survey experiments. *Journal of Experimental Political Science* 12(1):110–125. https://doi.org/10.1017/XPS.2024.1

Kelley, H. H. 1973. The processes of causal attribution. *American Psychologist* 28(2):107–128. https://doi.org/10.1037/h0034225

Kim, E. (Anna), M. Duffy, and E. Thorson. eds. 2021. Under the influence: social media influencers' impact on response to corporate reputation advertising. *Journal of Advertising* 50(2): 119–138. https://doi.org/10.1080/00913367.2020.1868026

Kirk, C. P., and J. Givi. 2025. The AI-authorship effect: Understanding authenticity, moral disgust, and consumer responses to AI-generated marketing communications. *Journal of Business Research* 186:114984. https://doi.org/10.1016/j.jbusres.2024.114984

Koning, B., and H. A. M. Voorveld. 2025. Disclaimer! This Content Is AI-Generated: How AI-disclosures influence trust in advertisements and organizations. *Journal of Interactive Advertising* 25(3):240–253. https://doi.org/10.1080/15252019.2025.2554149

Lakens, D., Meta-Analyses. 2017. Equivalence tests: a practical primer for t Tests, Correlations and meta-analyses. *Social Psychological and Personality Science* 8(4):355–362. https://doi.org/10.1177/1948550617697177

Landis, R. S., L. R. James, C. E. Lance, C. A. Pierce, and S. G. Rogelberg. 2014. When is nothing something? Editorial for the null results special issue of journal of business and psychology. *Journal Of Business And Psychology* 29(2):163–167. https://doi.org/10.1007/s10869-014-9347-8

Le, K. B. Q., H. Khan, F. Li, and W. H. Kunz. 2026. More than saying it's AI: How role disclosure transparency in AI-generated Ads influences persuasion. *Psychology and Marketing*. https://doi.org/10.1002/mar.70175

Lee, Y., S. Oh, T. Kim, and H. Bang. 2025. Human, you should help others': how AI disclosure, message assertiveness, and AI anxiety shape consumer responses to AI-generated charitable ads. *Journal of Marketing Management* 42(7–8):1056–1083. https://doi.org/10.1080/0267257X.2025.2596021

Longoni, C., and L. Cian. 2022. Artificial intelligence in Utilitarian vs. Hedonic contexts: the word-of-machine effect. *Journal of Marketing* 86(1):91–108. https://doi.org/10.1177/0022242920957347

Lowe, B., D. Laffey, and Y. Luo. eds. 2025. (Eddie) Advertising and generative AI: How can advertisers leverage new AI tools? Introducing a special issue on generative AI in advertising. *Journal of Advertising Research* 65(2): 146–149. https://doi.org/10.1080/00218499.2025.2494973

MacKenzie, S. B., and R. J. Lutz. 1989. An empirical examination of the structural antecedents of attitude toward the Ad in an advertising pretesting context. *Journal of Marketing* 53(2):48–65. https://doi.org/10.1177/002224298905300204

Moes, A., M. Fransen, T. Verhagen, and B. Fennis. 2022. A good reason to buy: justification drives the effect of advertising frames on impulsive socially responsible buying. *Psychology and Marketing* 39(12):2260–2272. https://doi.org/10.1002/mar.21733

Narayanan, P. 2025. Against the Green Schema: How Gen-AI negatively impacts green influencer posts. *Psychology and Marketing* 42(4):970–986. https://doi.org/10.1002/mar.22159

National Artificial Intelligence Centre. 2025. *Being clear about AI-generated content*. A guide for business.

Newell, S. J., and R. E. Goldsmith. 2001. The development of a scale to measure perceived corporate credibility. *Journal of Business*

Research 52(3):235–247. https://doi.org/10.1016/S0148-2963(99)00104-6

Ohanian, R. 1990. Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising* 19(3):39–52. https://doi.org/10.1080/00913367.1990.10673191

Petrescu, M., and A. S. Krishen. 2025. The significance of nothing: a call to discussion. *Journal of Marketing Analytics* 13(4):1033–1034. https://doi.org/10.1057/s41270-025-00447-x

Pierre, L., V.C.Z. Rodriguez, and J. Wang. eds. 2026. (Joy) Measuring consumer aversion toward AI in marketing communication and assessing its relationship to brand perceptions and purchase intention. *Journal of Consumer Behaviour* 25(4): 2193–2220. https://doi.org/10.1002/cb.70163

Pittman, M., and E. Haley. 2023. Cognitive load and social media advertising. *Journal of Interactive Advertising* 23(1):33–54. https://doi.org/10.1080/15252019.2022.2144780

Qiu, X., Y. Wang, Y. Zeng, and R. Cong. 2025. Artificial intelligence disclosure in cause-related marketing: a persuasion knowledge perspective. *Journal of Theoretical and Applied Electronic Commerce Research* 20(3):193. https://doi.org/10.3390/jtaer20030193

Reeder, G. D., R. Vonk, M. J. Ronk, J. Ham, and M. Lawrence. 2004. Dispositional attribution: multiple inferences about motive-related traits. *Journal of Personality and Social Psychology* 86(4):530–544. https://doi.org/10.1037/0022-3514.86.4.530

Reisenbichler, M., T. Reutterer, D. A. Schweidel, and D. Dan. 2022. Frontiers: supporting content marketing with natural language generation. *Marketing Science* 41(3):441–452. https://doi.org/10.1287/mksc.2022.1354

Rigdon, E. E. 2026. Statistical Stockholm syndrome: the statistical significance myth is the problem. *Journal of Marketing Analytics*. https://doi.org/10.1057/s41270-026-00474-2

Rogge, N. D., A. Lamovšek, and S. Batistič. 2025. AI meets IA: a typology of generative-AI-supported work through individual ambidexterity. *Business Horizons* 69(4):477–492. https://doi.org/10.1016/j.bushor.2025.06.006

Ruan, C., Y. Quan, X. Li, Y. Zheng, H. Deng, and X. Wei. 2026. Cost-cutting or trust building: consumer motive inference and purchase intention toward AI-produced food. *Foods* 15(8):1405. https://doi.org/10.3390/foods15081405

Ryoo, Y., M. Bakpayev, Y. A. Jeon, K. Kim, and S. Yoon. 2025. High hopes, hard falls: consumer expectations and reactions to AI-human collaboration in advertising. *International Journal of Advertising* 45(1):48–80. https://doi.org/10.1080/02650487.2025.2458996

Saeed, M. R., H. Khan, R. Lee, L. Lockshin, S. Bellman, J. Cohen, and S. Yang. 2024. Construal level theory in advertising research: a systematic review and directions for future research. *Journal of Business Research* 183:114870. https://doi.org/10.1016/j.jbusres.2024.114870

Sands, S., V. Demsar, C. Ferraro, S. Wilson, M. Wheeler, and C. Campbell. 2025. Easing AI-advertising aversion: how leadership for the greater good buffers negative response to AI-generated ads. *International Journal of Advertising* 45(1):27–47. https://doi.org/10.1080/02650487.2025.2457080

Sarstedt, M. 2026. Metrological uncertainty and the value of informative nulls in marketing research. *Journal of Marketing Analytics*. https://doi.org/10.1057/s41270-026-00465-3

Sarstedt, M., S. J. Adler, C. M. Ringle, G. Cho, A. Diamantopoulos, H. Hwang, and B. D. Liengaard. 2024. Same model, same data, but different outcomes: evaluating the impact of method choices in structural equation modeling. *Journal of Product Innovation Management* 41(6):1100–1117. https://doi.org/10.1111/jpim.12738

Schilke, O., and M. Reimann. 2025. The transparency dilemma: How AI disclosure erodes trust. *Organizational Behavior and Human Decision Processes* 188:104405. https://doi.org/10.1016/j.obhdp.2025.104405

Schreiner, T., and D. Baier. 2021. Online retailing during the COVID-19 pandemic: consumer preferences for marketing actions with consumer self-benefits versus other-benefit components. *Journal of Marketing Management* 37(17–18):1866–1902. https://doi.org/10.1080/0267257X.2022.2030784

Schreiner, T., A. Rese, and D. Baier. 2019. Multichannel personalization: Identifying consumer preferences for product recommendations in advertisements across different media channels. *Journal of Retailing and Consumer Services* 48:87–99. https://doi.org/10.1016/j.jretconser.2019.02.010

Spears, N., and S. N. Singh. 2004. Measuring attitude toward the brand and purchase intentions. *Journal of Current Issues & Research in Advertising* 26(2):53–66. https://doi.org/10.1080/10641734.2004.10505164

Sun, H., P. Xie, and Y. Sun. 2025. The inverted U-shaped effect of personalization on consumer attitudes in AI-generated ads: striking the right balance between utility and threat. *Journal of Advertising Research* 65(2):237–258. https://doi.org/10.1080/00218499.2025.2462405

TikTok. 2025. Community Guidelines. URL https://www.tiktok.com/community-guidelines/en/ (accessed 7.29.26).

Tintarev, N., and J. Masthoff. 2012. Evaluating the effectiveness of explanations for recommender systems: methodological issues and empirical studies on the impact of personalization. *User Modeling and User-Adapted Interaction* 22(4–5):399–439. https://doi.org/10.1007/s11257-011-9117-5

To, R. N., Y.-C. Wu, P. Kianian, and Z. Zhang. 2025. When AI doesn't sell prada: why using AI-generated advertisements backfires for luxury brands. *Journal of Advertising Research* 65(2):202–236. https://doi.org/10.1080/00218499.2025.2454120

Van Heerde, H. J., C. Moorman, C. P. Moreau, and R. W. Palmatier. 2021. Reality check: infusing ecological value into academic marketing research. *Journal of Marketing* 85(2):1–13. https://doi.org/10.1177/0022242921992383

Van Reijmersdal, E. A., E. Brussee, N. Evans, and B. W. Wojdynski. 2023. Disclosure-driven recognition of native advertising: a test of two competing mechanisms. *Journal of Interactive Advertising* 23(2):85–97. https://doi.org/10.1080/15252019.2022.2146991

Verlegh, P. W. J., G. Ryu, M. A. Tuk, and L. Feick. 2013. Receiver responses to rewarded referrals: the motive inferences framework. *Journal of the Academy of Marketing Science* 41(6):669–682. https://doi.org/10.1007/s11747-013-0327-8

Villegas Bravo, M. 2024. Demystifying generative AI disclosures. EPIC-electronic privacy information center. URL https://epic.org/demystifying-generative-ai-disclosures/ (accessed 7.29.26).

Vyas, Y., and P. Vishwakarma. 2026. Artificial intelligence in brand communication: a comprehensive literature review and research outlook. *International Journal of Consumer Studies* 50(1):1–36. https://doi.org/10.1111/ijcs.70170

Walther, C. 2026. Why AI disclosure matters at every level. Knowledge at Wharton. URL https://knowledge.wharton.upenn.edu/article/why-ai-disclosure-matters-at-every-level/ (accessed 7.29.26).

Weiner, B. 1985. An attributional theory of achievement motivation and emotion. *Psychological Review* 92(4):548–573. https://doi.org/10.1037/0033-295X.92.4.548

White, K., and J. Peloza. 2009. Self-benefit versus other-benefit marketing appeals: their effectiveness in generating charitable support. *Journal of Marketing* 73(4):109–124. https://doi.org/10.1509/jmkg.73.4.109

Wiebe, J., D. Z. Basil, and M. Runté. 2017. Psychological distance and perceived consumer effectiveness in a cause-related marketing context. *International Review on Public and Nonprofit Marketing* 14(2):197–215. https://doi.org/10.1007/s12208-016-0170-y

Wojdynski, B. W., and N. J. Evans. 2016. Going native: effects of disclosure position and language on the recognition and evaluation of online native advertising. *Journal of Advertising* 45(2):157–168. https://doi.org/10.1080/00913367.2015.1115380

Wortel, C., I. Vanwesenbeeck, and F. Tomas. 2024. Made with artificial intelligence: the effect of artificial intelligence disclosures in instagram advertisements on consumer attitudes. *Emerging Media* 2(3):547–570. https://doi.org/10.1177/27523543241292096

Wu, L., N. A. Dodoo, and T. J. Wen. 2025. Disclosing AI's involvement in advertising to consumers: a task-dependent perspective. *Journal of Advertising* 54(1):20–38. https://doi.org/10.1080/00913367.2024.2309929

Yang, D., Y. Lu, W. Zhu, and C. Su. 2015. Going green: How different advertising appeals impact green consumption behavior. *Journal of Business Research* 68(12):2663–2675. https://doi.org/10.1016/j.jbusres.2015.04.004

Youtube team. 2024. How we're helping creators disclose altered or synthetic content. URL https://blog.youtube/news-and-events/disclosing-ai-generated-content/ (accessed 7.29.26).

Yucel-Aybat, O., and M.-H. Hsieh. 2021. Consumer mindsets matter: Benefit framing and firm–cause fit in the persuasiveness of cause-related marketing campaigns. *Journal of Business Research* 129:418–427. https://doi.org/10.1016/j.jbusres.2021.02.051

Zhang, L., and C. Hur. 2025. The impact of generative AI images on consumer attitudes in advertising. *Administrative Sciences* 15(10):1–26. https://doi.org/10.3390/admsci15100395

Timo Schreiner Timo Schreiner is a Postdoctoral Researcher at the Nuremberg Institute for Market Decisions (NIM). His research examines how artificial intelligence and other digital technologies shape marketing effectiveness and consumer decision-making using experimental and data-driven approaches. He received his PhD from the University of Bayreuth, Germany, where he investigated how personalization and cause-related marketing influence the effectiveness of online advertising. Before joining NIM, he gained several years of experience in IT consulting in the field of Data & Analytics as well as in market research.

Yakov Bart Yakov Bart is a Professor of Marketing and Thomas E. Moore Faculty Fellow at Northeastern University. His research examining marketing implications of new digital technologies and business models has been funded by NSF, Amazon, Google, MSI, WPP, and published in leading marketing and management journals. Yakov holds a PhD and an MS in Business Administration from the University of California at Berkeley, an SM in Operations Research from MIT, and a Diploma in Mathematics from Moscow State University. Prior to joining Northeastern University, he was a faculty member at INSEAD and a visiting faculty at the Wharton School.

Koen Pauwels Koen Pauwels is the Associate Dean of Research and Distinguished Professor of Marketing at the D'Amore-McKim School of Business at Northeastern University. Formerly, he was Principal Research Scientist at Amazon Ads, with brand building and budget allocation recommendations reaching hundreds of thousands of advertisers. Koen received his Ph.D. from UCLA, where he was chosen Top 100 Inspirational Alumnus. After receiving tenure at the Tuck School of Business at Dartmouth, he helped build the startup Ozyegin University in Istanbul. Named a worldwide top 2% scientist, and 'The Best Marketing Academic on the Planet', Koen published over 100 articles on marketing effectiveness.

Cosima Schütze Cosima Schütze is a research assistant at the Nuremberg Institute for Market Decisions (NIM) and a master's student in Media Effects and Media Psychology at the University of Applied Sciences Ansbach, Germany. Her research interests focus on advertising effectiveness and consumer behavior, with a particular interest in how digital media and marketing influence consumer perceptions and decision-making. She holds a bachelor's degree in media psychology from the University of Applied Sciences Ingolstadt, Germany.

Fabian Buder Fabian Buder is Head of Marketing & Consumer Behavior at the Nuremberg Institute for Market Decisions, Germany, and an Affiliate Researcher at the D'Amore-McKim School of Business, Northeastern University, USA. His research focuses on the impact of emerging technologies, particularly generative AI and virtual environments, and societal trends on consumer behavior and marketing practice. He holds a PhD in Agricultural and Food Marketing from the University of Kassel. Before joining NIM, he worked in market research for the food industry at GfK SE's Household Panel division.