Transparent but Not Always Trusted: The Impact of AI Disclosure on Ad Credibility

Moritz Alexander Jung

Dissertation written under the supervision of professor Cristina Mendonça

Dissertation submitted in partial fulfilment of requirements for the MSc in Management with Specialisation in Strategic Marketing, at the Universidade Católica Portuguesa, 3rd January 2026.

Abstract

Title: Transparent but Not Always Trusted: The Impact of AI Disclosure on Ad Credibility

Author: Moritz Alexander Jung

The increasing adoption of generative artificial intelligence in advertising has heightened debates regarding transparency and consumer trust among both regulators and industry managers. In anticipation of the forthcoming EU AI Act, which will require disclosure of AI-generated content, this study examines the impact of AI disclosure in commercial advertising on perceived advertisement credibility and subsequent consumer responses. Utilizing persuasion and source credibility theory, a between-subjects experiment (N = 125) compared advertisements labeled as AI-generated with identical advertisements lacking such disclosure. The research also investigates whether perceived advertisement credibility affects brand attitude and purchase intention, and whether individual differences in AI aversion moderate the effects of disclosure.

The results indicate that AI disclosure does not significantly lower perceived advertisement credibility for commercial, medium-involvement products. Perceived advertisement credibility continues to be a strong predictor of both brand attitude and purchase intention, irrespective of disclosure. Contrary to initial expectations, AI aversion does not moderate the relationship between disclosure and advertisement credibility. These findings suggest that the effects of AI disclosure are present but not significant and context-dependent, rather than universally negative.

This research contributes to the literature on AI-generated advertising by offering empirical evidence from a standard commercial context prior to regulatory implementation. The findings suggest that, for managers, designing messages to enhance credibility is more critical for advertising effectiveness than the technological origin of the content. Consequently, AI disclosure should be regarded as a contextual design choice rather than an inherent risk.

Keywords: AI disclosure; Advertising credibility; Brand attitude; Purchase intention; AI aversion

Resumo

Título: Transparente, mas nem sempre confiável: o impacto da divulgação da IA na credibilidade dos anúncios

Autor: Moritz Alexander Jung

A crescente adoção da inteligência artificial generativa na publicidade intensificou os debates sobre transparência e confiança do consumidor entre reguladores e gestores. Antecipando a futura Lei da IA da União Europeia, que exigirá a divulgação de conteúdos gerados por IA, este estudo analisa o impacto da divulgação da IA em anúncios comerciais na credibilidade percebida da publicidade e nas respostas dos consumidores. Com base na teoria da persuasão e da credibilidade da fonte, foi conduzida uma experiência entre sujeitos (N = 125), comparando anúncios rotulados como gerados por IA com anúncios idênticos sem essa divulgação.

O estudo analisa ainda se a credibilidade percebida da publicidade influencia a atitude em relação à marca e a intenção de compra, bem como se a aversão individual à IA modera os efeitos da divulgação. Os resultados mostram que a divulgação da IA não reduz significativamente a credibilidade percebida da publicidade em anúncios de produtos comerciais de médio envolvimento. A credibilidade da publicidade permanece um forte preditor da atitude em relação à marca e da intenção de compra, independentemente da divulgação. Além disso, a aversão à IA não apresenta efeito moderador significativo.

Este estudo contribui para a literatura sobre publicidade gerada por IA ao fornecer evidências empíricas num contexto comercial padrão anterior à implementação regulatória. Para os gestores, os resultados indicam que reforçar a credibilidade das mensagens é mais determinante para a eficácia publicitária do que a origem tecnológica do conteúdo.

Palavras-chave: Divulgação da IA; Credibilidade da publicidade; Atitude em relação à marca; Intenção de compra; Aversão à IA

Application of Large Language Models

To enhance the clarity and fluency of the written content in this dissertation, AI tools were utilised for language refinement and to ensure grammatical accuracy. All suggestions were reviewed carefully and edited to maintain the originality of the work.

List of Tables

Glossary

α	Cronbach's index of reliability
β	Standardized regression coefficient
b	Unstandardized regression coefficient
AI	Artificial Intelligence
AIAv	AI aversion
ANCOVA	Analysis of covariance
BA	Brand attitude
CI	Confidence interval
Df	Degrees of freedom
EU	European Union
F	F statistic (Fisher's F ratio)
IV	Independent variable
LLM	Large language model
M	Mean
M (Mediator)	Mediating variable
N	Total number of cases
p	p value
PI	Purchase intention
R^2	Coefficient of determination
RQ	Research question
SD	Standard deviation
SE	Standard error

1 Introduction

"Users shall be informed that they are interacting with an AI system" (*European Union*, 2024). The recently adopted EU AI Act establishes transparency toward consumers as a legal requirement within the European Union. Although the regulation does not precisely define interaction, its focus on transparency and user awareness implies a broad interpretation that includes any contact with Artificial intellegence-generated content. Legal commentaries indicate that these transparency obligations cover the dissemination of AI-generated or AI-manipulated content, including text-based outputs, and may also apply to advertising messages created with artificial intelligence (*European Union*, 2024; Heuking, 2025).

AI already has a firm place in marketing today, from personalized product recommendations to automated content creation. Current data shows that 35% of marketers already use AI for text creation or optimization, and 36% generate image content with AI (*McKinsey*, 2024). The latest advances in generative AI even make it possible to automatically create complete advertising messages (Jung et al., 2025). While this promises enormous efficiency gains, it also raises a crucial question: How do consumers react when an advertisement is explicitly labeled as AI-generated?

Initial empirical studies paint a mixed picture. The studies show that AI disclosure triggers mixed reactions, ranging from reduced credibility and more negative ad attitudes in prosocial and social-media contexts (Baek et al., 2024; Wortel et al., 2024) and responses shaped by trust and perceived humanity in luxury settings (Jung et al., 2025), to largely unchanged credibility judgments in domains such as news and health information (Li & Yang, 2024).

These findings suggest that the effect of disclosure is highly context-dependent, and systematic evidence in the field of commercial advertising is currently lacking. This is precisely where the present study comes in.

The central question of this thesis is how AI disclosure in commercial advertising influences consumer perception, as ad credibility has often been framed in the literature as the central mechanism of influence (Baek et al., 2024). Furthermore, studies suggest that disclosure effects are not the same for all consumers. Individual differences such as AI aversion (i.e., general distrust of AI) can amplify negative effects, while the perceived humanity of the content can mitigate these effects (Baek et al., 2024; Jung et al., 2025). These aspects are considered as possible moderators in this study.

Although previous research has investigated AI disclosure in contexts such as charity campaigns and social-media advertising, there is limited evidence regarding its effects in standard commercial advertising, particularly for mid-involvement consumer products. Consequently, it remains uncertain whether mechanisms identified in these earlier contexts, such as credibility loss, also influence consumer responses in typical product advertising, where decisions are less complex yet personally relevant. Furthermore, most existing studies have neglected individual differences, despite the established role of AI aversion in shaping trust in algorithmic decisions and its potential to affect reactions to AI-labeled advertisements. Addressing these research gaps is essential for establishing a pre-regulatory baseline prior to the implementation of the EU's forthcoming disclosure requirements and for advancing understanding of how consumers evaluate AI-generated advertising in everyday commercial settings.

Based on this, the dissertation examines three central research questions:

1) Does the label "AI-generated" lead to lower perceived ad credibility?
2) Does ad credibility mediate the influence of disclosure on downstream variables such as ad attitude and purchase intention?
3) Are these effects moderated by individual differences such as AI aversion?

The significance of this research is twofold. First, the forthcoming EU AI Act will require AI disclosures in digital advertising, highlighting the need to determine whether such labels influence key advertising outcomes in standard commercial settings. Although prior studies have produced mixed findings (e.g., Baek et al., 2024; Wortel et al., 2024), there is insufficient evidence regarding the impact of AI disclosures on ad credibility, brand evaluations, and purchase intentions for mid-involvement consumer products. Generating this evidence is crucial for organizations aiming to anticipate consumer responses when disclosure becomes mandatory.

Second, this thesis advances advertising and AI-disclosure research by providing empirical evidence from a commercial context that has received limited scholarly attention. It extends persuasion and source-credibility theory (Eisend, 2006; Wilson & Sherrell, 1993) by investigating credibility as a mechanism influencing subsequent evaluations, such as brand attitude and purchase intention. By introducing AI aversion as a dispositional moderator and building on research in algorithm aversion (Dietvorst et al., 2015; Logg et al., 2019), the study incorporates an individual-difference perspective that is largely absent from current disclosure research (e.g., Baek et al., 2024; Jung et al., 2025). Furthermore, by capturing consumer responses in a pre-regulatory environment, the study establishes a timely empirical baseline prior to the implementation of the EU's disclosure requirements.

The structure of this work is as follows: Chapter 1 introduces the topic, the research problem, and its relevance. Chapter 2 discusses the topics of marketing, AI-generated content, and the effect of disclosure from a theoretical perspective. Chapter 3 describes the experimental design, including manipulation and measurement of the variables ad credibility, manipulative intent, ad attitude, and purchase intention. Chapter 4 presents the analyses and results. Chapter 5 summarizes implications, limitations, and suggestions for future research.

2 Literature Review

2.1 Marketing and Advertising

Marketing forms the basis of entrepreneurial value creation and describes the activities, institutions, and processes through which organizations create, communicate, deliver, and exchange offerings that have value for customers, partners, and society at large. This widely used definition is provided by the American Marketing Association (AMA) and was last updated by a scholarly panel in 2017 (American Marketing Association, 2025). It emphasizes that marketing cannot be reduced to short-term sales promotion, but rather represents a comprehensive management function with both economic and social relevance.

However, the discussion about the definition of marketing is ongoing. While the AMA's definition (American Marketing Association, 2017) offers a normative framework, scholars (Zhang & Watson IV, 2020) have noted that, in practice, it may remain too abstract to capture the dynamic realities faced by organizations. As marketing continuously adapts to changing markets, technologies, and social expectations, new questions regarding practical implementation and stakeholder responsibility come to the fore.

Research in the field of marketing science makes an important contribution to linking theory and practice. In their concept of the "marketing science value chain" (Roberts et al., 2014), they describe how academic findings are disseminated to the practical decision-making processes of marketing managers via intermediaries such as consultancies and market research companies. They show that marketing science tools have a demonstrable influence on corporate decisions, particularly in areas such as brand management, pricing, and product development. This makes it clear that marketing is both a scientific discipline and a practice-oriented field of activity.

Against the backdrop, recent reviews emphasize that innovations such as digitalization and the use of AI are increasingly shaping marketing (Haleem et al., 2022). While the fundamental function of marketing, creating value for customers and companies, remains unchanged, the means and strategies used to achieve this goal are constantly changing (Haleem et al., 2022). Such dynamics make it essential to closely integrate theory, empiricism, and practice.

2.2 Fundamentals of advertising and advertising effectiveness

Advertising is a key marketing communication tool designed to inform consumers, influence attitudes, and ultimately trigger purchase-related actions (Eisend, 2006). It structures communication between suppliers and consumers and attempts to deliberately shape perceptions and behavior regarding products, services, or brands (Eisend, 2006). Modern advertising effectiveness research examines how well such investments translate into performance outcomes (Sethuraman et al., 2011). A comprehensive meta-analysis shows that the average advertising elasticity is approximately 0.1, meaning that a 1 percent increase in advertising expenditure leads, depending on the performance measure, to an average increase of about 0.1 percent in either sales or market share (Sethuraman et al., 2011). However, this elasticity varies considerably across markets, product categories, and life cycle stages, highlighting that not only the extent of advertising exposure but also the quality and credibility of the message determine its effectiveness (Sethuraman et al., 2011).

A central framework for understanding how consumers evaluate and respond to advertising is the Persuasion Knowledge Model (PKM). Developed by Friestad and Wright (1994), the PKM conceptualizes consumers as active, learning agents who accumulate knowledge about persuasion tactics over time and use this knowledge to cope with influence attempts. The model distinguishes among persuasion knowledge, knowledge about the actors involved and their goals, and product knowledge. When consumers recognize a message as a persuasion attempt, they activate the principle of meaning change, reinterpreting the same information as interest-driven influence, which increases skepticism and triggers more critical processing (Friestad & Wright, 1994).

Building on this understanding of persuasion processing, research further emphasizes that the credibility of the advertising source plays a central role in shaping message effectiveness. Classic work shows that a message's impact strongly depends on perceptions of the source, particularly its expertise, trustworthiness, and dynamism, the latter referring to confidence, expressiveness, and communicative energy (Eisend, 2006; Hovland & Weiss, 1951; Wilson & Sherrell, 1993). Meta-analytic evidence indicates that source credibility accounts for approximately 7.4 percent of the variance in attitudes and purchase intentions, underscoring its statistical and practical relevance (Wilson & Sherrell, 1993). Eisend (2006) further refines this construct into inclination toward truth (honesty and sincerity), potential for truth (competence and expertise), and presentation (attractiveness and expressiveness), dimensions that have been shown to operate consistently across communication settings and source types. Together, these findings highlight credibility as a core determinant of how consumers interpret and evaluate advertising messages. To understand the downstream consequences of these evaluations, it is essential to examine the attitudinal outcomes that advertising produces.

Brand attitude reflects consumers' overall evaluation of a brand and represents a central component of brand knowledge structures (Keller, 1993). Such evaluations arise from the associations consumers hold about a brand, including perceived attributes and benefits, and influence how new brand-related information is processed. Favorable brand attitudes, therefore, provide an important basis for consumers' responses to marketing activities and communications. Spears and Singh (2004) conceptualize brand attitude as a relatively enduring evaluative judgment that integrates cognitive beliefs and affective reactions. Their work demonstrates that this judgment serves as a reliable predictor of downstream behavioral outcomes, particularly purchase intention. Further evidence shows that brand attitudes also guide how consumers process product information during decision-making: positive attitudes lead consumers to rely more on global brand evaluations rather than on detailed analytical comparisons (Park & Chang, 2022). This pattern underscores the centrality of brand attitude in structuring consumer judgments and shaping evaluative responses to advertising.

Given its role in shaping information processing and evaluative judgments, brand attitude functions as a key antecedent to purchase intention. Purchase intention is defined as a consumer's deliberate plan or motivation to purchase a brand in the future and represents a central behavioral outcome in advertising research (Spears & Singh, 2004). It reflects the consumer's readiness to act and is influenced by beliefs regarding the brand's ability to satisfy functional, symbolic, or experiential needs (Keller, 1993). Meta-analytic evidence further emphasizes that attitudes and related evaluative judgments are among the strongest predictors of online purchase intentions across diverse contexts(Ghosh, 2024). Spears and Singh (2004) show that more favorable brand attitudes consistently increase the likelihood of future purchase planning, a finding that aligns with Keller's (1993) understanding of brand attitudes as central elements of the associative structures guiding consumer judgment. Ghosh's (2024) meta-analysis reinforces this attitudinal pathway by demonstrating the robustness of the attitude–intention link.

In addition to attitudinal antecedents, credibility-based cues can further shape consumers' purchase intentions. Lafferty and Goldsmith (1999) show that corporate credibility, reflecting perceptions of a firm's trustworthiness and expertise, exerts a significant positive effect on purchase intention, whereas endorser credibility does not have a significant main effect. Their findings suggest that credibility serves as a contextual signal that consumers use to assess whether a brand can be trusted, strengthening or weakening their willingness to translate positive evaluations into behavioral intentions. Purchase intentions are therefore shaped by both internal brand evaluations and external cues embedded in the advertising environment.

2.3 Artificial Intelligence: General Perspective

Following the discussion of advertising fundamentals, it is essential to clarify how AI is shaping the next era of marketing. Artificial intelligence is defined by Davenport et al. (2020) as the ability of computer systems to perform tasks that typically require human intelligence, such as recognizing patterns, understanding language, making decisions, and learning from experience. In the marketing context, AI refers to using these systems to automate and optimize marketing processes as well as to create content. The core concept is not to replace humans, but to augment marketers' capabilities and accelerate processes (Davenport et al., 2020).

Haleem et al. (2022) did a comprehensive review of recent AI applications in marketing research and practice. Their review synthesizes studies on the use of AI for data analysis, pattern recognition, predictive modeling, and content generation within various marketing domains. The first type is analytical AI, which is designed to analyze large datasets, recognize patterns, and make predictions, a form widely used in personalizing advertising messages or forecasting consumer behavior. The second type is generative AI, which enables the creation of new content such as text, images, and media, moving beyond just processing existing information. This distinction is central to the present work, as research increasingly explores the implications of generative AI for marketing practice. While analytical AI determines which advertisements are shown to which consumers, generative AI can create the advertisements themselves (Haleem et al., 2022), as is the case of the scenario used in this thesis' study.

2.4 Innovation and digital transformation in marketing

The evolution of marketing from broad mass-orientation to today's data-driven individualization reflects how technological advances and academic insights have continually reshaped the discipline. Historically, marketing was characterized by undifferentiated, massoriented practices, a single message aimed at a wide, largely unsegmented audience, with measurements based on population averages and a predominantly one-way relationship from brand to consumer (Mudler, 2004). Over time, increased integration of academic marketing research into business practice led to greater professionalization and a more scientific approach, as documented in the marketing science value chain (Roberts et al., 2014).

This gradual but steady shift set the stage for an even more fundamental transformation with the rise of digital technologies and the Internet. As Verhoef et al. (2017) underscore, this new era exposed consumers to complex, interconnected environments where digital and physical touchpoints existed side by side and marketers gained tools to monitor customer behaviors across channels in greater detail (Verhoef et al., 2017). The introduction of web analytics, customer relationship management systems, and later social media platforms opened up new levels of transparency regarding consumer preferences and behaviors (Davenport et al., 2020).

Recent developments have enabled marketers to move beyond retrospective data analysis and to harness real-time signals, marking the advent of scalable personalization and refined target-group segmentation (Davenport et al., 2020). Given that consumers today navigate complex networks of virtual and physical touchpoints (Verhoef et al., 2017), marketers must adapt their communication to these multidimensional environments. As a result, instead of relying on a small number of generic messages delivered through mass media, they now craft more fragmented and individualized communications across a growing set of addressable, interactive channels (Roberts et al., 2014).

Building on the increasing digitalization and data-driven practices of the 2000s, the 2010s marked a significant leap in marketing sophistication through the rise of algorithmic decision-making. During this decade, machine learning and predictive analytics began to play a central role in identifying behavioral patterns at a scale that surpassed human analytical capabilities, enabling marketers to anticipate consumer actions and proactively optimize campaigns (Haleem et al., 2022). Marketing automation platforms further refined the timing and delivery of messages by dynamically responding to user behavior and customizing offers in real time (Haleem et al., 2022). Importantly, as Davenport et al. (2020) clarify, this period was defined by the automation of decision processes, not by the automation of content creation: the content itself continued to be developed by humans, while algorithms determined when, where, and to whom it was delivered (Davenport et al., 2020).

A qualitatively new development has emerged in the 2020s with the arrival of generative artificial intelligence. Unlike prior automation that simply optimized decisions, generative AI technologies such as large language models and diffusion models now enable the direct creation of marketing content by algorithms, fundamentally reshaping the creative aspects of marketing (Moharam & Tawalbeh, 2025). This enables various content-related marketing tasks, such as generating text elements, automating design processes, or supporting broader content creation workflows, to be increasingly assisted or executed by AI systems. Recent literature highlights the expanding role of AI-driven content generation and automation in marketing (Haleem et al., 2022).

2.5 **Artificial intelligence in marketing: possible applications**

AI has established itself as a key technology for marketing operations in recent years. According to Haleem et al. (2022), the spectrum of practical AI applications in marketing covers the following areas: (1) personalization and targeting, which enable customer interactions to be tailored based on individual preferences and behaviors; (2) recommendation systems that use historical data to suggest products or content relevant to specific customers; (3) chatbots and customer service automation that handle customer communication and significantly reduce response times; (4) dynamic pricing that adjusts prices in real time to demand, competition, and other factors; and (5) predictive analytics to forecast customer behavior, market trends, and churn rates (Haleem et al., 2022).

In addition, generative AI enables data-driven creativity by combining automated content production with the analysis of large consumer datasets, which allows highly personalized, context-specific messages to be produced at scale while freeing human teams to focus on higher-level strategic tasks (Moharam & Tawalbeh, 2025). These applications promise massive efficiency gains and scalability as AI systems automate human decisions and routine production processes, thereby reducing required working time exponentially (Haleem et al., 2022; Moharam & Tawalbeh, 2025).

This development marks a critical turning point: AI systems are no longer just taking over routine tasks but also creative and strategic functions traditionally the domain of human expertise. At the same time, this raises fundamentally new questions, particularly with regard to how consumers respond to content created by machines rather than humans.

2.6 AI-Generated Advertising (AIGC) & Consumer Responses

AI-Generated Content (AIGC) refers to advertising materials across all major media formats, ranging from text and static images to videos and even integrated campaigns, produced by generative AI models rather than human creators (Kim et al., 2023). As these technologies become increasingly accessible and economically viable, they are fundamentally transforming how advertising is produced and disseminated (Davenport et al., 2020). Recent work shows that generative AI now enables the end-to-end creation of marketing content, including social media posts, email, images, videos, landing pages, and full campaign concepts, effectively acting as a creative partner that can rapidly generate, adapt, and test content variations for different segments (Moharam & Tawalbeh, 2025). A central psychological question arises: How do consumers experience and evaluate advertising once they become aware that it is generated by an algorithm rather than a person?

Current research in consumer psychology demonstrates that this knowledge triggers distinctive patterns of cognitive and emotional processing that differ from responses to human-created advertising (Kim et al., 2023). To better understand how consumers evaluate AIGC, it is helpful it to relate to broader theories of technology-mediated consumer experience. According to Davenport et al. (2020), AI in marketing can be understood through a three-dimensional framework, focusing on the level of intelligence (from pure automation to contextual intelligence), types of tasks (analytical vs. creative), and the degree of physical embodiment (virtual vs. robotic). A key insight of this model is that AI is most effective when it complements, not replaces, human abilities, especially in creative and emotionally charged domains such as advertising. This helps explain why consumers remain skeptical of AIGC: advertising relies on creativity, empathy, and authentic expressiveness, elements that demand contextual understanding and human intuition rather than mere computational efficiency (Davenport et al., 2020).

Similarly, (Verhoef et al., 2017) emphasize that modern consumers inhabit technology-supported environments, interacting with people, intelligent objects, and digital spaces simultaneously. These environments can be described as multi-agent, technologically embedded contexts in which consumer experiences are shaped not only by human actors but also by AI systems and other digital agents (Puntoni et al., 2020). The distinction Verhoef et al. draw between active (conscious, intentional evaluation) and passive (automatic, incidental exposure) processing is particularly salient for understanding how AIGC is received and evaluated.

Within technology-enabled, multi-agent consumer environments, where individuals interact simultaneously with human actors, AI systems, digital interfaces, and intelligent objects, consumer responses to AIGC can be explained by several key psychological dimensions.
The first is perceived human-likeness; Jung et al. (2025) demonstrate that evaluations of AIGC depend less on objective content quality than on the extent to which advertising content appears human-like and emotionally resonant. Jung et al. (2025) show that perceived humanity and trust in AI mitigate the negative effects of disclosure. The more human and emotionally competent consumers perceive AI-generated advertising to be, the less negative their reaction to disclosure (Jung et al., 2025). The second is authenticity: Baek et al. (2024) show that consumers critically assess whether AI-generated advertising communicates genuine intentions or instead signals strategic manipulation. The third involves cognitive processing cues and their modest impact on content evaluation. Li and Yang (2024) found that AIGC labels function as a nudging intervention that helps users distinguish AI-generated content and deepens cognitive processing, leading to more cautious judgments. Their study also shows, however, that AIGC labels minimally affect perceived accuracy, message credibility, or sharing intention, but highlights their potential for fact-checking and governance.
These psychological processing patterns are further modulated by cognitive mode. In passive exposure, when consumers are unaware that they are viewing AI-generated content, AIGC is often evaluated based on conventional advertising criteria such as creativity, message clarity, and product fit even when its artificial origin is not disclosed (Kim et al., 2023; Baek et al., 2024). In contrast, the evaluation becomes more deliberative when consumers consciously recognize the machine-generated content, either through platform cues or prior knowledge (Baek et al., 2024; Li & Yang, 2024). Under these conditions, dimensions such as human-likeness (Jung et al., 2025), authenticity (Baek et al., 2024), anthropomorphism (Kim et al., 2023), and emotional warmth (Peng et al., 2022) become central drivers of consumer response.
While prior research has mainly examined how characteristics of AI-generated content shape consumer responses, an equally important source of variation lies in the consumers themselves. Beyond stimulus-related cues, individual attitudes toward AI play a critical role in determining how AI-generated advertising is evaluated

2.7 Consumer Attitudes Toward AI and AI Aversion

Consumer responses to artificial intelligence are influenced by both the characteristics of the technology and individuals' general attitudes toward AI. These attitudes represent relatively stable evaluative tendencies, including cognitive beliefs about AI competence and reliability, as well as affective reactions such as comfort or concern (Gillath et al., 2021; Li et al., 2024). Individuals approach AI with pre-existing psychological schemas, such as general trust propensity and interpersonal trust mechanisms, which inform their judgments about AI systems (Li et al., 2024). Previous research demonstrates systematic differences in the ease with which people develop trust in AI. For example, dispositional factors like attachment anxiety are negatively associated with trust in AI, while attachment avoidance does not show a reliable effect (Gillath et al., 2021). These findings suggest that attitudes toward AI are shaped by both technological features and underlying psychological dispositions, which are central to determining whether consumers perceive AI-based services as trustworthy and acceptable (Li et al., 2024).

A closely related construct is AI aversion, also referred to as algorithm aversion. This term describes the tendency to evaluate algorithmic or AI-based decisions more skeptically than human decisions, even when the algorithm demonstrates objective performance (Dietvorst & Bartels, 2022). Research reveals substantial heterogeneity in this tendency: some consumers readily rely on algorithmic outputs, while others systematically avoid them. Dietvorst and Bartels (2022) find that aversion is especially pronounced in domains involving morally relevant tradeoffs, where consumers believe algorithms apply rigid, outcome-maximizing strategies that overlook fairness or contextual nuance. Additional evidence suggests that consumers sometimes respond less favorably when algorithms, rather than humans, make advantageous decisions, as algorithmic judgments are more difficult to internalize and incorporate into self-related evaluations (Yalcin et al., 2022). Collectively, these findings define AI aversion as a relatively stable, negatively valenced disposition that shapes how individuals judge AI-generated decisions across contexts.

AI aversion can be conceptualized as a moderator that captures interindividual differences in how consumers respond to AI-based applications (Yalcin et al., 2022; Dietvorst & Bartels, 2022). While some individuals readily engage with algorithmic output, others approach AI-generated content with a baseline of skepticism. In AIGC contexts, this means that disclosure cues meet consumers with different trust predispositions. Individuals high in AI aversion may interpret algorithmic content creation as less authentic, more strategically motivated, or emotionally detached, thereby strengthening perceptions of manipulation (Li et al., 2024;

Yalcin et al., 2022; Dietvorst & Bartels, 2022). Those with lower aversion or higher trust in AI, in contrast, may view AI involvement as a signal of efficiency and competence. Building on these individual differences, the next section introduces AI disclosure as a central contextual cue that activates consumers' attitudes toward AI and may interact with AI aversion to shape advertising evaluations.

2.8 AI disclosure in advertising: mechanisms and effects

When advertising is openly labeled as AI-generated, consumers' psychological reactions often change in unexpected ways. Although transparency is considered a prerequisite for more favorable consumer reactions and greater persuasiveness in advertising effectiveness research, recent studies show that disclosures of AI involvement can significantly reduce the effectiveness of advertising. This seemingly paradoxical effect can be explained by the assumption that transparency is intended to build trust, but in the case of negative information (such as algorithmic control), it can provoke greater skepticism and critical scrutiny (Baek et al., 2024).

Research has identified two key psychological mechanisms that mediate these effects: First, the perceived credibility of the advertising offer decreases as soon as AI is disclosed as the source. Baek et al. (2024) demonstrate in a prosocial advertising study that disclosure of AI content significantly reduces the credibility of advertising, which in turn leads to lower intentions. Eisend (2006) provides a theoretical foundation for this and differentiates the source credibility dimensions into trust, competence, and presentation. In theory, transparency could increase trust in advertising, because disclosure is often associated with openness and honesty (Eisend, 2006). However, at the same time, the perception of competence suffers because consumers often judge advertising involving AI as algorithmic and less professional (Castelo et al., 2019). A second mechanism discussed is the reinforcement of perceived manipulative intentions. Wortel et al. (2024) show that disclosure of AI use in commercial contexts, such as Instagram ads, does not primarily influence credibility, but is increasingly interpreted as a signal of covert, manipulative corporate practices (Wortel et al., 2024).

These mechanisms are embedded in broader psychological processes. Castelo et al. (2019) and the algorithm aversion literature show that many consumers believe that algorithms lack the affective or emotional capabilities necessary to perform subjective tasks, thereby viewing algorithms as less effective for these tasks (Castelo et al., 2019).

Moderating factors have a major influence on disclosure effects. In addition, according to Peng et al. (2022), emotional advertising contexts reinforce resistance to algorithmic solutions, and disclosure effects become particularly apparent (Peng et al., 2022). Wortel et al. (2024) also show that high AI aversion leads to particularly negative reactions to disclosure labels, while low aversion does not cause such effects (Wortel et al., 2024).

2.9 Summary of Theoretical Background and Hypotheses

Prior research on the effects of disclosure cues in advertising shows that labels signaling the use of artificial intelligence can influence how consumers interpret persuasive messages. Studies in prosocial and social media contexts demonstrate that disclosure triggers increased cognitive vigilance, heightened skepticism, and more critical evaluations of the advertisement (Baek et al., 2024; Wortel et al., 2024). Research on technology-mediated content further suggests that disclosure cues increase processing depth and caution, even when effects on perceived accuracy remain limited (Li & Yang, 2024). In the context of AI-generated advertising (AIGC), consumers may therefore treat an “AI-generated” label as a diagnostic cue that the message was created without human judgment or emotional intentionality, which can undermine perceived credibility (Jung et al., 2025; Kim et al., 2023). Based on this body of work, a negative effect of AI disclosure on ad credibility is expected.

H1: Advertisements labeled as AI-generated are perceived as less credible than advertisements without such a disclosure.

Perceived ad credibility is a central determinant of downstream advertising outcomes. Research on persuasion and communication effectiveness demonstrates that credibility influences evaluative responses across contexts and explains substantial variance in attitudes and behavioral intentions (Eisend, 2006; Wilson & Sherrell, 1993). Brand attitude, in particular, represents a core evaluative judgment shaped by consumers’ beliefs and associations about a brand and has been shown to be a strong predictor of purchase intentions (Keller, 1993; Spears & Singh, 2004; Ghosh, 2024). When an advertisement is perceived as credible, consumers are more likely to form favorable beliefs and affective responses toward the brand, whereas lower credibility undermines these evaluations. Accordingly, perceived ad credibility is expected to exert a positive influence on brand attitude, independent of whether the advertisement contains an AI disclosure.

H2: Lower perceived advertising credibility leads to less favorable brand attitudes toward the advertised brand.

H3: Lower perceived advertising credibility leads to lower purchase intentions for the advertised product.

In addition to the main effects proposed above, consumers differ systematically in their predispositions toward artificial intelligence. General attitudes toward AI are shaped by cognitive beliefs about competence and reliability, affective reactions such as comfort or concern, and stable dispositional tendencies related to trust (Li et al., 2024; Gillath et al., 2021). AI aversion, a negatively valenced predisposition toward algorithmic decision-making, leads individuals to judge AI-based outputs more skeptically and to prefer human-generated decisions, particularly in contexts involving personalization or judgment (Dietvorst & Bartels, 2022; Yalcin et al., 2022). These individual differences suggest that the effect of AI disclosure is unlikely to be uniform across consumers. Instead, individuals high in AI aversion may interpret disclosure cues as signals of lower authenticity or competence, resulting in a stronger negative effect on perceived credibility.

H4: AI aversion moderates the effect of AI disclosure on ad credibility, such that the negative effect of the AI-generated label on perceived ad credibility is stronger for individuals high in AI aversion than for those low in AI aversion.

3 Methodology

3.1 Research design

This study examines the influence of transparency disclosures regarding the use of artificial intelligence (AI disclosures) in advertisements on consumer perception. The focus is on how such a disclosure influences ad credibility, brand attitude, and purchase intention, and to what extent individual aversion to artificial intelligence (AI aversion) moderates this relationship.

A quantitative between-subjects experiment was conducted to test this research question. Data was collected using an online questionnaire via Qualtrics. Participants were randomly assigned to one of two experimental conditions. In the first condition, with a disclosure about the use of artificial intelligence (AI Disclosure), and in the second condition, without this disclosure (No Disclosure).

The quantitative approach allows causal effects of the experimental manipulation to be identified and statistically verified (Viglia et al., 2021). The internal validity (Geuens & De Pelsmacker, 2017) of the design was ensured through the use of a fictitious brand and standardized stimuli.

The survey was distributed through various channels, including Instagram, WhatsApp group messages, direct messages, online survey exchange platforms (both paid and unpaid), as well as QR codes displayed in waiting rooms of physiotherapy practices in Vienna and Hamburg.

3.2 Procedure

Prior to the main study, a preliminary test with 20 participants was conducted to evaluate the procedure and the effectiveness of the experimental manipulation. Inspection of the manipulation check responses indicated that a substantial proportion of participants did not notice the reference to AI in the advertisement, with manipulation check failure rates being higher in the AI-disclosure condition than in the control condition. As the rating scales were displayed simultaneously with the ad, this suggested that perhaps some respondents focused primarily on the rating scales and did not pay sufficient attention to the ad itself. To increase visual attention to the stimulus and to ensure that the AI reference was noticed, a 30-second forced-exposure phase was implemented at the beginning of the experiment.

After reading the study information and providing informed consent, participants were randomly assigned to one of two experimental conditions (ad with AI-disclosure vs. ad without AI-disclosure). At the beginning of the experiment, they were first exposed to the 30-second forced-exposure page on which only the advertisement was displayed, together with a brief instruction asking them to carefully look at the ad. Once the 30 seconds had elapsed, the survey automatically advanced to the next page.

On this next page, the same advertisement remained visible at the top of the screen while participants completed the questionnaire. First, they evaluated the ad on three scales measuring ad credibility, brand attitude, and purchase intention. Afterwards, they answered the manipulation check item ("Did the ad mention anything about Artificial Intelligence?"), followed by the items assessing AI aversion and a series of demographic questions (age, gender, education level, country of residence). At the end of the survey, participants had the opportunity to provide optional open-ended feedback on the study. Participation was anonymous and voluntary, and participants could discontinue the survey at any point without negative consequences.

The average duration of participation was 7.36 minutes.

3.3 Participants

The target sample size of approximately 200–250 participants was determined a priori following the recommendations of Faul et al. (2009). This power analysis indicated that detecting medium-sized effects ($f^2 = 0.15$) with a statistical power of $1 - \beta = .80$ at $\alpha = .05$ in a between-subjects experiment with two conditions required a minimum of 200 participants. With 303 individuals completing the questionnaire, the sample exceeded this requirement. All participants took part voluntarily and anonymously, were required to be at least 18 years old, and needed sufficient English proficiency to complete the survey.

Prior to hypothesis testing, the dataset was subjected to a detailed quality assessment. During this process, unusually high correlations were observed among perceived advertising credibility, brand attitude, and purchase intention in the raw dataset ($r = .80$, $r = .76$, $r = .82$). Correlations of this magnitude between conceptually distinct constructs are atypical in behavioral survey research and may indicate non-organic or highly uniform response patterns, such as automated or mechanically generated responding.

This observation prompted a closer inspection of the data by recruitment source. Using submission timestamps as an objective, case-level criterion, responses obtained through an external recruitment platform were examined separately from organically collected responses. The activation of the survey on the recruitment platform was confirmed by an automated email, which provided a clear temporal reference point. Immediately after this activation, a large number of responses were submitted within a very short time window, resulting in a response cluster that coincided with the emergence of the anomalously high correlations.

To ensure the validity of the final dataset and to apply a transparent and reproducible exclusion rule, all responses submitted after the timestamp of the platform activation were excluded from the analysis. This procedure resulted in the removal of $n = 139$ purchased responses and yielded a final sample of $N = 125$ participants for all subsequent analyses. The exclusion criterion was based solely on submission timing and recruitment source and was independent of participants' responses to the study variables.

In a subsequent step, an attention check assessed whether participants correctly noticed the reference to artificial intelligence in the advertisement. Thirty-nine respondents (23.8% of the voluntary subsample) failed this item and were excluded, as only participants who recognized the experimental manipulation could meaningfully contribute to estimating disclosure effects. After these exclusions, the final analysis sample consisted of $N = 125$ participants.

The final sample ranged in age from 18 to 85 years ($M = 29.96$, $SD = 13.12$). Regarding gender, *58* participants identified as male (46.4%), 66 as female (52.8%), and *1* participant identified as non-binary or third gender (0.8%). Regarding country of residence, the majority lived in Germany ($n = 76$), followed by Denmark ($n = 10$), Portugal ($n = 9$), and Austria/Vienna ($n = 6$).

3.4 Stimuli and experimental conditions

The experimental stimuli consisted of two static display advertisements for a fictitious running shoe brand called NUVON. Both versions showed the same visual layout: a large image of the running shoe with the brand logo in the center, the headline "performance with support and comfort" at the top, and three short product benefit claims at the bottom ("Up to 85% energy return - Push harder, run further."; "20% less joint pressure - Comfort in every stride."; "40% recycled materials - Performance meets responsibility."). The background combined imagery of recreational runners in a park and a runner crossing a finish line in a race, in order to appeal to both leisure and performance-oriented consumers.

The ad was created using AI-based image generation and then edited to produce two experimental conditions that were strictly identical in all respects except for the AI disclosure. In the AI-disclosure condition, an additional line of text was placed in the top left corner, stating "Advertisement created using Artificial Intelligence". In the preliminary test the disclosure was placed in the bottom right corner. To increase visibility it was moved to the top left corner for the main study. In the no-disclosure condition, this text was removed while all other visual and textual elements remained unchanged. Thus, the presence versus absence of the AI disclosure served as the only manipulated feature and isolated the effect of transparency about AI involvement in ad creation.

3.5 Measuring instruments

Four measures were used to assess the central constructs of this study. All of these were established scales from marketing and communication research, either directly adopted from prior work or adapted from validated measurement frameworks. All items were presented in English and answered on a six-point Likert scale ranging from 1 = "strongly disagree" to 6 = "strongly agree".

First, ad credibility was measured using three items from Baek et al. (2024) and Wortel et al. (2024), for example: "This ad seems trustworthy." Second, brand attitude was measured using the scale developed by Spears and Singh (2004), with a sample item being: "I have a positive impression of this brand." Third, purchase intention was measured using three items adapted from Spears and Singh (2004), who conceptualize purchase intention as comprising purchase likelihood, deliberate intention, and purchase interest. Following this structure, participants rated their agreement with statements such as "I would consider buying this product," "I intend to buy this product in the future," and "I am interested in trying this product." Fourth, AI aversion was measured using four items from Logg et al. (2019), a sample item being: "I prefer advertising that is made by humans."

Furthermore, to verify the manipulation, the manipulation check asked whether the ad contained any reference to artificial intelligence ("Did the ad mention anything about artificial intelligence?"). Finally, the study also included measures for the demographic variables (age, gender, educational level, place of residence).

All scales used have been validated in the relevant literature and have shown high internal consistency in previous studies. The reliability of the scales was rechecked during data analysis.

4 Results

The statistical analysis of the collected data was performed using IBM SPSS Statistics (version 30) in combination with Hayes' PROCESS macro (version 4.3). All analyses were performed at a significance level of $\alpha = .05$.

4.1 Scale reliability and variable formation

The internal consistency of all multi-item scales was examined using Cronbach's α. Following (Tavakol & Dennick, 2011), values of $\alpha \geq .70$ were considered acceptable, $\alpha \geq .80$ good, and $\alpha \geq .90$ excellent. All scales met or exceeded these thresholds. For perceived advertising credibility, reliability was excellent in both experimental conditions ($\alpha = .91$ without disclosure, $\alpha = .89$ with disclosure). Brand attitude showed similarly high internal consistency ($\alpha = .93$ in the condition without disclosure, $\alpha = .94$ in the condition with disclosure). Purchase intention also demonstrated excellent reliability ($\alpha = .87$ without disclosure, $\alpha = .88$ with disclosure). Aversion toward artificial intelligence showed good internal consistency across the full sample ($\alpha = .72$; $N = 125$). Because all scales demonstrated satisfactory reliability, composite scores were calculated by averaging the respective items and used in all subsequent analyses.

4.2 Descriptive statistics and correlations

Before analyzing the specific hypotheses, descriptive statistics (*M*, *SD*, and frequency distributions for all variables) were computed to provide an initial overview of the sample and the study constructs. In addition, bivariate Pearson correlations were examined to obtain preliminary indications of the direction and strength of associations between the central variables (see Tables 1 and 2).

Participants reported moderately high perceptions of the advertisement's credibility and generally positive evaluations of the brand, while purchase intentions were more moderate.

Average levels of aversion toward artificial intelligence were also in the mid-range. Table 1 displays the minimum, maximum, mean, and standard deviation for each variable based on the cleaned dataset (N = 125).

The correlation matrix (Table 2) provides an overview of the bivariate relationships between the variables. As expected, perceived credibility was strongly and positively correlated with both brand attitude (r = .70, p < .001) and purchase intention (r = .70, p < .001). Brand attitude and purchase intention also showed a strong positive association (r = .73, p < .001). Aversion toward artificial intelligence correlated negatively with credibility, brand attitude, and purchase intention (r = –.18 to –.25, all p < .05). The disclosure condition showed small negative correlations with credibility (r = –.17, p = .061) and brand attitude (r = –.18, p = .041). Correlations with demographic variables were generally small.

Table 1 - Descriptive Statistics for Main Study Variables (N = 125)

Variable	**Min**	**Max**	***M***	***SD***
Perceived credibility	1.00	6.00	3.79	1.14
Brand attitude	1.00	6.00	3.90	1.06
Purchase intention	1.00	6.00	3.32	1.16
AI aversion	1.00	6.00	4.12	0.91
AI disclosure (1 = label)	0.00	1.00	0.46	N/A
Country (1 = Germany)	0.00	1.00	0.60	N/A
Higher education (1 = high)	0.00	1.00	0.33	N/A
Age (years)	18.00	85.00	29.96	13.12

Table 2 - Bivariate Pearson correlations of all study variables

Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1. Perceived credibility	—								
2. Brand attitude	**.70**	—							
3. Purchase intention	**.70**	**.73**	—						
4. AI aversion	**−.18**	**−.25**	**−.31**	—					
5. Gender	−.08	.03	−.07	−.07	—				
6. Country (1 = Germany)	.10	.10	.07	−.17	.13	—			
7. Higher education	−.05	−.05	.00	−.01	.04	.07	—		
8. Age (years)	.05	−.02	0.5	−.14	.12	**.23**	**.25**	—	
9. AI disclosure (1 = label)	−.17	**−.18**	−.12	.17	−.13	−.08	.04	−.12	—

Note. Correlation p < .05 (two-tailed) for all correlations

4.3 **Hypothesis Testing**

4.3.1 H1: Influence of the AI disclaimer on the perceived credibility of advertising

A linear regression was conducted to test whether the AI-disclosure reduces perceived credibility. In contrast to expectations, the model did not show a statistically significant effect of the disclosure, $b = -0.38$, $SE = 0.20$, $p = .061$, 95% CI [–0.79, 0.02]. Participants who viewed the ad without the disclosure rated the ad slightly more credible ($M = 3.96$, $SD = 1.14$) than participants who viewed the version with the disclosure ($M = 3.84$, $SD = 1.14$), but this difference did not reach statistical significance.

Thus, **H1** was not supported, although the descriptive pattern was in the expected direction, the AI-disclosure did not significantly decrease perceived credibility.

Table 3 - Regression testing the effect of the AI-disclosure on perceived credibility

Effect	**Estimate**	***SE***	**95% CI**		***p***
			LL	**UL**	
Intercept	3.76	0.13	3.70	4.22	< .001
AI-disclosure (1 = yes)	–0.38	0.20	–0.79	0.02	.061

4.3.2 H2: Relationship between Ad Credibility and Brand Attitude

To test the second hypothesis, a hierarchical multiple regression analysis was conducted with brand attitude as the dependent variable. In the first step, the experimental condition (AI disclosure vs. no disclosure) and demographic covariates (age, gender, education level, and country of residence) were entered. This model did not explain a significant proportion of variance in brand attitude, $R^2 = .046$, $F(5, 119) = 1.16$, $p = .335$.

In the second step, perceived advertising credibility was added to the model. The inclusion of credibility led to a substantial increase in explained variance, $\Delta R^2 = .461$, F-change(1, 118) = 110.43, $p < .001$, resulting in a total R^2 of .507 (adjusted $R^2 = .482$). Perceived advertising credibility emerged as a strong positive predictor of brand attitude, $b = 0.65$, $SE = 0.06$, $\beta = .70$, $t = 10.51$, $p < .001$. None of the control variables showed a significant effect in the final model. Taken together, these results support **H2**, indicating that lower perceived advertising credibility is associated with less favorable brand attitudes toward the advertised brand.

Table 4 - Hierarchical regression predicting brand attitude

Predictor	*b*	*SE*	*t*	*p*
Model 1				
AI disclosure (0 = no, 1 = yes)	−0.09	0.14	−0.63	.531
Age	0.01	0.01	0.78	.436
Gender (1 = male)	0.18	0.17	1.06	.292
High education	0.12	0.16	0.75	.456
Germany	−0.21	0.18	−1.17	.244
Model 2				
Advertising credibility	**0.65**	**0.06**	**10.51**	**< .001**
AI disclosure	−0.07	0.10	−0.72	.468
Age	0.00	0.01	0.34	.736
Gender	0.10	0.12	0.83	.410
High education	0.05	0.11	0.47	.641
Germany	−0.08	0.13	−0.61	.543

Note. $N = 124$. Model 1: $R^2 = .05$. Model 2: $R^2 = .51$, $\Delta R^2 = .46$, F-change(1, 118) = 110.43, $p < .001$. Unstandardized coefficients are reported.

4.3.3 H3 Relationship between Ad credibility and purchase intention

To test the third hypothesis, a hierarchical multiple regression analysis was conducted with purchase intention as the dependent variable. In the first step, the experimental condition (AI disclosure vs. no disclosure) and demographic covariates (age, gender, education level, and country of residence) were entered. This model did not explain a significant proportion of variance in purchase intention, $R^2 = .03$, $F(5, 119) = 0.72$, $p = .613$.

In the second step, perceived advertising credibility was added to the model. The inclusion of credibility resulted in a substantial increase in explained variance, $\Delta R^2 = .46$, F-change(1, 118) = 106.28, $p < .001$, yielding a total R^2 of .49 (adjusted $R^2 = .46$). Perceived advertising credibility emerged as a strong positive predictor of purchase intention, $b = 0.71$, $SE = 0.07$, $\beta = .70$, $t = 10.31$, $p < .001$. None of the control variables showed a significant effect in the final model.

Taken together, these results support H3, indicating that lower perceived advertising credibility is associated with lower purchase intentions for the advertised product.

Table 5 - Hierarchical regression predicting purchase intention

Predictor	*b*	*SE*	*t*	*p*
Model 1				
AI disclosure (0 = no, 1 = yes)	-0.29	0.22	-1.35	.181
Age	0.00	0.01	0.40	.688
Gender (1 = male)	-0.24	0.21	-1.11	.270
High education	-0.02	0.23	-0.07	.941
Germany	0.16	0.22	0.72	.475
Model 2				
Advertising credibility	**0.71**	**0.07**	**10.31**	**<.001**
AI disclosure	-0.01	0.16	-0.09	.933
Age	0.00	0.01	0.25	.802
Gender (1 = male)	-0.05	0.16	-0.32	.752
High education	0.06	0.17	0.37	.712
Germany	-0.01	0.16	-0.05	.961

Note. $N = 124$. Model 1: $R^2 = .03$. Model 2: $R^2 = .49$, $\Delta R^2 = .46$, F-change(1, 118) = 106.28, $p < .001$. Unstandardized coefficients are reported.

Supplementary mediation analyses (PROCESS Model 4) were conducted to explore whether AI disclosure indirectly influenced brand attitude and purchase intention via perceived advertising credibility. Detailed results are reported in the Appendix B and C.

4.3.4 H4 Moderation by AI aversion

To test the fourth hypothesis, a moderation analysis was conducted using PROCESS Model 1 with perceived advertising credibility as the dependent variable. AI disclosure (0 = no disclosure, 1 = AI disclosure), AI aversion, and their interaction term were specified as predictors. In line with the data analysis plan, gender, country of residence, education level, and age were included as covariates.

The overall model did not explain a significant proportion of variance in perceived advertising credibility, $R^2 = .07$, $F(7, 117) = 1.35$, $p = .234$. Importantly, the interaction between AI disclosure and AI aversion was not statistically significant, $b = -0.06$, $SE = 0.23$, $t = -0.28$, $p = .781$, and did not account for additional explained variance, $\Delta R^2 = .001$. Neither the main effect of AI disclosure nor the main effect of AI aversion reached statistical significance in the model. Taken together, these results do not support H4. A supplementary moderation analysis, excluding demographic covariates yielded the same non-significant interaction pattern and is reported in Appendix D.

Because AI aversion did not significantly moderate the effect of AI disclosure on perceived advertising credibility, no further conditional process effects were expected. Nevertheless, exploratory moderated mediation analyses combining the mediation models (H2 and H3) with AI aversion as a moderator (PROCESS Model 7) were conducted as a robustness check and are reported in Appendix E.

Table 6 - Moderation analysis predicting perceived advertising credibility

Predictor	*b*	*SE*	*t*	*p*
Intercept	4.12	0.62	6.65	< .001
AI disclosure (0 = no, 1 = yes)	−0.29	0.24	−1.21	.228
AI aversion	−0.08	0.12	−0.67	.505
AI disclosure × AI aversion	−0.06	0.23	−0.28	.781
Age	0.01	0.01	1.14	.258
Gender	−0.09	0.18	−0.49	.624
Education	0.04	0.11	0.36	.721
Country	0.12	0.20	0.60	.551

Note. *N* = 125. Entries are unstandardized regression coefficients. The moderation analysis was conducted using PROCESS Model 1. Gender, education level, country of residence, and age were included as covariates. ΔR^2 for the interaction term = .001.

To evaluate the robustness of the findings, all primary analyses were also performed using the full dataset before the exclusion of suspicious responses (*N* = 241). In this unfiltered dataset, a significant negative effect of AI disclosure on perceived advertising credibility emerged in the simple regression model corresponding to H1. However, this effect was not significant in the cleaned dataset (*N* = 125) and did not remain stable across models that included additional predictors.

The overall pattern of results was consistent across datasets. Perceived advertising credibility strongly predicted both brand attitude and purchase intention. AI aversion demonstrated a negative main effect on credibility but did not moderate the effect of disclosure. The stronger and more homogeneous effects observed in the full dataset indicate reduced response variability, which supports the decision to rely on the cleaned dataset for hypothesis testing. Detailed results of the robustness analyses are provided in Appendix F.

5 Discussion

5.1 Summary of results

This study investigated the impact of AI disclosure in advertising on perceived advertising credibility and subsequent consumer responses, as well as the potential moderating effect of individual AI aversion. Grounded in persuasion theory and source credibility research, the conceptual framework posited that disclosure cues indicating AI-generated content could diminish perceived credibility, thereby affecting brand attitude and purchase intention (Wilson & Sherrell, 1993; Eisend, 2006).

Contrary to H1, AI disclosure did not significantly reduce perceived advertising credibility. Although the effect was not statistically significant, the negative coefficient indicated a descriptive trend in the hypothesized direction. This result diverges from previous research suggesting that disclosure cues can trigger cognitive vigilance and skepticism in persuasive contexts (Baek et al., 2024; Wortel et al., 2024), yet aligns with studies demonstrating that transparency cues do not necessarily diminish perceived message quality or credibility when content quality is maintained (Li & Yang, 2024). Collectively, these findings indicate that AI disclosure may not serve as a universally negative diagnostic cue, and any effects are likely to be small or context-dependent (Kim et al., 2023).

The results provided strong support for H2 and H3. Perceived advertising credibility was a robust and substantial predictor of both brand attitude and purchase intention, even after controlling for AI disclosure and sociodemographic variables. This outcome is consistent with classical source credibility and persuasion models, which identify credibility as a central determinant of evaluative and behavioral responses to advertising messages (Wilson & Sherrell, 1993; Eisend, 2006). Furthermore, brand attitude and purchase intention were highly positively correlated in this sample ($r = .73$, $p < .001$), corroborating previous findings that favorable brand evaluations are linked to behavioral intentions (Keller, 1993; Spears & Singh, 2004).

H4, which posited a moderating role of AI aversion, was not supported. AI aversion did not significantly moderate the relationship between AI disclosure and perceived advertising credibility. Although previous research indicates that individuals with higher AI aversion generally evaluate algorithmic outputs more negatively (Dietvorst & Bartels, 2022; Yalcin et al., 2022), the current findings suggest that these predispositions did not intensify the effect of AI disclosure cues in this advertising context (Gillath et al., 2021; Li et al., 2024).

Supplementary mediation and moderated mediation analyses reinforced this interpretation. Neither a simple mediation of AI disclosure through perceived advertising credibility nor a moderated mediation involving AI aversion was detected. Across all models, the results remained consistent when controlling for sociodemographic variables, which did not exhibit significant or consistent effects. Overall, these findings indicate that perceived advertising credibility operates as an independent and central driver of downstream outcomes, while AI disclosure and AI aversion did not generate additional causal pathways in this study.

Non-significant effects should not be interpreted as definitive evidence for the absence of an effect. Methodological research highlights that null findings may result from either a true lack of effect or insufficient statistical power to detect small effects, especially in interaction and mediation models (Cohen, 1988). Given the sample size in this study, the analyses were primarily powered to detect moderate effects, indicating that any effects of AI disclosure are likely to be small and context-dependent rather than strong and universal.

5.2 Implications

The results of this study indicate that managerial decision-making in AI-supported advertising should prioritize design elements that enhance perceived credibility over the technological origin of advertising content. Previous research in persuasion and advertising demonstrates that credibility-enhancing cues, including clear argumentation, coherent messaging, and perceived authenticity, are essential drivers of consumer evaluations (Wilson & Sherrell, 1993; Eisend, 2006). Practically, this finding suggests that investments in AI-supported advertising should focus on message execution, rather than on concerns regarding potential consumer penalties for AI involvement.

Furthermore, the findings indicate that transparency regarding AI use does not necessitate extensive risk management or avoidance strategies. While disclosure cues may prompt increased scrutiny in certain contexts (Baek et al., 2024; Wortel et al., 2024), recent studies show that these cues do not consistently undermine persuasion outcomes when the overall message quality is high (Kim et al., 2023; Li & Yang, 2024). Therefore, practitioners can implement transparent communication about AI use while reinforcing credibility through effective message framing and high-quality content.

The lack of differential effects across consumer predispositions suggests that highly segmented disclosure strategies based on AI-related attitudes are unlikely to provide significant additional value. Although individual differences in AI aversion can influence general evaluations of algorithmic systems (Dietvorst & Bartels, 2022; Yalcin et al., 2022), advertising contexts are primarily shaped by universal credibility cues. As a result, a streamlined strategy emphasizing broadly effective persuasion principles may be more advantageous than complex audience segmentation based on AI attitudes (Gillath et al., 2021).

Finally, the findings highlight the importance of treating AI disclosure as a contextual design choice rather than a fixed rule. Prior work suggests that the effects of transparency cues depend on situational factors and audience expectations (Kim et al., 2023; Li & Yang, 2024). For practitioners, this underscores the value of iterative testing and contextual adaptation when deploying AI-supported advertising, rather than relying on one-size-fits-all disclosure policies.

5.3 Limitations and future Research

This research has limitations that need to be considered when analyzing the results. First, although the experimental design ensured that all participants were exposed to the advertisement, the study relied on a single controlled exposure. However, advertising contacts vary greatly in terms of attention, repetition, and context. Future research should examine whether the observed relationships also exist with repeated exposure or in more realistic environments where advertising messages compete with other content for attention.

Second, the study examined only consciously processed AI disclosures by analyzing participants who correctly identified the disclosure label in the manipulation check. While this approach enhances internal validity, it limits generalizability to contexts where AI disclosures are not consciously recognized. Since disclosure information may be ignored or only superficially processed, future research should investigate the effects of AI disclosure under conditions of superficial processing and assess whether self-reported purchase intention remains a valid outcome. Although purchase intention is a widely used and theoretically important measure, prior research indicates that intentions do not always translate into actual purchase behavior (Spears & Singh, 2004). Future studies should therefore complement self-reported measures with behavioral indicators to capture the downstream effects of AI-generated advertising more accurately.

Third, the sample size was reduced due to a rigorous data cleaning procedure that excluded responses based on objective criteria related to recruitment source and submission time. While this improved data quality, it also decreased statistical power, especially for detecting smaller interactions or indirect effects. Future research with larger samples could more precisely examine moderation and mediation processes.

Finally, the generalizability of the findings is constrained by the sample composition, as a large proportion of participants were from Germany. This may limit cultural generalizability. Variations in attitudes toward technology, data protection, and automation could influence responses to AI-generated advertising and disclosure notices. Future research should investigate these relationships in diverse national contexts to assess their robustness.

6 Conclusion

This study examined how labeling advertisements as “AI-generated” influences perceived ad credibility and downstream judgments such as brand attitude and purchase intent, and to what extent AI aversion moderates this relationship. Contrary to initial expectations, it was found that an explicit AI disclosure does not significantly reduce perceived credibility in a commercial, medium-involvement product context, while credibility itself has a clear, substantial influence on brand attitude and purchase intention. AI aversion also did not prove to be an effective moderator of the disclosure effect, suggesting that individual reservations about AI in this setting translate less strongly into specific reactions to transparency labels than assumed.

This work thus makes a twofold contribution: First, existing, sometimes contradictory findings on AI disclosure in prosocial and social media contexts are supplemented with evidence from an everyday commercial advertising context, showing that disclosure effects tend to be small and context-dependent rather than strong and universal. Second, it empirically confirms the central role of ad credibility in persuasion and source credibility theories by showing that it acts as a key factor in brand evaluation and purchase intention, regardless of whether the ad is labeled as AI-generated. For research and practice, this suggests that AI disclosure should be understood less as an inherent risk and more as a malleable contextual variable, and that the focus should be on designing AI-supported advertising in a way that fosters credibility.

7 References

American Marketing Association. (2025). *American Marketing Association*. https://www.ama.org/the-definition-of-marketing-what-is-marketing/

Baek, T. H., Kim, J., & Kim, J. H. (2024). Effect of disclosing AI-generated content on prosocial advertising evaluation. *International Journal of Advertising*, *0*(0), 1–22. https://doi.org/10.1080/02650487.2024.2401319

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

Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, *48*(1), 24–42. https://doi.org/10.1007/s11747-019-00696-0

Dietvorst, B. J., & Bartels, D. M. (2022). Consumers Object to Algorithms Making Morally Relevant Tradeoffs Because of Algorithms' Consequentialist Decision Strategies. *Journal of Consumer Psychology*, *32*(3), 406–424. https://doi.org/10.1002/jcpy.1266

Eisend, M. (2006). *Source Credibility Dimensions in Marketing Communication – A Generalized Solution*.

European Union: Transparency Obligations for Providers and Deployers of Certain AI Systems | EU Artificial Intelligence Act. (2024). https://artificialintelligenceact.eu/article/50/

Friestad, M., & Wright, P. (1994). The Persuasion Knowledge Model: How People Cope with Persuasion Attempts. *Journal of Consumer Research*, *21*(1), 1–31.

Geuens, M., & De Pelsmacker, P. (2017). Planning and Conducting Experimental Advertising Research and Questionnaire Design. *Journal of Advertising*, *46*(1), 83–100. https://doi.org/10.1080/00913367.2016.1225233

Ghosh, M. (2024). Meta-analytic review of online purchase intention: Conceptualising the study variables. *Cogent Business & Management*, *11*(1). https://doi.org/10.1080/23311975.2023.2296686

Gillath, O., Ai, T., Branicky, M. S., Keshmiri, S., Davison, R. B., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. *Computers in Human Behavior*, *115*, 106607. https://doi.org/10.1016/j.chb.2020.106607

Haleem, A., Javaid, M., Asim Qadri, M., Pratap Singh, R., & Suman, R. (2022). Artificial intelligence (AI) applications for marketing: A literature-based study. *International Journal of Intelligent Networks*, *3*, 119–132. https://doi.org/10.1016/j.ijin.2022.08.005

Heuking. (2025). *Artificial Intelligence: These Transparency Obligations must be observed*. Heuking. https://www.heuking.de/en/news-events/newsletter-articles/detail/artificial-intelligence-these-transparency-obligations-must-be-observed.html

Hovland, C. I., & Weiss, W. (1951). The Influence of Source Credibility on Communication Effectiveness. *The Public Opinion Quarterly*, *15*(4), 635–650.

Jung, T., Koghut, M., Lee, E., & Kwon, O. (2025). Artificial creativity in luxury advertising: How trust and perceived humanness drive consumer response to AI-generated content. *Journal of Retailing and Consumer Services*, *87*, 104403. https://doi.org/10.1016/j.jretconser.2025.104403

Keller, K. L. (1993). *Conceptualizing, Measuring, and Managing Customer-Based Brand Equity*.

Lafferty, B. A., & Goldsmith, R. E. (1999). Corporate Credibility's Role in Consumers' Attitudes and Purchase Intentions When a High versus a Low Credibility Endorser Is Used in the Ad. *Journal of Business Research*, *44*(2), 109–116. https://doi.org/10.1016/S0148-2963(98)00002-2

Li, T.-G., Zhang, C.-B., Chang, Y., & Zheng, W. (2024). The impact of AI identity disclosure on consumer unethical behavior: A social judgment perspective. *Journal of Retailing and Consumer Services*, *76*, 103606. https://doi.org/10.1016/j.jretconser.2023.103606

McKinsey. (2024). https://www.mckinsey.com/de/branchen/konsumguter-handel/akzente/akzente-2-2024/2024-2-ki-marketing

Moharam, M. M. R., & Tawalbeh, A. (2025). The role of gen AI in enhancing creativity and efficiency in content marketing creation: Scoping review and future insights. *International Journal of Innovative Research and Scientific Studies*, *8*(1), 2804–2814. https://doi.org/10.53894/ijirss.v8i1.5060

Mudler, D. (2004). *The evolution of marketing communication: From selling to integration*. https://doi.org/10.10520/EJC27634

Park, H. Y., & Chang, S. R. (2022). When and how brands affect importance of product attributes in consumer decision process. *European Journal of Marketing*, *56*(13), 1–25. https://doi.org/10.1108/EJM-09-2020-0650

Puntoni, S., Reczek, R., Giesler, M., & Botti, S. (2020). Consumers and Artificial Intelligence: An Experiential Perspective. *Journal of Marketing*, *85*, 002224292095384. https://doi.org/10.1177/0022242920953847

Roberts, J. H., Kayande, U., & Stremersch, S. (2014). From academic research to marketing practice: Exploring the marketing science value chain. *International Journal of Research in Marketing*, *31*(2), 127–140. https://doi.org/10.1016/j.ijresmar.2013.07.006

Sethuraman, R., Tellis, G. J., & Briesch, R. A. (2011). How Well Does Advertising Work? Generalizations from Meta-Analysis of Brand Advertising Elasticities. *Journal of Marketing Research*, *48*(3), 457–471. https://doi.org/10.1509/jmkr.48.3.457

Spears, N., & Singh, S. N. (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

Verhoef, P. C., Stephen, A. T., Kannan, P. K., Luo, X., Abhishek, V., Andrews, M., Bart, Y., Datta, H., Fong, N., Hoffman, D. L., Hu, M. M., Novak, T., Rand, W., & Zhang, Y. (2017). Consumer Connectivity in a Complex, Technology-enabled, and Mobile-oriented World with Smart Products. *Journal of Interactive Marketing*, *40*(1), 1–8. https://doi.org/10.1016/j.intmar.2017.06.001

Viglia, G., Zaefarian, G., & Ulqinaku, A. (2021). How to design good experiments in marketing: Types, examples, and methods. *Industrial Marketing Management*, *98*, 193–206. https://doi.org/10.1016/j.indmarman.2021.08.007

Wilson, E. J., & Sherrell, D. L. (1993). Source effects in communication and persuasion research: A meta-analysis of effect size. *Journal of the Academy of Marketing Science*, *21*(2), 101–112. https://doi.org/10.1007/BF02894421

Yalcin, Lim, Puntoni, & van Osselear. (2022). *Thumbs Up or Down: Consumer Reactions to Decisions by Algorithms Versus Humans*. https://doi.org/10.1177/00222437211070016

Zhang, J. Z., & Watson IV, G. F. (2020). Marketing ecosystem: An outside-in view for sustainable advantage. *Industrial Marketing Management*, *88*, 287–304. https://doi.org/10.1016/j.indmarman.2020.04.023

8 Appendix

8.1 **Appendix A -** Full Questionnaire of Study

A1. Introduction Text

Thank you for taking part in this academic study about how people perceive and evaluate product advertisements. This research is conducted as part of a Master's thesis at Católica Lisbon School of Business & Economics (Universidade Católica Portuguesa). The study aims to gain a better understanding of how consumers form impressions and attitudes toward modern advertising. You will be shown one short advertisement for a fictitious brand and then asked to answer a few short questions about your impressions and opinions. There are no right or wrong answers - we are only interested in your honest opinion. Participation will take approximately 3-4 minutes. The survey is completely anonymous and voluntary. No identifying or personal information will be collected, and all responses will be used solely for academic research purposes. The study does not involve any sensitive content. Data will be treated confidentially and processed in full compliance with the General Data Protection Regulation (GDPR). To ensure ethical compliance, participants must be at least 18 years old to take part. If you have any questions about the study, please contact the researcher at: s-majung@ucp.pt By clicking the button at the end of the page, you confirm that you are 18 years or older, that you have read the above information, and that you voluntarily agree to participate in this study

A2. Stimulus Exposure Screen

Please take a moment to carefully watch the following advertisement. This is the only part of the survey that includes a short time limit. After 30 seconds, you will be able to proceed to the next section and answer the related questions at your own pace.

A3 Advertisement without AI-Disclosure

A4 Advertisement with AI-Disclosure

A5 Ad Credibility (6-point Likert scale)

Please indicate how you evaluate the following advertisement

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Somewhat agree (4)	Agree (5)	Strongly agree (6)
This ad seems trustworthy (1)	○	○	○	○	○	○
The information in this ad appears reliable (2)	○	○	○	○	○	○
This ad seems honest (3)	○	○	○	○	○	○

A6 Brand Attitude (6-point Likert scale)

Please indicate your attitude toward the brand in the ad

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Somewhat agree (4)	Agree (5)	Strongly agree (6)
I have a positive impression of this brand (1)	○	○	○	○	○	○
I find this brand appealing (2)	○	○	○	○	○	○
Overall, I like this brand (3)	○	○	○	○	○	○

A7 Purchase Intention (6-point Likert scale)

Please take a close look at the advertisement below and answer the following questions based on your impression of it

Q15 Please indicate how likely you are to consider buying the product shown

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Somewhat agree (4)	Agree (5)	Strongly agree (6)
I would consider buying this product (1)	○	○	○	○	○	○
I intend to buy this product in the future (2)	○	○	○	○	○	○
I am interested in trying this product (3)	○	○	○	○	○	○

A8 Manipulation Check

Did the ad mention anything about Artificial Intelligence?

○ Yes (1)

○ Maybe (2)

○ No (3)

A9 AI Aversion

Please indicate your general attitude toward the use of AI in creative tasks (not about this specific ad)

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Somewhat agree (4)	Agree (5)	Strongly agree (6)
I feel uncomfortable when AI is used to create advertising (1)	○	○	○	○	○	○
I trust AI less than human creators (2)	○	○	○	○	○	○
I prefer advertising that is made by humans (3)	○	○	○	○	○	○
Using AI in creative work makes me uncomfortable (4)	○	○	○	○	○	○

A10 Demographics

Finally, a few background questions.

Q19 How old are you?

____

Q21 What is your gender?

- ○ Male (1)
- ○ Female (2)
- ○ Non-binary / third gender (3)
- ○ Prefer not to say (4)

Q23 Where do you live?

____

Q24 What is your highest level of education completed?

- ○ Less than high school (1)
- ○ High school diploma / secondary school (2)
- ○ Some college or vocational training (3)
- ○ Bachelor's degree (4)
- ○ Master's degree (5)
- ○ Doctorate (PhD or equivalent) (6)
- ○ Other (7)

Q21 Do you have any comments or feedback about this study or anything that felt unclear or confusing?

8.2 **Appendix B -** Supplementary mediation analysis predicting brand attitude

This appendix reports the results of a supplementary mediation analysis conducted using PROCESS Model 4 (Hayes, 2022) to examine whether AI disclosure indirectly influenced brand attitude via perceived advertising credibility. Indirect effects were estimated using 5,000 bootstrap samples.

Effect	***b***	***SE***	***t***	***p***	**95% CI**	
					LL	**UL**
Path a: Disclosure → Credibility	−0.38	0.20	−1.89	.061	−0.79	0.02
Path b: Credibility → Brand attitude	0.64	0.06	10.57	< .001	0.52	0.76
Direct effect (c′): Disclosure → Brand attitude	−0.14	0.14	−1.04	.303	−0.42	0.13
Indirect effect (a × b)	−0.25	0.14	—	—	−0.54	-0.01

Note. $N = 125$. Indirect effects were estimated using 5,000 bootstrap samples. Confidence intervals are bias-corrected.

8.3 **Appendix C - Supplementary mediation analysis predicting purchase intention (PROCESS Model 4)**

This appendix reports the results of a supplementary mediation analysis conducted using PROCESS Model 4 (Hayes, 2022) to examine whether AI disclosure indirectly influenced purchase intention via perceived advertising credibility. Indirect effects were estimated using 5,000 bootstrap samples.

Effect	***b***	***SE***	***t***	***p***	**95% CI**	
					LL	**UL**
Path a: Disclosure → Credibility	−0.38	0.20	−1.89	.061	−0.79	0.02
Path b: Credibility → Purchase intention	0.71	0.07	10.62	< .001	0.58	0.84
Direct effect (c′): Disclosure → Purchase intention	−0.01	0.15	−0.05	.960	−0.31	0.30
Indirect effect (a × b)	−0.27	0.15	—	—	−0.57	-0.01

Note. $N = 125$. Indirect effects were estimated using 5,000 bootstrap samples. Confidence intervals are bias-corrected.

8.4 **Appendix D -** Supplementary Moderation Analyses Examining AI Aversion as a Moderator of the Effect of AI Disclosure on Advertising Credibility (PROCESS Model 1)

Moderation Analysis Predicting Advertising Credibility (PROCESS Model 1, Without Covariates)

Predictor	*b*	*SE*	*t*	*p*
Intercept	4.18	0.52	8.04	< .001
AI disclosure (0 = no, 1 = yes)	−0.30	0.23	−1.32	.189
AI aversion	−0.09	0.11	−0.80	.424
AI disclosure × AI aversion	−0.00	0.23	−0.00	.999

Note. N = 125. Entries are unstandardized regression coefficients. The moderation analysis was conducted using PROCESS Model 1. No demographic covariates were included.

8.5 **Appendix E -** Exploratory moderated mediation analyses (PROCESS Model 7)

This appendix reports exploratory moderated mediation analyses (PROCESS Model 7) examining whether the indirect effect of AI disclosure on brand attitude and purchase intention via perceived advertising credibility is moderated by AI aversion. All models were estimated without demographic covariates.

Table E1
Moderated mediation analysis predicting brand attitude (PROCESS Model 7)

Mediator model (Outcome: Advertising credibility)

Predictor	*b*	*SE*	*t*	*p*
AI disclosure (X)	−0.32	0.96	−0.33	.738
AI aversion (W)	−0.20	0.15	−1.32	.188
X × W	−0.00	0.23	−0.00	.999

$R^2 = .05$, $F(3, 121) = 2.27$, $p = .084$

Outcome model (Outcome: Brand attitude)

Predictor	*b*	*SE*	*t*	*p*
AI disclosure (X)	−0.14	0.14	−1.04	.303
Advertising credibility (M)	0.64	0.06	10.57	< .001

$R^2 = .50$, $F(2, 122) = 59.88$, $p < .001$

Conditional indirect effects of AI disclosure on brand attitude via advertising credibility

AI aversion	Indirect effect	BootSE	Boot CI	
			LL	UL
Low (16th percentile)	−0.21	0.17	−0.55	0.12
Mean (50th percentile)	−0.21	0.13	−0.49	0.05
High (84th percentile)	−0.21	0.21	−0.65	0.18
Index of moderated mediation: −0.00, BootSE = 0.15, BootLLCI = −0.30, BootULCI = 0.29				

Note. N = 125. Entries are unstandardized regression coefficients. Bootstrap confidence intervals are based on 5,000 samples. Confidence level = 95%. AI aversion values correspond to the 16th, 50th, and 84th percentiles.

Table E2

Moderated mediation analysis predicting purchase intention (PROCESS Model 7)

Mediator model (Outcome: Advertising credibility)

Predictor	*b*	*SE*	*t*	*p*
AI disclosure (X)	−0.32	0.96	−0.33	.738
AI aversion (W)	−0.20	0.15	−1.32	.188
X × W	−0.00	0.23	−0.00	.999

$R^2 = .05$, $F(3, 121) = 2.27$, $p = .084$

Outcome model (Outcome: Purchase intention)

Predictor	*b*	*SE*	*t*	*p*
AI disclosure (X)	−0.01	0.15	−0.05	.960
Advertising credibility (M)	0.71	0.07	10.62	< .001

$R^2 = .49, F(2, 122) = 58.08, p < .001$

Conditional indirect effects of AI disclosure on purchase intention via advertising credibility

AI aversion	Indirect effect	BootSE	Boot CI	
			LL	UL
Low (16th percentile)	−0.23	0.18	−0.61	0.12
Mean (50th percentile)	−0.23	0.14	−0.52	0.04
High (84th percentile)	−0.23	0.22	−0.66	0.20
Index of moderated mediation: −0.00, BootSE = 0.17, BootLLCI = −0.33, BootULCI = 0.33				

Note. *N* = 125. Entries are unstandardized regression coefficients. Bootstrap confidence intervals are based on 5,000 samples. Confidence level = 95%. AI aversion values correspond to the 16th, 50th, and 84th percentiles.

Table E3
Moderated Mediation Analysis Predicting Brand Attitude Including Covariates (PROCESS Model 7)

Predictor	*b*	*SE*	*t*	*p*
Mediator model (Outcome: Advertising credibility)				
AI disclosure (X)	−0.07	0.98	−0.07	.942
AI aversion (W)	−0.16	0.16	−1.06	.292
X × W	−0.06	0.23	−0.28	.781
Gender	−0.28	0.21	−1.35	.180
Country (Germany)	0.21	0.22	0.97	.336
Higher education	−0.10	0.22	−0.46	.646
Age	0.00	0.01	0.15	.878
Outcome model (Outcome: Brand attitude)				
AI disclosure (X)	−0.13	0.14	−0.91	.363
Advertising credibility (M)	0.65	0.06	10.51	< .001
Gender	0.18	0.14	1.26	.209
Country (Germany)	0.07	0.15	0.46	.647
Higher education	−0.03	0.15	−0.17	.863
Age	−0.01	0.01	−1.08	.281
Index of moderated mediation: b = −0.04, BootSE = 0.16, 95% CI [−0.37, 0.26]				

Note. *N* = 125. Unstandardized regression coefficients are reported. Bootstrap confidence intervals are based on 5,000 samples. Gender, country of residence, education level, and age were included as covariates. None of the conditional indirect effects were statistically significant.

8.6 **Appendix F -** Robustness checks using the full dataset (N = 241)

Tabel F1
Regression Analysis Predicting Advertising Credibility (H1)

Predictor	*b*	*SE*	*t*	*p*
Intercept	4.25	0.10	43.63	< .001
AI disclosure (0 = no, 1 = yes)	−0.42	0.15	−2.76	.006

Note. *N* = 241. Unstandardized regression coefficients are reported. Dependent variable: advertising credibility.

Table F2
Hierarchical Regression Predicting Brand Attitude (H2)

Model 1 (Controls only)

Predictor	*b*	*SE*	*t*	*p*
AI disclosure	−0.40	0.15	−2.75	.006
Gender	−0.37	0.15	−2.51	.013
Germany	−0.22	0.16	−1.41	.160
Higher education	0.19	0.16	Jan 22	.226

$R^2 = .07$, $F(4, 236) = 4.31$, $p = .002$

Model 2 (Advertising credibility added)

Predictor	***b***	***SE***	***t***	***p***
Advertising credibility	0.78	0.04	19.41	< .001
AI disclosure	−0.07	0.09	−0.70	.487
Gender	−0.08	0.09	−0.88	.380
Germany	−0.02	0.10	−0.18	.857
Higher education	−0.04	0.10	−0.41	.685

$R^2 = .64$, $\Delta R^2 = .57$, $F(5, 235) = 84.28$, $p < .001$
Note. *N* = 241. Unstandardized coefficients reported. Dependent variable: brand attitude.

Table F3
Purchase Intention Regression (H3)

Model 1 (Controls only)

Predictor	***b***	***SE***	***t***	***p***
AI disclosure	−0.35	0.17	−2.09	.038
Gender	−0.54	0.17	−3.21	.002
Germany	−0.44	0.18	−2.46	.015
Higher education	0.35	0.18	1.96	.051

$R^2 = .10$, $F(4, 236) = 6.25$, $p < .001$

Model 2 (Advertising credibility added)

Predictor	***b***	***SE***	***t***	***p***
Advertising credibility	0.84	0.05	16.68	< .001
AI disclosure	0.01	0.12	0.12	.908
Gender	−0.23	0.12	−2.00	.047
Germany	−0.22	0.12	−1.83	.069
Higher education	0.10	0.12	0.86	.394

$R^2 = .59$, $\Delta R^2 = .49$, $F(5, 235) = 66.49$, $p < .001$
Note. *N* = 241. Unstandardized coefficients reported. Dependent variable: purchase intention.

Table F4

Moderation Analysis Predicting Advertising Credibility (H4, PROCESS Model 1)

Predictor	*b*	*SE*	*t*	*p*
AI disclosure (X)	−0.58	0.52	−1.11	.267
AI aversion (W)	−0.44	0.09	−4.96	< .001
X × W	0.05	0.13	0.36	.717
Gender	−0.32	0.14	−2.30	.022
Germany	−0.17	0.15	−1.19	.234
Higher education	0.09	0.15	0.60	.548
Age	0.01	0.01	1,15	.251

R^2 = .24, F(7, 233) = 10.36, p < .001

ΔR^2 (interaction) = .0004, p = .717

Note. N = 241. Unstandardized coefficients reported. AI aversion was mean-centered.