Lost in Summarization: How AI-Generated Review Summaries Shape Consumer Trust

Anna Loerakker

Student Number: 15623104

Dissertation written under the supervision of

Professor Paulo Romeiro.

Dissertation submitted in partial fulfilment of requirements for the MSc in Management with specialization in Strategic Marketing, at the Universidade Católica Portuguesa, 11.09.2025.

Abstract

Title: "Lost in Summarization: How AI-Generated Review Summaries Shape Consumer Trust."

Author: Anna Loerakker

Artificial intelligence (AI) is transforming e-commerce, including the way consumers interact with online reviews. Major platforms, such as Amazon, now use AI-generated review summaries (AGRS) to reduce information overload. However, the effect of AGRS on consumer trust remains unclear. This thesis examines whether AGRS enhance or diminish trust, how this effect depends on review valence (positive, negative, or two-sided), and whether perceived helpfulness mediates these relationships.

A 2×3 online experiment with 291 participants tested how consumers responded to reviews that differed in valence and whether or not they included an AGRS. The findings showed that the AGRS alone did not significantly influence trust or perceived helpfulness. However, helpfulness strongly predicted trust. Review valence had a partial effect: positive summaries were perceived as being less trustworthy than negative or two-sided ones.

These results extend trust transfer and information processing theories by showing that AGRS have little influence in high-trust contexts, except when summarizing positive reviews. Managers should adopt AGRS only when they clearly add value, ensure transparency, and position them as complements rather than substitutes. For researchers, the results underscore the need to test AGRS in more naturalistic and demanding settings, as well as across product types, to better understand their role in consumer decision-making.

Keywords: Online Consumer Reviews, AI-generated Review Summary, Review Summary, eCommerce, eWOM, Electronic Word of Mouth, Information Processing, Consumer Trust

SUMÁRIO

Título: “Perdido na síntese: como os resumos de avaliações gerados por IA moldam a confiança do consumidor.”

Autor: Anna Loerakker

A inteligência artificial (IA) está a transformar o comércio eletrónico, incluindo a forma como os consumidores interagem com as avaliações online. Grandes plataformas, como a Amazon, agora usam resumos de avaliações gerados por IA (AGRS) para reduzir a sobrecarga de informações. No entanto, o efeito do AGRS na confiança do consumidor ainda não está claro. Esta tese examina se o AGRS aumenta ou diminui a confiança, como esse efeito depende da valência da avaliação (positiva, negativa ou ambígua) e se a percepção de utilidade medeia essas relações.

Uma experiência online 2×3 com 291 participantes testou como os consumidores respondiam a avaliações que diferiam em valência e se incluíam ou não um AGRS. Os resultados mostraram que o AGRS por si só não influenciava significativamente a confiança ou a utilidade percebida. No entanto, a utilidade previa fortemente a confiança. A valência da avaliação teve um efeito parcial: os resumos positivos foram percebidos como menos confiáveis do que os negativos ou ambíguos.

Estes resultados ampliam as teorias de transferência de confiança e processamento de informação, mostrando que as AGRS têm pouca influência em contextos de alta confiança, exceto quando resumem avaliações positivas. Os gestores devem adotar as AGRS apenas quando elas claramente agregam valor, garantem transparência e posicionam-nas como complementos, em vez de substitutos. Para os investigadores, os resultados ressaltam a necessidade de testar as AGRS em ambientes mais naturalistas e exigentes, bem como em diferentes tipos de produtos, para compreender melhor o seu papel na tomada de decisão do consumidor.

Palavras-chave: Avaliações online de consumidores, Resumo de avaliações gerado por IA, Resumo de avaliações, Comércio eletrónico, eWOM, Propaganda boca a boca eletrónica, Processamento de informações, Confiança do consumidor

Acknowledgements

I would like to thank my supervisor, Professor Paulo Romeiro, for his guidance, feedback, and encouragement throughout the research process. His support and the materials he provided were very valuable to this thesis. I am also grateful to the professors at Católica I encountered, who inspired and challenged me to work harder. Their dedication to teaching has been a great source of motivation. Moreover, I am deeply grateful to CLSBE for an unforgettable period of growth, learning, and friendship. My time in Lisbon was transformative both academically and personally, and I will always cherish this experience.

I would also like to thank all the respondents of my survey and participants of my pre-studies for dedicating their time and effort to help me collect the necessary data. Without their contributions, this research would not have been possible.

Finally, I am beyond fortunate to have the support of my family, who gave me the opportunity to pursue an outstanding education, and my workplace, Heykes&Karstens, for enabling me to work part-time during my studies abroad, which made attending a private university possible. Without the help of these people, I would not have been able to reach the goals I have achieved so far.

Disclaimer on the Use of Artificial Intelligence

Artificial intelligence tools (OpenAI's ChatGPT and related applications) were used in a supportive capacity during the preparation of this thesis. Specifically, AI was employed for:

- rephrasing, shortening, and optimizing text passages,
- drafting and refining formulations,
- supporting the development and refinement of data collection procedures (e.g., survey and interview design) and experimental stimuli.

All uses of AI were limited to language and methodological assistance. The intellectual content, theoretical framing, research design, interpretation of results, and all final conclusions are the author's own work. AI tools were not used to generate original arguments, data, or literature, and all sources cited in the thesis were identified and verified by the author.

This disclosure is made in accordance with academic integrity guidelines requiring transparency about the use of artificial intelligence in scholarly work.

Table of contents

Table of Figures

Table of Tables

Table of Abbreviations

AGRS:	AI-generated review summary
AI:	Artificial intelligence
ELM:	Elaboration Likelihood Model
eWOM:	Electronic Word-of-Mouth
TAM:	Technology Acceptance Model
LLM:	Large Language Model

Glossary

AI-generated review summary (AGRS)	Concise summaries of consumer reviews created by large language models to reduce information overload and support decision-making.
Artificial Intelligence (AI)	Computer systems designed to perform tasks that typically require human intelligence, such as learning, reasoning, or language generation.
Consumer Trust	A consumer’s confidence in the reliability, objectivity, and sincerity of online review content.
Electronic Word-of-Mouth (eWOM)	User-generated online content that shares opinions online about products or services and influences others’ purchase decisions.
Information Diagnosticity	The extent to which information enables accurate evaluation of a product’s quality, typically through specific, factual content.
Large Language Model (LLM)	AI models trained on large amounts of text and data to understand, generate, summarize, and translate human language.
Perceived Helpfulness	The degree to which a review is seen as useful for decision-making, based on its clarity, relevance, and informational value.
Review Valence	The overall sentiment of a review (positive, negative, or two-sided).
Two-Sided Reviews	Reviews that present both positive and negative aspects of a product, often perceived as more balanced.

1 Introduction

1.1 Background

The way consumers make decisions has been fundamentally reshaped by digital technology (Lemon & Verhoef, 2016). Without the ability to physically inspect products, online shoppers face heightened uncertainty (K. T. Lee & Koo, 2012). To navigate this, consumers increasingly turn to peer-generated content, specifically online reviews, to inform their purchase decisions and mitigate risks.

Nine out of ten consumers read reviews before purchasing (Brand Rated, 2022), reflecting greater trust in peer recommendations than in traditional advertising (Nielsen, 2012). This electronic word of mouth (eWOM) is defined as “any positive or negative statement made by […] customers about a product or company, made available to a multitude of people and institutions via the internet” (Hennig-Thurau et al., 2004). By sharing experience-based information, reviewers act as “sales assistants for online retailers” (Y. Chen & Xie, 2008), making trust in such content a critical factor in e-commerce (M. Lee & Youn, 2009).

However, the growing volume of reviews has introduced challenges. Review sections are crowded with contradictory opinions, irrelevant details, and low-quality or spam content (C. C. Chen & Tseng, 2011; S. Lee & Choeh, 2014). This contributes to information overload, where even the most helpful reviews can get buried (Cao et al., 2011; Min & Park, 2012).

To restore clarity, many platforms now turn to artificial intelligence (AI). Large language models (LLMs) are being used to automatically generate review summaries, aiming to distill vast amounts of content into concise, digestible insights (Gambhir & Gupta, 2017). In e-commerce, AI-generated review summaries (in this study: AGRS) offer a promising solution to tackle the inefficiencies of information overload.

While promising, AGRS introduce a new layer of uncertainty. Unlike human-written reviews, they may lack emotional nuance, narrative richness and contextual detail (Jia et al., 2025). This raises a critical question: can consumers trust review content when the message is no longer written by another human, but crafted by an algorithm?

1.2 Problem Statement

In 2023, Amazon started integrating AGRS on product pages to support quick evaluation, explicitly disclosing AI involvement to prevent confusion (Levine, 2024). Although AGRS aim to enhance decision-making, their effect on customer trust remains unclear.

Building on these developments, this dissertation examines whether the presence of an AGRS strengthens or weakens consumer trust, and whether this effect depends on the sentiment of the underlying reviews: positive, negative, or two-sided. Six conditions were designed to reflect these combinations, inspired by Amazon's review format. Each sentiment is tested both with and without an AGRS, allowing for a structured comparison. Beyond the direct effect, the study examines whether perceived helpfulness of the review set serves as a mediating mechanism: does an AGRS enhance perceived helpfulness, and does this in turn increase consumer trust? The study is guided by the following research questions:

RQ1: How does the presence of an AGRS (present vs. absent) influence consumer trust in online reviews?

RQ2: Does perceived helpfulness act as a mediator in the relationship between AGRS and consumer trust?

RQ3: Does review valence moderate the effect of AGRS on consumer trust?

1.3 Relevance

From an academic perspective, extensive research has explored the relationship between online consumer reviews and trust, helpfulness, and review valence (C. M. K. Cheung & Thadani, 2012; Filieri, 2016; Mudambi & Schuff, 2010). However, the use of generative AI to produce into concise review summaries has received little empirical attention. In particular, the extent to which consumers trust review sets that combine individual reviews and AGRS, and the role of helpfulness in mediating this trust, have not yet been examined. This study addresses these gaps by drawing on consumer behavior, trust theory, and information-processing perspectives.

From a managerial perspective, this research is particularly relevant given the increasing use of summarization tools by platforms like Amazon and TripAdvisor. These tools aim to enhance the user experience and facilitate decision-making. However, if AI-generated content is perceived as untrustworthy or manipulative, the persuasive value of reviews could decline, reducing purchase intent. Therefore, understanding how AGRS shape trust and helpfulness is critical for optimizing review presentation and sustaining engagement. Finally, this study contributes to ongoing debates concerning the integration of AI into online retail environments by offering insights into trust mechanisms in hybrid review environments.

1.4 Research Methods

To address the research questions, this study adopted a mixed-methods approach, using a controlled online experiment as the primary method of data collection. This experiment tested the hypothesized relationships between AGRS, review valence, consumer trust and perceived helpfulness.

Prior to this, a comprehensive literature review was conducted to identify key theoretical concepts and validated measurement instruments. These insights informed the conceptual framework and the empirical design. To enhance ecological validity, qualitative interviews and two preliminary surveys were conducted to refine the experimental stimuli.

The final 2×3 between-subjects experiment randomly assigned participants to one of six conditions that varied in terms of review valence and the presence of an AGRS. The key dependent variables, trust and helpfulness, were measured using established scales. Data were collected via an online survey platform and analyzed in SPSS; mediation and moderation analyses were used to test the proposed relationships.

1.5 Dissertation Outline

The remainder of this work is structured as follows: First, the theoretical foundations are reviewed, and hypotheses are developed. Next, an overview of the research design is provided, covering the development of the experimental stimuli and the survey procedures. The results section presents the findings of the quantitative analyses of the hypothesis testing. The discussion section then interprets these findings in light of the theoretical framework and existing literature. Finally, the conclusion outlines the implications for management and research, acknowledges limitations and suggests directions for future research.

2 Theoretical Foundations and Related Work

This section provides an in-depth theoretical overview of the concepts of consumer trust, AGRS, perceived helpfulness, and review valence, as well as their interactions. It serves as the foundation for the development of testable research hypotheses for the quantitative analysis.

2.1 Consumer Trust in Reviews

Trust is a foundational element in consumer decision-making, especially in overcoming perceived risks and uncertainties (McKnight et al., 2002). In online marketplaces, where transactions often take place in an impersonal and anonymous environment (Pavlou & Gefen, 2004), trust plays an even more critical role in shaping consumers' willingness to engage in transactions (Pavlou & Fygenson, 2006). It allows consumers to assess a seller's intentions and capabilities (Khamitov et al., 2024) and mitigates the risks associated with the lack of physical product inspection. Here, trust is defined as "a consumer's confidence in the reliability and integrity of an entity" (De Wulf et al., 2001).

In the context of online reviews, trust expands beyond general trust in the platform or seller and additionally includes "the degree of confidence in the validity of the information, specifically regarding its objectivity and sincerity" (Reimer & Benkenstein, 2016). Unlike traditional retail, online transactions rely heavily on trust to compensate for the lack of physical verification. Without first-hand product experience, consumers turn to third-party information, such as user reviews, to reduce uncertainty and assess credibility (Y. D. Wang & Emurian, 2005). As a result, trust in reviews directly shapes consumer attitudes and behavior (Wolf & Muhanna, 2011).

This underscores the crucial impact of review trustworthiness on consumer responses. When reviews are perceived as untrustworthy, consumers actively discount them (Filieri, 2016). More so, trustworthiness serves as a moderator of review adoption, with consumers more likely to adopt highly trustworthy reviews to guide their judgment formation and decision-making (G. Huang & Liang, 2021). Review adoption occurs when consumers accept the recommendations of online reviews and act accordingly, for example by making a purchase (Filieri, 2014).

Despite its importance, not all reviews are, per se perceived as trustworthy. Identifying the factors that shape trustworthiness, and how consumers them, is essential for understanding the role of trust in review adoption. Consumers do not systematically evaluate each review in detail; instead, they rely on cognitive heuristics to assess trustworthiness. These heuristics can be classified into two main categories: source cues, such as the reviewer's expertise, reputation, and past review history (Machackova & Smahel, 2018), and content cues, such as the quality, length, factual depth, and tone of the review (L. Y. Pan & Chiou, 2011).

2.1.1 Dimensions of Trust in Online Reviews

Trust is a multidimensional concept essential to consumer decision making (McKnight et al., 2002). While various classifications exist (Mayer et al., 1995), a widely accepted framework (Ba & Pavlou, 2002; Doney & Cannon, 1997; Singh & Sirdeshmukh, 2000) distinguishes between two core dimensions: Benevolence and Credibility.

Benevolence refers to the belief that the trustee has good intentions and acts in the best interest of the trustor (Pavlou & Dimoka, 2006). In marketing, it is defined as "a buyer's belief that a seller will act fairly and not exploit consumers, even in unfavorable conditions" (J. C. Anderson & Narus, 1990). In online marketplaces, benevolence reflects the belief that a seller has beneficial motives and prioritizes consumer well-being over short-term profits (Pavlou & Dimoka, 2006).

Credibility reflects a buyer's belief that the seller is competent, reliable, and capable of fulfilling promises (Pavlou & Dimoka, 2006). With online reviews, credibility is related to the reviewer's expertise, the consistency of the information, and the factual basis of the review. Reviews that lack credibility, are being vague, overly promotional or poorly sourced, are often rejected by consumers (Filieri, 2016), while fact-based, detailed reviews are preferred (Dong et al., 2019).

In order to understand how trust in online reviews develops, both individual predispositions and situational cues should be examined. Trust in content arises not only from the characteristics of the review itself, but also from consumers' internal tendencies and their trust in the systems that present the information. Building on this, it is crucial to examine how trust in established entities such as e-commerce platforms transfers to emerging technologies like AGRS.

2.1.2 Trust Transfer Mechanisms

Trust Transfer Mechanisms describe the cognitive process by which trust in a familiar entity, such as a platform or seller, extends to a related, less familiar entity (Stewart, 2003). As trust in a platform grows, so does the likelihood that this trust will extend to associated features.

In online marketplaces, this often involves transferring trust from a well-known entity, such as a reputable platform, to a lesser known one, such as an AI tool presented by the former. Heuristic cues, such as brand reputation, platform reliability and social presence, provide guidance when evaluating unfamiliar content. Consumers may assume that the reliability of a platform also applies to the tools it offers. Supporting this, previous studies have shown that trust in a platform can positively affect trust in third-party services via heuristic cues such as reputation and perceived quality (X. Chen et al., 2015; Xiao et al., 2019).

However, trust is not transferred unconditionally. Lankton et al. (2015) argue that trust in human agents does not seamlessly extend to technological systems, particularly ones that lack human-like qualities. Similarly, Kim et al. (2008) observe that trust transfer weakens when new tools introduce additional uncertainty or perceived risk. For example, a consumer may trust Amazon, yet still distrust an AGRS if it is unclear, biased, or inaccurate. Further evidence is provided by Filieri (2016), demonstrating that in information-rich environments, consumers prioritize information quality over platform-based heuristics. This suggests that, when faced with complex or novel technologies, consumers shift from relying on transferred trust to critically evaluating the content itself.

In conclusion, trust transfer helps to explain initial openness to platform-based innovations such as AGRS. However, its effect is contingent on content-specific factors, which determine whether transferred trust is sustained or replaced by critical evaluation.

2.2 AI-Generated Review Summaries

Technological advancements have significantly reshaped the consumer decision-making journey (Lemon & Verhoef, 2016), with the widespread availability of artificial intelligence (AI) proving to be a major disruptor. While online reviews support informed decisions, they also present challenges such as information overload, conflicting opinions (Cao et al., 2011; Min & Park, 2012) and the prevalence of low-quality or spam reviews (S. Lee & Choeh, 2014). The sheer volume of reviews makes it difficult to identify relevant insights (Ghose & Ipeirotis, 2011), ultimately reducing decision-making efficiency (C. C. Chen & Tseng, 2011).

Traditionally, platforms have relied on aggregated metrics, such as star ratings, to simplify evaluation (Qiu et al., 2012). More recently, advances in large language models (LLMs) have enabled the automatic summarization of reviews. LLMs, such as those powering tools like ChatGPT, can generate coherent summaries by identifying patterns in the text (Telenti et al., 2024). Moreover, they have outperformed traditional natural language processing models in key tasks such as information extraction and sentiment analysis (Falatouri et al., 2024). Applied to online reviews, these tools condense complex inputs into concise, actionable insights (Gambhir & Gupta, 2017), thereby improving the accessibility and usability of consumer feedback (Jia et al., 2025). Notably, their performance has shown strong alignment with high-quality human summaries (T. Zhang et al., 2024).

However, these advancements also introduce challenges. Studies show that consumers struggle to distinguish between human-written and AI-generated reviews, raising concerns about declining trust in review systems (Kovács, 2024). AI-paraphrased reviews can furthermore confuse readers and reduce the perceived authenticity of real content (Xylogiannopoulos et al., 2024). This indistinguishability reveals a vulnerability to manipulation, reinforcing the need for stronger oversight and consumer protection. Beyond content-level concerns, LLMs also raise broader trust issues relating to transparency, alignment with human values and the potential for bias or misuse (Ferdaus et al., 2024). These risks emphasize the importance of ethical governance, industry accountability and public engagement. Against this backdrop, it becomes crucial to examine how AGRS influence consumer trust.

2.3 The Influence of AI-Generated Review Summaries on Consumer Trust

In online review settings, limited source information makes it difficult for consumers to assess the trustworthiness of reviews or platforms (Reimer & Benkenstein, 2016). Although AGRS are designed to enhance efficiency by condensing reviews into key insights, their impact on consumer trust remains largely unexplored. This reflects the recent emergence of large-scale AI summarization tools and underscores the need for empirical investigation.

Existing research suggests that consumers generally place greater trust in content created by humans than in AI-generated alternatives (Lim & Schmälzle, 2024; X. Luo et al., 2019). This is often due to a perceived lack of emotional resonance, contextuality and authenticity in AI outputs (Jia et al., 2025). Disclosing the AI-generated origin of text has been shown to reduce its perceived reliability (Lim & Schmälzle, 2024). However, consumer attitudes toward AI-generated content are dynamic and may evolve as users become more familiar with AI technologies (Venkatesh et al., 2003).

Furthermore, AGRS pose specific risks that can further undermine trust. Studies show that summarization algorithms can unintentionally reinforce dominant opinions and filter out different perspectives, thereby reducing information diversity (Falatouri et al., 2024). LLMs are also prone to 'hallucinations', whereby they produce content that is factually incorrect yet plausible, raising concerns about credibility (Telenti et al., 2024). The extent of these risks varies depending on the implementation, including measures to reduce bias and safeguard against hallucinations employed by platforms.

By contrast, human-written reviews often convey narrative richness, social cues and subjective nuances that make them more relatable and trustworthy (Jia et al., 2025). However, the quality of these reviews can vary widely, and not all of them provide the depth or authenticity typically associated with user-generated content.

Given these considerations, this study explores whether AI-generated summaries, despite their efficiency, reduce consumer trust by diminishing perceived authenticity, emotional depth, and informational transparency. Therefore, the following hypothesis is proposed:

> *H1: Consumers perceive review sets accompanied by an AGRS and two individual reviews as less trustworthy than review sets containing only two individual reviews.*

Beyond the content's origin (human vs. AI), the emotional valence of a review, whether positive, negative or two-sided, also influences trust and decision-making. The following section therefore explores how review sentiment shapes consumer trust.

2.4 Review Valence

The tone conveyed in online reviews, or review valence, shapes consumer perceptions by emphasizing product strengths (positive valence) or weaknesses (negative valence) (Ketelaar et al., 2015). Empirical research has repeatedly linked review valence and consumer behavior, particularly purchase decisions. However, the strength and direction of this effect can vary. While positive reviews are associated with increased sales (Chevalier & Mayzlin, 2006), high ratings do not always improve market performance. Furthermore, negative eWOM does not necessarily harm sales (Clemons et al., 2006; Maslowska et al., 2017). Although some studies emphasize the importance of review quantity, valence remains a key factor in determining purchasing response (Chintagunta et al., 2010; Dellarocas et al., 2007; Duan et al., 2008).

2.4.1 Negativity and Positivity Bias

While reviews generally shape perceptions, their impact varies depending on their valence. Research suggests that negative reviews often have a stronger influence on trust and decision-making than positive ones (Chevalier & Mayzlin, 2006). Grounded in prospect theory, this phenomenon known as negativity bias, suggests that consumers tend to give more weight to negative information (Ahluwalia, 2000; Brown & Reingen, 1987; Laczniak et al., 2001), perceiving it as more informative, diagnostic, and credible (Herr et al., 1991; Lynch, 2006). For example, one-star reviews tend to influence decision-making more than five-star reviews (Chevalier & Mayzlin, 2006), with consumers generally trusting negative reviews that appear balanced and factual more than positive ones. Conversely, overly positive reviews are often viewed with skepticism and perceived as less credible (Filieri, 2016; K. T. Lee & Koo, 2012).

While negativity bias is well-documented, some studies also highlight the opposite phenomenon, positivity bias, whereby positive information exerts a greater persuasive influence (Gershoff et al., 2003; Skowronski & Carlston, 1987, 1989) under certain conditions. Research suggests that positive messages can be more compelling than negative ones, particularly when they align with consumers' goals or expectations. For example, positively worded reviews can increase the purchase intent of consumers who are already inclined to view the product favorably (Guo et al., 2020).

2.4.1.1 Two-Sided Review Valence

In contrast to extremely positive or negative reviews, two-sided valence refers to moderate reviews that acknowledge both the strengths and weaknesses of a product. These types of reviews are generally perceived as more trustworthy than extreme, one-sided reviews (Filieri, 2016). By presenting both benefits and drawbacks, two-sided reviews help consumers to realistically assess whether a product's limitations are acceptable compared to its benefits (Purnawirawan et al., 2015). These reviews enhance credibility (Golden & Alpert, 1987; Smith & Hunt, 1978) while also lowering the likelihood that consumers will engage in counterarguing (Belch, 1981; Kamins & Assael, 1987). Two-sided messages are particularly persuasive when consumers have a neutral or negative attitude (Crowley & Hoyer, 1994), and are often considered more helpful than purely negative evaluations (Purnawirawan et al., 2015).

2.4.2 Contextual Factor: Product Type

Review valence effects differ between utilitarian and hedonic products. Consumers rely more on negative reviews for utilitarian products, as they provide diagnostic information about objective performance (Sen & Lerman, 2007). In contrast, negative reviews for hedonic products are often attributed to reviewer preferences rather than actual product flaws, making them less persuasive.

2.5 Review Valence as Moderator

Review valence plays a critical role in how individuals evaluate and trust product information. When source cues are limited, consumers often rely on the tone and extremity of reviews to assess credibility (Filieri, 2016). Given these findings, review valence is expected to shape how AGRS influence trust.

H2: The effect of AGRS on consumer trust is moderated by review valence; specifically, the difference in trust between review sets with and without an AI summary varies depending on whether the review valence is positive, negative, or two-sided.

2.5.1 Negative vs. Positive Reviews and Trust

Negative reviews are often perceived as more diagnostic, credible, and informative than positive ones (Chatterjee, 2001; Filieri, 2016; Herr et al., 1991). This is particularly true for utilitarian products, where consumers interpret negative feedback as highlighting objective product flaws (Sen & Lerman, 2007).

However, AGRS can mitigate this effect by condensing and mixing sentiments, sometimes omitting the contextual details that give negative feedback its diagnostic value (Jia et al., 2025). Conversely, summaries based predominantly on positive reviews run the risk of being perceived as less trustworthy, especially if they lack critical balance or appear overly enthusiastic (Chatterjee, 2001; Filieri, 2016; Herr et al., 1991). Therefore, while positive sentiment can increase a product's appeal, AGRS that only reflect positive content may undermine trust when consumers are looking for balanced, informative feedback.

H2a: AGRS based on predominantly negative reviews are perceived as more trustworthy than summaries based on predominantly positive reviews, with review valence influencing trustworthiness perceptions.

2.5.2 Two-Sided Reviews and Trust

Two-sided reviews are consistently perceived as the most credible and trustworthy (Filieri, 2016). By presenting a balanced perspective, these reviews reduce skepticism, foster authenticity, and help consumers weigh trade-offs more effectively (Purnawirawan et al., 2015).

When AGRS reflect the balance of two-sided reviews, they are more likely to maintain perceived fairness and objectivity. This, in turn, strengthens trust without appearing biased through increasing persuasiveness and credibility.

H2b: Consumers perceive AGRS based on two-sided reviews as more trustworthy than summaries based on predominantly positive or negative reviews, with review valence influencing trustworthiness perceptions.

2.6 Review Helpfulness and Consumer Information Processing

Information helpfulness is a key construct in adoption behavior (Sussman & Siegal, 2003). By offering valuable information, e-commerce websites with more helpful reviews maximize customer satisfaction (Kohli et al., 2004). Helpful reviews also improve decision-making and optimize consumer value since they encourage review adoption (J. Lee & Hong, 2019).

A helpful review is defined as "a peer-generated product evaluation that facilitates a consumer's purchase decision" (Mudambi & Schuff, 2010). Perceived helpfulness refers to the value assigned to a review by readers, based on its usefulness in informing their judgments (Mudambi & Schuff, 2010). In other words, it is the degree to which people find a review useful in supporting their decision-making. This includes both individual opinions and the overall usefulness as judged by many users, such as when reviews receive high helpfulness votes[1] (Baek et al., 2012; Cao et al., 2011; Li, Huang, et al., 2013).

2.6.1 The Elaboration Likelihood Model

The Elaboration Likelihood Model (ELM), developed by Petty & Cacioppo (1986), is a widely used framework in consumer psychology for understanding how individuals process persuasive information. The model distinguishes between two cognitive routes: the central route, which involves thoughtful evaluation of message content, and the peripheral route, which relies on superficial cues such as source credibility, emotional tone, or visual indicators like star ratings (C. M. Y. Cheung et al., 2012; Petty & Cacioppo, 1986).

Whether an individual engages in central or peripheral processing depends on their motivation and ability to interpret the message (Petty et al., 1983). When both are high, such as when the message is personally relevant and cognitively accessible, consumers are more likely to use the central route (Chaiken, 1980; Eagly & Chaiken, 1993). In contrast, low motivation or limited ability leads to reliance on heuristics, activating the peripheral route (Gupta & Harris, 2010). Attitudes formed through central-route persuasion are more stable, whereas those formed through peripheral-route persuasion have weaker, short-term effects (Petty & Cacioppo, 1986). Figure 1 below illustrates this dual-process framework:

[1] user-generated ratings indicating whether a review was perceived as useful; typically phrased as "Was this review helpful?"

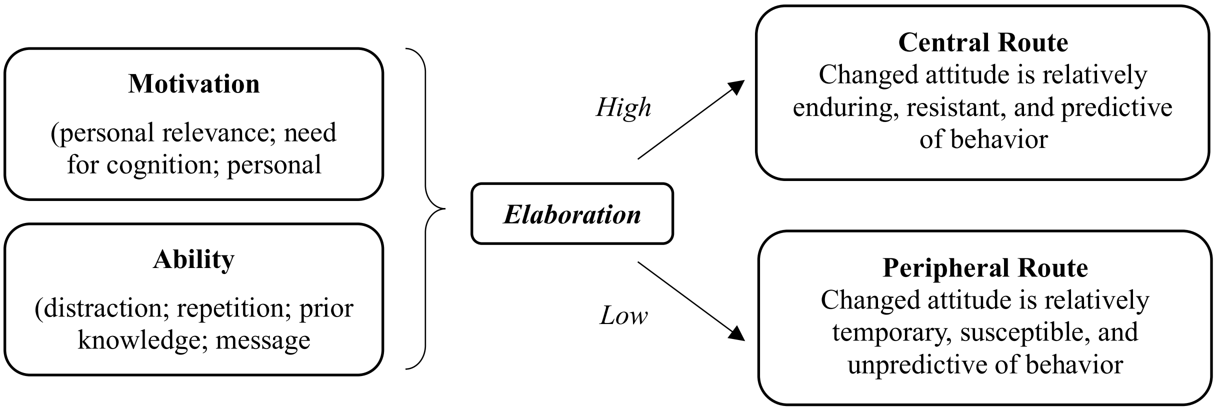


Figure 1: Elaboration Likelihood Model, adapted from Petty and Cacioppo (1986)

In online reviews, some consumers engage deeply with content, evaluating arguments, while others rely on superficial cues like helpfulness votes (C. M. K. Cheung & Thadani, 2012). Highly involved consumers might read several detailed reviews, while less motivated consumers may rely on average star ratings. Notably, elaboration is not binary; individuals may process information through both routes simultaneously, depending on context and cognitive capacity (C. M. Y. Cheung et al., 2012).

Prior eWOM studies have widely applied the ELM to explain review adoption and purchase intent, showing that consumers rely on both content and contextual cues when evaluating persuasive reviews (Brand & Reith, 2022; J. Lee et al., 2008; D. H. Park et al., 2007). Within this framework, information quality emerges as the strongest predictor of review adoption under conditions of high elaboration likelihood (C. M. K. Cheung & Thadani, 2012; Filieri & McLeay, 2014; Zhou et al., 2016). This underscores the central role of argument quality in shaping decision-making when individuals are motivated and able to engage deeply with review content.

2.6.2 Information Diagnosticity

Perceived helpfulness is closely linked to the concept of information diagnosticity, being the degree to which a review enables consumers to evaluate a product's true quality (Jiang & Benbasat, 2004). Diagnosticity is determined by the perceived correlation between available information and consumer decision-making, often conceptualized as the degree of helpfulness (Dick et al., 1990; Qiu et al., 2012; Skowronski & Carlston, 1987). In online environments where physical inspection is impossible, diagnosticity compensates by providing detailed cues about product attributes (Kirmani & Rao, 2000). In this context, diagnostic reviews with highly factual, specific, and evaluative content have been shown to positively influence consumer decisions (Ghose & Ipeirotis, 2011).

This link is further supported by the ELM. In central-route processing, consumers pay close attention to the diagnostic value of a review, i.e., how well it supports an informed evaluation of the product. Information-rich reviews are therefore considered more persuasive and helpful. Even negative reviews highlighting product drawbacks can enhance decision quality if they are seen as objective and informative (Guo et al., 2020). Empirical research confirms that the more relevant and specific the information is, the more diagnostic it will be and the more helpful it is perceived to be (Chua & Banerjee, 2016).

2.6.3 Determinants of Review Helpfulness

Beyond information diagnosticity, the helpfulness of online reviews is influenced by multiple factors, including content characteristics (Ghose & Ipeirotis, 2011; S. Lee & Choeh, 2014; Li, Huang, et al., 2013), reviewer attributes (Hong et al., 2017; A. H. Huang et al., 2015), and product type (Mudambi & Schuff, 2010; Nelson, 1970; Y. Pan & Zhang, 2011; C. Park & Lee, 2009). Baek et al. (2012) show that both peripheral cues, including review rating and reviewer's credibility, and central cues, such as the content of reviews, influence the helpfulness of reviews.

Content quality is consistently found to be a key determinant. Reviews that are detailed, fact-based, and specific are generally rated as more helpful than those that are vague, emotional, or overly subjective (Filieri, 2014; Ghose & Ipeirotis, 2011; Li, Choi, et al., 2013). However, while review depth increases perceived helpfulness, this effect plateaus; overly lengthy reviews can reduce readability and user engagement, as readers may struggle to process too much information (A. H. Huang et al., 2015). Conversely, evidence also suggests that informal tone and conversational language enhance engagement and make the review feel more authentic (Schindler & Bickart, 2012).

Lastly, the influence of review content depends on the type of product. Consumers prioritize diagnostic and objective information for utilitarian products (Mudambi & Schuff, 2010; Y. Pan & Zhang, 2011), such as electronics, while emotional and experiential elements play a larger role for experience goods, whose quality can only be judged after use (Nelson, 1970, 1974).

2.6.4 Technology Acceptance (TAM)

While the ELM explains how consumers process review information, it does not fully account for how new technologies, such as AI-generated summaries, influence user attitudes and behaviors.

The Technology Acceptance Model (TAM) complements this by focusing on perceived usefulness and ease of use as key drivers of user acceptance (Davis, 1989). In the context of review summaries, perceived helpfulness can be interpreted as a form of perceived usefulness, as it reflects the degree to which consumers believe the summary facilitates their decision-making. This parallels Davis's (1989) finding that individuals are more likely to accept and engage with systems that they believe will improve their job performance. Ease of use is equally important, as summaries that are clear and effortless to process are more likely to be adopted (Venkatesh & Davis, 2000).

As such, helpfulness acts not only as an outcome of cognitive processing (ELM), but also as a determinant of trust in AI-generated content (TAM). The following section examines this relationship in more detail by exploring how perceived helpfulness mediates consumer trust.

2.7 Helpfulness and Trust

According to the ELM, when peripheral cues like helpfulness ratings are unavailable or less salient, consumers are likely to engage in central-route processing, critically evaluating the argument quality of the information presented (C. M. K. Cheung & Thadani, 2012; Petty & Cacioppo, 1986).

In this context, information diagnosticity becomes a key determinant of trust (Jiang & Benbasat, 2004). Reviews that are fact-based, specific and well-structured increase perceived credibility by facilitating informed decision making (Ghose & Ipeirotis, 2011; Sparks & Browning, 2011). This effect is particularly strong for attribute-based reviews that provide clear assessments of product features and thus are considered more trustworthy than vague or emotional narratives (Folse et al., 2016; K. T. Lee & Koo, 2012). Consumers who engage in deeper elaboration are expected to place greater weight on the helpfulness of the summary, while others may lean more heavily on peripheral indicators like average ratings. The TAM underpins this perspective by identifying perceived usefulness as one of the most important factors for system acceptance (Davis, 1989). Even when peripheral cues such as star ratings are present, consumers are more likely to trust summaries they perceive as useful.

Thus, this study proposes that the diagnostic quality of the review summary, operationalized as perceived helpfulness, acts as a primary driver of consumer trust.

> *H3: The perceived helpfulness of AGRS positively influences consumer trust in the reviews.*

2.8 AI-Generated Review Summaries and Helpfulness

The use of LLMs to distill large volumes of content into concise summaries presents notable advantages for improving the perceived helpfulness of review sets. By efficiently extracting central content and structuring information, LLMs can improve readability and reduce cognitive effort for consumers (Ishtiaq et al., 2024). Especially in contexts involving high information complexity, such as medical documentation, AGRS have even outperformed human experts in clarity and completeness (Chien et al., 2024; Liu et al., 2023). These findings suggest that AI summarization can improve the diagnosticity of review content. In other words, it can increase the extent to which consumers can extract meaningful and actionable information, thus increasing perceived helpfulness.

However, the use of LLMs in generating AGRS also presents challenges. A key concern is their tendency to generate hallucinated or factually inconsistent content, which can compromise summary reliability (Z. Luo et al., 2024). Moreover, over-coordination in language, leading to AI outputs appearing overly uniform or generic, may reduce linguistic authenticity and perceived individuality, characteristics often valued in user-generated content (Liu et al., 2023). As a result, summaries may be perceived as less trustworthy or helpful. Studies also show that AGRS often lack contextual specificity and emotional nuance, particularly when compared to human-written reviews that contain rich, experience-based narratives (Namvar & Chua, 2023). This tendency to oversimplify can diminish the emotional depth and unique insights that contribute to the perceived value of a review.

Despite these concerns, AGRS offer significant advantages, primarily increased clarity and reduced information overload. When designed to emphasize key attributes and retain relevant details, AGRS can improve the perceived helpfulness of online reviews by enhancing information diagnosticity. Although is not directly measured in this study, diagnosticity provides a theoretical rationale for why AGRS may increase helpfulness by offering content-rich representations of consumer opinions. This relationship forms the basis for the following hypothesis.

> *H4: Consumers perceive reviews accompanied by an AGRS as more helpful than reviews presented without summaries.*

Beyond this direct relationship, perceived helpfulness is theorized to mediate the effect of AGRS on consumer trust. The TAM (Davis, 1989) highlights perceived usefulness as a driver of acceptance, while the ELM (Petty & Cacioppo, 1986) emphasizes the persuasive impact of diagnostic information. Together, these perspectives suggest that when AGRS are perceived as helpful, trust in the content and platform increases. Therefore, the following mediation hypothesis is proposed:

H5: The effect of AGRS on consumer trust is mediated by their perceived helpfulness.

2.9 Conceptual Framework

The conceptual model visualizing the proposed relationships and variables is shown below.

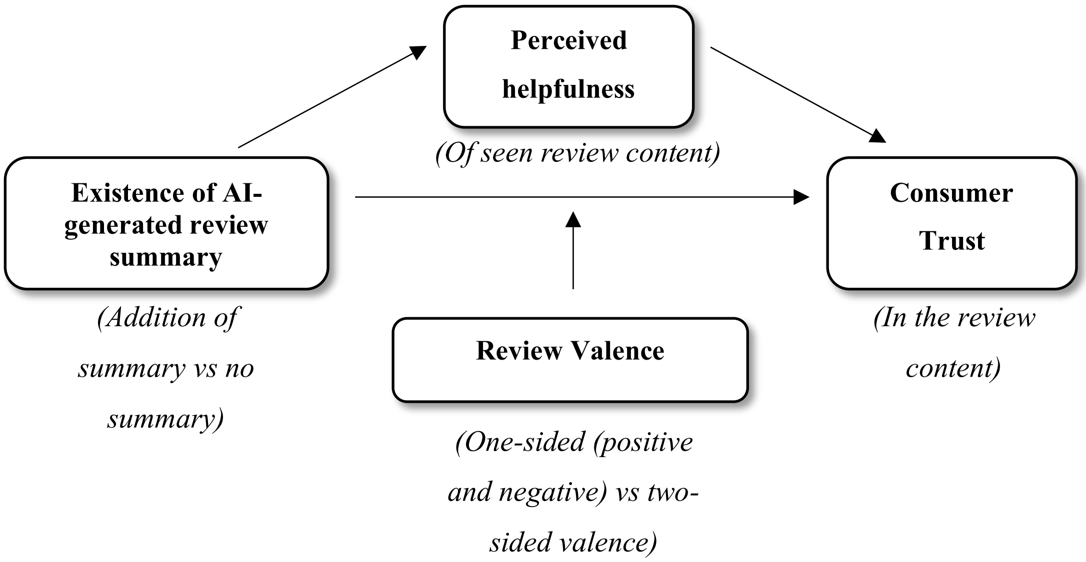


Figure 2: Conceptual Framework

3 Methodology

This chapter outlines the methodology used to examine the research topic and test the hypotheses introduced in Chapter 2.

3.1 Research Approach

To examine the variable interdependencies, a combination of exploratory, descriptive, and hypothesis-testing methods was employed (Malhotra et al., 2017).

An exploratory approach was first undertaken to refine the research problem and identify relationships among the key variables. This involved conducting a thorough literature review to establish theoretical foundations, define core concepts, and assess existing empirical evidence on dependencies within the conceptual framework. The aim was to improve understanding of the variables and their interactions. In addition, consumer interviews and two preliminary surveys were conducted to develop the stimuli. A mixed approach was chosen here.

Subsequently, an online survey was conducted using a cross-sectional design with an experimental approach that allowed for the manipulation of key variables to determine causal relationships. This enabled the effects of AGRS on helpfulness and trust to be isolated, while considering the moderating influence of review valence. The collected data were analyzed using an inferential quantitative approach to validate predictions derived from theoretical models through empirical testing (de Vaus, 2002).

3.2 Data Collection

3.2.1 Data Type and Collection Method

This study employs a sequential exploratory mixed-methods design (Creswell, 2009) in two phases, combining qualitative exploration to refine concepts and variables, followed by a quantitative phase to empirically test the hypotheses.

3.2.1.1 Phase 1: Qualitative Data Collection and Analysis

The first phase focused on exploring how consumers perceive and evaluate online reviews, aiming to refine the research problem (Malhotra et al., 2017). Structured interviews with predetermined questions ensured consistency (Kothari, 2004), while allowing for rich insights into consumer motivations, decision-making processes, and underlying values. This approach also revealed subconscious drivers often missed in surveys (Malhotra et al., 2017). The findings informed the design of the quantitative data collection.

3.2.1.2 Phase 2: Quantitative Data Collection and Analysis

Secondly, a structured online survey was conducted to statistically test the research hypotheses and generate generalizable insights (Malhotra et al., 2017). Participants responded to a series of questions to measure their attitudes towards the stimuli, as well as providing demographic information. A systematic and standardized design minimized potential biases and improved response consistency (Kothari, 2004). The resulting data were analyzed to identify significant relationships between the variables.

3.2.2 Sampling

In the qualitative phase, participants were recruited through convenience sampling to focus on accessibility and willingness rather than randomization (Malhotra et al., 2017). This method enabled the efficient recruitment of participants who met the research criteria and were willing to cooperate. The target population consisted of adults residing in Germany, and efforts were made to ensure variation in household structure, age and educational background, to enhance representativeness.

For the quantitative phase, participants were recruited using a structured convenience sampling approach that combined personal contacts with respondents recruited through SurveySwap.io and TikTok. Although SurveySwap skewed the sample towards young adults (18–25 years old) from Western Europe, North America and the UK (SurveySwap Helpcenter, n.d.), this approach provided efficient access to a broader and more diverse pool, thereby improving generalizability while saving time and resources.

To ensure relevance, participants in both phases were screened based on their usage habits regarding reviews, thereby reducing sampling frame error and aligning the sample with the study objectives (Malhotra et al., 2017). Respondents were informed about the purpose of the study and measures taken to protect their data. All data were anonymized to ensure compliance with data protection regulations (Sekaran & Bougie, 2016).

3.2.3 Variable Measurement

This section outlines the measurement of key constructs.

3.2.3.1 Mediator: Perceived Helpfulness

The mediating variable was measured using a 7-point Likert scale adapted from Wu (2013). Participants rated the helpfulness of the review set based on the following items:

The review was…

- Informative
- Useful
- helpful

The scale's Cronbach's alpha is 0.89, which is considered good by Peterson (1994).

3.2.3.2 IV: Consumer Trust in the Review

The dependent variable was assessed using an adapted version of Ohanian's (1990) trustworthiness scale, originally published in the Journal of Advertising. This scale has been widely applied in research and was refined by Reimer & Benkenstein (2016) to four items measured by a 7-point Likert scale:

The writer of the review is …

- Honest – Dishonest
- Untrustworthy – Trustworthy
- Unreliable – Reliable
- Insincere – Sincere

The Cronbach's alpha for this scale is 0.87, which is considered good by Peterson (1994).

The final operationalized model is the following:

Framework	Measure	Items	Scale	Reference	α
IV	AGRS	Stimuli	Not applicable	*na*	*na*
Moderator	Review Valence	Stimuli	Not applicable	*na*	*na*
Mediator	Perceived Helpfulness	4*	7-point Likert Scale	(Wu, 2013)	0.89
DV	Consumer Trust	3	7-point Likert Scale	(Ohanian, 1990)	0.87

Table 1: Variable Measurement Scales

**Ohanian (1990); Adapted from original 5 Items*

3.3 Stimuli Development

The multimodal stimuli comprised a combination of static visual and text-based elements to present AGRS alongside human-written reviews. Stimuli development followed four phases: (1) category identification, (2) composition selection, (3) stimuli validation, and (4) the main study.

3.3.1.1 Category Identification

A utilitarian search good was selected based on literature showing that consumers attribute negative review sentiment more to the product than to the reviewer when evaluating utilitarian rather than hedonic goods (Sen & Lerman, 2007). Ten consumer interviews were conducted to validate and refine this choice (see Appendix A.1 Pre-test Interview: Category Identification & Characteristics of the Review & A.2 Pre-test Interview Results Category Identification). The considered product categories included typical search goods with extensive review data, including consumer electronics, home appliances, and electronics. The final product had to meet the following criteria: it required careful purchasing evaluation, involved technical complexity, avoided well-established brands to reduce pre-existing brand trust, and was not a low-cost impulse purchase.

A vacuum cleaner was selected as it met all these requirements. Its functional variety and technical features make detailed reviews valuable for evaluation purposes. To avoid brand bias, no brand name was shown in the stimuli. Participants were shown the product page (see Appendix A.3 below) before being exposed to the stimuli.

3.3.1.2 Composition Selection

A pre-survey was conducted to determine the composition and information load of the stimuli (see Appendix A.4 below). Participants were randomly assigned to one of two conditions: one group viewed individual reviews, summaries, and a combined format (summary and two reviews), while the other group viewed the same formats with added star ratings and helpfulness votes. Each participant assessed the helpfulness and information sufficiency of all three formats.

Across both conditions, the combination of an AGRS with two individual reviews was rated most helpful and informative. In the star-rating condition, 79% of participants noticed the ratings and reported higher trust, confirming that numerical ratings serve as key heuristic cues and enhance ecological validity by reflecting real-world settings. Conversely, helpfulness votes had limited impact: fewer than half of the respondents noticed them, and only 18% stated that they influenced their trust in the reviews. Given their low explanatory power, helpfulness votes were excluded.

The final design featured two individual reviews, while the experimental condition additionally included an AGRS. Star ratings were displayed, but helpfulness votes were omitted. Based on feedback, richer technical product details were integrated to improve realism, and as mobile devices are the preferred choice for online shopping (Natarajan et al., 2018), all stimuli were designed in a mobile-friendly view.

The graphical development process consisted of three steps:

Step 1		Step 2		Step 3
Content Identification		**Stimuli Generation**		**Visual Design**
Real amazon.com user reviews were analyzed to identify key product and review features to include in the stimuli.	→	Six reviews were generated using ChatGPT, varying only in specific attributes (valence) while holding all other content constant.	→	Amazon's layout and branding were recreated using Canva to enhance ecological validity.

Table 2: Stimuli Creation Process

3.3.1.3 Validating the Stimuli

The stimuli were tested for validity, to ensure they accurately reflected the review format and valence, and for reliability, to ensure participants consistently recognized these manipulations (Malhotra et al., 2017; Sekaran & Bougie, 2016).

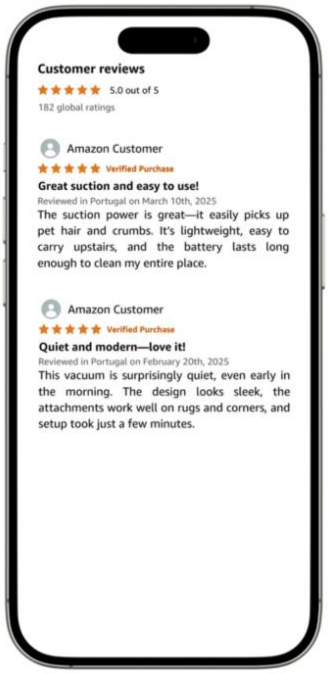
The first pre-survey included a manipulation check to ensure that review valence was correctly perceived. A second pre-survey validated the final stimulus composition and manipulations across positive, negative and two-sided conditions (see Appendix A.5 below). Participants rated two randomly selected stimuli on a seven-point Likert scale (1 = very negative, 7 = very positive) to confirm alignment with the intended sentiment. The results showed high validity: 92.2% of participants in the first pre-survey and 95.9% in the second correctly identified the intended sentiment.

The main study included two manipulation checks to ensure validity. Participants were asked to indicate whether they had seen only reviews, an AGRS plus reviews, or whether they were unsure. They were also asked to classify the overall sentiment as positive, negative, or two-sided. Only responses from participants who correctly identified both manipulations were included in the final dataset.

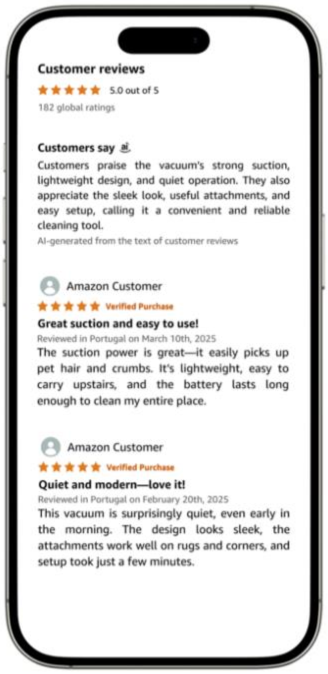
3.3.1.4 Final Stimuli

The final experimental design included six conditions: each valence (positive, negative, two-sided) was presented with and without the AGRS. The final stimuli are presented below:

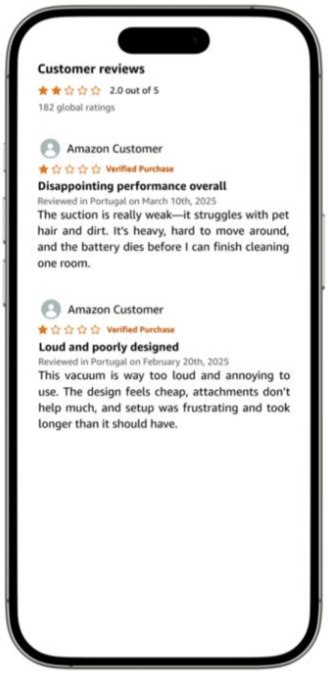
Positive Valence – Without AGRS



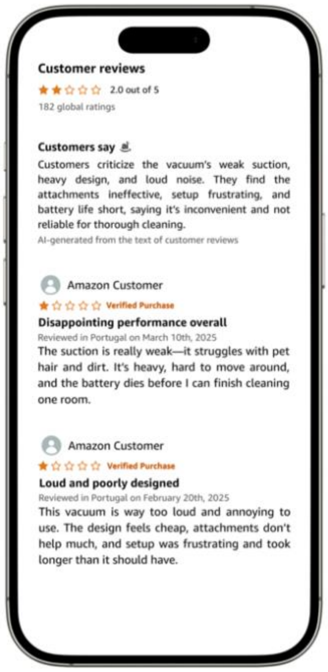
Positive Valence – With AGRS



Negative Valence – Without AGRS



Negative Valence – With AGRS



Two-Sided Valence – Without AGRS	Two-Sided Valence – With AGRS

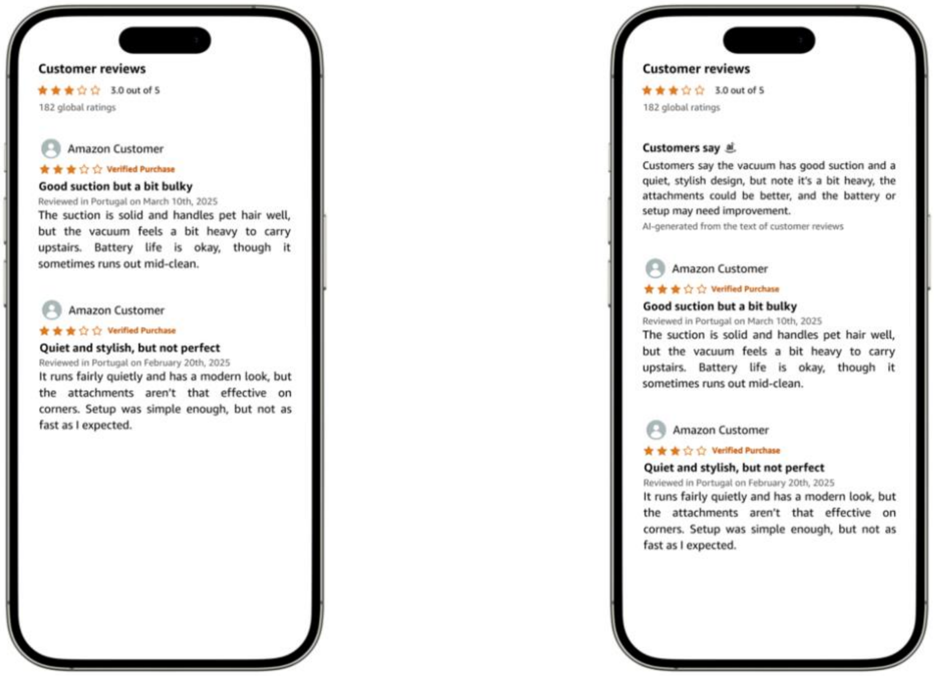


Figure 3: Final Stimuli 2x3 Experimental Design

3.4 Questionnaire Design

The final survey employed a 2×3 between-subjects design with two AGRS presence conditions and three valence levels. Respondents were randomly assigned to one of six groups and exposed to a single stimulus to prevent carry-over effects and reduce demand characteristics (Charness et al., 2012). The experimental setup is summarized in Table 3.

	Review Valence		
Review Summary Presence	**Positive**	**Negative**	**Two-Sided**
Without AGRS	01 WOS-P	02 WOS-N	03 WOS-2S
With AGRS	04 WS-P	05 WS-N	06 WS-2S

Table 3: Survey Research Design: 2x3 Between-subjects Factorial Design

The survey flow was structured using a block design, grouping related sections for modular sequencing and randomization. The blocks measuring helpfulness and trust were randomized to limit the effects of order. To avoid influencing participants' initial impressions, the manipulation checks were placed after these measures. While this approach minimized demand bias, it introduced the risk of recall issues. To compensate for this, a larger sample size was collected to ensure a minimum of 30 valid responses per condition. Table 10 (See Appendix) provides a detailed overview of the survey logic, including condition assignment and randomization.

3.5 Data Analysis

The quantitative data collected were obtained via Qualtrics and analyzed using IBM SPSS Statistics 30.0 to test the proposed hypotheses and assess interactions between variables. The data was thoroughly cleaned, and only participants who passed both manipulation checks and completed the entire survey were retained for the analysis. Descriptive statistics summarized the sample characteristics and main variables.

A range of statistical tests was applied for hypothesis testing. Independent samples t-tests were used to test the hypotheses and compare group differences, while a linear regression was used to examine the relationship between helpfulness and trust. To assess more complex effects, Hayes' PROCESS macro was applied to analyze moderation (Model 1), mediation (Model 4), and the full model (Model 5). The statistical assumptions for all models were tested to ensure validity, while Cronbach's alpha was used to evaluate scale reliability. The experimental conditions (AGRS presence and review valence) were randomly assigned and treated as categorical variables. All statistical tests applied a 5% significance level. The table below illustrates the statistical tests used to test the hypotheses.

Hypotheses	**Statistical Test**
H1: Presence of AGRS leads to lower trust	Independent-samples t-test
H2: Review valence moderates the effect of AGRS on trust **H2a**: Negative valence in AGRS leads to higher trust than positive valence **H2b**: Two-sided valence in AGRS leads to higher trust (vs. positive or negative)	Hayes Process Macro Model 1 (Moderation analysis) & Contrast analysis (Simple slopes via Hayes PROCESS)
H3: Higher helpfulness of AGRS leads to higher trust	Simple linear regression
H4: Presence of AGRS leads to higher helpfulness	Independent-samples t-test
H5: The effect of AGRS on trust is mediated by their helpfulness	Hayes Process Macro Model 4 (Mediation analysis)
Full Model: Combined moderated mediation model with trust as outcome, AGRS as predictor, helpfulness as mediator, and valence as moderator	Hayes Process Macro Model 5

Table 4: Statistical Tests used for Analysis

The moderation mediation model used to identify the statistical effects of the full model is shown in Figure 4:

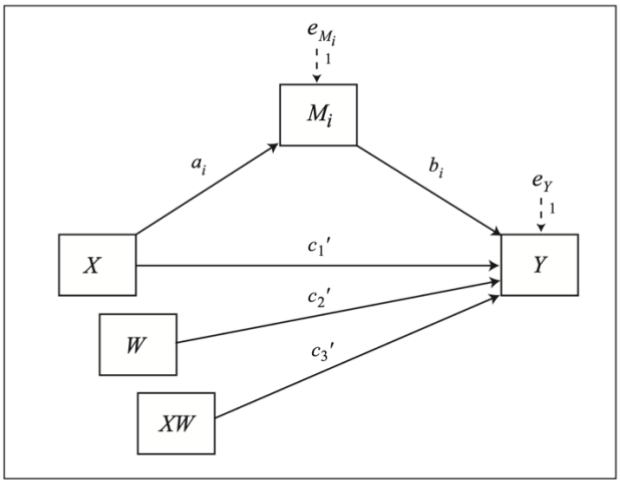


Figure 4: Statistical Model: Hayes' PROCESS Model 5

4 Results

This section presents the results of the quantitative analysis. First, the sample's characteristics are described, followed by an assessment of measurement reliability. Finally, the formulated hypotheses are tested, and the related statistical results are presented.

4.1 Data Preparation

Several actions were performed to prepare the dataset for analysis. System-generated variables from Qualtrics were omitted to protect participant privacy and retain only relevant data. Variable names were changed for clarity, and appropriate value labels were assigned.

4.1.1 Missing Data

Out of 649 total responses, 90 (13.9%) were excluded due to system-missing values. Nine participants were excluded because they had never shopped online (2 respondents) or never read reviews (7 respondents). Additionally, 40 respondents (6.2%) dropped out before completing at least 95% of the survey. Overall, 559 valid responses were retained before conducting manipulation checks.

4.1.2 Manipulation Checks

Two manipulation checks ensured participants perceived the independent variable (AGRS present vs. not present) and the moderator (review valence) as intended.

For the AGRS manipulation, 61.4% correctly identified their condition (59.2% the summary; 63.3% its absence), with significant group differences ($\chi^2(2, N = 559) = 131.98$, $p < 0.001$; see Table 12 & Table 13). For valence, correct identification rates were 76.6% (positive), 90.4% (negative), and 80.3% (two-sided), with significant differences ($\chi^2(4, N = 559) = 630.48$, $p < .001$; see Table 16), indicating effective manipulation. Overall, 82.5% recalled the valence correctly (see Table 14).

In total, 52.8% correctly identified both manipulations. Only these participants were retained to ensure data quality and reduce issues from inattention or misunderstanding.

4.1.3 Outlier Identification

To detect multivariate outliers, composite scores for trust and helpfulness were used to calculate Mahalanobis distances. Cases exceeding the critical value ($\chi^2 = 13.82$, $p < 0.001$, $df = 2$) were flagged as outliers. Eight such cases were identified and subsequently excluded.

After removing missing data, outliers, and inattentive responses, only participants who completed the questionnaire and passed both manipulation checks were retained, yielding a final sample of N = 291. Thus, no data imputation was necessary (Hair et al., 2019). The number of valid responses per stimulus is shown in Table 5 below.

		Frequency	*Percent*	*Valid Percent*	*Cumulative Percent*
Valid	01 WOS-P	36	12,4	12,4	12,4
	02 WOS-N	59	20,3	20,3	32,6
	03 WOS-2S	58	19,9	19,9	52,6
	04 WS-P	45	15,5	15,5	68,0
	05 WS-N	57	19,6	19,6	87,6
	06 WS-2S	36	12,4	12,4	100,0
	Total	**291**	**100,0**	**100,0**	

Table 5: Valid responses per Stimulus

4.2 Measurement Overview

This section outlines how the key variables were operationalized and assessed for reliability. Multi-item scales were tested for internal consistency to ensure that they reliably measure the intended constructs.

The independent variable, presence of AGRS, was coded as a categorical variable with two levels, based on stimulus assignment:

1. No Summary (Stimuli 1–3)
2. With Summary (Stimuli 4–6)

The moderator, review valence, was also coded categorically with three levels:

1. Positive (Stimuli 1, 4)
2. Negative (Stimuli 2, 5)
3. Two-Sided (Stimuli 3, 6)

The dependent variable, consumer trust, was measured using a 7-point Likert-type scale adapted from Ohanian (1990), refined by Reimer & Benkenstein (2016). Participants rated four bipolar items, which were averaged to form a composite trust score. The reliability analysis yielded a Cronbach's alpha value of $\alpha = 0.86$ (see Table 18), indicating good reliability (George

& Mallery, 2019). The corrected item-total correlations were above 0.66, and deleting items did not improve reliability (see Table 17). Therefore, all four items were retained in the final scale.

The mediator perceived helpfulness was measured using a 7-point Likert scale adapted by (Wu, 2013). Participants rated three statements, which were averaged to produce a composite helpfulness score. The Cronbach's alpha coefficient was α = 0.78 (see Table 19), indicating acceptable reliability (George & Mallery, 2019). All corrected item-total correlations were above 0.57, and no improvement was observed through item deletion (see Table 22). Therefore, all items were retained in the final scale.

4.2.1 Measurement Summary

The models' key variables are presented in the following summary table:

Variable	**Description**	**Values**	**Measure**
Ai_summary	Predictor dummy variable indicating exposure to an AI-generated summary	1 = No Summary 2 = With Summary	Nominal
Valence	Predictor categorical variable representing the sentiment of the reviews	1 = Positive 2 = Negative 3 = Two-Sided	Nominal
Trust_scale	Dependent variable measuring mean trust score across multiple items	1 to 7	Scale
Helpfulness_ scale	Dependent variable measuring mean helpfulness score across multiple items	1 to 7	Scale

Table 6: Model Variables

The descriptive statistics for the main variables are as follows:

Variable	**Type**	**Min**	**Max**	**Mean**	**Std. Deviation**	**α**
Consumer Trust	DV	2,00	7,00	5,6143	,91682	.860
Perceived Helpfulness	Mediator	2,67	7,00	5,9679	,72538	.780
Ai Summary	IV	1	2	1,47	,500	*na*
Valence	Moderator	1	3	2,04	,776	*na*

Table 7: Descriptive Statistics and Reliability of Main Constructs

4.3 Sample Characterization

Following data preparation, the final sample consisted of N=291 participants. The gender distribution was predominantly female (72.9%), followed by male (23.4%), with a smaller proportion identifying as non-binary or choosing not to disclose (3.8%) (see Table 24). The majority were young adults: 47.4% were aged 18–24, and 35.7% were aged 25–34. Participants under 18 accounted for 7.6%, while older age groups (35 and older, 9.3%) were less represented (see Table 23). In terms of education, 40.5% of respondents held an undergraduate degree, followed by 24.1% with a graduate degree and 23.0% with a high school diploma (see Table 25). The sample included respondents from over 30 countries, the largest groups being from Germany (49.1%), the United Kingdom (13.7%), the Netherlands (6.9%), and the United States (6.2%) (see Table 26). Despite numerous responses being excluded, the demographic characteristics of the stimulus groups remained well balanced, indicating that random assignment worked as intended.

4.4 Hypothesis Testing

This section presents the statistical analyses employed to test the study's hypotheses. The effects of the experimental manipulations on the outcome variables were examined using appropriate inferential statistical methods.

Before testing the hypotheses, diagnostics were conducted to assess multicollinearity and interdependence among the predictor variables. The diagnostic measures, including the variance inflation factor (VIF), eigenvalues, and condition index, met the established thresholds (VIF < 2.5, Eigenvalue > 0.01, Condition Index < 30.0) for acceptable multicollinearity. All VIF values for the predictors were well below the critical threshold, with the highest value being 1.013 (see Table 27), indicating minimal shared variance (Hair et al., 2019). The maximum condition index was 24.578, remaining below the conventional threshold of 30 for severe multicollinearity (Belsley et al., 1980). Although the lowest eigenvalue was 0.006 (slightly below the ideal threshold value of 0.01), variance share analysis showed that no two predictors within a dimension with a high state index exceeded the critical value of 0.5 (see Table 28). These results suggest that multicollinearity did not compromise the stability or interpretability of the regression coefficients.

A significance level of $\alpha = 0.05$ was applied, in line with social science standards (APA, 2020). The between-subjects design ensured independence of observations, as each participant was exposed to only one condition. The following table provides an overview of the tested hypotheses, their corresponding null hypotheses (H_0) and the statistical methods used.

Hypothesis Type	**Hypothesis**	**Statistical Test**	**Null Hypothesis (H_0)**
Group Difference	H1	Independent-Samples t-test	$\mu_1 = \mu_2$
Moderation (Interaction)	H2	Moderation Analysis (PROCESS Model 1)	β_interaction = 0
Simple Contrast (Moderation)	H2a	Moderation Analysis (PROCESS Model 1, contrast)	β_diff (pos vs. neg) = 0

Hypothesis Type	Hypothesis	Statistical Test	Null Hypothesis (H_0)
Simple Contrast (Moderation)	H2b	Moderation Analysis (PROCESS Model 1, contrast)	β_diff (two-sided vs. others) = 0
Direct Effect	H3	Linear Regression	$\beta = 0$
Group Difference	H4	Independent-Samples t-test	$\mu_1 = \mu_2$
Mediation (Indirect Effect)	H5	Mediation Analysis (PROCESS Model 4, bootstrapping)	$a \times b = 0$

Table 8: Null Hypotheses and Statistical Tests used

4.4.1 AI-Generated Review Summaries and Trust

H1: Consumers perceive review sets accompanied by an AGRS and two individual reviews as less trustworthy than review sets containing only two individual reviews.

To evaluate the impact of an AGRS on consumer trust, an independent samples t-test was performed. Prior to the analysis, the underlying assumptions were evaluated. Although the Shapiro-Wilk test indicated non-normality in both groups ($p < 0.01$; see Table 29), the large sample size ($n > 30$ per group) justified the application of the central limit theorem. Levene's test confirmed the homogeneity of variances ($p = 0.691$; see Table 30). The t-test is therefore considered robust for these data.

The results showed no statistically significant difference in perceived trustworthiness between evaluation sets with and without AGRS ($p = 0.096 > 0.05$, ibid.). Consequently, **H_0 could not be rejected and H1 was not supported**

4.4.2 The Moderating Role of Review Valence

H2: The effect of AGRS on consumer trust is moderated by review valence; specifically, the difference in trust between review sets with and without an AI summary varies depending on whether the review valence is positive, negative, or two-sided.

To test this hypothesis, Hayes' PROCESS macro (Model 1) was performed to examine whether the relationship between the presence of an AGRS (0 = absent, 1 = present) and consumer trust was moderated by review valence (positive, negative, two-sided).

All model assumptions were tested. To assess residual normality and homoscedasticity, the model was replicated using standard linear regression with the same predictors and interaction terms. A Shapiro–Wilk test revealed a significant deviation from normality ($p < .001$; see Table 31). However, given the large sample size, the central limit theorem suggests that the results remain robust. Visual inspection of the residual plot (Figure 6) revealed no apparent patterns or funnel-shaped distribution, suggesting that the assumption of homoscedasticity was met.

The overall moderation model was significant ($F(5, 285) = 3.41$, $p = .005$), explaining 5.6% of the variance in trust ($R^2 = .056$; see Table 33). AGRS had a significant negative main effect on trust ($B = -0.439$, $p = .030$). However, the interaction with review valence was not significant ($F(2, 285) = 2.32$, $p = .100$), indicating that moderation was inconsistent across valence levels.

Simple contrasts suggested possible effects in specific cases; thus, H_0 cannot be fully rejected and **H2 is partially supported**.

H2a: Consumers perceive AGRS based on predominantly positive reviews as less trustworthy than summaries based on predominantly negative reviews, with review valence influencing trustworthiness perceptions.

A contrast comparing negative and positive valences showed a significant interaction ($B = 0.550$, $p = .036$), indicating that the negative effect of AGRS on trust was weaker in the negative condition. Simple slopes analysis showed that AGRS significantly decreased trust in the positive review condition ($B = -0.439$, $p = .030$), while the effect in the negative condition was nonsignificant and slightly positive ($B = +0.111$, $p>.05$; see Table 33 & Table 34).

Therefore, summaries based on positive reviews were perceived as less trustworthy than those based on negative reviews. **H_0 is rejected, thus supporting H2a**.

H2b: Consumers perceive AGRS based on two-sided reviews as more trustworthy than summaries based on predominantly positive or negative reviews, with review valence influencing trustworthiness perceptions.

The interaction between AGRS presence and two-sided valence was not statistically significant ($B = 0.222$, $p = .425$), indicating no evidence that trust in summaries changes meaningfully in the two-sided condition. The conditional effect of the AGRS in this condition was negative but nonsignificant ($B = -0.217$, $>.05$), suggesting that trust in summaries remained low and did not improve when reviews were two-sided.

Since there was no evidence that a two-sided review valence improved trust in AGRS, **H_0 cannot be rejected, and H2b is not supported.**

4.4.3 The Effect of Perceived Helpfulness on Consumer Trust

H3: The perceived helpfulness of AGRS positively influences consumer trust in the reviews.

To test H3, a linear regression analysis was conducted in which perceived helpfulness predicted consumer trust (N = 291).

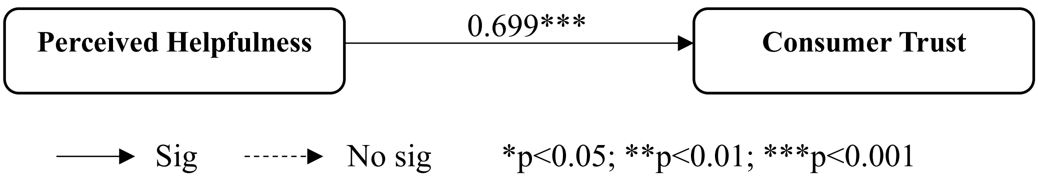
$$Trust_i = \beta_o + \beta_1 \, x \, Helpfulness_i + \epsilon$$

$$i = 1, \ldots, N$$

Assumption testing revealed several violations. The residuals showed signs of positive autocorrelation (Durbin–Watson = 0.818; see Table 36), and the Shapiro–Wilk test indicated non-normality ($p < .001$; Table 35). Visual inspection of the residual scatterplot plot (Figure 7) also suggested a lack of clear linearity and possible heteroscedasticity. These findings indicate that the core assumptions for linear regression were not met, and the results should therefore be interpreted with caution.

Despite the violations, the regression model was statistically significant, $F(1, 298) = 127.59$, $p < .001$, indicating a strong positive association between helpfulness and trust ($R = 0.553$). Perceived helpfulness accounted for 30.6% of the variance in trust scores ($R^2 = .306$; see Table 36). The standardized regression coefficient ($\beta = 0.699$, $p < .001$) suggests that higher helpfulness is associated with consumer trust. Specifically, a one-unit increase in helpfulness was associated with a 0.699-unit increase in trust, holding other factors constant.

Therefore, **H3 was confirmed, and the null hypothesis H_0 of no effect was rejected**.



4.4.4 The Effect of AI-Generated Summaries on Perceived Helpfulness

H4: Consumers perceive reviews accompanied by an AGRS as more helpful than reviews presented without summaries.

An independent samples t-test was used to compare the helpfulness of reviews with and without AGRS. The assumptions were met (Central Limit Theorem applicable; Levene's test, $p = 0.691$; see Table 38). Helpfulness was slightly higher without summaries ($M = 6.04$) than with summaries ($M = 5.89$; see Table 39), but this difference was not significant ($p = 0.087$) and the effect size was small (Cohen's d = 0.20; see Table 40) ,suggesting that the presence of an AGRS had only a minimal practical impact on perceived helpfulness.

Consequently, **H_0 could not be rejected and H4 was not supported.**

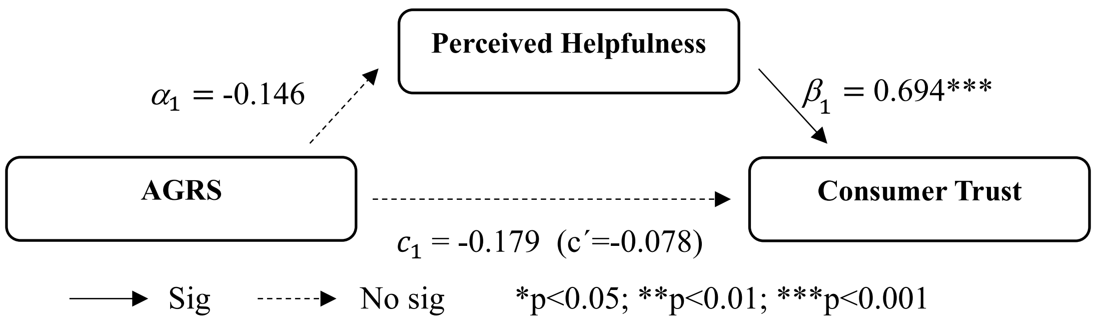
4.4.5 Mediation Test

H5: The effect of AGRS on consumer trust is mediated by their perceived helpfulness.

A mediation analysis using Hayes PROCESS Model 4 was conducted to test whether perceived helpfulness mediated the effect of AGRS on trust. The assumptions were met; although the residuals deviated from normality (Shapiro–Wilk $p < .001$; see Table 44), the large sample size indicates robustness, and the residual plots showed homoscedasticity (see Figure 8).

The indirect effect was non-significant, as were the total effect ($b = –0.179$, $p = .096$) and the direct effect ($b = –0.078$, $p = .387$). Although summaries were linked to a non-significant decrease in helpfulness ($b = −0.146$, $p = 0.087$), helpfulness significantly predicted trust ($b = 0.694$, $p < 0.001$; see Table 41, Table 42 & Table 43.

Therefore, perceived helpfulness did not mediate the relationship between AGRS and trust. **H_0 could not be rejected and H5 was not supported.**



4.4.6 Full Model Test

A moderated mediation analysis was conducted using PROCESS Model 5. AGRS presence was the independent variable (X), perceived helpfulness was the mediator (M), consumer trust was the dependent variable (Y), and review valence was the moderator of the direct effect.

The results revealed that the presence of an AGRS had no significant impact on helpfulness ($B = –0.146$, $p = .087$), though helpfulness was a strong predictor of trust ($B = 0.72$, $p < .001$). The direct effect of summaries on trust was non-significant ($B = –0.104$, $p = .541$), as was the interaction with valence ($p > .05$). The indirect effect via helpfulness was also non-significant

(B = –0.106, 95% CI [–0.229, 0.017]), indicating no support for moderated mediation (see Table 45 & Table 46).

Overall, the full model provided no evidence that AGRS directly or indirectly influenced trust, nor did it interact with review valence or helpfulness. Although helpfulness was a strong predictor of trust, neither summaries nor valence had a significant impact on this relationship.

The results of hypothesis testing and their implications with the full model test are summarized in the following table:

H	Test used	H Result	Full Model Result
1	Independent samples t-test	**Rejected**	Rejected
2	PROCESS Model 1 (Moderation)	**Partially supported**	Rejected
2a&b	Simple slopes & interaction contrast	**Supported**	Partially supported
3	Linear regression	**Supported**	Supported
4	Independent samples t-test	**Rejected**	Rejected
5	PROCESS Model 4 (Mediation)	**Rejected**	Rejected

Table 9: Summary of Hypothesis Testing Results

4.5 Further Results: Cluster Analysis

To complement the tested hypotheses and explore whether online shopping and review behaviors influence trust dynamics, a cluster analysis was conducted. Three user clusters with fair quality emerged: "High-Frequency Shoppers", "Review Reliant Shoppers" and "Ambivalent Reviewers." Their detailed profiles are described in Appendix C.3 below.

A chi-square test of independence was conducted to examine whether consumer responses to AGRS differed across the identified user clusters. The association was not statistically significant ($\chi^2(2, N = 291) = 5.11$, $p = .078$ see Table 47), indicating that the presence of an AGRS did not significantly affect trust among different types of shoppers. However, a significant linear-by-linear association ($\chi^2 = 4.98$, $p = .026$), indicated a potential directional trend. A binary logistic regression was conducted to test the interaction effect of cluster membership and AGRS exposure on trust. However, the model was not statistically significant (see Table 48) and no significant group differences were found (see Table 49). Overall, these findings suggest that shoppers and reviewers across user clusters perceive AGRS similarly in terms of trust.

4.6 Key Findings and Discussion

This study examined five hypotheses to determine whether AGRS that subtly disclose AI involvement and are formatted like Amazon's reviews influence consumer trust. The results showed no significant effect of AGRS on trust or helpfulness, nor a mediation effect through helpfulness. However, helpfulness emerged as a strong predictor of trust, suggesting that consumers still rely heavily on the quality of information provided. Review valence had a partial effect: summaries based on positive reviews were perceived as less trustworthy than those based on negative or two-sided reviews.

4.6.1 The role of AI in Trust Formation

A key finding of this study is that the inclusion of an AGRS did not significantly impact consumer trust. This challenges the common assumption in existing literature that AI-generated content is less trustworthy due to its perceived lack of authenticity, emotional nuance or contextual depth (Jia et al., 2025; Lim & Schmälzle, 2024; X. Luo et al., 2019).

This outcome aligns with trust transfer theory (Stewart, 2003), which suggests that trust in a credible entity, such as a well-known e-commerce platform, can be transferred to its less familiar features. In this case, the AGRS was embedded within Amazon-style interfaces, likely triggering institutional trust (McKnight & Chervany, 2001). When institutional trust is high, consumers may rely on the platform’s overall credibility rather than scrutinizing each piece of content individually.

This effect is further supported by heuristic processing. In information-rich environments such as online marketplaces, consumers often rely on peripheral cues such as layout, branding and design when assessing trustworthiness (L. Y. Pan & Chiou, 2011). Because the summary was seamlessly integrated into the familiar structure of a trusted platform, participants likely did not critically examine its origin or authorship. Instead, they viewed it as a neutral efficiency tool, especially since it was accompanied by genuine individual reviews.

Notably, prior research has not fully examined how trust operates in hybrid review formats. While real-world platforms increasingly present AI-generated and human-written content together, existing studies typically compare them in isolation. This study contributes to existing research by demonstrating that, in blended contexts, AGRS may not be noticeable enough to significantly impact trust.

Overall, these findings suggest that, when subtly disclosed and integrated into a trusted, familiar interface, AGRS neither significantly enhance nor undermine trust. This emphasizes the context-dependent nature of AI credibility and highlights the need for further research into the circumstances in which AI-generated content becomes influential, particularly in hybrid review environments.

4.6.2 Perceived helpfulness as a Trust cue

This study confirms previous research indicating that perceived helpfulness is a strong predictor of trust in review content. This is because specific, evaluative, and diagnostic information facilitates informed decision-making (Ghose & Ipeirotis, 2011; Mudambi & Schuff, 2010). These findings are consistent with the ELM, which posits that motivated and able consumers prioritize content quality over superficial cues (Petty & Cacioppo, 1986).

However, the AGRS did not improve perceptions of helpfulness or mediate the relationship between AGRS and trust. Participants already considered the reviews to be helpful and trustworthy, and the summary added little informational value. This contradicts earlier findings that summaries can enhance decision-making by reducing cognitive effort (Chua & Banerjee, 2016; Filieri, 2014; Jia et al., 2025).

One possible explanation for this discrepancy is the controlled nature of the study design: participants were required to engage with the stimuli for at least 30 seconds, encouraging central-route processing and reliance on diagnostic information rather than peripheral cues (Petty & Cacioppo, 1986). Here, the AGRS may have seemed unnecessary, as it did not present new information or reduce cognitive load. Drawing on the Technology Acceptance Model, although not perceived negatively, AGRS failed to reach the threshold of usefulness required to influence trust. In real-world contexts of information overload, however, well-designed summaries may play a greater role by reducing cognitive burden. This is particularly true if they incorporate the emotional authenticity, nuance or narrative depth that consumers associate with helpful content (Z. Luo et al., 2024; Namvar & Chua, 2023).

However, this study shows that merely including a summary is insufficient. Future research should explore how elements such as tone, specificity, and contextual considerations (e.g., time constraints and review quantity) influence the perceived helpfulness of AGRS.

4.6.3 Valence as a Moderator

One of the most significant findings of this study is that AGRS demonstrated a reduction in trust exclusively in the positive valence condition. When positive reviews were accompanied by a summary, participants expressed significantly lower trust than when they were presented alone, thereby supporting H2a. This is consistent with previous research indicating that overly positive content may appear biased or promotional (Chatterjee, 2001; K. T. Lee & Koo, 2012). A possible explanation is that, rather than enhancing informativeness, summarization appeared to dilute the evaluative richness of the reviews. This effect is reinforced under central-route processing, whereby consumers pay close attention to depth and relevance (Petty & Cacioppo, 1986). The lack of nuance in overly positive summaries may invite consumer skepticism by glossing over details consumers rely on to make balanced judgments (Filieri, 2016; K. T. Lee & Koo, 2012). Negativity bias further explains why positive reviews are vulnerable to trust erosion, as consumers typically assign more weight to critical content, perceiving it as more authentic and informative (Herr et al., 1991). This suggests that positive valence may be incompatible with the current form of summarization.

In the negative valence condition, AGRS exerted no significant effect. Critical reviews appeared to demonstrate a greater degree of resilience when summarized, a phenomenon that can be attributed to the notion that negativity is often associated with authenticity and objectivity (Chevalier & Mayzlin, 2006; Filieri, 2016). However, the existence of a AGRS did not improve trust either, raising questions about its added value when the content is already credible.

Contrary to expectations, two-sided summaries also failed to increase trust. Despite prior research into the persuasive power of balanced messaging (Filieri, 2016; Purnawirawan et al., 2015), AGRS might have blurred the line between positive and negative aspects, weakening the balance that usually boosts authenticity. Participants may have found the summaries overly polished, reducing their persuasive impact. This suggests that the mere presence of opposing viewpoints is not enough and that they must be recognizably preserved.

Overall, AGRS interact differently depending on the valence of the reviews they are based on: positive reviews lose trustworthiness when summarized, whereas negative and two-sided reviews appear to be more resilient to trust erosion. However, without direct measurement of summary quality, it is unclear whether the null effects in these conditions are due to technological limitations or greater tolerance of summarized critical content. Future research should explore this distinction by varying the perceived informativeness and balance of critical details in AGRS to better understand how review valence interacts with summary design to shape trust.

5 Conclusion and Limitations

This study examined whether and under what conditions AI-generated review summaries (AGRS) influence consumer trust in online reviews in a realistic retail setting. Three key objectives were addressed: to compare trust in reviews with and without a summary; to examine the moderating role of review valence; and to test perceived helpfulness as a potential mediator. To investigate these objectives, the study employed a six-condition, Amazon-style design, presenting each valence level with and without a summary.

This final chapter synthesizes the main findings concerning each research question and discusses their implications. It also outlines the study's practical and academic contributions, acknowledges its limitations, and suggests avenues for future research.

5.1 Main Findings

RQ1: How does the presence of an AGRS influence consumer trust in online reviews?

The presence of an AGRS had no significant effect on consumer trust in the set of reviews. This null effect likely reflects the transfer of trust to the familiar interface of the platform, as well as the fact that heuristic processing reduced the prominence of AI authorship. Under the tested conditions, the summary neither increased nor decreased trust.

RQ2: Does perceived helpfulness act as a mediator in the relationship between AGRS and consumer trust?

Perceived helpfulness did not mediate the relationship, as the summary did not significantly improve helpfulness and thus could not affect trust. However, helpfulness still predicted trust independently and positively, indicating that diagnostic content remained a primary driver of trust within the study conditions.

RQ3: Does review valence moderate the effect of AGRS on consumer trust?

Review valence moderated the effect of AGRS on consumer trust, but only in the positive condition. When reviews were positive, adding a summary significantly reduced trust. In negative and two-sided conditions, however, no significant differences emerged. This suggests that positive reviews are more vulnerable to trust erosion when summarized, likely due to a perceived promotional tone and reduced evaluative richness.

5.2 Theoretical Implications

This study advances theory on AI-generated content and consumer trust in hybrid review settings. The presence of an AGRS was not found to influence trust, reinforcing the idea that consumers rely more on the helpfulness of human-written reviews than on the mere presence of AI. This expands trust-transfer theory (Stewart, 2003) by showing that platform credibility can extend to AI-generated elements when the interface is familiar, and AI involvement is disclosed. In such cases, human reviews seem to anchor trust, allowing AGRS to be passively accepted even when they provide little additional value.

From a mechanistic perspective, these patterns reflect how consumers process information in different situations. The requirement to engage with the content for at least 30 seconds likely activated central-route processing (Petty & Cacioppo, 1986), prompting careful evaluation of the reviews. Because the AGRS offered little novelty or efficiency, its perceived usefulness did not surpass the threshold in the TAM (Davis, 1989) necessary to enhance helpfulness and, consequently, trust. By contrast, under peripheral-route conditions such as time pressure or information overload, reducing effort becomes a dominant cue. In such contexts, an AGRS that is contrast-rich, and effort-saving is more likely to enhance perceived usefulness, which should increase perceived helpfulness and foster trust.

Beyond contextual and processing dynamics, this study advances theory by showing that AGRS effects are valence sensitive. Summarizing positive reviews reduced trust, which is consistent with the idea that excessive positivity appears biased or promotional. In contrast, negative and two-sided content did not produce this effect, suggesting that trust erosion is specific to perceived over-positivity. While participants read both reviews and summaries without time pressure or purchase intent, positivity bias did not appear to influence participants' judgments. However, it may occur in real-world decision-making contexts during heuristic processing, challenging the assumption that AGRS are neutral tools and highlighting the risk of losing trust when summarizing uniformly positive content.

Overall, the influence of AGRS on trust appear to be both contextual and threshold based. When platform trust is high and disclosure is subtle, AGRS have little effect. However, when they clearly enhance clarity or efficiency, their impact may become more pronounced.

5.3 Managerial Implications

Although AGRS can support consumer decision-making, they should only be embedded when clearly adding value by reducing effort or increasing helpfulness. Otherwise, they offer little benefit and may even undermine trust. This study provides managers in e-commerce and other review-based sectors with actionable insights by testing AGRS in a realistic shopping interface that mirrors actual consumer experiences.

The results suggest that AGRS should be implemented selectively, based on evidence that they enhance helpfulness. Where this threshold is not met, trust remains unchanged, and no added benefit can be expected. Managers are therefore advised to run small-scale pilots and track helpfulness ratings, time spent, and review engagement before scaling up deployment. Individual reviews should always be kept visible, with summaries presented as a complement rather than a substitute. AGRS should be accompanied by clear, non-intrusive disclosure of their AI origin to maintain transparency without distracting users. To maximize value, summaries can be optimized for efficiency; for example, they could be presented as scannable bullet points or contain links to the original reviews that highlight key product features.

Special attention should be paid to summaries based on positive reviews. This study found that such AGRS reduced trust, suggesting that they should either be avoided or carefully redesigned to include specific, balanced, and informative content. Until further evidence supports their effectiveness, AGRS may be more effective when applied to negative or two-sided reviews.

Finally, decisions about implementation should not be based solely on perceived trust or helpfulness. Evaluation should be guided by broader performance metrics, such as conversion rates, checkout times, order value and return rates. Realistic testing is recommended, which implies mobile-first trials, high review volumes, time-pressure variants and pilots with explicit success and rollback thresholds before scaling.

5.4 Further Analysis

Although this study examined AGRS in a realistic e-commerce context, there are still several areas to be explored in future research. Firstly, the central-route processing assumed to underlie the observed neutrality of AGRS should be tested more directly to confirm that the ELM adequately explains the findings. Future studies should measure user motivation and ability while manipulating factors such as time pressure, review complexity or cognitive load. This helps to identify when consumers shift to peripheral processing, in which summaries may exert a greater influence by reducing effort. Similarly, while platform trust likely facilitated trust transfer in this study, its role may be weaker in unfamiliar or lower-trust contexts. Varying interface familiarity, disclosure prominence, and visual integration could clarify the circumstances in which AI-generated content becomes more noticeable or scrutinized.

The design of AGRS content also warrants closer examination. Future work should explore how tone, specificity, and balance influence trust, while evaluating summary quality using metrics such as informativeness and clarity. Testing AGRS in more naturalistic, high-load environments, such as mobile shopping, crowded review sections, or time-sensitive decision-making, may better capture its functional value. With these adaptations, the current research design could produce different outcomes.

Particular attention should be given to review valence: Incorporating mild criticism, emotional cues, or verified buyer signals to positive summaries could improve authenticity. For two-sided content, clearer separation of the pros and cons may help to preserve balance and credibility. Additionally, future research could examine how processing depth and reliance on summaries are shaped by involvement level or device type.

Finally, little is known about how AGRS influence user behavior. Future research should examine whether summaries reduce the number of reviews that users read and if they are perceived as genuinely timesaving. As AI adoption increases, longitudinal studies can assess the effects of repeated exposure to AGRS on trust, engagement, and decision-making over time. Together, these areas of research provide a roadmap for refining AGRS design, and supporting more robust, user-centered applications of generative AI.

5.5 Limitations

While this study provides valuable insights into the impact of AGRS on trust and helpfulness, it is important to acknowledge several limitations.

First, the use of a convenience sample resulted in an overrepresentation of young, educated, Western women, which limits the study's generalizability. The modest sample size and uneven group distributions suggest that some conditions may not reflect broader attitudes. Additionally, the simulated Amazon-style interface focused on a single utilitarian B2C product, which makes the results less applicable to hedonic or high-involvement purchases and B2B contexts, where trust dynamics differ. In addition, the stimuli were static and non-interactive, which diverges from real-world responses to dynamic systems. The absence of time pressure and information overload likely encouraged central-route processing, likely reducing the influence of summaries.

Theoretical models such as the ELM, TAM, and trust-transfer theory were applied, though they were not directly measured. Constructs such as motivation, ability, and platform trust were inferred, leaving explanatory boundaries open. Furthermore, the study relied on one-time exposure and self-reported data, and there were violations of regression assumptions, all of which require cautious interpretation. Finally, the cross-sectional design cannot capture how trust in AGRS evolves over time. Longitudinal studies could examine how repeated exposure, familiarity, and shifting digital norms affect consumer trust and behavior.

Overall, these limitations underscore the necessity of more diverse samples and real-world testing under varying conditions to better grasp the role of AI-generated review summaries in digital decision-making processes.

References

Ahluwalia, R. (2000). Examination of psychological processes underlying resistance to persuasion. *Journal of Consumer Research*, *27*(2), 217–232. https://doi.org/10.1086/314321

Anderson, J. C., & Narus, J. A. (1990). A Model of Distributor Firm and Manufacturer Firm Working Partnerships. *Journal of Marketing*, *54*(1), 42–58.

APA. (2020). *Publication manual of the American psychological association 2020*. American psychological association.

Ba, S., & Pavlou, P. A. (2002). Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior. *MIS Quarterly*, *26*(3), 243–268.

Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. In *International Journal of Electronic Commerce*, 17(2), pp. 99–126). https://doi.org/10.2753/JEC1086-4415170204

Belch, G. E. (1981). An examination of comparative and noncomparative television commercials: The effects of claim variation and repetition on cognitive response and message acceptance. *Journal of Marketing Research*, *18*(3), 333-349.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons Inc.

Brand, B. M., & Reith, R. (2022). Cultural differences in the perception of credible online reviews – The influence of presentation format. *Decision Support Systems*, *154*. https://doi.org/10.1016/j.dss.2021.113710

Brand Rated. (2022, January 13). *Nine out of ten customers read reviews before buying a product*. GlobeNewswire.com

Brown, J. J., & Reingen, P. H. (1987). Social Ties and Word-of-Mouth Referral Behavior. In: *Journal of Consumer Research.* 14(3). https://www.jstor.org/stable/2489496

Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems*, *50*(2), 511–521. https://doi.org/10.1016/j.dss.2010.11.009

Chaiken, S. (1980). Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion. In *Journal of Personality and Social Psychology*. 39(5).

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior and Organization*, *81*(1), 1–8. https://doi.org/10.1016/j.jebo.2011.08.009

Chatterjee, P. (2001). *Online Reviews: Do Consumers Use Them?* (pp. 129–134). Association for Consumer Research.

Chen, C. C., & Tseng, Y. De. (2011). Quality evaluation of product reviews using an information quality framework. *Decision Support Systems*, *50*(4), 755–768. https://doi.org/10.1016/j.dss.2010.08.023

Chen, X., Huang, Q., Davison, R. M., & Hua, Z. (2015). What Drives Trust Transfer? The Moderating Roles of Seller-Specific and General Institutional Mechanisms. *International Journal of Electronic Commerce*, *20*(2), 261–289. https://doi.org/10.1080/10864415.2016.1087828

Chen, Y., & Xie, J. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Science*, *54*(3), 477–491. https://doi.org/10.1287/mnsc.1070.0810

Cheung, C. M. K., & Thadani, D. R. (2012). The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decision Support Systems*, *54*(1), 461–470. https://doi.org/10.1016/j.dss.2012.06.008

Cheung, C. M. Y., Sia, C. L., & Kuan, K. K. Y. (2012). Is this review believable? A study of factors affecting the credibility of online consumer reviews from an ELM perspective. *Journal of the Association for Information Systems*, *13*(8), 618–635. https://doi.org/10.17705/1jais.00305

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, *43*(3), 345–354. https://doi.org/10.1509/JMKR.43.3.345

Chien, A., Tang, H., Jagessar, B., Chang, K. W., Peng, N., Nael, K., & Salamon, N. (2024). AI-Assisted Summarization of Radiologic Reports: Evaluating GPT3davinci, BARTcnn, LongT5booksum, LEDbooksum, LEDlegal, and LEDclinical. *American Journal of Neuroradiology*, *45*(2), 244–248. https://doi.org/10.3174/ajnr.A8102

Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets. *Marketing Science*, *29*(5), 944–957. https://doi.org/10.1287/mksc.ll00.0572

Chua, A. Y. K., & Banerjee, S. (2016). Helpfulness of user-generated reviews as a function of review sentiment, product type and information quality. *Computers in Human Behavior*, *54*, 547–554. https://doi.org/10.1016/j.chb.2015.08.057

Clemons, E. K., Gao, G., & Hitt, L. M. (2006). When online reviews meet hyperdifferentiation: A study of the craft beer industry. In *Journal of Management Information Systems,* 23(2), pp. 149–171). https://doi.org/10.2753/MIS0742-1222230207

Creswell, J. W. (2009). *Research design- Qualitative, quantitative, and mixed methods approaches* (3rd ed.). SAGE Publications Inc.

Crowley, A. E., & Hoyer, W. D. (1994). An Integrative Framework for Understanding Two-Sided Persuasion. In*: Journal of Consumer Research*. 20(4), 561-574.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information. In*: MIS Quarterly*, 13(3), 319-340.

de Vaus, D. (2002). Surveys in Social Research. In *Surveys in Social Research, 6th Edition*. Allen & Unwin. https://doi.org/10.4324/9780203519196/SURVEYS-SOCIAL-RESEARCH-DAVID-DE-VAUS-DAVID-DE-VAUS

De Wulf, K., Odekerken-Schröder, G., & Iacobucci, D. (2001). Investments in consumer relationships: A cross-country and cross-industry exploration. *Journal of Marketing*, *65*(4), 33–50. https://doi.org/10.1509/JMKG.65.4.33.18386

Dellarocas, C., Zhang, X., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, *21*(4), 23–45. https://doi.org/10.1002/dir.20087

Dick, A., Chakravarti, D., & Biehal, G. (1990). Memory-Based Inferences during Consumer Choice. In*: Journal of Consumer Research*, 17(1). https://about.jstor.org/terms

Doney, P. M., & Cannon, J. P. (1997). An Examination of the Nature of Trust in Buyer-Seller Relationships. In*: Journal of Marketing*, 61(2).

Dong, B., Li, M., & Sivakumar, K. (2019). Online Review Characteristics and Trust: A Cross-Country Examination. *Decision Sciences*, *50*(3), 537–566. https://doi.org/10.1111/deci.12339

Duan, W., Gu, B., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales-An empirical investigation of the movie industry. *Journal of Retailing*, *84*(2), 233–242. https://doi.org/10.1016/j.jretai.2008.04.005

Eagly, A., & Chaiken, S. (1993). *The psychology of attitudes.* Harcourt brace Jovanovich college publishers. https://psycnet.apa.org/record/1992-98849-000

Falatouri, T., Hrusecka, D., & Fischer, T. (2024). Harnessing the Power of LLMs for Service Quality Assessment from User-Generated Content. *IEEE Access*, *12*, 99755–99767. https://doi.org/10.1109/ACCESS.2024.3429290

Ferdaus, M., Abdelguerfi, M., Ioup, E., Niles, K. N., Pathak, K., & Sloan, S. (2024). Towards trustworthy ai: A review of ethical and robust large language models. *PROCEEDINGS OF THE IEEE*.

Filieri, R. (2014). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research*, *68*(6), 1261–1270. https://doi.org/10.1016/j.jbusres.2014.11.006

Filieri, R. (2016). What makes an online consumer review trustworthy? *Annals of Tourism Research*, *58*, 46–64. https://doi.org/10.1016/j.annals.2015.12.019

Filieri, R., & McLeay, F. (2014). E-WOM and Accommodation: An Analysis of the Factors That Influence Travelers' Adoption of Information from Online Reviews. *Journal of Travel Research*, *53*(1), 44–57. https://doi.org/10.1177/0047287513481274

Folse, J. A. G., Porter, M., Godbole, M. B., & Reynolds, K. E. (2016). The Effects of Negatively Valenced Emotional Expressions in Online Reviews on the Reviewer, the Review, and the Product. *Psychology and Marketing*, *33*(9), 747–760. https://doi.org/10.1002/mar.20914

Gambhir, M., & Gupta, V. (2017). Recent automatic text summarization techniques: a survey. *Artificial Intelligence Review*, *47*(1), 1–66. https://doi.org/10.1007/s10462-016-9475-9

George, D., & Mallery, P. (2019). *IBM SPSS Statistics 25 Step by Step: A Simple Guide and Reference* (15th ed.). Routledge.

Gershoff, A. D., Mukherjee, A., & Mukhopadhyay, A. (2003). Consumer Acceptance of Online Agent Advice: Extremity and Positivity Effects. In *Journal of Consumer Psychology,* 13(2). https://doi.org/10.1207/S15327663JCP13-1&2_14

Ghose, A., & Ipeirotis, P. G. (2011). Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. In *IEEE Transactions on Knowledge and Data Engineering,* 23(10), 1498–1512. https://doi.org/10.1109/TKDE.2010.188

Golden, L. L., & Alpert, M. L. (1987). Comparative analysis of the relative effectiveness of one- and two- sided communication for contrasting products. *Journal of Advertising*, *16*(1), 18–68. https://doi.org/10.1080/00913367.1987.10673056

Guo, J., Wang, X., & Wu, Y. (2020). Positive emotion bias: Role of emotional content from online customer reviews in purchase decisions. *Journal of Retailing and Consumer Services*, *52*. https://doi.org/10.1016/j.jretconser.2019.101891

Gupta, P., & Harris, J. (2010). How e-WOM recommendations influence product consideration and quality of choice: A motivation to process information perspective. *Journal of Business Research*, *63*(9–10), 1041–1049. https://doi.org/10.1016/j.jbusres.2009.01.015

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning. www.cengage.com/highered

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, *18*(1), 38–52. https://doi.org/10.1002/dir.10073

Herr, P. M., Kardes, F. R., & Kim, J. (1991). Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective. In*: Journal of Consumer Research*, 17(4). https://about.jstor.org/terms

Hong, H., Xu, D., Wang, G. A., & Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems*, *102*, 1–11. https://doi.org/10.1016/j.dss.2017.06.007

Huang, A. H., Chen, K., Yen, D. C., & Tran, T. P. (2015). A study of factors that contribute to online review helpfulness. *Computers in Human Behavior*, *48*, 17–27. https://doi.org/10.1016/j.chb.2015.01.010

Huang, G., & Liang, H. (2021). Uncovering the effects of textual features on trustworthiness of online consumer reviews: A computational-experimental approach. *Journal of Business Research*, *126*, 1–11. https://doi.org/10.1016/j.jbusres.2020.12.052

Ishtiaq, A., Munir, K., Raza, A., Samee, N. A., Jamjoom, M. M., & Ullah, Z. (2024). Product Helpfulness Detection With Novel Transformer Based BERT Embedding and Class Probability Features. *IEEE Access*, *12*, 55905–55917. https://doi.org/10.1109/ACCESS.2024.3390605

Jia, S. J., Chi, O. H., & Chi, C. G. (2025). Unpacking the impact of AI vs. human-generated review summary on hotel booking intentions. *International Journal of Hospitality Management*, *126*. https://doi.org/10.1016/j.ijhm.2024.104030

Jiang, Z., & Benbasat, I. (2004). Virtual Product Experience: Effects of Visual and Functional Control of Products on Perceived Diagnosticity and Flow in Electronic Shopping. In *Journal of Management Information Systems*, 21(3). Winter.

Kamins, M. A., & Assael, H. (1987). Two-Sided Versus One-Sided Appeals: A Cognitive Perspective on Argumentation, Source Derogation, and the Effect of Disconfirming Trial on Belief Change. *Journal of Marketing Research*, *24*, 29–39.

Ketelaar, P. E., Willemsen, L. M., Sleven, L., & Kerkhof, P. (2015). The Good, the Bad, and the Expert: How Consumer Expertise Affects Review Valence Effects on Purchase Intentions in Online Product Reviews. *Journal of Computer-Mediated Communication*, *20*(6), 649–666. https://doi.org/10.1111/jcc4.12139

Khamitov, M., Rajavi, K., Huang, D. W., & Hong, Y. (2024). Consumer Trust: Meta-Analysis of 50 Years of Empirical Research. *Journal of Consumer Research*, *51*(1), 7–18. https://doi.org/10.1093/JCR/UCAD065

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision Support Systems*, *44*(2), 544–564. https://doi.org/10.1016/j.dss.2007.07.001

Kirmani, A., & Rao, A. R. (2000). No Pain, No Gain: A Critical Review of the Literature on Signaling Unobservable Product Quality. *Journal of Marketing*, *64(2)*, 66–79.

Kohli, R., Devaraj, S., & Mahmood, M. A. (2004). Understanding Determinants of Online Consumer Satisfaction: A Decision Process Perspective. In*: Journal of Management Information Systems*, 21(1).

Kothari, C. R. (2004). *Research methodology: Methods & techniques*. New Age International (P) Ltd.

Kovács, B. (2024). The Turing test of online reviews: Can we tell the difference between human-written and GPT-4-written online reviews? *Marketing Letters*. https://doi.org/10.1007/s11002-024-09729-3

Laczniak, R. N., DeCarlo, T. E., & Ramaswami, S. N. (2001). Consumers' Responses to Negative Word-of-Mouth Communication: An Attribution Theory Perspective. *Journal of Consumer Psychology*, *11*(1), 57–73. https://doi.org/10.1207/s15327663jcp1101_5

Lankton, N. K., Harrison Mcknight, D., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. In *Journal of the Association for Information Systems* (Vols. 16, Fascicolo 10, Issue 10, pp. 880–918). Association for Information Systems. https://doi.org/10.17705%2F1jais.00411

Lee, J., & Hong, I. B. (2019). Consumer's Electronic Word-of-Mouth Adoption: The Trust Transfer Perspective. *International Journal of Electronic Commerce*, *23*(4), 595–627. https://doi.org/10.1080/10864415.2019.1655207

Lee, J., Park, D. H., & Han, I. (2008). The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic Commerce Research and Applications*, *7*(3), 341–352. https://doi.org/10.1016/j.elerap.2007.05.004

Lee, K. T., & Koo, D. M. (2012). Effects of attribute and valence of e-WOM on message adoption: Moderating roles of subjective knowledge and regulatory focus. *Computers in Human Behavior*, *28*(5), 1974–1984. https://doi.org/10.1016/j.chb.2012.05.018

Lee, M., & Youn, S. (2009). Electronic word of mouth (eWOM): How eWOM platforms influence consumer product judgement. *International Journal of Advertising*, *28*(3), 473–499. https://doi.org/10.2501/S0265048709200709

Lee, S., & Choeh, J. Y. (2014). Predicting the helpfulness of online reviews using multilayer perceptron neural networks. *Expert Systems with Applications*, *41*(6), 3041–3046. https://doi.org/10.1016/j.eswa.2013.10.034

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, *80*(6), 69–96. https://doi.org/10.1509/jm.15.0420

Levine, I. (2024, July 30). *How to use Amazon's AI-generated review*. https://www.aboutamazon.com/news/retail/amazon-ai-generated-review-highlights

Li, M., Choi, T. Y., Rabinovich, E., & Crawford, A. (2013). Self-service operations at retail stores: The role of inter-customer interactions. *Production and Operations Management*, *22*(4), 888–914. https://doi.org/10.1111/j.1937-5956.2012.01321.x

Li, M., Huang, L., Tan, C. H., & Wei, K. K. (2013). Helpfulness of online product reviews as seen by consumers: Source and content features. In *International Journal of Electronic Commerce*, 17(4), pp. 101–136). M.E. Sharpe Inc. https://doi.org/10.2753/JEC1086-4415170404

Lim, S., & Schmälzle, R. (2024). The effect of source disclosure on evaluation of AI-generated messages. *Computers in Human Behavior: Artificial Humans*, *2*(1), 100058. https://doi.org/10.1016/j.chbah.2024.100058

Liu, A. X., Wang, Y., & Zhang, J. (2023). The Effects of Linguistic Coordination on Perceived Quality of Consumer Reviews: A Dual Process Perspective. *Journal of Interactive Marketing*, *58*(4), 361–379. https://doi.org/10.1177/10949968221150933

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, *38*(6), 937–947. https://doi.org/10.1287/mksc.2019.1192

Luo, Z., Xie, Q., & Ananiadou, S. (2024). Factual consistency evaluation of summarization in the Era of large language models. *Expert Systems with Applications*, *254*. https://doi.org/10.1016/j.eswa.2024.124456

Lynch, J. G. (2006). Accessibility-Diagnosticity and the Multiple Pathway Anchoring and Adjustment Model. In *Journal of Consumer Research,* (33). https://academic.oup.com/jcr/article/33/1/25/1822595

Machackova, H., & Smahel, D. (2018). The perceived importance of credibility cues for the assessment of the trustworthiness of online information by visitors of health-related websites: The role of individual factors. *Telematics and Informatics*, *35*(5), 1534–1541. https://doi.org/10.1016/j.tele.2018.03.021

Malhotra, N. K., Nunan, D., & Birks, D. F. (2017). *Marketing Research: an applied approach* (5th ed.). Pearson Education Limited. www.pearson.com/uk

Maslowska, E., Malthouse, E. C., & Bernritter, S. F. (2017). Too good to be true: The role of online reviews' features in probability to buy. *International Journal of Advertising*, *36*(1), 142–163. https://doi.org/10.1080/02650487.2016.1195622

Mayer, R. C., Davis, J. H., & David Schoorman, F. (1995). *An Integrative Model of Organizational Trust*, 20(3). https://www.jstor.org/stable/258792?seq=1&cid=pdf-

Mcknight, D. H., & Chervany, N. L. (2001). What Trust Means in E-Commerce Customer Relationships: An Interdisciplinary Conceptual Typology. In *International Journal of Electronic Commerce*, 6(2).

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research*, *13*(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

Min, H. J., & Park, J. C. (2012). Identifying helpful reviews based on customer's mentions about experiences. *Expert Systems with Applications*, *39*(15), 11830–11838. https://doi.org/10.1016/j.eswa.2012.01.116

Mudambi, S. M., & Schuff, D. (2010). Research Note: What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Quarterly*, *34*(1), 185–200. www.travelpost.com

Namvar, M., & Chua, A. Y. K. (2023). The impact of context clues on online review helpfulness. *Internet Research*, *33*(3), 1015–1030. https://doi.org/10.1108/INTR-02-2021-0093

Natarajan, T., Balasubramanian, S. A., & Kasilingam, D. L. (2018). The moderating role of device type and age of users on the intention to use mobile shopping applications. *Technology in Society*, *53*, 79–90. https://doi.org/10.1016/j.techsoc.2018.01.003

Nelson, P. (1970). Information and Consumer Behavior. *Journal of Political Economy*, *78*(2), 311–329.

Nelson, P. (1974). Advertising as Information. *Journal of Political Economy*, *82*(4), 729–754. https://www.jstor.org/stable/1837143

Nielsen. (2012, April). *Consumer Trust in Online, Social and Mobile Advertising Grows*. Nielsen.

Ohanian, R. (1990). Construction and Validation of a Scale to Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness. *Source: Journal of Advertising*, *19*(3), 39–52.

Pan, L. Y., & Chiou, J. S. (2011). How Much Can You Trust Online Information? Cues for Perceived Trustworthiness of Consumer-generated Online Information. *Journal of Interactive Marketing*, *25*(2), 67–74. https://doi.org/10.1016/j.intmar.2011.01.002

Pan, Y., & Zhang, J. Q. (2011). Born Unequal: A Study of the Helpfulness of User-Generated Product Reviews. *Journal of Retailing*, *87*(4), 598–612. https://doi.org/10.1016/j.jretai.2011.05.002

Park, C., & Lee, T. M. (2009). Information direction, website reputation and eWOM effect: A moderating role of product type. *Journal of Business Research*, *62*(1), 61–67. https://doi.org/10.1016/j.jbusres.2007.11.017

Park, D. H., Lee, J., & Han, I. (2007). The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement. In *International Journal of Electronic Commerce*, 11(4), pp. 125–148). M.E. Sharpe Inc. https://doi.org/10.2753/JEC1086-4415110405

Pavlou, P. A., & Dimoka, A. (2006). The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation. *Informations Sytems Research.* *17*(4), 392–414. https://doi.org/10.1287/isre

Pavlou, P. A., & Fygenson, M. (2006). Understanding and Predicting Electronic Commerce Adoption: An Extension of the Theory of Planned Behavior. In*: MIS Quarterly*, 30(1).

Pavlou, P. A., & Gefen, D. (2004). Building Effective Online Marketplaces with Institution-Based Trust. *Information Systems Research*, *15*(1), 37–59. https://doi.org/10.1287/isre

Peterson, R. A. (1994). A Meta-analysis of Cronbach's Coefficient Alpha. *Journal of Consumer Research*, *21*(2), 381–391.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*(C), 123–205. https://doi.org/10.1016/S0065-2601(08)60214-2

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement. In *Journal of Consumer Research*, 10(2). https://www.jstor.org/stable/2488919

Purnawirawan, N., Eisend, M., De Pelsmacker, P., & Dens, N. (2015). A Meta-analytic Investigation of the Role of Valence in Online Reviews. *Journal of Interactive Marketing*, *31*, 17–27. https://doi.org/10.1016/j.intmar.2015.05.001

Qiu, L., Pang, J., & Lim, K. H. (2012). Effects of conflicting aggregated rating on eWOM review credibility and diagnosticity: The moderating role of review valence. *Decision Support Systems*, *54*(1), 631–643. https://doi.org/10.1016/j.dss.2012.08.020

Reimer, T., & Benkenstein, M. (2016). When good WOM hurts and bad WOM gains: The effect of untrustworthy online reviews. *Journal of Business Research*, *69*(12), 5993–6001. https://doi.org/10.1016/j.jbusres.2016.05.014

Schindler, R. M., & Bickart, B. (2012). Perceived helpfulness of online consumer reviews: The role of message content and style. *Journal of Consumer Behaviour*, *11*(3), 234–243. https://doi.org/10.1002/cb.1372

Sekaran, U., & Bougie, R. (2016). *Research methods for business- A skill-building approach* (5th ed.). John Wiley and Sons Ltd.

Sen, S., & Lerman, D. (2007). Why are you telling me this? An examination into negative consumer reviews on the web. *Journal of Interactive Marketing*, *21*(4), 76–94. https://doi.org/10.1002/dir.20090

Singh, J., & Sirdeshmukh, D. (2000). Agency and Trust Mechanisms in Consumer Satisfaction and Loyalty Judgments. In *Journal of the Academy of Marketing Science*. 28(1).

Skowronski, J. J., & Carlston, D. E. (1987). Social Judgment and Social Memory: The Role of Cue Diagnosticity in Negativity, Positivity, and Extremity Biases. *Journal of Personality and Social Psychology*, *52*(4), 689–699. https://doi.org/10.1037/0022-3514.52.4.689

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*, 131. https://psycnet.apa.org/record/1989-15444-001

Smith, R. E., & Hunt, S. D. (1978). Attributional Processes and Effects in Promotional Situations. In*: Journal of Consumer Research*, 5(3). https://www.jstor.org/stable/2488663

Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management*, *32*(6), 1310–1323. https://doi.org/10.1016/j.tourman.2010.12.011

Stewart, K. J. (2003). Trust Transfer on the World Wide Web. *Organization Science*, *14*(1), 5–17. https://www.jstor.org/stable/3086029

Sussman, S. W., & Siegal, W. S. (2003). Informational influence in organizations: An integrated approach to knowledge adoption. *Information Systems Research*, *14*(1), 47–65. https://doi.org/10.1287/isre.14.1.47.14767

Telenti, A., Auli, M., Hie, B. L., Maher, C., Saria, S., & Ioannidis, J. P. A. (2024). Large language models for science and medicine. In *European Journal of Clinical Investigation*, 54(6). John Wiley and Sons Inc. https://doi.org/10.1111/eci.14183

Venkatesh, V., & Davis, F. D. (2000). Theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425–478.

Wang, Y. D., & Emurian, H. H. (2005). An overview of online trust: Concepts, elements, and implications. *Computers in Human Behavior*, *21*(1), 105–125. https://doi.org/10.1016/j.chb.2003.11.008

What audience can I reach on SurveySwap? - SurveySwap Helpcenter. (n.d.). Retrieved July 1, 2025, from https://knowledge.surveyswap.io/article/98-what-audience-can-i-reach-here

Wolf, J. R., & Muhanna, W. A. (2011). Feedback Mechanisms, Judgment Bias, and Trust Formation in Online Auctions*. *Decision Sciences*, *42*(1), 43–68. https://doi.org/10.1111/J.1540-5915.2010.00301.X

Wu, P. F. (2013). In Search of Negativity Bias: An Empirical Study of Perceived Helpfulness of Online Reviews. *Psychology & Marketing*, *30*(11), 971–984.

Xiao, L., Zhang, Y., & Fu, B. (2019). Exploring the moderators and causal process of trust transfer in online-to-offline commerce. *Journal of Business Research*, *98*, 214–226. https://doi.org/10.1016/j.jbusres.2019.01.069

Xylogiannopoulos, K. F., Xanthopoulos, P., Karampelas, P., & Bakamitsos, G. A. (2024). ChatGPT paraphrased product reviews can confuse consumers and undermine their trust in genuine reviews. Can you tell the difference? *Information Processing and Management*, *61*(6). https://doi.org/10.1016/j.ipm.2024.103842

Zhou, T., Lu, Y., & Wang, B. (2016). Examining online consumers' initial trust building from an elaboration likelihood model perspective. *Information Systems Frontiers*, *18*(2), 265–275. https://doi.org/10.1007/s10796-014-9530-5

Appendix

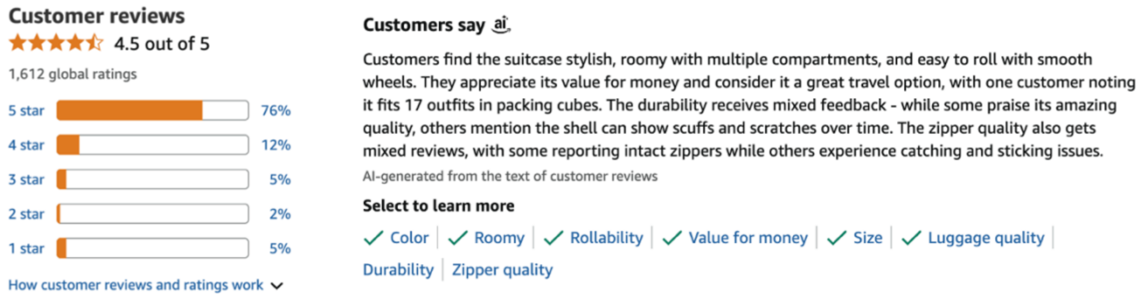


Figure 5: Example AGRS for a suitcase

Appendix A: Stimuli Development

A.1 Pre-test Interview: Category Identification & Characteristics of the Review

"Thank you for agreeing to this quick interview. The goal is to identify a suitable product category for my master's thesis. My research looks at how AI-generated review summaries—a feature where platforms like Amazon automatically summarize customer reviews—might influence how much people trust a product. This interview will take about 8–10 minutes, and all your responses are anonymous and confidential. Do you agree to these conditions? "

1. General Shopping Habits

"When you shop online, what kinds of products do you usually check customer reviews for?" (*Follow-up:* "Why those products?")

2. Product Familiarity Check

"Out of the following product types, which ones do you feel familiar with or have shopped for before?"

- Bluetooth headphones or earbuds
- Coffee maker or kitchen appliance
- Vacuum cleaner
- Electric toothbrush
- Laptop or Phone
- Portable power bank or phone charger

Follow-up: "For which of these do you feel most comfortable judging product quality just by reading reviews?"; "Why do you feel that way?"

“Can you tell me about a product you recently bought where you relied on customer reviews before deciding?”
(*Follow-up:* “Do you remember what convinced you?”)

4. Stimulus Preference

“If you were participating in a short study and had to read some reviews about a product, which of these would you prefer to read about?”
(*List top 2–3 categories they mentioned.*)

(*Optional follow-up:*)

“Would it make a difference to you if the reviews were summarized in one paragraph first?”

Wrap-up: "Thank you for your time and honest answers! Your input will help me choose a product category that feels realistic and relevant to online shoppers like you."

A.2 Pre-test Interview Results Category Identification

In a series of 10 pre-test interviews, participants shared their preferences for product categories that would be suitable for evaluating review summary formats. The main findings from these interviews were as follows:

Participants favored higher-involvement purchases, such as products that require more time to evaluate and consider before making a purchase. These products are often more complex or expensive, and customers are more likely to rely on reviews to guide their decision-making process. When it came to technical complexity, products with multiple functions or features that are harder to understand were highlighted as strong candidates. Vacuum cleaners and air fryers were specifically mentioned as examples of these types of products, where detailed reviews would be particularly valuable to help clarify how various features work in practice. On the other hand, low brand trust was another critical factor that influenced the decision-making process. Phones, for example, were quickly ruled out, as participants felt they already trusted established brands like Apple and didn't need to rely on reviews, focusing instead on technical specifications.

Finally, products that are typically bought on impulse with a lower price point, such as portable power banks, were not considered ideal. Participants stated that these types of products are usually purchased in-store based on immediate need, with little to no foreplanning or detailed online research. As a result, reviews were seen as unnecessary for these products. In conclusion, the most suitable product categories for the study were those that require more detailed evaluation, involve some level of technical complexity, require a longer decision time as compared to low-involvement products and where brand trust is not already established. Products like vacuum cleaners, air fryers, and electric toothbrushes fit these criteria well, while phones and power banks were less appropriate for the research.

A.3 Final detailed Product Page

A.4 Quantitative Pre-test Survey for Developing Final Stimuli Composition

Introduction (excerpt)

Participants were told the study was part of a master's thesis and that they would be shown different review formats (review only, summary only, summary + reviews), with or without star ratings and helpfulness votes. The survey took ~5–7 minutes, and responses were anonymous and confidential.

Stimuli & Questions

For each stimulus (Review only / Summary only / Summary + Reviews):

- Q1: How helpful did you find the information in understanding the product? (5-point Likert: 1 = Not helpful at all … 5 = Very helpful)

- Q2: Did you feel you had enough information to form a good opinion about the product? (5-point Likert: 1 = Definitely not … 5 = Definitely yes)

Additional Items (Condition 2 only: with star ratings & helpfulness votes)

- Q3: Did you notice an average star rating? (Yes/No/Not sure)
- Q4: If yes, how much did the rating influence your trust? (5-point Likert: None at all → A great deal)
- Q5: Did you notice helpfulness votes? (Yes/No/Not sure)
- Q6: If yes, how much did they influence your trust? (5-point Likert: None at all → A great deal)
- Q7: While reading, what did you focus on most? (Star rating / Helpfulness votes / Review text)

Manipulation Check

- Q8: How would you describe the sentiment of the reviews? (5-point Likert: Very positive → Very negative + "Not sure")

Post-Exposure Questions

- Q9: What part was most helpful in evaluating the product? (Review / Summary / Both equally / None)
- Q10: If shopping online, which option would you prefer? (Review only / Summary only / Multiple reviews / Mix of summary + reviews / Doesn't matter)
- Q11: Was anything missing or unclear? (Open text)

Demographics

- Gender (Male / Female / Non-binary / Prefer not to say)
- Age group (Under 18 → 65+)
- Education (Secondary / Apprenticeship / Bachelor's / Master's / Doctorate / Other)

A.5 Quantitative Pre-test Survey for Validating Final Stimuli

Introduction (excerpt)

Participants were shown the detailed product page and two random review formats (out of six conditions) and asked two short questions after each. The survey took ~5 minutes, responses were anonymous, and conducted as part of a master's thesis at Católica Lisbon SBE.

Stimuli & Questions (repeated for each of two displayed conditions)

- Q1.1: How would you rate the overall sentiment (tone) of the reviews? (7-point Likert: 1 = Very negative … 7 = Very positive)
- Q1.2: How trustworthy did you find these reviews? (7-point Likert: 1 = Not at all trustworthy … 7 = Extremely trustworthy)

Post-Exposure Questions

- Q8: What part was most helpful in evaluating the product? (Review(s) / Summary / Both equally / None)
- Q10: Was anything missing or unclear? (Open text)
- Demographics (Same items as in Pre-test Survey 1 (gender, age group, education, income))

Appendix B: Main Study Questionnaire

B.1 Questionnaire Design Logic Flow

Introduction (1 Question)

Screening Questions (2 Questions)

- Q0.1 Online purchase frequency.
  - *If "Never" → End of Survey.*
- Q0.2 Review usage frequency.
  - *If "Rarely" or "Never" → End of Survey*

Product Introduction (1 Question)

Set Embedded Data: Completion_Code

Randomizer: Evenly present 1 element only

- Group: 01 WOS-Positive
  - *Set Embedded Data: Completion_Code = 02*
  - Show Block: 02 WOS-Negative
- Group: 02 WOS-Negative
  - *Set Embedded Data: Completion_Code = 03*
  - Show Block: 03 WOS-2-Sided
- Group: 03 WOS-2-Sided
  - *Set Embedded Data: Completion_Code = 04*
  - Show Block: 04 WS-Positive

	Group: 04 WS-Positive
	Set Embedded Data: Completion_Code = 05
	Show Block: 05 WS-Negative
	Group: 05 WS-Negative
	Set Embedded Data: Completion_Code = 06
	Show Block: 06 WS-2-Sided
	Group: 06 WS-2-Sided
	Set Embedded Data: Completion_Code = 06
	Show Block: 06 WS-2-Sided
Randomizer: Evenly present 2 elements	
	Trust Questions (1 Question)
	Helpfulness Questions (1 Question)
Manipulation Checks (2 Questions)	
After all Stimuli (1 Question)	
Demographics (5 Questions)	
End of Survey & Karma for SurveySwap (2 Questions)	

Table 10: Logical Survey Flow

B.2 Online Survey Questionnaire

Block 1: Introduction

Welcome!

Thank you for taking part in this Master Thesis research on how consumers evaluate product information in online shopping environments.

All answers collected are completely confidential and will be used solely for academic purposes. Your responses are fully anonymized and cannot be linked back to you.

Your participation is voluntary, and you may withdraw at any time without penalty. There are no right or wrong answers. I am interested in your honest opinions and perceptions.

The survey will take about 5 minutes to complete.

As a thank you: This survey includes Karma points for SurveySwap.io. Plus, I'm giving away two €15 Wunschgutschein gift cards.

This survey is conducted by Anna Loerakker, M.Sc. student of Católica Lisbon School of Business and Economics. If you have any questions or need further clarification, feel free to contact s-aloerakker@ucp.pt. By proceeding, you agree to participate under the above-mentioned conditions.

To begin, please click the arrow below to proceed.

Block 2: Screening Questions

Q0.1 How often do you purchase products online?

○ Several times a week ○ About once a week ○ A few times a month

○ About once a month ○ Less than once a month ○ Never

Q0.2 How often do you read customer reviews before buying something online?

○ Always ○ Often ○ Sometimes ○ Rarely ○ Never

Block 3: Product Introduction

Q0 You're about to see customer reviews for the product shown below. Take your time to look at the product briefly. You do not need to remember any technical details.

Blocks 4–9: Stimuli Presentation (Review Types)

Instruction: *Take your time to carefully read the reviews for the product shown. The 'Next' button will appear after 30 seconds so that you can read through all the information first.*

Q01 (WOS: Positive)

Q02 (WOS: Negative)

Q03 (WOS: Two-sided)

Q04 (WS: Positive)

Q05 (WS: Negative)

Q06 (WS: Two-sided)

Timing recorded: First Click, Last Click, Page Submit, Click Count.

Block 10: Trust Questions

Q1 Please evaluate the review content you just read by selecting a point between the two attributes.

Dishonest ○ ○ ○ ○ ○ ○ ○ Honest

Untrustworthy ○ ○ ○ ○ ○ ○ ○ Trustworthy

Unreliable ○ ○ ○ ○ ○ ○ ○ Reliable

Insincere ○ ○ ○ ○ ○ ○ ○ Sincere

Block 11: Helpfulness Questions

Q2 Please rate how you would evaluate the review content you just read:

(7-point Likert scale: Strongly Disagree → Strongly Agree)

○ The review content was informative.

○ The review content was useful for evaluating the product.

○ The review content was helpful for making a purchase decision.

Block 12: Manipulation Checks

Q3 Please think back to what you just saw. What best describes the review content?

○ Only reviews written by individuals ○ An AI summary and reviews ○ I'm not sure

Q4 How would you best describe the overall sentiment of the review content you just read?

○ Negative ○ Neutral or mixed ○ Positive

Block 13: After All Stimuli

Q10 Was anything missing, confusing, or unclear for you in this survey?

(Open text response)

Block 14: Demographics

Q7.1 What is your gender?

○ Male ○ Female ○ Non-binary / third gender ○ Prefer not to say

Q7.2 How old are you?

○ Under 18 ○ 18–24 ○ 25–34 ○ 35–44 ○ 45–54 ○ 55–64 ○ 65 or older

Q7.3 What is the highest level of education you've completed?

○ Secondary school (High School) ○ Apprenticeship or Some Undergraduate

○ Undergraduate Degree (e.g., Bachelors) ○ Graduate Degree (e.g., Masters, PhD or Diploma)

○ Other, please specify: ____________

Q7.4 Which country do you currently live in?

Dropdown with full country list

Block 15: SurveyCircle & SurveySwap

Q59 Thank you for your time! You've completed the main part of the study. If you'd like to enter the giveaway for one of two €15 Wunschgutscheine, please enter your email address below:

(Open text box)

Q58 If you are participating through SurveySwap, you can use the following code to redeem points:

Link: surveyswap.io/sr/AWC1-9JTP-47S6

Code: AWC1-9JTP-47S6

Appendix C: Statistical Output

AI_Summary * MC1_num Crosstabulation

			MC1_num		
			Fail	Pass	Total
AI_Summary	No Summary	Count	100	175	275
		% within AI_Summary	36,4%	63,6%	100,0%
		% within MC1_num	46,3%	51,0%	49,2%
		% of Total	17,9%	31,3%	49,2%
	With Summary	Count	116	168	284
		% within AI_Summary	40,8%	59,2%	100,0%
		% within MC1_num	53,7%	49,0%	50,8%
		% of Total	20,8%	30,1%	50,8%
Total		Count	216	343	559
		% within AI_Summary	38,6%	61,4%	100,0%
		% within MC1_num	100,0%	100,0%	100,0%
		% of Total	38,6%	61,4%	100,0%

Table 11: Crosstabulation of Experimental Condition and aggregated response of the Manipulation Control

AI_Summary * MC1_Ai_individual Crosstabulation

			MC1_Ai_individual			
			Individual Reviews	AI summary + Individual Reviews	Not sure	Total
AI_Summary	No Summary	Count	175	37	63	275
		% within AI_Summary	63,6%	13,5%	22,9%	100,0%
		% within MC1_Ai_individual	71,7%	18,0%	57,3%	49,2%
		% of Total	31,3%	6,6%	11,3%	49,2%
	With Summary	Count	69	168	47	284
		% within AI_Summary	24,3%	59,2%	16,5%	100,0%
		% within MC1_Ai_individual	28,3%	82,0%	42,7%	50,8%
		% of Total	12,3%	30,1%	8,4%	50,8%
Total		Count	244	205	110	559
		% within AI_Summary	43,6%	36,7%	19,7%	100,0%
		% within MC1_Ai_individual	100,0%	100,0%	100,0%	100,0%
		% of Total	43,6%	36,7%	19,7%	100,0%

Table 12: Crosstabulation of Experimental Condition and Manipulation Check Response for the Independent Variable

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	131,978[a]	2	<,001
Likelihood Ratio	140,426	2	<,001
Linear-by-Linear Association	26,345	1	<,001
N of Valid Cases	559		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 54,11.

Table 13: Chi-square Test of Independence for Manipulation Check Responses for the Independent Variable

*Valence * MC2_pass_fail Crosstabulation*

			MC2_pass_fail		
			Fail	Pass	Total
Valence	Positive	Count	43	141	184
		% within Valence	23,4%	76,6%	100,0%
		% within MC2_pass_fail	43,9%	30,6%	32,9%
		% of Total	7,7%	25,2%	32,9%
	Negative	Count	18	169	187
		% within Valence	9,6%	90,4%	100,0%
		% within MC2_pass_fail	18,4%	36,7%	33,5%
		% of Total	3,2%	30,2%	33,5%
	2-Sided	Count	37	151	188
		% within Valence	19,7%	80,3%	100,0%
		% within MC2_pass_fail	37,8%	32,8%	33,6%
		% of Total	6,6%	27,0%	33,6%
Total		Count	98	461	559
		% within Valence	17,5%	82,5%	100,0%
		% within MC2_pass_fail	100,0%	100,0%	100,0%
		% of Total	17,5%	82,5%	100,0%

Table 14: Crosstabulation of Experimental Condition and aggregated response of the Manipulation Control: Moderator Valence

*Valence * MC2_Sentiment Crosstabulation*

			MC2_Sentiment			
			Negative	Neutral	Positive	Total
Valence	Positive	Count	2	41	141	184
		% within Valence	1,1%	22,3%	76,6%	100,0%
		% within MC2_Sentiment	1,1%	19,9%	84,9%	32,9%
		% of Total	0,4%	7,3%	25,2%	32,9%
	Negative	Count	169	14	4	187
		% within Valence	90,4%	7,5%	2,1%	100,0%
		% within MC2_Sentiment	90,4%	6,8%	2,4%	33,5%
		% of Total	30,2%	2,5%	0,7%	33,5%
	2-Sided	Count	16	151	21	188
		% within Valence	8,5%	80,3%	11,2%	100,0%
		% within MC2_Sentiment	8,6%	73,3%	12,7%	33,6%
		% of Total	2,9%	27,0%	3,8%	33,6%
Total		Count	187	206	166	559
		% within Valence	33,5%	36,9%	29,7%	100,0%
		% within MC2_Sentiment	100,0%	100,0%	100,0%	100,0%
		% of Total	33,5%	36,9%	29,7%	100,0%

Table 15: Crosstabulation of Experimental Condition and Manipulation Check Response: Moderator Valence

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	630,484[a]	4	<,001
Likelihood Ratio	633,039	4	<,001
Linear-by-Linear Association	76,306	1	<,001
N of Valid Cases	559		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 54,64.

Table 16: Chi-square Test of Independence for Manipulation Check Response: Moderator Valence

Reliability Statistics

Cronbach's Alpha	N of Items
,857	4

Table 18: Reliability Statistics Trust items

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
T1_Dishonest:Honest	16,62	8,401	,660	,834
T2_Untrustworthy: Trustworthy	16,82	8,078	,737	,804
T3_Unreliable:Reliable	17,04	7,836	,693	,821
T4_Insincere:Sincere	16,89	7,388	,720	,811

Table 17: Item-Total Statistics Trust items

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
trust_score	291	2,00	7,00	5,6143	,91682
Valid N (listwise)	291				

Table 21: Determination of trust_scale mean

Reliability Statistics

Cronbach's Alpha	N of Items
,780	3

Table 19: Reliability Statistics Helpfulness items

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
H1_informative	12,01	2,586	,571	,752
H2_useful	11,89	2,342	,686	,634
H3_helpful	11,91	2,079	,613	,720

Table 22: Item-Total Statistics Helpfulness items

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
helpfulness_score	291	2,67	7,00	5,9679	,72538
Valid N (listwise)	291				

Table 20: Determination of helpfulness_scale mean

C.1 Sample Characterization

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	68	23,4	23,4	23,4
	Female	212	72,9	72,9	96,2
	Non-binary / third gender	7	2,4	2,4	98,6
	Prefer not to say	4	1,4	1,4	100,0
	Total	291	100,0	100,0	

Table 24: Sample Characterization: Gender

Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Under 18	22	7,6	7,6	7,6
	18-24	138	47,4	47,4	55,0
	25-34	104	35,7	35,7	90,7
	35-44	17	5,8	5,8	96,6
	45-54	8	2,7	2,7	99,3
	55-64	2	,7	,7	100,0
	Total	291	100,0	100,0	

Table 23: Sample Characterization: Age

Education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Secondary school (High School)	67	23,0	23,0	23,0
	Apprenticeship or Some Undergraduate	26	8,9	8,9	32,0
	Undergraduate Degree (e.g. Bachelors)	118	40,5	40,5	72,5
	Graduate Degree (e.g. Masters, PhD or Diploma)	70	24,1	24,1	96,6
	Other, please specify	10	3,4	3,4	100,0
	Total	291	100,0	100,0	

Table 25: Sample Characterization: Education

List of Countries

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Australia	7	2,4	2,4	2,4
	Belgium	3	1,0	1,0	3,4
	Brazil	1	,3	,3	3,8
	Bulgaria	1	,3	,3	4,1
	Canada	6	2,1	2,1	6,2
	Central African Republic	1	,3	,3	6,5
	China	3	1,0	1,0	7,6
	Finland	1	,3	,3	7,9
	France	4	1,4	1,4	9,3
	Georgia	1	,3	,3	9,6
	Germany	143	49,1	49,1	58,8
	Hong Kong (S.A.R.)	1	,3	,3	59,1
	India	13	4,5	4,5	63,6
	Ireland	1	,3	,3	63,9
	Israel	1	,3	,3	64,3
	Italy	3	1,0	1,0	65,3
	Japan	4	1,4	1,4	66,7
	Malaysia	2	,7	,7	67,4
	Mali	1	,3	,3	67,7
	Malta	1	,3	,3	68,0
	Netherlands	20	6,9	6,9	74,9
	New Zealand	1	,3	,3	75,3
	Philippines	2	,7	,7	75,9
	Poland	1	,3	,3	76,3
	Portugal	1	,3	,3	76,6
	Singapore	2	,7	,7	77,3
	South Africa	2	,7	,7	78,0
	Spain	3	1,0	1,0	79,0
	Sweden	1	,3	,3	79,4
	Switzerland	1	,3	,3	79,7
	United Kingdom of Great Britain and Northern Ireland	40	13,7	13,7	93,5
	United States of America	18	6,2	6,2	99,7
	Viet Nam	1	,3	,3	100,0
	Total	291	100,0	100,0	

Table 26: Sample Characterization: Country

C.2 Hypothesis testing

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1,005	,430		2,340	,020		
	helpfulness_score	,707	,061	,559	11,584	<,001	,987	1,013
	AI_Summary	-,032	,089	-,017	-,354	,723	,971	1,030
	Valence	,215	,057	,182	3,747	<,001	,979	1,022

a. Dependent Variable: trust_score

Table 27: Variance Inflation Factor (VIF) and Tolerance Values for Predictors

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	helpfulness_score	AI_Summary	Valence
1	1	3,805	1,000	,00	,00	,01	,01
	2	,130	5,404	,00	,00	,32	,54
	3	,058	8,094	,03	,08	,56	,39
	4	,006	24,578	,97	,92	,11	,06

a. Dependent Variable: trust_score

Table 28: Collinearity Diagnostics: Eigenvalues, Condition Indices, and Variance Proportions

Tests of Normality

	AI_Summary	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
trust_score	No Summary	,097	153	,001	,943	153	<,001
	With Summary	,097	138	,003	,964	138	,001

a. Lilliefors Significance Correction

Table 29: Tests of Normality H1

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
trust_score	Equal variances assumed	,158	,691	1,672	289	,048	,096	,17942	,10730	-,03177	,39061
	Equal variances not assumed			1,677	288,222	,047	,095	,17942	,10702	-,03121	,39005

Table 30: Independent samples t-test and Levene's Test H1

Tests of Normality

	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
Unstandardized Residual	,076	561	<,001	,974	561	<,001

a. Lilliefors Significance Correction

Table 31: Tests of Normality H2

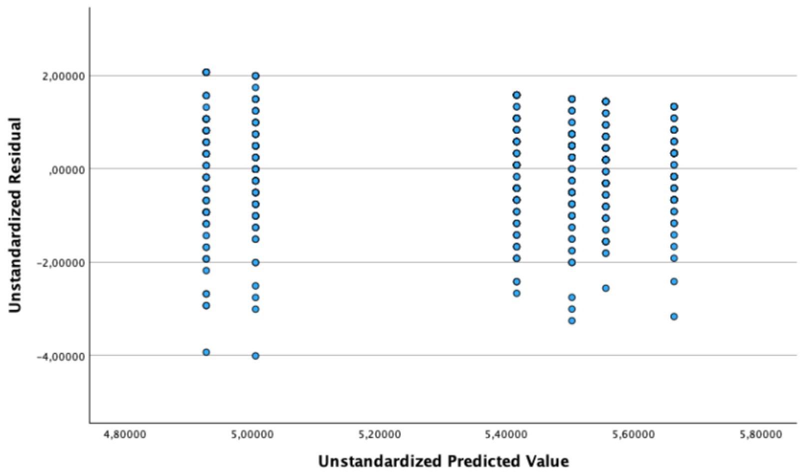


Figure 6: Residual Plot for Homoscedasticity Check H2

Levene's Test of Equality of Error Variances[a,b]

		Levene Statistic	df1	df2	Sig.
trust_score	Based on Mean	1,333	5	285	,250
	Based on Median	1,108	5	285	,356
	Based on Median and with adjusted df	1,108	5	230,805	,357
	Based on trimmed mean	1,245	5	285	,288

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Dependent variable: trust_score

b. Design: Intercept + AI_Summary + Valence + AI_Summary * Valence

Table 32: Levene's Test Moderation Effect H2

```
Run MATRIX procedure:

Copyright 2013-2025 by Andrew F. Hayes. ALL RIGHTS RESERVED.
This version of PROCESS requires SPSS version 26 or later
Workshop schedule available at haskayne.ucalgary.ca/CCRAM
In SPSS 29 and later, change default output font to Courier New for tidier
output. More information about PROCESS at processmacro.org/faq.html.
This beta release has not been completely tested. Use at your own risk.

***************** PROCESS Procedure for SPSS Version 5.0 *****************
          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2022). www.guilford.com/p/hayes3
**************************************************************************
  Model: 1
      Y: trust_sc
      X: AI_Sum
      W: Valence

Sample
Size:  291

Coding of categorical W variable for analysis:
 Valence      W1      W2
   1,000    ,000    ,000
   2,000   1,000    ,000
   3,000    ,000   1,000

**************************************************************************

OUTCOME VARIABLE:
 trust_sc

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      ,2375      ,0564      ,8071     3,4081     5,0000   285,0000      ,0052

Model
              coeff         se          t          p       LLCI       ULCI
constant     6,0222      ,3280    18,3584      ,0000     5,3765     6,6679
AI_Sum       -,4389      ,2009    -2,1848      ,0297     -,8343     -,0435
W1           -,4762      ,4201    -1,1336      ,2579    -1,3031      ,3507
W2            ,0099      ,4309      ,0229      ,9817     -,8383      ,8580
Int_1         ,5497      ,2611     2,1049      ,0362      ,0357     1,0636
Int_2         ,2215      ,2769      ,7997      ,4245     -,3236      ,7665

Product terms key:
 Int_1    :        AI_Sum   x        W1
 Int_2    :        AI_Sum   x        W2


Test(s) of highest order unconditional interaction(s):
       R2-chng          F        df1        df2          p
X*W      ,0153     2,3173     2,0000   285,0000      ,1004
----------
    Focal predict: AI_Sum   (X)
          Mod var: Valence  (W)

Data for visualizing the conditional effect of the focal predictor:
Paste text below into a SPSS syntax window and execute to produce plot.
```

Table 33: Hayes Process Model 1, Moderation

```
DATA LIST FREE/
   AI_Sum      Valence     trust_sc    se          LLCI        ULCI        .
BEGIN DATA.
     1,0000      1,0000      5,5833      ,1497      5,2886      5,8780
     2,0000      1,0000      5,1444      ,1339      4,8808      5,4080
     1,0000      2,0000      5,6568      ,1170      5,4266      5,8870
     2,0000      2,0000      5,7675      ,1190      5,5333      6,0018
     1,0000      3,0000      5,8147      ,1180      5,5825      6,0468
     2,0000      3,0000      5,5972      ,1497      5,3025      5,8919
END DATA.
GRAPH/SCATTERPLOT=
 AI_Sum   WITH     trust_sc BY       Valence  .

*********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95,0000

------ END MATRIX -----
```

Table 34: Hayes Process Model 1, Moderation II

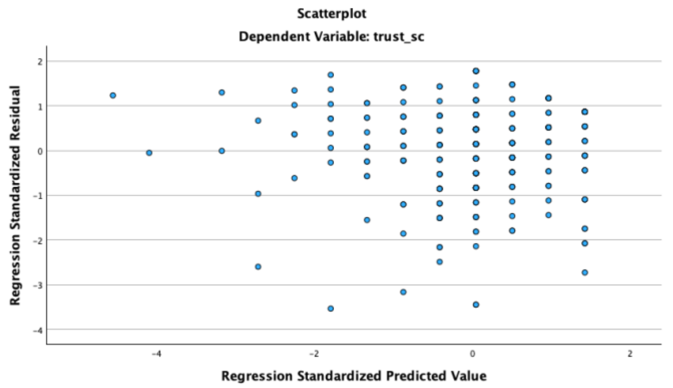


Figure 7: Scatterplot Unstandardized Residuals Trust x Helpfulness

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	,096	291	<,001	,956	291	<,001

a. Lilliefors Significance Correction

Table 35: Tests of Normality Linear Regression H3

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,553[a]	,306	,304	,76495	,818

a. Predictors: (Constant), helpfulness_score

b. Dependent Variable: trust_sc

Table 36: Linear Regression Results, H3

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1,440	,372		3,868	<,001
	helpfulness_score	,699	,062	,553	11,296	<,001

a. Dependent Variable: trust_sc

Table 37: Coefficients Results Linear Regression, H3

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
helpfulness_score	Equal variances assumed	1,908	,168	1,717	289	,044	,087	,145733	,084873	-,021315	,312781
	Equal variances not assumed			1,700	265,590	,045	,090	,145733	,085736	-,023076	,314541

Table 38: Independent Samples t-test and Levene's Test H4

Group Statistics

	AI_Sum	N	Mean	Std. Deviation	Std. Error Mean
helpfulness_score	No Summary	153	6,03704	,651758	,052692
	With Summary	138	5,89130	,794510	,067633

Table 39: Group Statistics Independent Samples t-test H4

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
helpfulness_score	Cohen's d	,722952	,202	-,029	,432
	Hedges' correction	,724835	,201	-,029	,431
	Glass's delta	,794510	,183	-,048	,414

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control (i.e., the second) group.

Table 40: Independent Samples Effect Sizes H4

```
Run MATRIX procedure:

Copyright 2013-2025 by Andrew F. Hayes. ALL RIGHTS RESERVED.
This version of PROCESS requires SPSS version 26 or later
Workshop schedule available at haskayne.ucalgary.ca/CCRAM
In SPSS 29 and later, change default output font to Courier New for tidier
output. More information about PROCESS at processmacro.org/faq.html.
This beta release has not been completely tested. Use at your own risk.



***************** PROCESS Procedure for SPSS Version 5.0 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2022). www.guilford.com/p/hayes3

**************************************************************************
  Model: 4
```

Table 41: Hayes Process Model 4, Mediation

```
      Y: trust_sc
      X: AI_Sum
      M: helpf_sc

Sample
Size:  291

**************************************************************************

OUTCOME VARIABLE:
 helpf_sc

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      ,1005      ,0101      ,5227     2,9483     1,0000   289,0000      ,0870

Model
              coeff         se          t          p       LLCI       ULCI
constant     6,1828      ,1321    46,8020      ,0000     5,9228     6,4428
AI_Sum       -,1457      ,0849    -1,7171      ,0870     -,3128      ,0213

**************************************************************************

OUTCOME VARIABLE:
 trust_sc

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      ,5550      ,3081      ,5856    64,1152     2,0000   288,0000      ,0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1,5876      ,4096     3,8760      ,0001      ,7814     2,3938
AI_Sum       -,0783      ,0903     -,8668      ,3868     -,2560      ,0995
helpf_sc      ,6941      ,0623    11,1464      ,0000      ,5715      ,8166

************************** TOTAL EFFECT MODEL ****************************

OUTCOME VARIABLE:
 trust_sc

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      ,0979      ,0096      ,8354     2,7959     1,0000   289,0000      ,0956

Model
              coeff         se          t          p       LLCI       ULCI
constant     5,8788      ,1670    35,1992      ,0000     5,5500     6,2075
AI_Sum       -,1794      ,1073    -1,6721      ,0956     -,3906      ,0318


************** TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y **************

Total effect of X on Y
     Effect         se          t          p       LLCI       ULCI
     -,1794      ,1073    -1,6721      ,0956     -,3906      ,0318
Direct effect of X on Y
     Effect         se          t          p       LLCI       ULCI
     -,0783      ,0903     -,8668      ,3868     -,2560      ,0995

Indirect effect(s) of X on Y:
```

Table 42: Hayes Process Model 4, Mediation II

```
              Effect     BootSE   BootLLCI   BootULCI
helpf_sc     -,1011      ,0604     -,2262      ,0138

**************************************************************************
Bootstrap estimates were saved to a file

Map of column names to model coefficients:
          Conseqnt Antecdnt
 COL1     helpf_sc constant
 COL2     helpf_sc AI_Sum
 COL3     trust_sc constant
 COL4     trust_sc AI_Sum
 COL5     trust_sc helpf_sc

*********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95,0000

Number of bootstrap samples for bias-corrected bootstrap confidence
intervals:
  5000

------ END MATRIX -----
```

Table 43: Hayes Process Model 4, Mediation III

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	,092	291	<,001	,957	291	<,001

a. Lilliefors Significance Correction

Table 44: Hayes Process Model 4, Tests of Normality Residuals H5

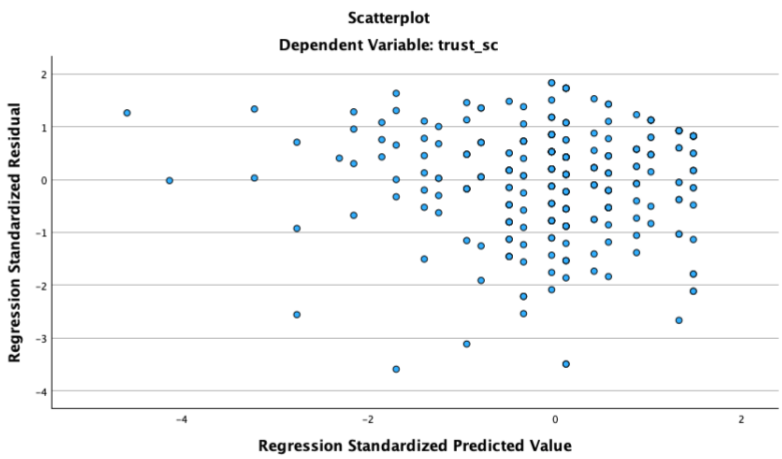


*Figure 8: Scatterplot zresid*zpred H5*

```
Run MATRIX procedure:

Copyright 2013-2025 by Andrew F. Hayes. ALL RIGHTS RESERVED.
This version of PROCESS requires SPSS version 26 or later
Workshop schedule available at haskayne.ucalgary.ca/CCRAM
In SPSS 29 and later, change default output font to Courier New for tidier
output. More information about PROCESS at processmacro.org/faq.html.
This beta release has not been completely tested. Use at your own risk.


***************** PROCESS Procedure for SPSS Version 5.0 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2022). www.guilford.com/p/hayes3

**************************************************************************
  Model: 5
      Y: trust_sc
      X: AI_Sum
      M: helpf_sc
      W: Valence

Sample
Size:  291

Coding of categorical W variable for analysis:
 Valence      W1      W2
   1,000    ,000    ,000
   2,000   1,000    ,000
   3,000    ,000   1,000

**************************************************************************

OUTCOME VARIABLE:
 helpf_sc

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      ,1005      ,0101      ,5227     2,9483     1,0000   289,0000      ,0870

Model
              coeff         se          t          p       LLCI       ULCI
constant     6,1828      ,1321    46,8020      ,0000     5,9228     6,4428
AI_Sum       -,1457      ,0849    -1,7171      ,0870     -,3128      ,0213

**************************************************************************

OUTCOME VARIABLE:
 trust_sc

Model Summary
          R       R-sq        MSE          F        df1        df2          p
      ,5920      ,3505      ,5575    25,5405     6,0000   284,0000      ,0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     1,2765      ,4995     2,5556      ,0111      ,2933     2,2597
AI_Sum       -,1037      ,1696      -,6118      ,5411     -,4375      ,2300
helpf_sc      ,7239      ,0638    11,3391      ,0000      ,5983      ,8496
W1           -,2863      ,3496      -,8192      ,4134     -,9744      ,4017
W2            ,5290      ,3611     1,4651      ,1440     -,1817     1,2397
```

Table 45: Hayes Process Model 5, Full Model

```
Int_1         ,2550      ,2186     1,1668      ,2443     -,1752      ,6853
Int_2        -,0850      ,2317     -,3669      ,7139     -,5412      ,3711

Product terms key:
 Int_1    :        AI_Sum   x        W1
 Int_2    :        AI_Sum   x        W2

Test(s) of highest order unconditional interaction(s):
       R2-chng          F       df1        df2          p
X*W      ,0066     1,4515     2,0000   284,0000      ,2359


****************** DIRECT AND INDIRECT EFFECTS OF X ON Y ******************

Conditional direct effects of X on Y
    Valence     Effect         se          t          p       LLCI       ULCI
     1,0000     -,1037      ,1696     -,6118      ,5411     -,4375      ,2300
     2,0000      ,1513      ,1387     1,0907      ,2763     -,1217      ,4244
     3,0000     -,1888      ,1584    -1,1914      ,2345     -,5006      ,1231

Indirect effect(s) of X on Y:
             Effect     BootSE   BootLLCI   BootULCI
helpf_sc     -,1055      ,0619     -,2289      ,0168

***************************************************************************
Bootstrap estimates were saved to a file

Map of column names to model coefficients:
          Conseqnt Antecdnt
 COL1     helpf_sc constant
 COL2     helpf_sc AI_Sum
 COL3     trust_sc constant
 COL4     trust_sc AI_Sum
 COL5     trust_sc helpf_sc
 COL6     trust_sc W1
 COL7     trust_sc W2
 COL8     trust_sc Int_1
 COL9     trust_sc Int_2

*********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95,0000

Number of bootstrap samples for bias-corrected bootstrap confidence
intervals:
  5000

------ END MATRIX -----
```

Table 46: Hayes Process Model 5, Full Model II

C.3 Cluster Analysis Characterization

Cluster 1: High-Frequency Shoppers

These participants shop online frequently (at least weekly), but only moderately engage with reviews. They likely rely on habit or prior experience to make decisions, indicating efficient, low-effort shopping behavior.

Cluster 2: Review-Reliant Shoppers

This group shops infrequently but consistently reads reviews. The majority reports reading reviews for most purchases, suggesting a high need for information and risk-averse decision-making, with strong reliance on peer opinions.

Cluster 3: Ambivalent Reviewers

Ambivalent reviewers shop moderately often and show inconsistent review behavior. They consult reviews situationally, reflecting a flexible or pragmatic approach to online decision-making without strong reliance on user feedback.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	5,108[a]	2	,078
Likelihood Ratio	5,127	2	,077
Linear-by-Linear Association	4,978	1	,026
N of Valid Cases	291		

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 21,34.

Table 47: Chi-Square Test of Independence Cluster Analysis

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	6,762	5	,239
	Block	6,762	5	,239
	Model	6,762	5	,239

Table 48: Binary Logistic Regression: Omnibus Tests of Model Coefficients

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 1[a]	Ai_Dummy(1)	,318	,608	,274	1	,600	1,375	,418	4,528
	Cluster_2(1)	,800	,541	2,186	1	,139	2,226	,771	6,431
	Cluster_3(1)	,651	,518	1,580	1	,209	1,917	,695	5,286
	AIxC2	-1,149	,720	2,548	1	,110	,317	,077	1,299
	AIxC3	-,502	,704	,508	1	,476	,605	,152	2,407
	Constant	-,318	,465	,470	1	,493	,727		

a. Variable(s) entered on step 1: Ai_Dummy, Cluster_2, Cluster_3, AIxC2, AIxC3.

Table 49: Binary Logistic Regression: Variables in the Equation