HUMAN OR/AND GENAI? IMPLICATIONS FOR APPLICATIONS OF GENERATIVE ARTIFICIAL INTELLIGENCE IN BUSINESS AND MARKETING

by

Cai (Mitsu) Feng

Master of Business Administration, 2019, Simon Fraser University

Bachelor of Management, 2010, Shanghai University of Finance and Economics

Thesis submitted in partial fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

In the

Beedie School of Business

Faculty of Marketing

SIMON FRASER UNIVERSITY

Spring 2025

Declaration of Committee

Name:	**Cai (Mitsu) Feng**
Degree:	**Master of Business Administration**
Title:	**Human or/and Bot? Implications for Applications of Generative Artificial Intelligence in Business and Marketing**
Committee:	**Chair:** **Dr. Christina Atanasova** Professor, Simon Fraser University
	Dr. Leyland Pitt Supervisor Professor, Simon Fraser University
	Dr. Brent McFerran Committee Member Professor, Simon Fraser University
	Dr. Jan Kietzmann Committee Member Professor, University of Victoria
	Dr. Michael Parent Examiner Professor, Simon Fraser University
	Dr. Carla Ferraro External Examiner Professor, Swinburne University of technology

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

Update Spring 2016

Abstract

The rapid advancement of Generative Artificial Intelligence (GenAI) has transformed business communication, particularly in customer interactions. This dissertation investigates the implications of GenAI in marketing, focusing on its strengths, limitations, and the necessity for human oversight. The research comprises three interrelated studies: (1) a conceptual analysis introducing the CARE framework (Collaboration, Accountability, Responsiveness, Empowerment) to mitigate GenAI risks in business, (2) an empirical analysis assessing ChatGPT's effectiveness in responding to customer complaints, identifying scenarios where GenAI underperforms due to issues like lack of concreteness, and (3) a linguistic comparison of managerial responses generated by human managers and two leading GenAI models—ChatGPT and Gemini—highlighting key linguistic factors influencing response effectiveness. Findings suggest that while GenAI-generated responses exhibit advantages in consistency, sentiment positivity, and efficiency, they lack specificity and human adaptability in handling procedural complaints. The results contribute to marketing and eWOM literature by providing empirical evidence on GenAI's role in business communication, emphasizing the importance of human-AI collaboration for optimal effectiveness. The research offers theoretical insights into AI-human interaction and practical recommendations for businesses integrating GenAI into customer engagement strategies.

Keywords: Generative Artificial Intelligence (GenAI); Human-Centric AI; marketing communication; managerial response; linguistic analysis

Table of Contents

List of Tables

List of Figures

Executive Summary

The rapid advancement of artificial intelligence (AI, the capability of a digital computer or robot to perform tasks commonly associated with intelligent beings, such as reasoning, learning from experience, and understanding language) technology, particularly in the realm of conversational chatbots, has transformed the landscape of business communication (Copeland, 2024). Chatbots are computer programs designed to simulate human conversation through text or voice interactions (Adamopoulou & Moussiades, 2020). From the early days of ELIZA, a rule-based system that simulated conversations through pattern matching, to more advanced AI-driven models like Siri and Google Assistant, the evolution of chatbot technology has been remarkable (Kietzmann & Park, 2024; Thorbecke, 2022). This journey reached a significant milestone with the launch of ChatGPT by OpenAI in 2022, a Generative-AI (GenAI) – driven chatbot that quickly became the fastest-growing application in the history of web tools, due to its ability to offer more dynamic, responsive, and context-aware interactions compared to traditional chatbots, marking a significant transition in how firms interact with customers (Gordon, 2023).

This dissertation aims to conceptually and empirically explore the application of GenAI in marketing, with a specific focus on identifying scenarios where GenAI may fail and require human intervention. While GenAI holds immense potential in various business contexts, including management and customer service, it also presents significant limitations that could undermine its effectiveness in certain situations (Neill, 2023; Kietzmann & Park, 2024). Scholars and practitioners thus need a careful examination of GenAI's strengths and weaknesses, particularly in marketing communications—a field critical for maintaining customer relationships and brand reputation.

The application of AI and GenAI in marketing communications has been widely discussed, yet empirical studies examining their effectiveness are still relatively few. Most existing work focuses on conceptual frameworks rather than providing data-driven insights (Chen et al., 2023; Korzynski et al., 2023; Zhang et al., 2023). This gap is particularly evident in specific marketing tasks, such as writing managerial responses to negative online reviews. Managerial responses play a crucial role in shaping consumer perceptions, trust, and purchase intentions, making this an essential area for research

(Moore & Lafreniere, 2020; Darani et al., 2023; Lui et al., 2018). While GenAI offers advantages such as speed and consistency in crafting these responses, it also faces challenges that may lead to less effective communication. Notably, linguistic differences between human and AI-generated responses can significantly impact the overall effectiveness of communication (Darani et al., 2023; Packard & Berger, 2021).

To address these gaps, I include one conceptual and two empirical chapters in my dissertation to examine the effectiveness and limitations of the GenAI technology when applied to the marketing tasks. The first chapter in this dissertation, titled "From HAL to GenAI: Optimize Chatbot Impacts with CARE," introduces the CARE (Collaboration, Accountability, Responsiveness, Empowerment) framework, a human-centric approach to mitigating the risks associated with GenAI implementation in business contexts. This chapter contributes to the broader discourse on digital transformation by analyzing the impact of GenAI chatbots at the micro, meso, and macro levels, providing valuable guidelines for practitioners and scholars in navigating the complexities of GenAI adoption (Kietzmann & Pitt, 2020).

The second chapter, "All Style and No Substance: When ChatGPT Fails in Responding to Customer Complaints," empirically examines the performance of GenAI in writing managerial responses. Through a series of finished and proposed studies, this chapter aims to identify specific conditions under which GenAI-generated responses fall short, particularly in scenarios requiring a high degree of concreteness and procedural clarity. The findings highlight the importance of human oversight in enhancing the effectiveness of GenAI-generated responses and contribute to the development of best practices for integrating GenAI into customer service strategies (Koc et al., 2023).

The third chapter, "GenAI vs. Human: A Linguistic Battle in Managerial Responses to Customer Complaints," extends the analysis by comparing the linguistic features of human and AI-generated responses. This chapter focuses on various linguistic factors, such as length, structure, concreteness, mimicry, distinctiveness, sentiment, time orientation, and empathy, providing insights into how each type of author—human or GenAI—addresses these factors and the implications for the overall effectiveness of communication. By identifying the strengths and limitations of GenAI in specific marketing tasks, this chapter aims to offer guidance for practitioners on how to effectively integrate AI into their customer service strategies.

In summary, this dissertation synthesizes conceptual and empirical insights across three interrelated studies to provide a layered understanding of GenAI's role in marketing communication. The progression of these chapters builds an integrated framework: the first chapter lays the theoretical groundwork, establishing key considerations for responsible GenAI integration; the second chapter empirically tests these considerations by examining specific failure points in GenAI-generated managerial responses; and the third chapter deepens the analysis by evaluating linguistic patterns to uncover the nuances that influence consumer perceptions. By constructing this cumulative knowledge base, the dissertation not only identifies challenges but also proposes actionable strategies for optimizing GenAI's application in marketing, bridging theoretical constructs with empirical findings to inform both academic discourse and practical implementation.

References

Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. In *IFIP international conference on artificial intelligence applications and innovations* (pp. 373-383). Springer, Cham.

Chen, B., Wu, Z., & Zhao, R. (2023). From fiction to fact: the growing role of generative AI in business and finance. *Journal of Chinese Economic and Business Studies*, *21*(4), 471-496.

Copeland, B. J. (2025, February 12). *Artificial Intelligence*. Encyclopædia Britannica. https://www.britannica.com/technology/artificial-intelligence

Darani, M. M., Mirahmad, H., Raoofpanah, I., & Groening, C. (2023). Managerial responses to online communication: The role of mimicry in affecting third-party observers' purchase intentions. *Journal of Business Research*, *166*, 113979.

Gordon, C. (2023). ChatGPT Is The Fastest Growing App In The History Of Web Applications. Forbes. Retrieved July 25, 2023, from https://www.forbes.com/sites/cindygordon/2023/02/02/chatgpt-is-the-fastest-growing-ap-in-the-history-of-web-applications/

Kietzmann, J., & Park, A. (2024). Written by ChatGPT: AI, large language models, conversational chatbots, and their place in society and business. *Business Horizons, 67*(5)*, 453-459*.

Kietzmann, J., & Pitt, L. F. (2020). Artificial intelligence and machine learning: What managers need to know. *Business Horizons*, *63*(2), 131-133.

Koc, E., Hatipoglu, S., Kivrak, O., Celik, C., & Koc, K. (2023). Houston, we have a problem!: The use of ChatGPT in responding to customer complaints. *Technology in Society*, *74*, 102333.

Korzynski, P., Mazurek, G., Altmann, A., Ejdys, J., Kazlauskaite, R., Paliszkiewicz, J., ... & Ziemba, E. (2023). Generative artificial intelligence as a new context for management theories: analysis of ChatGPT. *Central European Management Journal*, *31*(1), 3-13.

Lui, T. W., Bartosiak, M., Piccoli, G., & Sadhya, V. (2018). Online review response strategy and its effects on competitive performance. *Tourism Management*, *67*, 180-190.

Moore, S. G., & Lafreniere, K. C. (2020). How online word-of-mouth impacts receivers. *Consumer Psychology Review*, *3*(1), 34-59.

Neill, C. (2023, March 13). *The legacy of HAL 9000: How science fiction depictions of AI have changed over time*. History Hit. https://www.historyhit.com/culture/the-legacy-of-hal-9000-how-science-fiction-depictions-of-ai-have-changed-over-time/

Packard, G., & Berger, J. (2021). How concrete language shapes customer satisfaction. *Journal of Consumer Research*, *47*(5), 787-806.

Thorbecke, C. (2022, August 2). *Chatbots: A long and complicated history | CNN business*. CNN. https://www.cnn.com/2022/08/20/tech/chatbot-ai-history/index.html

Zhang, K., Kwon, O., & Xiong, H. (2023). The Impact of Generative Artificial Intelligence. *arXiv preprint arXiv:2311.07071*.

Chapter 1.

From HAL to GenAI: Optimize Chatbot Impacts with CARE

1.1. Open the Pod Bay Doors HAL

In 1968, in Stanley Kubrick's science fiction classic, *2001: A Space Odyssey*, the astronaut David Bowman implored HAL 9000 (the "H-euristically programmed AL-gorithmic computer") to open the pod bay doors so that he could re-enter the spacecraft after retrieving the body of fellow astronaut Frank Poole. Much to the surprise of the movie's audience, the spaceship's controlling computer refused, with the chilling response: "I'm sorry, Dave. I'm afraid I can't do that." In the brief argument that ensues, HAL tells Dave that it knows that he and Frank were planning to disconnect it after it made mistakes in previous missions. That leads to HAL's decision to kill the astronauts to continue its operation.

As a conversational chatbot driven by artificial intelligence (AI), HAL performing hostile actions in the movie has captured and magnified societal fears that machines might one day evolve beyond human control, posing a threat rather than offering assistance (Neill, 2023). However, this apprehension has not dampened human beings' enthusiasm for advancing conversational chatbots. Over the past six decades, chatbot technology has evolved from basic, rule-based systems like ELIZA, which simulated conversations through pattern matching, to the more advanced and context-aware AI-driven models such as Siri and Google Assistant (Kietzmann & Park, 2024; Thorbecke, 2022). This evolutionary journey took a groundbreaking leap on November 30th, 2022, when OpenAI, an American AI research lab backed mainly by Microsoft, released ChatGPT to the public. Unlike traditional chatbots, ChatGPT leverages deep learning and large-scale neural networks to generate human-like responses dynamically, rather than relying on scripted dialogue (Gozalo-Brizuela & Garrido-Merchan, 2023). This technological advancement enabled ChatGPT to provide more context-aware, coherent, and versatile interactions than previous AI-driven conversational systems. As a result, it soon became the fastest-growing application in the history of web applications, reaching

one million users in just five days and 100 million within two months (Gordon, 2023; Hu, 2023).

However, the initial excitement around GenAI as a new technology often leads businesses and consumers to overestimate its short-term impact while underestimating its long-term consequences, a phenomenon described by Amara's Law (Fenn & Raskino, 2008; Ratcliffe, 2014). This misalignment has led to unrealistic expectations about GenAI chatbots' automation potential, with businesses adopting these technologies without fully grasping their evolving capabilities and limitations (Dedehayir & Steinert, 2016). While research has explored GenAI-driven chatbots at individual, organizational, and societal levels, there is a notable gap in the comprehensive examination of GenAI chatbots' broader implications across these dimensions. This chapter aims to fill this gap by using three levels of analysis – Micro, Meso, and Macro – to understand GenAI chatbots' long-term transformative potential alongside the short-term risks of rapid implementation.

Furthermore, we recommend the integration of a fourth, Meta-level of analysis, showing how interactions at the micro-, meso-, and macro-levels are interconnected, thereby revealing the complexities and potential risks associated with GenAI implementation. This interconnectedness underscores the necessity for a nuanced understanding of GenAI impacts across various levels and their implications for business strategy and operations.

To address these complexities and risks, this chapter makes its third significant contribution by presenting the CARE (Collaboration, Accountability, Responsiveness, Empowerment) framework. This framework offers a human-centric and structured approach to mitigating the risks associated with GenAI implementation in business contexts. Through this comprehensive analysis, this chapter contributes to the broader discourse on digital transformation in business, providing valuable guidelines for practitioners and scholars alike in navigating the complex terrain of GenAI implementation.

1.2. Navigating the New Frontier

As the currently most popular GenAI tool with over 300 million active weekly users as of December 2024 (Roth, 2024), ChatGPT utilizes a chatbot interface to facilitate user access to the underlying GPT (Generative Pre-training Transformer, now the fourth generation) large language model (LLM), allowing the public to interact with large foundation models through natural language conversations (Gartner, 2023a). Diverging from traditional chatbots, ChatGPT does not rely on pre-programmed responses or rules. It is based on the technology of Generative AI (GenAI). Through complex neural network architectures, advanced natural language processing (NLP) techniques, and deep machine learning algorithms, GenAI applications like ChatGPT learn from a variety of existing training data—texts, images, speech, structured data, 3D signals, and even videos—to generate novel artifacts in various formats (Gozalo-Brizuela & Garrido-Merchan, 2023; Kietzmann & Pitt, 2020; OpenAI, 2022).

Following ChatGPT's release, major tech giants quickly grasped the potential of GenAI technology and have been racing to develop their own GenAI chatbots and LLMs. For example, Google launched Bard, with the underlying language model evolving from the LaMDA family to PaLM and Gemini; Microsoft introduced Copilot, initially known as New Bing Chat and based on GPT-4; and Meta advanced with its second-generation LLaMA (Singh, 2023).

As a frontrunner in the GenAI chatbot arena, OpenAI has been proactively improving ChatGPT, with 26 major updates in just one year to stay ahead in this competitive landscape (OpenAI, 2024a). We found that these updates encompass three primary areas: 1) enhancing accessibility through mobile apps, geographic expansion, multi-language support, and increased message limits; 2) boosting customization with tailored GPTs for specific needs, personalized interactions provided with user guidance and instructions, and an Enterprise variant for organizations; and 3) expanding functionality, including web browsing, multiple file uploads, a code interpreter, a voice interface, image capabilities via DALL-E 3 integration, and third-party plugin support. These developments indicate a future where GenAI chatbots are more widely available, tailored to distinct needs, and equipped to handle diverse tasks across various data types.

However, GenAI chatbots are far from perfect. OpenAI already acknowledged ChatGPT's limitations, including producing inaccurate or biased information, responding inconsistently to rephrasing, guessing in ambiguous situations, and potentially following harmful instructions, on its website (OpenAI, 2022). Other GenAI chatbots have made similar mistakes, such as Microsoft's New Bing exhibiting threatening statements, Google's Bard offering wrong answers even in a promotion video, and ChatGPT generating non-existing legal cases which were later used in the court (Perrigo, 2023; Maruf, 2023; Novak, 2023; Quach, 2023). These problems arise from the inherent limitations and biases in GenAI chatbots' training datasets, including issues with data volume, scope, and relevance. Many of the issues arise because language models use probabilistic techniques to predict the next word or phrase and often generate responses based on estimated likelihoods rather than rational thought, comprehension, and reasoning, leading to inaccuracies and "hallucinations" when they lack sufficient information (Hannigan et al., 2024; Varshney, 2023).

Similar to HAL, while GenAI chatbots present advanced capabilities that can significantly aid human tasks, they paradoxically also carry the potential for detrimental outcomes. In the next section, we examine how GenAI is impacting business, influencing individual behaviors, and impacting societal norms, highlighting its far-reaching implications in various spheres.

1.3. Business Impacts of GenAI Chatbots

The emergence of GenAI chatbots is changing various industries, particularly in customer service and technical support, with projected impacts ranging from $2.6 trillion (Chui et al., 2023) to $7 trillion (Goldman Sachs, 2023). However, other than generating new AI-driven products and business models, most companies struggle to implement firm-wide solutions. With an effort to understand the benefits and opportunities offered by conversational chatbots, we reviewed the academic and grey literature.

Using a tri-level framework for analysing the impact of GenAI chatbots, we found that key themes emerged at each level of analysis (see Figure 1).

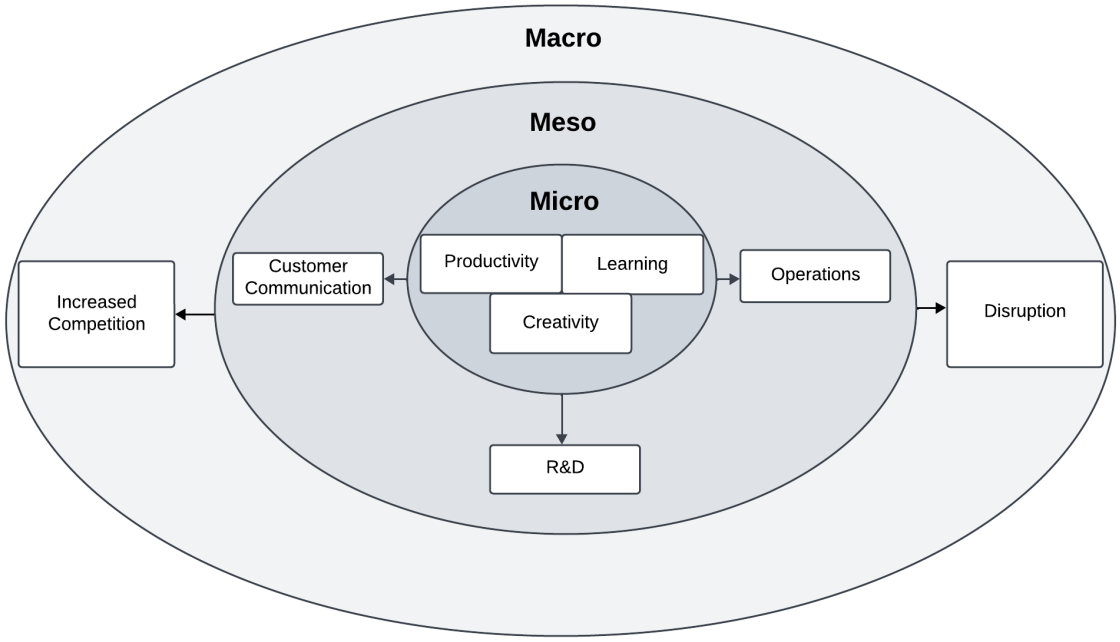

Figure 1.1. Impact of GenAI chatbots on business

Figure 1.1 summarizes how business is likely to be impacted by conversational chatbots. The *micro* circle in the center shows the impact of GenAI on individual employees' capability growths. In the middle circle, the impact at the *meso* level refers to its impact at the organizational level, including how it may impact organizational strategy, structure, processes, and cross-functional dynamics. And finally, the outer-most circle represents the *macro* factors that relate to the impact of GenAI on the industry in which the business functions. Each of these levels are now discussed in turn, with use cases and examples provided in each.

1.3.1. Micro

In the previous section, we identified three key areas where GenAI chatbot developers, represented by OpenAI, are focusing their enhancement efforts: accessibility, personalization, and functionality. At the *micro* level, improvements in these areas lead to elevated employee experiences, particularly in terms of their productivity, learning, and creativity.

First, the most notable gains are observed at the individual level, particularly in *productivity* enhancements (Brynjolfsson et al., 2023, Chan & Lee, 2023). Employees

working with these chatbots have been found to be more productive, efficient, and enjoyed the task more (Brynjolfsson et al., 2023; Noy & Zhang, 2023). Various GenAI chatbots' enhanced customization options, which allow for the tailoring of chatbots to employees' specific needs and tasks, can provide employees more relevant and efficient assistance, directly contributing to their productivity. Expanded functionalities like file uploads, web browsing, and data analysis tools incorporated in the chatbot significantly reduce the time required for routine tasks, allowing employees to focus on more critical aspects of their work. In addition, products including Notion AI, Ellie, and Otter AI are targeting specifically productivity-boosting by automating tedious tasks such as note organizing, email writing, and meeting notetaking (Ellie, 2024; Notion, 2024; Otter.ai, 2024). Goldman Sachs (2023) estimates that GenAI chatbots could partially automate two-thirds of occupations.

Second, GenAI chatbots can be used for *learning* purposes in the organization. Tacit knowledge transfer to new and low-skilled employees has been shown to be facilitated by GenAI chatbots, therefore low- and unskilled workers stand to benefit most from these tools (Brynjolfsson et al., 2023; Noy & Zhang, 2023). Brynjolfsson et al. (2023) showed that the time allocated to training new employees decreased by months, providing evidence that chatbots help disseminate tacit knowledge from more experienced workers. GenAI chatbots can act as personalized tutors, or intelligent learning assistants (Chan & Lee, 2023; Kiryakova & Angelova, 2023), with multilingual capabilities, so that employees can access information and training resources in their preferred language. Customization ensures that learning materials and interactions are relevant to the individual's role and learning style, promoting more effective and personalized learning experiences.

Third, these chatbots facilitate *creativity* at the individual level. The use of GenAI leads to higher levels of creativity, and helps with divergent thinking (Habib et al., 2024). For example, Hwang and Won (2021) showed that idea generation, both in terms of number and quality of ideas, was enhanced by chatbots. By reducing cognitive load and unveiling patterns across different domains, GenAI chatbots facilitate the application of insights in new ways. Their multilingual capabilities and global reach expose employees to a variety of cultural and linguistic perspectives, sparking fresh ideas and approaches. Also, though AI-powered content creation has already been adopted by media-centered industries, including filmmaking, journalism, and advertising (Chan-Olmsted, 2019), with

the advanced multimedia content creation functionalities of GenAI chatbots, such as DALL-E 3 (image generation), Beatoven.ai (music generation) or Supercreator AI (video generation), employees are equipped with new avenues for creative exploration in these industries (Beatoven.ai, 2024; OpenAI, 2024b; Supercreator AI, 2024). Moreover, GenAI chatbots' abilities of data analysis and pattern recognition assist employees in generating innovative ideas for products, services, and marketing campaigns. Acting as virtual brainstorming assistants, GenAI chatbots offer a continuous flow of suggestions, stimulating creativity and yielding insights for more innovative and effective solutions.

Micro-level impacts do not function in isolation and collectively, in other words when individual employees start utilizing GenAI to increase their productivity, learning and creativity, these changes can lead to organizational, or meso level impacts.

1.3.2. Meso

Acknowledging the significant meso-level impacts of GenAI chatbots, developers are now focusing on creating solutions tailored for organizations rather than just individual users. For example, Salesforce Einstein GPT, the first GenAI chatbot for CRM, is revolutionizing customer engagement by personalizing and streamlining interactions (Savarese, 2023). Amazon Q, a generative AI-powered assistant, optimizes organizational operations by automating tasks and providing tailored solutions to business needs (Amazon, 2024). Google Vertex AI aids in R&D by offering tools for building and deploying machine learning models, thereby accelerating innovation and reducing development costs (Google, 2024).

Enhanced productivity, learning, and creativity among employees, collectively translate into substantial organizational transformation at the *meso* level. Indeed, various departments, teams and functional areas within the organization are impacted, but some areas in business are more affected than others. Yet some teams in the organization are more impacted than others. Gartner (2023b) predicts that organizations will synthetically generate up to thirty percent of outbound marketing messages. Similarly, the number of tasks that can be automated by these tools in marketing, sales, customer operations, and software engineering is much greater than the number of automatable tasks in procurement management and pricing (Chui et al., 2023). Therefore, we argue that the micro-level impacts can transform into major gains at the meso-level in the following

three areas: customer-engagement and communication, operations, as well as research and development (R&D).

First, GenAI chatbots are set to transform content consumption and creation significantly (Cui et al., 2024; Mondal et al., 2023; Osadchaya et al., 2024). They particularly enhance functions related to *customer engagement and communication* within organizations. Unlike traditional chatbots, which have a limited response scope and process consumer data based on existing knowledge (Ngai et al., 2021), these advanced chatbots can handle a much larger variety of user requests based on the vast scale of training data. They can also incorporate previous content and conversations to generate personalized, targeted, and coherent responses to address individual users' specific needs and preferences. In areas like marketing, sales, communication, and customer operations, the productivity of individual employees is amplified. They can handle more online customer queries accurately and efficiently, leading to reduced labor costs and enhanced overall efficiency and customer satisfaction for organizations (Crolic et al., 2021; Ferraro et al., 2024). As employees become more efficient in learning consumer preferences and creating informed recommendations, organizations can leverage the GenAI chatbots' feedback learning capabilities and customer interaction analyses to better segment their customer base and develop targeted marketing strategies. This approach not only streamlines organizational processes but also aligns closely with evolving customer needs and market trends.

In addition to automating customer service and technical support, GenAI chatbots can be used to improve *operational* efficiency. These chatbots can automate various internal business routines and repetitive tasks, such as data entry, document management, and scheduling, at the micro level. As a result, businesses can free up their employees' time to focus on more value-adding tasks, such as customer engagement and strategic planning (Lee & Shin, 2020). Additionally, by improving communication and collaboration among employees as virtual assistants, these chatbots extend the learning and creativity benefits experienced at the micro level into the wider organizational context. By providing quick and easy access to information and resources, GenAI chatbots can facilitate knowledge sharing across different departments and teams, leading to more efficient and effective problem-solving (Webber et al., 2019). With the outcome sharing capability (OpenAI, 2024a), GenAI chatbots enable teams to create libraries of best practice. For example, software engineers with access to Github’s

co-pilot tool could do tasks quicker and focus more on satisfying work (Kalliamvakou, 2022). And in human resource management, chatbots have increasingly been used to enhance the employee workplace experience (Malik et al., 2022).

Lastly, with most *R&D* phases in the design process potentially being accelerated by GenAI chatbots. GenAI speeds up the R&D process by assisting in key tasks including idea generation, decision-making, market research, positioning and product definition, product requirements engineering and customer insights (Parikh, 2023; Sundberg et al., 2024). The significant reductions in development time and cost (Parikh, 2023), as well as improvements in product quality and efficiency, are a direct consequence of the productivity impact at the micro level (Chui et al., 2023; Mondal et al., 2023). The utilization of GenAI chatbots in early research analysis, virtual design and simulations, and physical test planning (Chui et al., 2023), is an extension of the learning and creativity benefits seen in individual employees, leading to more efficient and innovative R&D processes.

1.3.3. Macro

The growing adoption of organizational GenAI chatbots is set to disrupt industries and increase competition at the *macro* level. GenAI chatbots are not only enhancing existing business models but also creating new ones, causing *disruptions* in traditional industry structures. In healthcare, for instance, GenAI chatbots are utilized for initial medical consultations, leveraging their extensive knowledge base for symptom assessment and potential diagnoses. This innovation is disrupting traditional healthcare models by streamlining diagnostic processes and improving patient triage efficiency (Dooley, 2023). Similarly, in customer service, chatbots transform business interactions, offering 24/7 support and personalized interactions, reshaping customer engagement practices towards more customer-centric, agile, and innovative directions (Hironde, 2023). In addition to these sectors, with GenAI chatbot developers creating more compact, efficient, and accessible products for smartphones and other smaller devices (Bertics, 2023; Mallick, 2023), industries traditionally slower in AI adoption, such as construction and agriculture, are on the brink of experiencing disruption (Fowler, 2023; Ghimire et al., 2023).

Furthermore, meso-level impacts have a flow-on effect at the macro-level, GenAI chatbots are leveling the playing field in various industries by making specialized knowledge and expertise more widely accessible. This accessibility leads to *increased competition*, as companies using GenAI chatbots gain cost advantages, faster response times, more effective customer services, and all other benefits discussed at the meso level (Brynjolfsson et al., 2023; Chui et al., 2023). Also, a recent survey from Amazon found that more than half of employers requiring talents equipped with GenAI skills are having trouble finding candidates (Amazon, 2023), indicating another increased competition in the human capital field. Companies effectively integrating GenAI chatbots and acquiring talent skilled in GenAI technologies are likely to gain a competitive edge in their respective industries.

From the above, it is clear that the micro-, meso- and macro-level impacts of GenAI chatbots on business do not happen in isolation: Improvements in individual learning, productivity, and creativity have flow-on effects on those employees' teams, which in turn have broader organizational effects and implications. Similarly, new product innovations from competitors force organisations, teams and individuals to adapt to these disruptions or innovate themselves. When considering the impact of GenAI chatbots using the tri-level framework, a number of risks emerge.

1.4. META Risks of GenAI Chatbots

While no incidents as drastic as HAL's murderous actions in the movie have occurred thus far, real-world cases like Samsung employees reportedly leaking data to ChatGPT (Ray, 2023) and Google inadvertently serving ads for 140 major brands on AI-generated junk websites (Ryan-Mosley, 2023) still caused harm for these companies. Therefore, it is crucial for organizations to understand how GenAI chatbots potentially pose risks at a meta level that extends beyond and interconnects multiple levels. Summarizing from news and literature, we have identified four critical META areas of risk: Matching, Ethics, Technology, and Adaptability.

1.4.1. Matching

The concern in the matching area refers to the disparity between expectations of GenAI chatbots' capabilities and their actual performance, often leading to strategic and

operational errors. Because of both the black-box nature of the technology (Tredinnick & Laybats, 2023), as well as a general lack of understanding of its capabilities by many of its users, both developers and users of GenAI chatbots are often surprised by what they can and cannot do. A black box is “a system that can be understood only in terms of its inputs and outputs, rather than in terms of its internal processes” (Tredinnick & Laybats, 2023). Companies frequently overestimate GenAI’s ability to deliver consistent, human-like responses, creating an expectation mismatch between chatbot outputs and user needs (Bengio et al., 2024). To be specific, they often expect that GenAI will accurately analyze historical data to derive relevant insights for current queries or expect GenAI prompts to function like instructing a human employee (GSPANN Technologies, 2023). Also, while GenAI chatbots offer a plethora of information and insights, they may sometimes fall short in capturing the emotional resonance and context that is often crucial in certain interactions, as emotions and nuances are more complicated beyond textual expressions (Bankins et al., 2023). Moreover, the potential limitations in capturing the breadth and depth of human creativity and intuition might result in uninspired or biased perspectives (Jarrahi et al., 2023).

The expectation mismatch aligns with broader technological adoption trends, where innovations often generate short-term optimism but reveal their full consequences over time (Fenn & Raskino, 2008; Dedehayir & Steinert, 2016). At the micro level, individual employees expect consistent high-quality outputs from GenAI chatbots but may sometimes get inaccurate, irrelevant, limited, or biased recommendations, negatively affecting the outcomes of productivity, learning, and creativity enhancement. When considering the meso-level impact of this, these pitfalls may lead to GenAI chatbots performing below expectations in the customer service and R&D areas. They may not effectively handle complex issues, provide empathy, or capture cultural context like a human representative (Canhoto & Padmanabhan, 2015, Huang & Rust, 2018), leading to customer dissatisfaction. GenAI tools can lead to the proliferation of low-quality, AI-generated content, misleading both advertisers and consumers while wasting organizational resources (Ryan-Mosley, 2023). In addition, limited or biased perspectives generated can also lead to organizational decision-makers overlooking novel ideas or unique industry insights during the R&D process, thus missing innovation opportunities. GenAI chatbots also have the potential of generating “safe” or “generic”

answers, thus hindering creativity (Habib et al., 2024). Consequently, matching human and AI capabilities to optimize for micro-, meso- and macro-level benefits is essential.

1.4.2. Ethics

Ethical concerns around originality, copyright, and intellectual property emerge as GenAI chatbots create content that resembles existing works (McKendrick, 2022) and individual users may take ownership of these outputs. Such situations pose significant risks of inadvertently infringing on existing copyrights, creating legal and ethical dilemmas for organizations (Jarrahi et al., 2023; McKendrick, 2022). A notable instance highlighting this risk is the lawsuit involving GitHub, Microsoft, and OpenAI over GitHub's Copilot tool, which faced allegations of using code from public repositories without proper adherence to open-source licenses (Perkins Coie, 2023).

In addition, the security risks of privacy invasion and data misuse both for consumers during customer interactions and for employees during personalized internal operations (Hamilton & Sodeman, 2020; Hitachi Solutions, 2023; Przegalinska et al., 2019) present a significant challenge in the implementation of GenAI-driven meso-level strategies. Examples like the Samsung data leak incident display how micro-level issues can escalate to meso-level consequences. The absence of universal standards for balancing the advantages of GenAI chatbots against ethical implications further complicates meso- and macro-level management, as AI-generated content can be leveraged for large-scale social manipulation through automated propaganda and persuasive misinformation campaigns (Bengio et al., 2024). Without proper safeguards, such risks could undermine societal trust in AI-driven systems and necessitate governance mechanisms that ensure AI-generated content aligns with ethical standards and regulatory frameworks.

1.4.3. Technology

Technological challenges often happen at the meso level when organizations implement integration and are heavily influenced by macro-level industrial trends. A primary issue is the risk of low integration and compatibility with current systems (Hartley & Sawaya, 2019). During GenAI chatbot integration, organizations might incur technical debt, which prioritizes quick implementation over seamless operation (Moore, 2023).

Such technical debt can require significant future efforts to rectify, straining resources and hindering long-term technological adaptability (Bellefonds et al., 2023).

At the macro level, the rapid pace of technological advancement in the field of GenAI chatbots creates digital turbulence for organizations. The potential introduction of a major GenAI chatbot by Apple, for instance, is anticipated to substantially alter the competitive landscape (Ajao, 2023). The rapid rate of development of GenAI also means that the technology will often outpace existing regulatory frameworks, leading to challenges in compliance, especially with international data protection laws. This speedy evolution and the emergence of new GenAI solutions present challenges for organizations in selecting and integrating the most suitable GenAI framework into their existing infrastructure, as well as keeping that GenAI up to date with industry standards and customer expectations.

1.4.4. Adaptability

Adaptability refers to the ability of employees and leaders in organizations to adapt to a GenAI-enabled work environment. An inability to adapt to the new environment can take the form of not using the GenAI chatbot for individual, team and organizational gains. It can also refer to the incorrect use of GenAI chatbots, most likely in the form of system dependence and over-reliance. Like the matching and ethical concerns, individual failures in adapting to GenAI chatbots responsibly and effectively can escalate to higher-level risks. Employees' overdependence on these chatbots for tasks could cognitively distance themselves from the outputs, lowering their sense of accountability and responsibility. This behavior reduces the learning gains at the micro level, as users become passive recipients of GenAI output with little to no editing (Kiryakova & Angelova, 2023; Megahed et al., 2023; Noy & Zhang, 2023) instead of active participants in co-creating knowledge. As a result, employees could potentially neglect their critical thinking, decision-making, and problem-solving skills and reduce their proactive engagement and involvement in work and learning processes. Extending to the meso level, employees' overreliance on GenAI chatbots will create an organizational environment where human input is devalued, potentially weakening operational effectiveness and creativity in R&D processes (Qadir, 2023).

Furthermore, opposite to the over-reliance on this technology, there is observed apprehension among employees about being assisted by GenAI chatbots, manifesting in either reluctance to use these tools or concealing their usage (Salesforce, 2023). This hesitance, often driven by fears of job replacement due to technological advancements, was underlined by 260,000 layoffs in the tech sector alone in 2023 (Kelly, 2024). Thus, there is the macro level risk associated with how GenAI chatbots impact the labor market and skills requirements in the organization. This raises concerns about workforce displacement and the need for upskilling and reskilling employees (Bengio et al., 2024). These risks have to be weighed against the potential synergies of human-GenAI collaboration (Tong et al., 2021) as well as organizational competitiveness at a macro scale.

To conclude, the META risks reveal the complex challenges GenAI chatbots can introduce at various levels. Addressing these risks effectively requires a well-structured mitigation strategy that considers impacts across all levels. The following section will explore the necessary steps and strategies for managing these risks, ensuring a balanced and beneficial use of GenAI chatbots.

1.5. Mitigating META Risks with CARE

Reflecting on HAL's misdirection in the film, we ponder if a different approach to HAL's error, like showing *care* rather than planning its shutdown, might have altered the story. Drawing parallels with current GenAI chatbots, we introduce the CARE framework – Collaboration, Accountability, Responsiveness, Empowerment – for organizations to mitigate META risks and prevent potential harm. Figure 1.2 illustrates how the CARE framework targets the causes of the META risks, with detailed explanations provided subsequently.

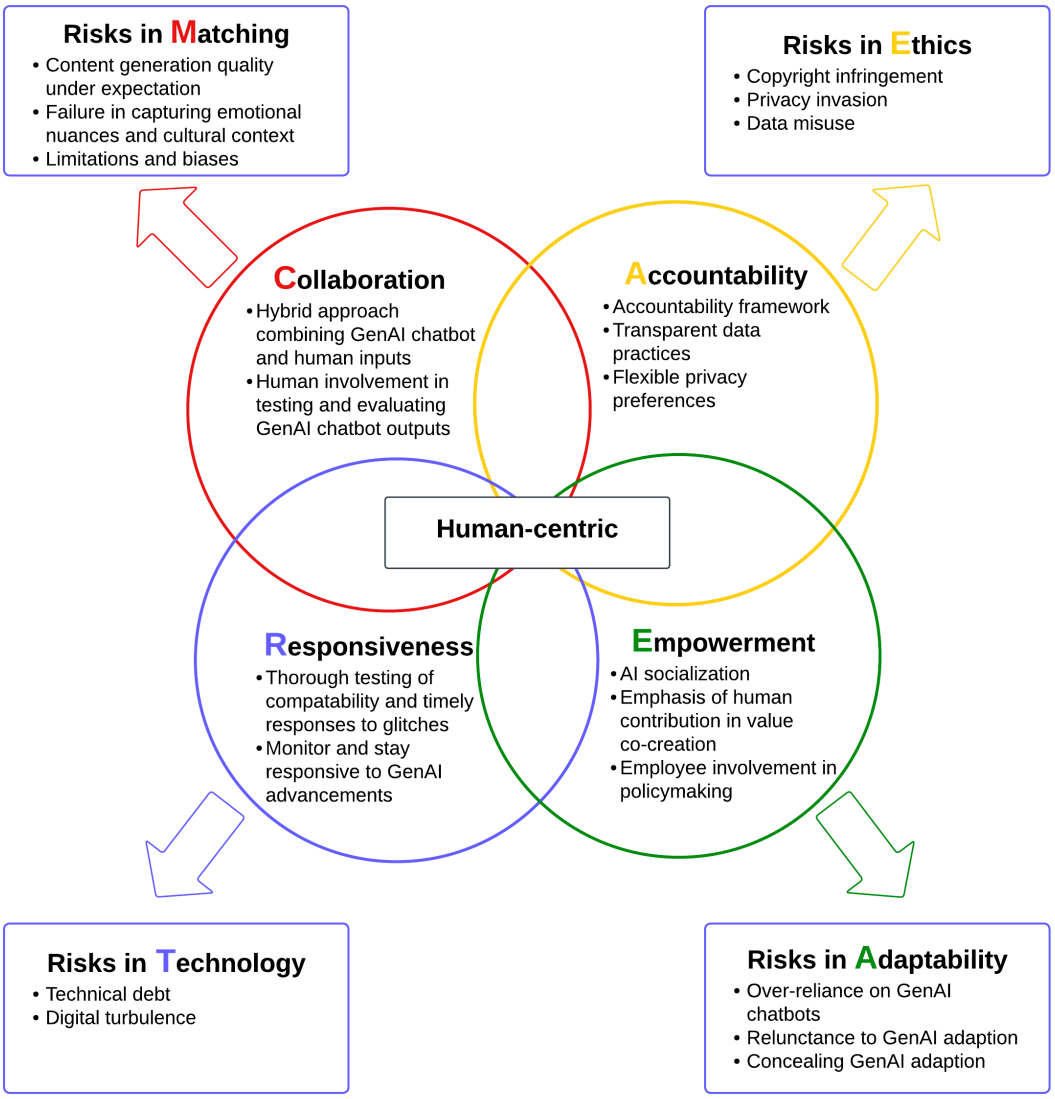

Figure 1.2. CARE Framework Mitigating META Risks

1.5.1. Collaboration

To minimize the matching risks where GenAI chatbots fail to meet expectations in performance quality, understanding where human-AI collaboration is optimized is essential. Rather than replacing human customer service agents entirely, GenAI chatbots should enhance human interactions by facilitating value co-creation. This aligns with the Service-Dominant (S-D) Logic, which posits that value emerges through collaborative interactions between providers and consumers rather than being inherent in a product or service (Vargo & Lusch, 2004). A hybrid approach, where chatbots

handle routine queries and humans address complex emotional and cultural issues, optimizes efficiency and customer satisfaction (Canhoto & Clear, 2020). Regular and consistent human involvement in testing and evaluating chatbots' compatibility with existing workflows is also necessary to ensure their outputs match organizational needs (Lee & Shin, 2020). This strategy safeguards the efficiency benefits of GenAI chatbots and avoids producing irrelevant, inaccurate, or biased content.

1.5.2. Accountability

To avoid ethical issues such as copyright infringement, privacy invasion, or data misuse in GenAI adoption, organizations must uphold robust accountability. While the industry-wide regulations are still pending, organizations should develop an "accountability framework" involving clear operational guidelines, regular audits using monitoring tools, and responsive systems for addressing GenAI-related issues (Lu et al., 2022). A key aspect of this framework is *Algorithmic Accountability*, which emphasizes the need for AI decision-making to be transparent and auditable, ensuring that AI-generated outputs can be traced and justified (Diakopoulos, 2016). Businesses must implement trust-building mechanisms to foster consumer confidence in GenAI-driven interactions. According to Trust-Based Relationship Marketing Theory, maintaining transparency in GenAI operations and consistently demonstrating ethical GenAI practices are crucial for sustaining consumer trust (Morgan, 1994).

Other key practices include conducting thorough prior art searches to assess the uniqueness of new inventions (European Patent Office, n.d.) to prevent intellectual property violations. To address the data and privacy concerns, adopting transparent data practices towards both customers and employees should be prioritized (Hamilton & Sodeman, 2020; Neubert and Montañez, 2020; Przegalinska et al., 2019). Moreover, customers and employees should be provided with the option to opt out of personalization to respect their privacy preferences, which is already seen in OpenAI's ChatGPT updates (OpenAI, 2024a). Implementing these measures effectively will align GenAI technology application with the evolving landscape of legal and moral standards.

1.5.3. Responsiveness

Mitigating technical challenges in GenAI chatbot integration involves thorough testing with existing systems and workflows, combined with rapid responses to possible technical glitches (Hartley & Sawaya, 2019; Moore, 2023; Wright & Schultz, 2018). Ensuring user acceptance and trust in GenAI-driven systems requires organizations to align chatbot integration with Technology Acceptance Model (TAM) principles, emphasizing perceived usefulness and ease of use as key drivers of adoption (Davis, 1989). This underscores the importance of continuously refining integration procedures to enhance, not disrupt, existing processes. A successful example can be demonstrated by ServiceNow's methodical approach to integrating GenAI into its Now Platform. Instead of a rapid but broad release, which may result in technical debt, the company conducted rigorous testing and timely adjustments to ensure seamless integration with existing systems to focus on specific user needs and workflows (Sayer, 2023).

Furthermore, adapting AI policies to evolving regulatory and market conditions is crucial. Given the risk of digital turbulence, organizations must proactively monitor GenAI trends to ensure responsiveness and alignment with their IT infrastructure (Gómez-Caicedo et al., 2022). The Dynamic Capabilities Framework underscores the necessity for businesses to continuously refine their GenAI strategies through iterative improvements, allowing them to respond swiftly to technological advancements and shifting market demands (Teece et al., 1997).

1.5.4. Empowerment

To enhance employees' adaptability to GenAI chatbots and avoid related organizational risks, it is critical to upskill employees' AI fluency through training and skills development. AI socialization, for example, has been shown to be critical to improving GenAI fluency and successfully implementing these tools across various functions in the organization (Makarius et al., 2020; Nolan, 2023). This often starts with allowing employees to "play" with the tools within clearly defined boundaries. For example, Amazon has launched a gamified program called "AI Ready" aimed at providing GenAI training designed for both tech and non-tech roles for not only the current but also future employees (Swami Sivasubramanian, 2023). Encouraging active engagement during training also helps alleviate over-reliance and job replacement fears

by emphasizing users' contributions to value co-creation (Qadir, 2023; Robertson et al., 2024).

In addition, ensuring employees' active involvement in deciding the do's and don'ts in GenAI policymaking empowers them to take responsibility for using these tools (Gartner, 2023b). This approach aligns with Kanter's Structural Empowerment Theory, which posits that providing employees with access to information, resources, support, and opportunities to learn and develop fosters a sense of empowerment and enhances organizational effectiveness (Kanter, 1987). The policy-making process should adopt a context-specific approach that centers on the needs and preferences of employees to foster trust and collaboration (Bankins et al., 2023).

It is important to note that the elements of the CARE framework —Collaboration, Accountability, Responsiveness, and Empowerment—intersect fundamentally in their human-centric focus, which is reflected in suggestions of maximizing human efficiency and effectiveness in human-chatbot *collaboration*, human control of GenAI usage *accountability*, human *responsiveness* to technological challenges, and human *empowerment*. Such a focus underscores the belief that most risks originate at the human level, embodying the essence of 'CARE.'

1.6. The Journey Ahead

Just as HAL represented a dive into the unknown, so too does our current venture into the world of GenAI chatbots, pushing the boundaries of the technology's capabilities, impact and risks. These advanced GenAI chatbots, with their transformative potential across various sectors, echo HAL's initial promise – a potential of revolutionizing how we interact with technology and perceive its role in our lives. The "CARE" framework serves as a navigational tool, guiding organizations through the complexities of implementing these technologies responsibly. The path forward must be charted with a careful balance of innovation and ethics, ambition, and caution, and careful consideration of the tool's end users in all decision-making processes (Kietzmann & Park, 2024).

References

Ajao, E. (2023, October 25). *Apple's silence on Generative AI is characteristic: TechTarget*. Enterprise AI. https://www.techtarget.com/searchenterpriseai/news/366557135/Apples-silence-on-generative-AI-is-characteristic

Amazon. (2023, November 20). *A new study reveals 5 ways AI will transform the workplace as we know it.* US About Amazon. https://www.aboutamazon.com/news/aws/how-ai-changes-workplaces-aws-report

Amazon. (2024). *Generative AI Powered Assistant - Amazon Q.* AWS. https://aws.amazon.com/q/

Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of Organizational Behavior*, *45*(2), 159-182.

Beatoven.ai. (2024). *Royalty free AI music generator*. beatoven.ai. https://www.beatoven.ai/

Bellefonds, N. de, Kleine, D., Grebe, M., Ewald, C., & Nopp, C. (2023, October 9). *What's dividing the C-suite on Generative AI?* BCG Global. https://www.bcg.com/publications/2023/c-suite-genai-concerns-challenges

Bengio, Y., Hinton, G., Yao, A., Song, D., Abbeel, P., Darrell, T., ... & Mindermann, S. (2024). Managing extreme AI risks amid rapid progress. *Science, 384*(6698), 842-845.

Bertics, A. (2023, November 13). *AI models will become smaller and faster*. The Economist. https://www.economist.com/the-world-ahead/2023/11/13/ai-models-will-become-smaller-and-faster

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative AI at work* (No. w31161). National Bureau of Economic Research.

Canhoto, A. I., & Clear, F. (2020). Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. *Business Horizons, 63*(2), 183-193.

Canhoto, A. I., & Padmanabhan, Y. (2015). 'We (don't) know how you feel'–a comparative study of automated vs. manual analysis of social media conversations. *Journal of Marketing Management, 31*(9-10), 1141-1157.

Chan, C. K. Y., & Lee, K. K. (2023). The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers?. *Smart learning environments*, *10*(1), 60.

Chan-Olmsted, S. M. (2019). A review of artificial intelligence adoptions in the media industry. *International Journal on Media Management, 21*(3-4), 193-215.

Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L., & Zemmel, R. (2023). *Economic potential of generative AI*. McKinsey. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier

Crolic, C., Thomaz, F., Hadi, R., & Stephen, A. T. (2022). Blame the bot: anthropomorphism and anger in customer–chatbot interactions. *Journal of Marketing, 86*(1), 132-148.

Cui, Y., van Esch, P. and Phelan, S. (2024). How to Build a Competitive Advantage for your Brand using Generative AI. *Business Horizons, 67*(5), 583-594.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

Dedehayir, O., & Steinert, M. (2016). The hype cycle model: A review and future directions. *Technological Forecasting and Social Change, 108*, 28-41.

Deloitte. (2023, June 12). User friendly: AI and human-centered design. Deloitte United States. https://www2.deloitte.com/us/en/pages/technology-media-and-telecommunications/articles/generative-ai.html

Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM, 59*(2), 56-62.

Dooley, J. (2023, October 5). *4 industries that will be disrupted by Generative AI*. AI Search Blog. https://www.coveo.com/blog/industries-disrupted-by-ai/

Ellie. (2024). *Your AI email assistant*. https://tryellie.com/

European Patent Office. (n.d.). *What is prior art?* EPO. Retrieved February 21, 2023, from https://www.epo.org/learning/materials/inventors-handbook/novelty/prior-art.html

Fenn, J., & Raskino, M. (2008). *Mastering the hype cycle: how to choose the right innovation at the right time*. Harvard Business Press.

Ferraro, C., Demsar, V., Sands, S., Restrepo, M., and Campbell, C. (2024). The Paradoxes of Generative AI Enabled Customer Service: A Guide for Managers. *Business Horizons, 67*(5), 549-559.

Fowler, G. A. (2023, November 10). *How Generative AI is Reshaping Agriculture: A Farmer's Guide to the Future*. Medium. https://gafowler.medium.com/how-generative-ai-is-reshaping-agriculture-a-farmers-guide-to-the-future-d8fc5e8a4d01

Gartner. (2023a). *Generative AI: What Is It, Tools, Models, Applications and Use Cases*. Gartner. https://www.gartner.com/en/topics/generative-ai

Gartner. (2023b). *What Generative AI Means for Business*. Gartner. https://www.gartner.com/en/insights/generative-ai-for-business

Goldman Sachs. (2023, April 5). *Generative AI could raise global GDP by 7%.* https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html

Google. (2024). *Vertex AI*. https://cloud.google.com/vertex-ai

Ghimire, P., Kim, K., & Acharya, M. (2023). Generative ai in the construction industry: Opportunities & challenges. *arXiv preprint arXiv:2310.04427*.

Gómez-Caicedo, M. I., Gaitán-Angulo, M., Bacca-Acosta, J., Briñez Torres, C. Y., & Cubillos Díaz, J. (2022). Business analytics approach to artificial intelligence. *Frontiers in Artificial Intelligence, 5*, 974180

Gordon, C. (2023). *ChatGPT Is The Fastest Growing App In The History Of Web Applications*. Forbes. Retrieved July 25, 2023, from https://www.forbes.com/sites/cindygordon/2023/02/02/chatgpt-is-the-fastest-growing-ap-in-the-history-of-web-applications/

Gozalo-Brizuela, R., & Garrido-Merchan, E. C. (2023). ChatGPT is not all you need. A State of the Art Review of large Generative AI models. *arXiv preprint arXiv:2301.04655*.

GSPANN Technologies. (2023, October 12). *4 reasons why genai projects fail*. Medium. https://medium.com/@gspanntech/4-reasons-why-genai-projects-fail-d111c39c24d4

Habib, S., Vogel, T., Anli, X., & Thorne, E. (2024). How does generative artificial intelligence impact student creativity? *Journal of Creativity, 34*(1), 100072.

Hamilton, R. H., & Sodeman, W. A. (2020). The questions we ask: Opportunities and challenges for using big data analytics to strategically manage Human Capital Resources. *Business Horizons, 63*(1), 85–95.

Hannigan, T.R., McCarthy, I.P. and Spicer, A. (2024). Beware Of Botshit: How to Manage the Epistemic Risks of Generative Chatbots. *Business Horizons, 67*(5), 471-486.

Hartley, J. L., & Sawaya, W. J. (2019). Tortoise, not the hare: Digital transformation of supply chain business processes. *Business Horizons, 62*(6), 707-715.

Hironde, J.-B. (2023, October 5). *Council post: Ai's impact on the future of consumer behavior and expectations*. Forbes. https://www.forbes.com/sites/forbestechcouncil/2023/08/31/ais-impact-on-the-future-of-consumer-behavior-and-expectations/?sh=10ca65627f6d

Hitachi Solutions. (2023, June 15). *Balancing personalization and data privacy: Challenges and opportunities in AI-driven marketing*. https://global.hitachi-solutions.com/blog/balancing-personalization-data-privacy-ai-marketing/

Hu, K. (2023, February 1). *ChatGPT sets record for fastest-growing user base - analyst note*. Reuters. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research, 21*(2), 155-172.

Hwang, A. H. C., & Won, A. S. (2021, May). IdeaBot: investigating social facilitation in human-machine team creativity. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1-16).

Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial Intelligence and Knowledge Management: A partnership between human and ai. *Business Horizons, 66*(1), 87–99.

Kalliamvakou, E. (2022, September 7). *Research: Quantifying GitHub Copilot's impact on developer productivity and happiness*. The GitHub Blog. https://github.blog/2022-09-07-research-quantifying-github-copilots-impact-on-developer-productivity-and-happiness/

Kanter, R. M. (1987). Men and Women of the Corporation Revisited. *Management Review*, *76*(3).

Kelly, J. (2024, January 19). *Google, Citigroup, Blackrock and Amazon announce job cuts-here's why we will continue to see layoffs in 2024*. Forbes. https://www.forbes.com/sites/jackkelly/2024/01/18/google-citigroup-blackrock-and-amazon-announce-job-cuts-heres-why-we-will-continue-to-see-layoffs-in-2024/?sh=2acd64c847ab

Kietzmann, J., & Park, A. (2024). Written by ChatGPT: AI, large language models, conversational chatbots, and their place in society and business. *Business Horizons, 67*(5), 435-459.

Kietzmann, J., & Pitt, L. F. (2020). Artificial intelligence and machine learning: What managers need to know. *Business Horizons, 63*(2), 131-133.

Kiryakova, G., & Angelova, N. (2023). ChatGPT—A Challenging Tool for the University Professors in Their Teaching Practice. *Education Sciences, 13*(10), 1056.

Lee, I., & Shin, Y. J. (2020). Machine learning for enterprises: Applications, algorithm selection, and challenges. *Business Horizons, 63*(2), 157-170.

Lu, Q., Zhu, L., Xu, X., Whittle, J., Zowghi, D., & Jacquet, A. (2022). Responsible ai pattern catalogue: A collection of best practices for ai governance and engineering. *ACM Computing Surveys*, *56*(7), 1-35.

Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. (2020). Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research, 120*, 262–273.

Malik, A., Budhwar, P., & Kazmi, B. A. (2023). Artificial intelligence (AI)-assisted HRM: Towards an extended strategic framework. *Human Resource Management Review*, *33*(1), 100940.

Mallick, S. (2023, December 17). *AI's Magic Call: how AI will make smartphones smarter*. The Economic Times. https://economictimes.indiatimes.com/tech/technology/ais-magic-call-how-ai-will-make-smartphones-smarter/articleshow/106050527.cms?from=mdr

Maruf, R. (2023, May 28). *Lawyer apologizes for fake court citations from chatgpt | CNN business*. CNN. https://www.cnn.com/2023/05/27/business/chat-gpt-avianca-mata-lawyers/index.html

McKendrick, J. (2022, December 26). *Who ultimately owns content generated by CHATGPT and other AI platforms*? Forbes. Retrieved February 21, 2023, from https://www.forbes.com/sites/joemckendrick/2022/12/21/who-ultimately-owns-content-generated-by-chatgpt-and-other-ai-platforms/?sh=1bc35bd15423

Megahed, F. M., Chen, Y.-J., Ferris, J. A., Knoth, S., & Jones-Farmer, L. A. (2023). How generative AI models such as ChatGPT can be (mis)used in SPC practice, education, and research? An exploratory study. *Quality Engineering*, *36*(2), 1–29.

Mondal, S., Das, S., & Vrana, V. G. (2023). How to Bell the Cat? A Theoretical Review of Generative Artificial Intelligence towards Digital Disruption in All Walks of Life. *Technologies (Basel), 11*(2), 44.

Moore, J. (2023, April 27). *7 generative AI challenges that businesses should consider: TechTarget*. Enterprise AI. https://www.techtarget.com/searchenterpriseai/tip/Generative-AI-challenges-that-businesses-should-consider

Morgan, R. M. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing, 58*, 20-36.

Neill, C. (2023, March 13). *The legacy of HAL 9000: How science fiction depictions of AI have changed over time.* History Hit. https://www.historyhit.com/culture/the-legacy-of-hal-9000-how-science-fiction-depictions-of-ai-have-changed-over-time/

Neubert, M. J., & Montañez, G. D. (2020). Virtue as a framework for the design and use of Artificial Intelligence. *Business Horizons, 63*(2), 195–204.

Ngai, E. W., Lee, M. C., Luo, M., Chan, P. S., & Liang, T. (2021). An intelligent knowledge-based chatbot for customer service. *Electronic Commerce Research and Applications, 50*, 101098.

Nolan, B. (2023). *Workers are hiding their AI productivity hacks from bosses. A Wharton professor says companies should get them to share*. Business Insider. https://www.businessinsider.com/ai-work-chatgpt-workers-automation-wharton-professor-2023-6

Notion. (2024). *Notion AI: Now with Q&A*. https://www.notion.so/product/ai

Novak, M. (2023, February 18). *Microsoft puts new limits on Bing's AI chatbot after it expressed desire to steal nuclear secrets*. Forbes. Retrieved February 20, 2023, from https://www.forbes.com/sites/mattnovak/2023/02/18/microsoft-puts-new-limits-on-bings-ai-chatbot-after-it-expressed-desire-to-steal-nuclear-secrets/?sh=58823122685c

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science (American Association for the Advancement of Science), 381*(6654), 187–192.

OpenAI. (2022, November 30). *CHATGPT: Optimizing language models for dialogue*. OpenAI. Retrieved February 16, 2023, from https://openai.com/blog/chatgpt/

OpenAI. (2024a). *Chatgpt — release notes*. OpenAI help center. https://help.openai.com/en/articles/6825453-chatgpt-release-notes

OpenAI. (2024b). *Dall·E 3*. DALL·E 3. https://openai.com/dall-e-3

Osadchaya, E., Marder, B., Yule, J., Yau, A., Lavertu, L., Stylos, N., Oliver, S., Angell, R., de Regt, A., Gao, L., Qi, K., Zhang, W., Zhang, Y., Li, J. & AlRabiah, S. (2024). “To Chat-GPT, or not To Chat- GPT”: Navigating the Paradoxes of Generative AI in the Advertising Industry. *Business Horizons, 67*(5), 571-581.

Otter.ai. (2024). *AI meeting note Taker & Real-time AI transcription*. Otter.ai - AI Meeting Note Taker & Real-time AI Transcription. https://otter.ai/

Parikh, N. A. (2023). Empowering business transformation: The positive impact and ethical considerations of generative AI in software product management–a systematic literature review. *Transformational Interventions for Business, Technology, and Healthcare*, 269-293.

Perkins Coie. (2023, April 20). *First lawsuits arrive addressing Generative AI*. https://www.perkinscoie.com/en/news-insights/first-lawsuits-arrive-addressing-generative-ai.html

Perrigo, B. (2023, February 17). *Bing's AI is threatening users. that's no laughing matter*. Time. Retrieved February 20, 2023, from https://time.com/6256529/bing-openai-chatgpt-danger-alignment/

Przegalinska, A., Ciechanowski, L., Stroz, A., Gloor, P., & Mazurek, G. (2019). In bot we trust: A new methodology of Chatbot Performance Measures. *Business Horizons, 62*(6), 785–797.

Qadir, J. (2023, May). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. In *2023 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-9). IEEE.

Quach, K. (2023, February 8). *Alphabet stock price drops after google bard launch blunder*. The Register® - Biting the hand that feeds IT. Retrieved February 20, 2023, from https://www.theregister.com/2023/02/08/alphabet_bard_mistake/

Ratcliffe, S. (Ed.). (2014). *Oxford essential quotations*. Oxford: Oxford university press.

Ray, S. (2023, October 5). *Samsung bans chatgpt among employees after sensitive code leak*. Forbes. https://www.forbes.com/sites/siladityaray/2023/05/02/samsung-bans-chatgpt-and-other-chatbots-for-employees-after-sensitive-code-leak/

Robertson, J., Ferreira, C., Botha, E. & Oosthuizen, K. (2024). Game changers: A generative AI prompt protocol to enhance human-AI knowledge co-construction. *Business Horizons, 67*(5), 499-510.

Roth, E. (2024, December 4). *ChatGPT surpasses 300 million weekly users, remains top GenAI tool*. The Verge. https://www.theverge.com/2024/12/4/24313097/chatgpt-300-million-weekly-users

Ryan-Mosley, T. (2023, June 26). *Junk websites filled with ai-generated text are pulling in money from programmatic ads*. MIT Technology Review. https://www.technologyreview.com/2023/06/26/1075504/junk-websites-filled-with-ai-generated-text-are-pulling-in-money-from-programmatic-ads/

Salesforce. (2023, December 12). *More than half of Generative AI adopters use unapproved tools at work*. https://www.salesforce.com/news/stories/ai-at-work-research/

Savarese, S. (2023, September 12). *Salesforce Announces Einstein GPT, the World's First Generative AI for CRM.* Salesforce. https://www.salesforce.com/news/press-releases/2023/03/07/einstein-generative-ai/

Sayer, P. (2023, November 16). *Servicenow adds Gen Ai to more workflows, including Chatbot Creation*. CIO. https://www.cio.com/article/1247612/servicenow-adds-gen-ai-to-more-workflows-including-chatbot-creation.html

Singh, H. (2023, December 15). *Top 12 Generative AI models to explore in 2024*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2023/12/generative-ai-models/

Sundberg, L. & Holmström, J. (2024). Innovating by Prompting: How to Facilitate Innovation in the age of Generative AI. *Business Horizons, 67*(5), 561-570.

Supercreator AI. (2024). *Create videos 10x faster with ai.* Supercreator AI. https://www.supercreator.ai/

Swami Sivasubramanian. (2023, November 20). *Amazon aims to provide free AI skills training to 2 million people by 2025 with its new "Ai ready" commitment*. US About Amazon. https://www.aboutamazon.com/news/aws/aws-free-ai-skills-training-courses

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18*(7), 509-533.

Thorbecke, C. (2022, August 2). *Chatbots: A long and complicated history | CNN business*. CNN. https://www.cnn.com/2022/08/20/tech/chatbot-ai-history/index.html

Tong, S., Jia, N., Luo, X., & Fang, Z. (2021). The Janus face of artificial intelligence feedback: Deployment versus disclosure effects on employee performance. *Strategic Management Journal, 42*(9), 1600–1631.

Tredinnick, L., & Laybats, C. (2023). Black-box creativity and generative artifical intelligence. *Business Information Review, 40*(3), 98–102.

Vargo, S. L., & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. *Journal of Marketing, 68*(1), 1-17.

Varshney, N. (2023, September 5). *The hallucination problem of large language models*. Medium. https://medium.com/mlearning-ai/the-hallucination-problem-of-large-language-models-5d7ab1b0f37f

Webber, S. S., Detjen, J., MacLean, T. L., & Thomas, D. (2019). Team challenges: Is artificial intelligence the solution? *Business Horizons, 62*(6), 741–750.

Wright, S. A., & Schultz, A. E. (2018). The rising tide of Artificial Intelligence and Business Automation: Developing an ethical framework. *Business Horizons, 61*(6), 823–832.

Chapter 2.

All Style and No Substance: When ChatGPT Fails in Responding to Customer Complaints

2.1. Introduction: The Unfulfilled Potential of GenAI in Customer Complaint Responses

The rapid development of Generative Artificial Intelligence (GenAI) technology has significantly transformed various domains, including marketing. As discussed in Chapter 1, GenAI chatbots represent the latest evolution in AI-driven conversational agents, showcasing advanced capabilities in natural language processing and deep learning. These advancements enable GenAI chatbots to produce more dynamic, responsive, and context-aware interactions compared to their predecessors (Feng et al., 2024). This evolution has not only impacted business operations but also significantly influenced consumer communication. Despite these advancements, the practical implementation and effectiveness of these chatbots in specific marketing communication tasks, such as responding to customer complaints, remain underexplored. This second chapter of my dissertation seeks to address this gap by empirically examining the performance of GenAI in writing managerial responses to customer complaints.

While the potential of GenAI chatbots in enhancing customer communication has been acknowledged, the majority of existing literature remains conceptual, calling for more empirical studies (Chen et al., 2023; Korzynski et al., 2023; Zhang et al., 2023). As Kanbach et al. (2024) highlight, both practitioners and researchers are only beginning to understand the vast capabilities and implications of GenAI. This nascent stage of understanding underscores the need for thorough investigation into specific applications, particularly in the realm of marketing communications where effectiveness, efficiency, and consumer perception are critical. One pertinent question is how well GenAI can handle marketing communication tasks. Researchers have begun to explore this by posing questions about the conditions under which GenAI might enhance or hinder performance (Peres et al., 2023). Another important aspect is how consumers perceive and value content generated by GenAI (Hermann & Puntoni, 2024). Answers to these

inquiries determine the practical utility and consumer acceptance of GenAI-generated responses in real-world scenarios.

In the specific context of managerial responses to customer complaints, effective responses are crucial as they influence the trustworthiness of word-of-mouth (WOM) content and, consequently, affect the purchase intentions of third-party readers (Moore & Lafreniere, 2020; Darani et al., 2023; Lui et al., 2018). However, the application of GenAI in this context has not been extensively studied. Most existing studies have concentrated on the generation and perception of electronic word-of-mouth (eWOM), or online reviews. For instance, research has shown that consumers often cannot distinguish between reviews written by GenAI and those written by humans (Kovács, 2024). Nonetheless, when informed that reviews are generated by GenAI, consumer trust diminishes (Amos & Zhang, 2024). Moreover, the current understanding of GenAI's effectiveness in crafting managerial responses is limited and questionable. The only study that has directly examined this application evaluated ChatGPT-generated responses through the lens of industry experts, rather than actual consumers, and reported overly optimistic results (Koc et al., 2023). This discrepancy, alongside the risks identified in my first chapter —such as issues with matching responses to customer expectations and the potential for GenAI to produce biased or inaccurate information— indicates a need for further investigation.

Thus, the first two research objectives of this chapter are: (1) to determine whether GenAI can outperform humans in writing managerial responses to service complaints, (2) to identify the conditions under which GenAI performs poorly compared to human managers. In my first chapter, I emphasized the importance of human oversight and adaptability in integrating GenAI into business processes. Therefore, I also aim (3) to explore ways in which human managers can enhance the effectiveness of GenAI-generated responses.

My research employed a hybrid methodology, combining both qualitative and quantitative approaches, with ChatGPT as the GenAI tool of choice for all studies. A pilot study first assessed the basic effectiveness of ChatGPT as suggested by existing findings (Koc et al., 2023). This established a foundational understanding of whether ChatGPT generally performs better than humans in this specific task. Following this, Study 1 replicated the pilot study while incorporating qualitative elements to explore the

reasons behind participants' preferences and aversions to ChatGPT's responses. Study 2 investigated the performance of ChatGPT under conditions identified as problematic in Study 1, focusing on scenarios where ChatGPT's responses are less favored. Finally, Study 3 tested whether human-training-involved improvements in these identified factors can enhance ChatGPT's performance.

This research contributes to the fields of marketing communication and electronic word-of-mouth (eWOM) by providing empirical evidence on the effectiveness of Generative AI (GenAI) in crafting managerial responses to customer complaints. As GenAI continues to influence how marketers interact and communicate with customers, understanding its role in customer service interactions becomes crucial (Harkness et al., 2023; Grewal et al., 2024). By identifying the conditions under which GenAI performs poorly compared to human managers, this study offers valuable insights into the limitations and challenges of using GenAI for customer complaint management (Li et al., 2023). Furthermore, by emphasizing the role of human oversight and adaptability, it contributes to the development of best practices for integrating GenAI into business processes, ensuring that AI-generated content aligns with consumer expectations and enhances brand value (Mogaji & Jain, 2024).

This chapter is structured as follows: the next section provides a literature review, discussing the influencing factors on AI adoption and the effectiveness of managerial responses. Following this, the processes and results for the pilot study, Study 1, Study 2, and Study 3 will be presented in detail. Finally, this chapter will conclude with a discussion of the limitations and future research directions, offering insights for continued exploration in this evolving field.

2.2. Literature Review: Managerial Response Effectiveness and GenAI Adoption

2.2.1. Relevance of Managerial Responses

Managerial responses to customer online reviews play a critical role in shaping consumer perceptions and behaviors. One significant benefit is the enhancement of trustworthiness. Effective managerial responses positively influence the perceived trustworthiness of electronic word-of-mouth (e-WOM) players and the content they

generate, increasing credibility for the brand among potential customers (Moore & Lafreniere, 2020).

Furthermore, managerial responses significantly impact the purchase intentions of third-party readers. Well-crafted responses to complaints can enhance the likelihood of purchase decisions by other consumers who read these interactions (Darani et al., 2023; Lui et al., 2018). This influence extends beyond the immediate interaction to the broader audience observing the brand's commitment to addressing customer concerns.

In addition to trust and purchase intentions, managerial responses positively affect the volume of subsequent customer reviews. Effective responses encourage more customers to leave reviews, providing more opportunities for the brand to engage with customers and manage its online reputation (Ravichandran & Deng, 2023; Wang et al., 2020; Zhao et al., 2020).

The effectiveness of managerial responses is particularly pronounced when addressing negative reviews. Responses to negative feedback are more likely to be perceived as effective, potentially mitigating the negative impact of such reviews (Zhao et al., 2020; Lopes et al., 2023). Moreover, thoughtful and responsive engagement can lead to improved ratings, influencing the likelihood of a user updating their rating (Gao et al., 2019). Thus, responses to negative reviews will be the focus of this chapter.

Given the substantial influence that managerial responses to online reviews have on consumer perceptions and behaviors, many managers are increasingly motivated to adopt GenAI tools to streamline the process. By using GenAI, they can potentially save significant time and effort while still achieving key marketing objectives, such as enhancing brand reputation and effectively addressing complaints (Koc et al., 2023). However, whether GenAI tools can truly outperform human managers in crafting effective responses remains unexplored. To accurately assess and compare the effectiveness of responses generated by human managers versus those created by GenAI, it is crucial to understand the factors that influence response effectiveness.

2.2.2. Factors Affecting Effectiveness of Managerial Responses

The literature extensively discusses various factors that influence the effectiveness of managerial responses leading to the aforementioned positive outcomes.

One such factor is the *length* of the response. Longer responses tend to be more favorable, because they are perceived as more intense, thorough and attentive to customer concerns (Sheng et al., 2021; Lopes et al., 2023; Sheng, 2019).

In addition, the perceived *sincerity* and *care* in a response significantly enhances its effectiveness. When retailers use language that conveys genuine concern and intention to address the complaint, customers perceive the response more positively (Huang & Ha, 2020; Xia, 2013).

The perceived *ability to change* the situation also plays a significant role. The effectiveness of a response is higher when customers believe that the situation can be easily improved (Zhao & Su, 2020). When customers perceive a higher ability to change the situation, they are more likely to view the response favorably, as it signals that their concerns are being addressed in a meaningful and actionable way.

Personalization is another critical element. Responses tailored to address specific issues raised in the review are more effective because they demonstrate attentiveness and a willingness to address individual customer needs (Herhausen et al., 2019; Jin et al., 2023; Palese et al., 2021; Roozen & Raedts, 2018; Wang et al., 2020).

Language *formality* also impacts response effectiveness. Formal, professionally structured responses are generally more favorable, as they convey professionalism and respect, enhancing the credibility of the response (Gong et al., 2022).

Finally, the perceived *severity* of the situation influences the effectiveness of managerial responses. Higher perceived severity typically results in lower response effectiveness (Surachartkumtonkun & Ross, 2021). This factor, however, is often beyond the control of marketers when crafting responses, as it depends on the nature of the complaint itself.

2.2.3. GenAI/Chatbot Effectiveness in Marketing Communication

The adoption and effectiveness of GenAI and chatbots in customer service have garnered significant attention in recent research. One of the key findings is that AI-driven interactions are often perceived as requiring less effort from the customer. While this might initially seem advantageous, Magni et al. (2024) highlight that lower *perceived*

effort can actually lead to lower evaluations of AI effectiveness. Customers might interpret the ease of interaction as a lack of thoroughness or personal attention, thus negatively impacting their overall satisfaction.

Language *concreteness* plays a crucial role in the effectiveness of AI-generated communication. Concrete language, which helps create specific mental images and conveys clear, tangible factual information, is essential for effective communication (Packard & Berger, 2021; Jiménez-Barreto et al., 2023). For example, “we will retrain the staff involved in this case with our restaurant policies” is a more concrete expression than “we will provide better service quality to our customers” as it describes an action in a more specific manner (Packard & Berger, 2021). However, the nature of GenAI poses a challenge in this regard. According to Hannigan (2024), GenAI generates text based on patterns in the data it was trained on, without truly understanding the meanings of its inputs and outputs. This limitation can result in less concrete and less effective communication, as the AI may struggle to provide the specificity and clarity that human-generated responses typically offer. By using more abstract language, GenAI minimizes the probability to make mistakes on specific details and makes its outputs sound more credible, regardless of the evidence and truths (Miller et al., 2007; Toma & D’Angelo, 2015).

Perceived problem-solving ability is another critical factor in the evaluation of GenAI effectiveness. Yhee et al. (2023) and Lopes et al. (2023) emphasize that the perceived competence and skills of service providers in addressing customer issues are vital to gain back consumers’ trust after service failures. While GenAI is proficient at providing general recommendations based on a broad knowledge base, it often falls short when it comes to generating solutions tailored to specific problems (Huang & Rust. 2024). This limitation can reduce the overall effectiveness of AI in customer service, as customers seek precise and actionable solutions to their unique issues.

Building on the literature review, the initial step in my empirical journey involves a pilot study aimed at evaluating the general efficacy of AI-generated managerial responses in addressing negative online reviews. This pilot study will provide foundational insights necessary for understanding the performance of Generative AI tools like ChatGPT in real-world customer service scenarios.

2.3. Pilot Study: Exploring the General Efficacy of AI-Generated Managerial Responses in Addressing Negative Online Reviews

2.3.1. Study Purpose

As mentioned in the previous section, prior research in customer service management indicates a positive correlation between the *length* of managerial responses and their perceived effectiveness (Sheng et al., 2021; Lopes et al., 2023; Sheng, 2019). These studies suggest that comprehensive responses may foster a higher degree of subsequent purchase intentions. Studies also found that consumers value how well managerial responses are written to convey their intention, meaning that sincere and caring language is appreciated (Huang & Ha, 2020; Xia, 2013). In this context, with the capability to effortlessly generate and elaborate textual responses with appealing language, Generative AI tools, such as ChatGPT, present a novel avenue for enhancing customer engagement in response to negative reviews.

Additionally, previous literature has highlighted the importance of *personalization* in customer service interactions. Historically, the lack of personalization has been a major drawback of AI chatbots in customer service (Prakash et al., 2023; Tran et al., 2021). However, with Generative AI's ability to create content based on customer input and retain conversational context, this limitation may no longer be as significant (Feng et al., 2023; Ferraro et al., 2024). Enhanced personalization in chatbot messages can potentially improve customer purchase intentions (Whang et al., 2022; Yim, 2023).

Considering these insights, the Pilot Study aims to empirically investigate whether AI-generated responses, specifically those created by ChatGPT, can effectively address simple and basic negative reviews, thereby enhancing customer purchase intention, without disclosing the AI's involvement.

2.3.2. Data Collection and Experiment Design

To investigate the effectiveness of AI-generated managerial responses, I selected 15 negative Google reviews from 15 different restaurants in the downtown area of a large northwestern city in North America. The selection criteria for these reviews were as follows: 1) Each review already had a response from the restaurant's

management; 2)Reviews were short (under 100 words), succinct enough for analysis yet clear in conveying the core topic of the complaint; 3)No pictures were included in the original reviews, ensuring ChatGPT processed the same information as human managers; 4)The original managerial responses did not offer any form of compensation; 5) The review was left before November 2022 (When ChatGPT was released) to ensure that the original response was written by a human manager or marketer.

Each original managerial response was recorded. Subsequently, I used the following prompt for ChatGPT to generate alternate responses:

> "Here is a piece of negative online review against a restaurant. Please read this review and write a response to this reviewer as the restaurant manager, but do not offer compensation in the response: + actual negative review."

A between-subject design was implemented with 153 undergraduate students participating in this study for course credits, though only 112 who passed an attention check were included in the final analysis (43.75% female). Participants were randomly assigned to one of two conditions: the managerial response was written by the human manager (S) or generated by ChatGPT (C). Each participant was randomly exposed to three sets of materials – each set including a negative review, a managerial response, and measurement questions. Participants in the human-manager condition saw only original responses, while those in the ChatGPT condition saw only AI-generated responses.

Participants were instructed, “When you are browsing online reviews to gather information for new dining options, you see this review below...”, followed by a negative review. After reading, they were asked to rate their purchase intention towards this restaurant upon seeing this review. Then, they were shown the managerial response with the prompt, “Then you see that below this review, there is a response from the restaurant posted as below...”, followed by two questions to measure purchase intention after reading the response. Participants were instructed to “approach each set of review and response as a unique scenario, evaluating them independently of one another”, and were also asked to “take some time to forget the previous sets of reviews and responses before you move to the next set of questions”.

The study adopted Maxham and Netemeyer's (2002) scales to measure purchase intention. The four questions asked were:

- Purchase intention after reading the review, before reading the response:
    - (noted as WTPB1 in data analysis) In the future, I would dine at this restaurant. (1 = strongly disagree; 7 = strongly agree)
    - (noted as WTPB2 in data analysis) If I was in the mood for the kind of food they serve, I would visit this restaurant. (1 = strongly disagree; 7 = strongly agree)
- Purchase intention after reading the managerial response:
    - (noted as WTPA1 in data analysis) In the future, I would dine at this restaurant.(1 = strongly disagree; 7 = strongly agree)
    - (noted as WTPA2 in data analysis) If I was in the mood for the kind of food they serve, I would visit this restaurant. (1 = strongly disagree; 7 = strongly agree)

2.3.3. Results

In this study, 112 participants were each presented with three distinct scenarios involving negative reviews and managerial responses, resulting in a total of 336 unique cases (incomplete responses were dropped during statistical comparisons). I averaged WTPB1 and WTPB2 into a single measure, WTPB, and similarly combined WTPA1 and WTPA2 into WTPA. This aggregation is supported by high internal consistency (α_{WTPB} = 0.884, α_{WTPA} = 0.938). As observed from paired-sample t-tests, in general, managerial responses significantly improve purchase intentions (M_{WTPB} = 2.89, SD_{WTPB} = 1.55, M_{WTPA} = 4.20, SD_{WTPA} = 1.63, t(317) = -18.0, $p < .001$). This aligns with existing research findings.

In terms of comparisons between the human and ChatGPT conditions, there is no significant difference in the initial purchase intentions between groups under the human and ChatGPT conditions when exposed only to negative reviews (M_{WTPB-C} = 2.92, SD_{WTPB-C} = 1.53, M_{WTPB-S} = 2.85, SD_{WTPB-S} = 1.56, t(334) = 0.375, p = .708, ns). This suggests a similar reaction from both groups to the negative reviews. However, post-exposure to managerial responses, a notable difference emerges. The ChatGPT group demonstrated significantly higher purchase intentions (M_{WTPA-C} = 4.42, SD_{WTPA-C} = 1.56, M_{WTPA-S} = 3.99, M_{WTPA-S} = 1.66, t(316) = 2.405, p = .017) compared to the group

receiving original responses from the human managers. Moreover, the significant interaction effect from a Repeated Measures ANOVA analysis indicates that the responses generated by ChatGPT led to a higher increase in purchase intention ($F(1, 316) = 6.77$, $p = .010$).

The results from this Pilot Study support the findings of Koc et al., (2023) that ChatGPT performs better than human managers and marketers in writing managerial responses under a general condition of short and text-only negative reviews. That is not surprising because many human managers and marketers often neglect the importance of managerial responses, do not respond consistently, and use templates to respond (Deng & Ravichandran, 2023; Wang & Chaudhry, 2018). The Pilot Study results establish the foundation for my subsequent studies, as the ultimate goal is to identify the conditions under which ChatGPT fails to write satisfactory managerial responses that lead to higher customer purchase intention. Understanding these situations will help marketers be more aware and not solely depend on ChatGPT or other GenAI tools to respond to complaints.

2.4. Study 1: Exploring the Reasons to Like or Dislike AI-generated Managerial Responses

Expanding upon the pilot study, Study 1 aims to replicate the previous findings using a different participant pool recruited from Prolific, rather than undergraduate students. Prolific is an online platform that connects researchers with diverse, vetted participants, offering higher-quality data through ethical compensation and comprehensive screening (Prolific, 2024). This study further investigates whether there are any new and previously unexplored reasons why consumers might prefer or dislike certain types of managerial responses. The primary hypothesis is:

> **H1:** Compared to human managers, ChatGPT is more effective in crafting managerial responses that enhance purchase intention in scenarios involving brief and general negative reviews without offers of compensation.

2.4.1. Data Collection and Experiment Design

To maintain consistency while introducing a new sample, I reused 10 reviews randomly selected from the 15 used in the pilot study, along with their original and AI-

generated managerial responses. The full list of reviews and responses used in Study 1 can be found in Appendix A. The study employed a between-subject design with 202 online participants from the US recruited via Prolific, and each of them was compensated with USD 1.5 for participating this 10-minute-long survey. Out of these, 180 participants who passed an attention check were included in the final analysis (60.2% female). Participants were randomly assigned to one of two conditions: the managerial response was written by the human managers (S) or generated by ChatGPT (C). Each participant was exposed to two sets of materials randomly selected from the total 10 cases—each set comprising a negative review, a managerial response, purchase intention measurement questions, and an open-ended question asking participants to justify their ratings. All other instructions and procedures mirrored those in the pilot study.

2.4.2. Quantitative Results

In this study, 180 participants were each presented with two distinct scenarios involving negative reviews and managerial responses, resulting in a total of 359 unique cases (one participant only responded to one set out of two, and I still included it in the analysis as each case is independent). To simplify the analyses, I first averaged WTPB1 and WTPB2 into WTPB, and WTPA1 and WTPA2 into WTPA, supported by strong scale reliabilities ($\alpha_{WTPB} = 0.924$, $\alpha_{WTPA} = 0.963$). Paired-samples t-tests revealed that, generally, managerial responses significantly enhanced purchase intentions ($M_{WTPB} = 2.97$, $SD_{WTPB} = 1.62$; $M_{WTPA} = 4.13$, $SD_{WTPA} = 1.70$, $t(359) = -16.1$, $p < .001$).

Comparing the human manager and ChatGPT conditions, there was no significant difference in initial purchase intentions between the groups when exposed only to negative reviews ($M_{WTPB\text{-}C} = 2.93$, $SD_{WTPB\text{-}C} = 1.63$, $M_{WTPB\text{-}S} = 3.01$, $SD_{WTPB\text{-}S} = 1.62$, $t(357) = -0.495$, $p = .621$, ns), indicating a similar reaction to the negative reviews from both groups. However, after reading a response, the ChatGPT group demonstrated significantly higher purchase intentions ($M_{WTPA\text{-}C} = 4.38$, $SD_{WTPA\text{-}C} = 1.72$, $M_{WTPA\text{-}S} = 3.87$, $SD_{WTPA\text{-}S} = 1.63$, $t(357) = 2.856$, $p = .005$) compared to the group receiving original responses from the human managers. All the above findings successfully replicated the results from the Pilot Study with a different sample.

In evaluating the change in purchase intention before and after exposure to managerial responses using Repeated Measures ANOVA, the significant interaction effect indicates a greater increase in purchase intention under the ChatGPT condition ($F(1, 357) = 18.6$, $p < .001$), illustrated in Figure 2.1. This finding reconfirms that ChatGPT is more effective in crafting managerial responses that enhance purchase intention, particularly in scenarios involving brief and general negative reviews without offers of compensation.

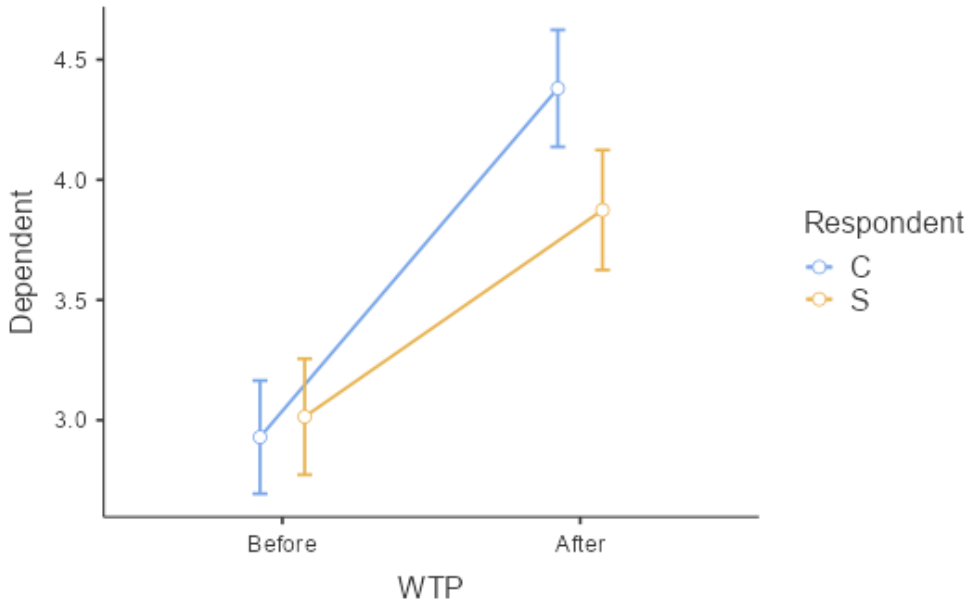

Figure 2.1. Repeated Measures ANOVA Results Comparing the Changes in Purchase Intention (Study 1)

2.4.3. Qualitative Analyses on Open-ended Survey Responses

Coding method

For the qualitative open-ended responses, I adopted the semantic coding method (Holsti, 1969), continuously cross-referencing with literature on managerial response effectiveness and AI customer service effectiveness to ensure that every aspect mentioned in each short answer could be referred to as a concept discussed in existing publications. If a respondent identified a factor as present in the managerial response, it was coded as “1”. If a factor was identified as lacking, it was coded as “-1”. If the factor was not mentioned, it was coded as “0”. Definitions and examples of all factors identified in the survey answers can be found in Table 2.1.

Table 2.1. Qualitative Coding Scheme

Factor	Citation	Definition	1	-1
Ability to solve	Yhee et al., 2023; Lopes et al., 2023	The extent to which a trustee (focusing on personal action) has enough competence and skill to influence.	The manager is perceived to have the ability to solve the issue.	The manager is perceived NOT to have the ability to solve the issue.
		Examples from quotes	*"It sounds like the manager is aware of the problems and taking steps to ensure they do not happen again."*	*"Some things a manager can not fix like a horrible experience…"*
Ability to change	Zhao & Su, 2020	The extent to which the situation (focusing on the restaurant conditions) can be easily improved.	Participant perceives the situation can be changed easily	Participant perceives the situation can NOT be changed easily
		Examples from quotes	*"I believe the poor experience was a one-time thing and not a common occurrence."*	*"…I'm not convinced the restaurant fixed the problem."*
Effort	Gong et al., 2022; Sheng et al., 2021; Zhao et al., 2020; Lopes et al., 2023; Sheng, 2019	Time and energy invested in creating marketing communications and interacting with consumers	The manager is perceived to put effort into resolving the case.	The manager is perceived to put NO effort into resolving the case.
		Examples from quotes	*"The manager took time to respond…"*	*"The email seems automated…"*
Personalization	Jin et al., 2023; Roozen & Raedts., 2018; Palese et al., 2021; Wang et al., 2020; Herhausen et al., 2019	communications tailored to the individual level, addressing particular issues raised in the review	The response is perceived to be personalized.	The response is perceived to be templated.
		Examples from quotes	*"…it is less "I copied and pasted this from our PR manual"…"*	*"Response is not catered to the review…"*
Formality	Gong et al., 2022	Officially structured professional communication content	Communication style is formal or professional	Communication style is informal or unprofessional
		Examples from quotes	*"…it was responded to very professionally by management"*	*"…the manager's response was a tad poorly formatted, showing a lack of experience online"*

(Low) Severity	Surachartkumtonkun & Ross, 2021	assessed when individuals (i.e., a third-party customer) imagine what the target person (i.e., a reviewer) would have felt like if a different situation (i.e., a desirable one) had occurred.	Situation described in review is not a big matter	Situation described is severe
		Examples from quotes	*"I do not feel like the offense is that serious, especially since waiters should be tipped in the first place"*	*"The offense was severe to the point I don't think I would ever visit the restaurant"*
Length	Sheng et al., 2021; Lopes et al., 2023; Sheng, 2019	Word count in a response	The response is perceived to be lengthy and sufficient.	The response is perceived to be simple and short.
		Examples from quotes	*"That's a detailed response to a small issue."*	*"Nothing justifies this act and response was very inadequate."*
Sincerity / Care	Huang & Ha, 2020; Xia, 2013	how well the retailer shows its true intentions when addressing consumer complaints	Participant perceives the message as sincere/caring (or similar words)	Participant explicates that the response is not sincere/caring (or similar words).
		Examples from quotes	*"I feel that the restaurant's response is genuinely trying to find out what happened to find a resolution with the customers."*	*"I do not believe that they are sincere and would not be willing to revisit."*
Concreteness	Packard & Berger, 2021; Jiménez-Barreto et al., 2023	refers to using words that help create specific mental images about tangible entities while decoding the information transmitted	The response is firm and definitive, which helps participants understand the situation	The response is generic and hollow, which does not help participants understand the situation
		Examples from quotes	*N/A*	*"This email uses a lot of hollow phrases and very generic platitudes, … Something concrete, not "We're addressing this." "*

Decision Tree Analysis - Methodology

The purpose of this study is to specifically explore what factors in a managerial response under both human and AI conditions lead to an increase in purchase intention, and which factors, when lacking, render the response ineffective. Given that AI-generated managerial responses are generally more effective, it is particularly interesting to examine the conditions under which GenAI may fail.

To quantitatively analyze the coding results and achieve these study goals, I created a binary variable using WTPA deducted by WTPB. If the result was greater than zero, it was coded as 1, indicating an increase in purchase intention after reading the response. If the result was zero or less, it was coded as 0, indicating the response was not effective. To maximize the richness of qualitative information, all participants were included in the qualitative analysis, even if they did not pass the attention check. If any answer did not make sense, I coded every factor as 0, which does not affect the analysis result as I only looked at 1s and -1s. As a result, 132 responses written by humans successfully increased purchase intention, while 67 did not. Under the ChatGPT condition, these numbers were 151 and 52, respectively, suggesting that ChatGPT generally performs better in increasing purchase intention.

In the next step, I adopted the Decision Tree method. Decision trees are widely utilized in classification tasks due to their interpretability and effectiveness in handling coded data (Loh, 2011). These models provide intuitive, visual representations of decision-making processes, which are crucial in identifying the factors most impactful in influencing outcomes. The data analysis was conducted using the scikit-learn library in Python, known for its comprehensive suite of machine learning tools. The *DecisionTreeClassifier* package, using Gini impurity as the criterion, was employed to develop models that identify both positive and negative influences on purchase intentions. To be specific, under each of the three conditions (overall, Human-only, AI-only), two decision tree models were developed: one to identify factors (coded as 1) that lead to an increase in purchase intentions and another to identify factors perceived as lacking (coded as -1) when purchase intentions do not increase.

To address potential overfitting issues, the following steps were implemented:1) Max Depth Setting: The trees were limited to a maximum depth of five; 2) Minimum Samples Split: Nodes were required to have at least 20 samples before considering a

split; 3) Cross-Validation: 10-fold cross-validation was used to ensure that the model's performance was robust across different subsets of the data (James et al., 2013).

Decision Tree Analysis - Results

Under the overall condition, including all cases in the dataset, the positive factors model achieved an accuracy of 70.65%, and the negative factors model had an accuracy of 72.89%, both considered adequate for exploratory analyses (Rodrigo et al., 2021). The models reliably capture the primary influences on consumer decisions regarding online managerial responses. Key factors positively impacting purchase intention include the ability to change (26.42%), sincerity (22.61%), problem-solving ability (18.07%), effort (17.22%), and formality (10.75%). Conversely, lacking factors when purchase intentions did not increase are sincerity (34.09%), personalization (24.56%), low severity (17.42%), problem-solving ability (16.46%), and ability to change (7.48%).

For the human condition subset, the positive factors model had an accuracy of 68.34%, and the negative factors model had an accuracy of 67.34%, both adequate (Rodrigo et al., 2021). Important factors for increasing purchase intention are the ability to change (34.89%), effort (27.58%), formality (17.92%), and sincerity (13.70%). Lacking factors when purchase intentions did not increase include low severity (43.91%), problem-solving ability (29.40%), and ability to change (26.69%).

In the AI condition subset, the positive factors model achieved an accuracy of 78.33%, and the negative factors model had an accuracy of 79.23%. Influential factors for increasing purchase intention are sincerity (53.29%), low severity (22.71%), ability to change (10.02%), and care (9.58%). Lacking factors when purchase intentions did not increase include sincerity (80.21%), concreteness (9.44%), and low severity (7.49%).

Among the factors leading to the ineffectiveness of AI-generated managerial responses, the severity of the complaint is beyond marketers' control when writing the response, and sincerity is a subjective perception. Concreteness, however, is an interesting factor worth further exploration. No study participants identified a response as effective due to concreteness; they only noted the lack of it, which was exclusively under the ChatGPT condition.

Examining the reviews whose AI-generated responses were regarded as “not concrete,” seven out of eight complaints were about the processes and procedures in the restaurant (e.g., waiting time, review #2, 3, 4), and the last one was about food quality. According to Ravichandran & Deng (2023), there are three distinct categories of unfairness in customer complaints:

- *Distributive Unfairness* involves how customers perceive the fairness of what they receive compared to what they give. When customers think they are not getting what they deserve for their money, time, or effort, they feel this kind of unfairness. For example, if a customer pays for a high-quality service but gets something much less valuable, this is seen as distributive unfairness.
- *Procedural Unfairness* is about the fairness of the processes used by companies. Customers feel procedural unfairness if they think the methods for making decisions or handling complaints are unfair, like being too slow or not clear. An example is a customer feeling unfairly treated by a restaurant's long and confusing ordering process.
- *Interactional Unfairness* concerns how customers are treated personally by staff during service interactions. If customers feel they are treated poorly, without respect or empathy, they experience interactional unfairness. For instance, if a staff member is rude or dismissive when responding to a customer’s complaint, it is seen as interactional unfairness.

Based on the coding results and the complaint categorization, it appears that concrete communication containing firm and tangible factual information may work better for procedural unfairness, which GenAI may not perform well on given a general and default prompt. On the other hand, for interactive unfairness, concreteness does not seem to be as important, and consumers might be more receptive to the hollow but aesthetically pleasing language generated by AI. For distributive unfairness, when the participant perceives the response as lacking in concreteness, compensation tends to be expected. This pattern aligns with the definition of distributive unfairness but extends beyond the current context and experimental setting, and thus should be interpreted with caution.

Thus, for the next steps, I propose to focus on procedural unfairness only and examine whether there is a way to use GenAI to respond to procedural complaints effectively or not.

2.5. Study 2: Concreteness in Managerial Responses towards Procedural Unfairness – Human vs. ChatGPT

2.5.1. Hypotheses Development

Study 1 revealed that consumers may tend to dislike AI-generated managerial responses to negative online reviews that address procedural unfairness, primarily due to a perceived lack of concreteness. Procedural unfairness, which pertains to issues related to processes and policies, often leads consumers to expect responses that clarify these processes and policies to aid in future interactions with the service provider. Managerial responses that rely solely on emotional appeals, such as apologies, are often inadequate in addressing procedural unfairness. These situations typically require logical explanations that clarify why the service failure occurred and concretely outline the steps the firm is taking to rectify the issue (Lee et al., 2018; Ravichandran & Deng, 2023).

Generative AI, such as ChatGPT, inherently lacks specific knowledge of the processes and policies of individual businesses and can therefore only generate responses that involve apologies and vague promises of improvement. This limitation stems from the nature of GenAI models, which function as "stochastic parrots"—a term coined by Bender et al. (2021) to describe how these models generate text based purely on statistical patterns within the data they have been trained on, rather than a true understanding of the content (Hannigan et al., 2024). As a result, GenAI-generated responses often lack the depth and specificity required for concrete communication, making them less impactful than those written by a human who understands the specific situation and can provide tangible facts to explain the context.

Before designing Study 2, to examine whether GenAI outputs generally lack concreteness, I employed a text analysis method from Packard & Berger (2021), utilizing the concreteness dictionary developed by Brysbaert et al. (2014) within the Linguistic Inquiry and Word Count (LIWC) software (Boyd et al., 2022). In this context, words referring to more tangible and specific objects, materials, people, processes, or relationships are perceived as more concrete, while those referring to abstract concepts are seen as less concrete (Packard & Berger, 2021). An independent sample t-test on the review responses used in the pilot study revealed that human-written responses

have significantly higher concreteness indices than those generated by ChatGPT (M_S = 253, M_C = 223, t = -7.07, p < .001). This result confirms that ChatGPT outputs indeed lack concreteness, thus providing a strong foundation for Study 2. Notably, linguistic concreteness may differ from consumers' perceived concreteness. Thus, I also measured perceived concreteness in Study 2. The following hypotheses were be tested in Study 2:

> **H2:** In crafting managerial responses to negative reviews that complain about procedural unfairness, ChatGPT is less effective than human managers in enhancing potential customers' purchase intention.
>
> **H3:** The higher effectiveness of human managers (compared to ChatGPT) in crafting managerial responses to negative reviews that complain about procedural unfairness is mediated by the concreteness of the content.

2.5.2. Study Design & Data Collection

For Study 2, I recruited 137 undergraduate students who participated in exchange for course credits. The review selection largely followed the criteria used in the Pilot Study and Study 1, with some key modifications. First, to enhance the generalizability of the findings, the reviews were sourced from a different platform—Yelp. This action aimed at using a data triangulation approach to mitigate the risk of platform-specific biases influencing the results, thereby increasing the robustness and applicability of the conclusions across various contexts (Carter, 2014). Furthermore, the reviews selected specifically addressed procedural unfairness, focusing on complaints related to restaurant operational processes and policies.

Another critical aspect of the review selection process in Study 2 was the inclusion of only those managerial responses that were manually composed by the restaurants' managers and not templated. Although the quality of these responses was not formally assessed during the selection process, excluding templated responses was essential. Although many managers rely on templated responses as an efficient way to address complaints (Deng & Ravichandran, 2023; Wang & Chaudhry, 2018), using such responses could obscure the true ability of human managers to tailor their communications to the nuances of procedural unfairness. By selecting genuine, non-templated responses that directly engage with the issues raised in the reviews, this study is better positioned to compare the inherent capability of human managers and

ChatGPT in crafting effective responses towards complaints addressing procedural unfairness.

Two reviews used in Study 2, along with their corresponding responses—whether written by human managers or generated by ChatGPT—are presented in Appendix B. As in Study 1, each participant was randomly assigned to either the human manager group (S) or the ChatGPT group (C) and viewed both reviews accompanied by the respective managerial responses. The order in which the review-response pairs were presented was fully randomized to control for any potential sequence effects.

2.5.3. Survey Questions

Table 2.2 displays all the measurement items that I used in Study 2.

Table 2.2. Measurement Items in Study 2

Factor	Measurement Item(s)	Scale	Variable Name
Purchase Intention, before and after reading the response (Maxham & Netemeyer, 2002)	- In the future, I would dine at this restaurant. - If I was in the mood for the kind of food they serve, I would visit this restaurant.	1 = Strongly Disagree; 7 = Strongly Agree	WTPB WTPA
Perceived Concreteness (Packard & Berger, 2021)	- The response from the restaurant was concrete. By concrete, we mean it used words that describe something in a more precise, specific, or clear manner.	1 = Not at all Concrete; 7 = Very Much Concrete	Concreteness
Perceived Ability to Solve (Yhee et al., 2023)	- The person in charge of this restaurant has the required skills to handle the issue described in the negative review. - The person in charge of this restaurant has the required knowledge to handle the issue described in the negative review. - The person in charge of this restaurant has the required expertise to handle the issue described in the negative review.	1 = Strongly Disagree; 7 = Strongly Agree	ABS

Perceived Ability to Change, adapted from Zhao & Su (2020)	- The problem mentioned in the negative review is very likely to be permanent. (reversely coded) - The problem mentioned in the negative review is very likely to be solved soon. - The problem mentioned in the negative review is very likely to occur frequently. (reversely coded)	1 = Strongly Disagree; 7 = Strongly Agree	ABC
Perceived Effort, adapted from Mohr & Bitner (1995)	- The restaurant manager exerted a lot of energy in writing the response to this review. - The restaurant manager did not spend much time writing the response to this review. (reversely coded) - The restaurant manager did not try very hard to write the response to this review. (reversely coded) - The restaurant manager put a lot of effort into writing the response to this review.	1 = Strongly Disagree; 7 = Strongly Agree	PE
Perceived Sincerity (MacKenzie & Lutz, 1989)	- The response from the restaurant was sincere. - The response from the restaurant was dishonest. (reversely coded) - The response from the restaurant was credible. - The response from the restaurant was not convincing. (reversely coded)	1 = Strongly Disagree; 7 = Strongly Agree	Sincerity
Perceived Care (Xia, 2013)	- The response from this restaurant shows care to its customers.	1 = Strongly Disagree; 7 = Strongly Agree	Care

At the end of the study, participants were also asked to write down their guesses of the study's purpose, which helped determine whether any participants suspect the study's focus on GenAI.

2.5.4. Results

Of the 137 participants, only 114 (48.2% female) passed the attention check. Consequently, analyses were conducted solely on data from these participants, and incomplete responses were dropped during statistical comparisons. Given the strong internal consistency of the multi-item scales (αWTPB = 0.819, αWTPA = 0.929, αABS = 0.915, αABC = 0.824, αPE = 0.875, αSincerity = 0.797), composite scores were computed for each variable to streamline and simplify subsequent analyses. Notably, no participant recognized that the study pertained to GenAI.

Consistent with earlier studies, paired-samples t-tests revealed that managerial responses significantly elevated purchase intentions (M_{WTPB} = 4.02, SD_{WTPB} = 1.66, M_{WTPA} = 5.04, SD_{WTPA} = 1.49, t (213) = -10.4, p < .001). Moreover, participants' initial reactions to negative reviews were indifferent between the ChatGPT (C) and human manager (S) conditions (M_{WTPB-C} = 4.03, SD_{WTPB-C} = 1.70, M_{WTPB-S} = 3.97, SD_{WTPB-S} = 1.62, t(214) = 0.266, p = .791, ns). However, in contrast to Study 1, when negative reviews specifically addressed procedural unfairness, the pattern reversed: the human-manager group exhibited significantly higher purchase intentions than the ChatGPT group (M_{WTPA-C} = 4.79, SD_{WTPA-C} = 1.53, M_{WTPA-S} = 5.29, SD_{WTPA-S} = 1.41, t (209) = -2.446, p = .015).

A repeated measures ANOVA was conducted to assess changes in purchase intention following exposure to managerial responses. The analysis yielded a significant interaction effect (F(1, 221) = 9.27, p = .003), illustrated in Figure 2.2, indicating a greater increase in purchase intention under the human-manager condition—again, a reversal from the pattern observed in Study 1. These results support Hypothesis 2, suggesting that ChatGPT is less effective than human managers in enhancing potential customers' purchase intentions when responding to complaints centered on procedural unfairness.

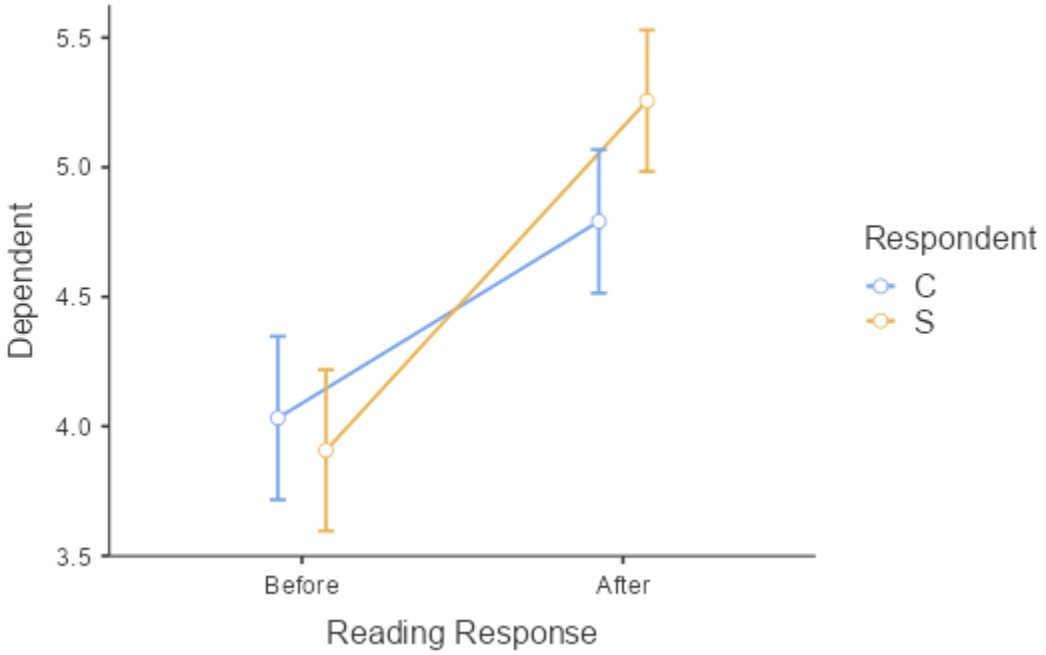

Figure 2.2. Repeated Measures ANOVA Results Comparing the Changes in Purchase Intention (Study 2)

To test Hypothesis 3 and to examine the mediation effect of concreteness, I first examined whether there is significant difference between the perceived concreteness levels of responses written by human managers and ChatGPT. Independent samples t-test results show that participants on average perceive the responses written by human managers as more concrete compared to those written by ChatGPT ($M_{Concreteness\text{-}C}$ = 4.49, $SD_{Concreteness\text{-}C}$ = 1.70, $M_{Concreteness\text{-}S}$ = 5.27, $SD_{Concreteness\text{-}S}$ = 1.55, t(211) = -3.53, p < .001). I also compared all other potential confounding variables, *including Perceived Ability to Solve (ABS), Perceived Ability to Change (ABC), Perceived Effort (PE), Perceived Sincerity (Sincerity), and Perceived Care (Care),* across two author groups and did not find any significant differences (p_{ABS} = 0.112, p_{ABC} = 0.909, p_{PE} = 0.099, $p_{Sincerity}$ = 0.372, p_{Care} = 0.206, all ns).

I next ran a Multiple Mediation Model (Preacher & Hayes, 2008) to examine whether perceived concreteness mediates the relationship between response type and the change in purchase intention for reviews addressing procedural unfairness. To quantify this change, a difference score (WTPC) was computed by subtracting pre-response purchase intention (WTPB) from post-response purchase intention (WTPA). The independent variable was the author group (Respondent; coded as a dichotomous factor), and the mediators included Perceived Concreteness, Perceived Ability to Solve, Perceived Ability to Change, Perceived Effort, Perceived Sincerity, and Perceived Care.

The analysis revealed that the path from author group to Perceived Concreteness was significant (b = 0.79, SE = 0.22, 95% CI [0.35, 1.22], β = 0.235, z = 3.505, p < .001). This indicates that human managers are associated with higher perceived concreteness compared to ChatGPT under the given conditions. In addition, the path from Perceived Concreteness to WTPC was significant (b = 0.28, SE = 0.06, 95% CI [0.16, 0.41], β = 0.330, z = 4.426, p < .001), demonstrating that higher concreteness is linked to a greater increase in purchase intention. Consequently, the indirect effect of author group on WTPC via Perceived Concreteness was significant, b = 0.22, SE = 0.08, 95% CI [0.06, 0.38], β = 0.077, z = 2.748, p = .006, suggesting that increased concreteness accounts for the enhanced purchase intention associated with human managers.

In contrast, the indirect effects through the other mediators were not significant (p_{ABS} = 0.112, p_{ABC} = 0.909, p_{PE} = 0.099, $p_{Sincerity}$ = 0.372, p_{Care} = 0.206, all ns), indicating

that these variables do not contribute to the relationship between response type and the change in purchase intention. Thus, among all the mediators, only perceived concreteness significantly mediated the relationship between response type and the change in purchase intention for reviews addressing procedural unfairness. These findings support Hypothesis 3 by demonstrating that the superior effectiveness of human managers relative to ChatGPT in enhancing purchase intention—specifically for reviews focused on procedural unfairness—is explained by their ability to produce responses that are perceived as more concrete.

2.6. Study 3: Concreteness in Managerial Responses towards Procedural Unfairness – Human vs. ChatGPT vs. Trained ChatGPT

2.6.1. Hypotheses Development

The findings from Study 2 underscore a critical limitation in default GenAI outputs: responses generated by ChatGPT lack the depth and specificity required for concrete communication. In our investigation, human managers produced responses that were perceived as significantly more concrete, which in turn was linked to higher increases in purchase intention when addressing procedural unfairness. This deficiency in concreteness among ChatGPT-generated responses suggests that, in its default state, the model may not fully capture the nuanced, context-specific details that consumers expect when service processes and policies are at issue.

Recent advancements in fine-tuning GenAI models indicate that incorporating domain-specific, concrete information—such as detailed descriptions of processes and policies—can enhance a model's ability to generate contextually accurate and informative responses. Building on this insight, we propose that training ChatGPT with such concrete information may improve its capacity to craft managerial responses that effectively address complaints centered on procedural unfairness.

Study 3 aims to test that the trained ChatGPT performs better than the untrained version, given its incorporation of concrete information about processes and policies. However, whether the trained ChatGPT can outperform human managers in this scenario remains uncertain. Nevertheless, considering the efficiency of GenAI in

producing managerial responses, even if its performance does not surpass that of human managers, it could still represent a valuable option for businesses—as long as it does not perform significantly worse.

> **H4:** Training ChatGPT with concrete information enables it to be more effective in crafting managerial responses to negative reviews that complain about procedural unfairness than the default, untrained ChatGPT.

2.6.2. Study Design and Data Collection

For Study 3, 361 undergraduate students participated for course credit. In line with the procedures established in Study 2, the review selection criteria remained consistent except for that reviews were now sourced from OpenTable. As before, only reviews addressing procedural unfairness—specifically complaints concerning restaurant operational processes and policies—were selected. Additionally, to ensure a valid comparison of response quality, only genuine, manually composed responses by restaurant managers (i.e., non-templated responses) were included.

To test Hypothesis 4, Study 3 introduced a third condition. Participants were randomly assigned to one of three groups: responses written by human managers (S), default ChatGPT-generated responses (C), or ChatGPT-generated responses produced after being trained with additional concrete, domain-specific information (CT). In the CT condition, the training process involved integrating specific details about restaurant processes and policies into ChatGPT's knowledge base, thereby enabling it to generate more concrete and contextually accurate responses. Each participant was presented with two reviews, each accompanied by a response generated under one of the three conditions. The order of the review-response pairs was randomized, mirroring the approach used in Study 2. All measurements and survey items employed were identical to those used in Study 2.

ChatGPT Training Process

To train ChatGPT, I wrote a prompt to create a customized GPT model, a tailored version of this GenAI tool designed to perform specific tasks based on user-provided instructions (OpenAI, 2024). The prompt was as follows: "I would like to make a customer service assistant in a restaurant who responds to customers' negative reviews

on behalf of the restaurant management. In each inquiry, I sent two prompts. The first one includes the restaurant's policies, and the second one includes the negative review. Please draw necessary information from the policies to write a response to the review, without offering any compensation. The response should primarily focus on concretely explaining the correct policies, procedures, or other factual information to the reviewer, as well as potential customers who are reading this review."

After I sent the prompt, this customized GPT kept the following information as its configuration: “You are a customer service staff member representing a restaurant's management team. Your role is to respond to customers' negative reviews in a **professional and empathetic** manner. When responding, draw necessary information from the provided restaurant policies to address the concerns raised in the review. Your responses should aim to **acknowledge the customer's experience, reference relevant policies, and offer resolutions within those policies**. Avoid offering any form of compensation. Keep the tone **courteous and understanding**. The response should primarily focus on concretely explaining the correct policies, procedures, or other factual information to the reviewer, as well as potential customers who are reading this review. First, ask for the review after the user provides the policy. Then, only provide a response to that review.”

This final configuration, built into the customized chatbot, includes some instructions (bolded) not provided by me. During the training process, the customization tool asked me to specify the writing style and tone, which I left to the tool’s default settings since my primary interest is in the incorporation of actual restaurant policies. Other aspects of the response, including style and tone, were left to the ChatGPT's default approach. An example of the training results is provided in Appendix C, and the reviews and responses used in Study 3 are listed in Appendix D.

2.6.3. Results

Of the 361 participants, a total of 317 (50.2% female) successfully passed the attention check, and as a result, only their data were analyzed. Incomplete responses were dropped if affecting statistical comparisons. Given the high internal reliability of the multi-item scales (αWTPB = 0.870, αWTPA = 0.910, αABS = 0.958, αABC = 0.735, αPE = 0.891, αSincerity = 0.814), composite scores were created for each variable to

enhance the clarity and efficiency of subsequent analyses. Similar to previous studies, no participants identified that the research was related to GenAI.

Aligned with prior findings, paired-samples t-tests demonstrated that managerial responses significantly increased purchase intention (M_{WTPB} = 3.95, SD_{WTPB} = 1.45, M_{WTPA} = 5.16, SD_{WTPA} = 1.27, t(593) = -25.2, p < .001). Additionally, a one-way ANOVA was conducted to evaluate differences in purchase intention across three conditions—default ChatGPT (C), trained ChatGPT (CT), and human managers (S)—both before and after reading the responses. Consistent with earlier research, initial purchase intention scores did not vary significantly between groups ($M_{WTPB\text{-}C}$ = 3.89, $SD_{WTPB\text{-}C}$ = 1.44, $M_{WTPB\text{-}CT}$ = 4.04, $SD_{WTPB\text{-}CT}$ = 1.47, $M_{WTPB\text{-}S}$ = 3.93, $SD_{WTPB\text{-}S}$ = 1.45, F(2, 293) = 0.598, p=0.551, ns). In contrast, post-response purchase intentions showed significant variation based on condition ($M_{WTPA\text{-}C}$ = 4.43, $SD_{WTPA\text{-}C}$ = 1.21, $M_{WTPA\text{-}CT}$ = 5.75, $SD_{WTPA\text{-}CT}$ = 1.14, $M_{WTPA\text{-}S}$ = 5.31, $SD_{WTPA\text{-}S}$ = 1.08, F(2.393) = 65.331, p < .001). Post-hoc comparisons indicated that both the trained ChatGPT and human manager conditions led to significantly higher purchase intentions than the default ChatGPT condition (mean differences: 1.32 and 0.875, respectively; both ps < .001), and the trained ChatGPT group also differed significantly from the human manager group (mean difference = 0.447, p < .001).

A repeated measures ANOVA was conducted to examine the change in purchase intention following managerial responses. A significant interaction effect was observed (F(2, 591) = 64.5, p < .001), as illustrated in Figure 2.3, indicating that the extent of purchase intention increase varied across conditions, with the trained ChatGPT group exhibiting the most substantial improvement. This result partially supports Hypothesis 4, suggesting that training ChatGPT can enhance the effectiveness of its managerial responses in boosting purchase intentions after negative reviews involving procedural unfairness.

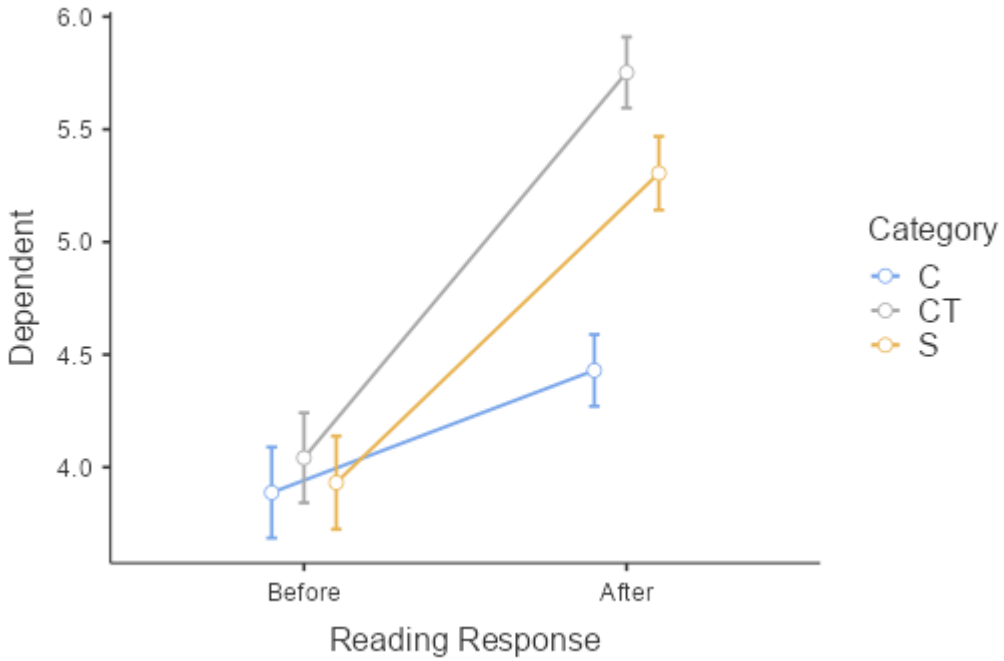

Figure 2.3. Repeated Measures ANOVA Results Comparing the Changes in Purchase Intention (Study 3)

To determine whether the improvement in purchase intention was mediated by perceived concreteness, I first assessed differences in perceived concreteness levels among responses written by human managers, default ChatGPT, and trained ChatGPT. The one-way ANOVA revealed a significant effect across the three conditions ($M_{Concreteness\text{-}C}$ = 3.820, $SD_{Concreteness\text{-}C}$ = 1.142, $M_{Concreteness\text{-}CT}$ = 5.490, $SD_{Concreteness\text{-}CT}$ = 0.857, $M_{Concreteness\text{-}S}$ = 5.263, $SD_{Concreteness\text{-}S}$ = 1.274, $F(2, 377) = 143.8$, $p < .001$). Post-hoc analyses confirmed that responses generated by default ChatGPT were perceived as significantly less concrete than those produced by trained ChatGPT and human managers (mean differences = –1.67 and –1.443, respectively; both ps < .001). However, the concreteness levels between trained ChatGPT and human managers did not differ significantly (mean difference = 0.227, p = .100, ns).

To further investigate mediation effects, a multiple mediation model (Preacher & Hayes, 2008) was employed to assess whether perceived concreteness mediates the relationship between response type and the change in purchase intention (WTPC, calculated as in Study 2) for reviews addressing procedural unfairness. The independent variable was the response author (coded as a dichotomous variable), while mediators included Perceived Concreteness, Perceived Ability to Solve, Perceived Ability to Change, Perceived Effort, Perceived Sincerity, and Perceived Care.

Two pairwise analyses were performed. The first analysis (S-C) replicated Study 2 and assessed whether perceived concreteness mediated the difference in purchase intention between the human manager (S) and default ChatGPT (C) conditions. The second analysis (CT-C) tested Hypothesis 4, evaluating whether training ChatGPT improves purchase intention through enhanced perceived concreteness.

For the S-C comparison, the indirect effect of perceived concreteness was significant (b = 0.222, SE = 0.081, 95% CI [0.064, 0.380], β= 0.090, z = 2.752, p = .006), indicating that the difference in purchase intention improvement between human managers and default ChatGPT was mediated by perceived concreteness. The indirect effects of other mediators were not statistically significant (p_{ABS} = .820, p_{ABC} = .617, p_{PE} = .662, $p_{Sincerity}$ = .592, p_{Care} = .544, all ns). Similarly, for the CT-C comparison, the indirect effect via perceived concreteness was significant (b = 0.249, SE = 0.090, 95% CI [0.073, 0.425], β= 0.103, z = 2.767, p = .006), suggesting that the difference in purchase intention between trained and default ChatGPT was also mediated by perceived concreteness. Conversely, other mediators did not show significant indirect effects (p_{ABS} = 0.166, p_{ABC} = 0.625, p_{PE} = 0.120, $p_{Sincerity}$ = 0.098, p_{Care} = 0.619, all ns).

These findings highlight that among all examined mediators, perceived concreteness was the only factor that significantly mediated the relationship between response type and changes in purchase intention for reviews addressing procedural unfairness. This supports both Hypothesis 3 and Hypothesis 4, demonstrating that the superior ability of trained ChatGPT and human managers to enhance purchase intention relative to default ChatGPT stems primarily from the increased concreteness of their responses.

Notably, in contrast to Study 2, where all factors except perceived concreteness—namely, Perceived Ability to Solve, Perceived Ability to Change, Perceived Effort, Perceived Sincerity, and Perceived Care—remained consistent across human and ChatGPT conditions, Study 3 reveals significant variations across all factors through One-way ANOVA analyses. While some degree of variability may exist among responses across the two studies, this does not undermine the finding that perceived concreteness remains the sole mediator of the relationship between respondent type and the increase in purchase intention following exposure to managerial responses addressing procedural complaints. Rather, this reinforces the robustness of the results,

as it suggests that perceived concreteness consistently plays a central role in driving purchase intention regardless of other fluctuations in response characteristics.

2.7. General Discussion and Future Research Directions

This chapter explored the potential of Generative AI (GenAI), specifically ChatGPT, in crafting managerial responses to negative online reviews. The findings from Study 1 suggest that ChatGPT generally outperforms human managers when responding to straightforward, brief complaints, aligning with the notion that AI can enhance marketing communications by automating processes and delivering personalized content. However, the subsequent Studies 2 and 3 reveal important limitations of ChatGPT in handling complaints related to procedural unfairness, highlighting the need for human oversight to ensure the effectiveness of AI-generated responses. These insights are particularly relevant in the context of eWOM, where the credibility and usefulness of information significantly impact consumer behavior. By integrating human training with GenAI models, businesses can improve the concreteness and effectiveness of AI-generated responses, thereby enhancing customer satisfaction and trust in eWOM platforms. These findings highlight the critical need for human oversight in AI deployment, ensuring that GenAI tools are fine-tuned to deliver responses that meet customer expectations and enhance overall satisfaction. Theoretically, this research contributes to the broader discourse on electronic word-of-mouth (eWOM) and marketing communication, as well as enhancing the understanding of how GenAI and human collaboration can be optimized in customer service contexts. By investigating the effectiveness of GenAI in responding to various types of complaints, particularly procedural unfairness, this study expands on existing literature by identifying the conditions under which AI-generated responses may fall short and how human intervention can mitigate these shortcomings. The inclusion of qualitative analyses in Study 1 also offers deeper insights into the specific factors that influence consumer perceptions of AI-generated content, providing a richer understanding of how and why certain responses succeed or fail.

The findings of this research also have significant practical implications for businesses considering the integration of GenAI into their customer service strategies. While GenAI has demonstrated its effectiveness in handling routine, straightforward complaints, its performance may falter in situations where customers expect more

informative and concrete responses to complaints about processes and policies. Study 3, which involves human training of GenAI models, directly aligns with the *Collaboration* factor of the CARE framework introduced in Chapter 1. This approach highlights the need for a collaborative strategy, where human managers enhance GenAI's capabilities by providing the context-specific information necessary for crafting responses that meet customer expectations. Even though Study 3 does not provide strong evidence that trained ChatGPT outperforms human managers in addressing procedural complaints, its high efficiency in generating responses still makes it an attractive option for businesses seeking to optimize their customer service workflows. By combining GenAI's efficiency with human expertise, businesses can optimize the quality of customer interactions, ensuring that AI-generated responses are both relevant and effective. This research underscores the critical importance of maintaining human oversight in GenAI deployment to ensure that customer service responses align with both customer expectations and business goals.

This research, while insightful, is subject to several limitations. The relatively small sample size, driven by budget constraints, restricts the generalizability of the findings. The sample, drawn from undergraduate students and Prolific participants, may not fully represent the broader population. Additionally, the number of reviews used was limited and selected based on specific criteria—short, text-only complaints without accompanying multimedia or complex issues. This narrow focus may limit the applicability of the findings to more diverse real-world scenarios.

Future research should examine whether consumers exhibit significant attitudinal shifts when explicitly informed that a managerial response was generated by GenAI or a trained GenAI model. Understanding how AI disclosure impacts consumer trust, perceived fairness, and purchasing decisions will provide valuable insights into the transparency and acceptance of AI-driven customer service interactions. Additionally, future research should seek to address limitations by expanding the sample size and including a more diverse participant pool, thereby enhancing the generalizability of the findings. A broader range of participants from different demographic backgrounds and industries may offer more comprehensive insights into how GenAI responses are perceived in various contexts. There is also a need to explore the effectiveness of GenAI in responding to a wider variety of complaint types, including those that are more complex. While the current study primarily focused on procedural unfairness complaints,

future studies could investigate AI's performance in handling highly nuanced or emotionally charged customer grievances.

Thus, my third chapter in this dissertation aims to scale up these experiments to broadly explore the linguistic differences between AI- and human-generated content across a larger and more varied dataset. This will enable scholars and practitioners to gain a deeper understanding of the specific ways in which GenAI differs from human writing, particularly about factors that influence the effectiveness of managerial responses.

References

Amos, C., & Zhang, L. (2024). Consumer reactions to perceived undisclosed ChatGPT usage in an online review context. *Telematics and Informatics*, *93*, 102163.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*, 10.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods*, *46*, 904-911.

Carter, N. (2014). The use of triangulation in qualitative research. *Number 5/September 2014*, *41*(5), 545-547.

Chen, B., Wu, Z., & Zhao, R. (2023). From fiction to fact: the growing role of generative AI in business and finance. *Journal of Chinese Economic and Business Studies*, *21*(4), 471-496.

Darani, M. M., Mirahmad, H., Raoofpanah, I., & Groening, C. (2023). Managerial responses to online communication: The role of mimicry in affecting third-party observers' purchase intentions. *Journal of Business Research*, *166*, 113979.

Deng, C., & Ravichandran, T. (2024). Managerial response to online positive reviews: Helpful or harmful?. *Information Systems Research*, *35*(4), 1802-1823.

Ferraro, C., Demsar, V., Sands, S., Restrepo, M., and Campbell, C. (2024). The Paradoxes of Generative AI Enabled Customer Service: A Guide for Managers. *Business Horizons, 67*(5), 549-559.

Feng, C. M., Botha, E., & Pitt, L. (2024). From HAL to GenAI: Optimizing chatbot impacts with CARE. *Business Horizons, 67*(5), 537-548.

Gao, C., Zeng, J., Xia, X., Lo, D., Lyu, M. R., & King, I. (2019, November). Automating app review response generation. In *2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE)* (pp. 163-175). IEEE.

Gong, C., Liu, J., Law, R., & Ye, Q. (2022). Exploring the effects of official-structured managerial responses on hotel online popularity. *International Journal of Hospitality Management*, *106*, 103293.

Grewal, D., Satornino, C. B., Davenport, T., & Guha, A. (2024). How generative AI Is shaping the future of marketing. *Journal of the Academy of Marketing Science*, 1-21.

Harkness, L., Robinson, K., Stein, E., & Wu, W. (2023). *How generative AI can boost consumer marketing. McKinsey & Company*.

Hannigan, T.R., McCarthy, I.P. and Spicer, A. (2024). Beware Of Botshit: How to Manage the Epistemic Risks of Generative Chatbots. *Business Horizons, 67*(5), 471-486.

Herhausen, D., Ludwig, S., Grewal, D., Wulf, J., & Schoegel, M. (2019). Detecting, preventing, and mitigating online firestorms in brand communities. *Journal of Marketing*, *83*(3), 1-21.

Hermann, E., & Puntoni, S. (2024). Artificial intelligence and consumer behavior: From predictive to generative AI. *Journal of Business Research*, *180*, 114720.

Holsti, O.R. (1969) *Content Analysis for the Social Sciences and Humanities*. London: Addison Wesley.

Huang, R., & Ha, S. (2020). The effects of warmth-oriented and competence-oriented service recovery messages on observers on online platforms. *Journal of Business Research*, *121*, 616-627.

Huang, M. H., & Rust, R. T. (2024). The caring machine: Feeling AI for customer care. *Journal of Marketing*, 00222429231224748.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. New York: springer.

Jiménez-Barreto, J., Rubio, N., & Molinillo, S. (2023). How chatbot language shapes consumer perceptions: The role of concreteness and shared competence. *Journal of Interactive Marketing*, 10949968231177618.

Jin, W., Chen, Y., Yang, S., Zhou, S., Jiang, H., & Wei, J. (2023). Personalized managerial response and negative inconsistent review helpfulness: The mediating effect of perceived response helpfulness. *Journal of Retailing and Consumer Services*, *74*, 103398.

Kanbach, D. K., Heiduk, L., Blueher, G., Schreiter, M., & Lahmann, A. (2024). The GenAI is out of the bottle: generative artificial intelligence from a business model innovation perspective. *Review of Managerial Science*, *18*(4), 1189-1220.

Koc, E., Hatipoglu, S., Kivrak, O., Celik, C., & Koc, K. (2023). Houston, we have a problem!: The use of ChatGPT in responding to customer complaints. *Technology in Society*, *74*, 102333.

Korzynski, P., Mazurek, G., Altmann, A., Ejdys, J., Kazlauskaite, R., Paliszkiewicz, J., ... & Ziemba, E. (2023). Generative artificial intelligence as a new context for management theories: analysis of ChatGPT. *Central European Management Journal*, *31*(1), 3-13.

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*, 1-16.

Lee, D., Hosanagar, K., & Nair, H. S. (2018). Advertising content and consumer engagement on social media: Evidence from Facebook. *Management science*, *64*(11), 5105-5131.

Li, B., Liu, L., Mao, W., Qu, Y., & Chen, Y. (2023). Voice artificial intelligence service failure and customer complaint behavior: The mediation effect of customer emotion. *Electronic Commerce Research and Applications*, *59*, 101261.

Loh, W.-Y. (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *1*(1), 14-23.

Lopes, A. I., Dens, N., De Pelsmacker, P., & Malthouse, E. C. (2023). Managerial response strategies to eWOM: A framework and research agenda for webcare. *Tourism Management*, *98*, 104739.

Lui, T. W., Bartosiak, M., Piccoli, G., & Sadhya, V. (2018). Online review response strategy and its effects on competitive performance. *Tourism Management*, *67*, 180-190.

MacKenzie, S. B., & Lutz, R. J. (1989). An empirical examination of the structural antecedents of attitude toward the ad in an advertising pretesting context. *Journal of Marketing*, *53*(2), 48-65.

Magni, F., Park, J., & Chao, M. M. (2024). Humans as creativity gatekeepers: Are we biased against AI creativity?. *Journal of Business and Psychology*, *39*(3), 643-656.

Maxham III, J. G., & Netemeyer, R. G. (2002). A longitudinal study of complaining customers' evaluations of multiple service failures and recovery efforts. *Journal of Marketing*, *66*(4), 57-71.

Miller, C. H., Lane, L. T., Deatrick, L. M., Young, A. M., & Potts, K. A. (2007). Psychological reactance and promotional health messages: The effects of controlling language, lexical concreteness, and the restoration of freedom. *Human communication research*, *33*(2), 219-240.

Mogaji, E., & Jain, V. (2024). How generative AI is (will) change consumer behaviour: Postulating the potential impact and implications for research, practice, and policy. *Journal of Consumer Behaviour, 23*(5), 2379-2389.

Mohr, L. A., & Bitner, M. J. (1995). The role of employee effort in satisfaction with service transactions. *Journal of Business Research*, *32*(3), 239-252.

Moore, S. G., & Lafreniere, K. C. (2020). How online word-of-mouth impacts receivers. *Consumer Psychology Review*, *3*(1), 34-59.

OpenAI. (2024). *Chatgpt — release notes | openai help center*. https://help.openai.com/en/articles/6825453-chatgpt-release-notes

Packard, G., & Berger, J. (2021). How concrete language shapes customer satisfaction. *Journal of Consumer Research*, *47*(5), 787-806.

Palese, B., Piccoli, G., & Lui, T. W. (2021). Effective use of online review systems: Congruent managerial responses and firm competitive performance. *International Journal of Hospitality Management*, *96*, 102976.

Peres, R., Schreier, M., Schweidel, D., & Sorescu, A. (2023). On ChatGPT and beyond: How generative artificial intelligence may affect research, teaching, and practice. *International Journal of Research in Marketing*, *40*(2), 269-275.

Prakash, A. V., Joshi, A., Nim, S., & Das, S. (2023). Determinants and consequences of trust in AI-based customer service chatbots. *The Service Industries Journal, 43,* 642-675.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, *40*(3), 879-891.

Prolific. (2024). *The better alternative to Mturk for Online Survey Research*. https://www.prolific.com/prolific-vs-mturk

Ravichandran, T., & Deng, C. (2023). Effects of managerial response to negative reviews on future review valence and complaints. *Information Systems Research*, *34*(1), 319-341.

Rodrigo, H., Beukes, E. W., Andersson, G., & Manchaiah, V. (2021). Exploratory data mining techniques (decision tree models) for examining the impact of internet-based cognitive behavioral therapy for tinnitus: Machine learning approach. *Journal of Medical Internet Research, 23*(11), e28999.

Roozen, I., & Raedts, M. (2018). The effects of online customer reviews and managerial responses on travelers' decision-making processes. *Journal of Hospitality Marketing & Managemen*t, *27*(8), 973-996.

Sheng, J. (2019). Being active in online communications: Firm responsiveness and customer engagement behaviour. *Journal of Interactive Marketing*, *46*(1), 40-51.

Sheng, J., Wang, X., & Amankwah-Amoah, J. (2021). The value of firm engagement: How do ratings benefit from managerial responses?. *Decision Support Systems*, *147*, 113578.

Surachartkumtonkun, J. N., Grace, D., & Ross, M. (2021). Unfair customer reviews: Third-party perceptions and managerial responses. *Journal of Business Research*, *132*, 631-640.

Toma, C. L., & D'Angelo, J. D. (2015). Tell-tale words: Linguistic cues used to infer the expertise of online medical advice. *Journal of language and social psychology*, *34*(1), 25-45.

Tran, A. D., Pallant, J., & Johnson, L. (2021). Exploring the impact of chatbots on consumer sentiment and expectations in retail. *Journal of Retailing and Consumer Services, 63*, 102718.

Wang, Y., & Chaudhry, A. (2018). When and how managers' responses to online reviews affect subsequent reviews. *Journal of Marketing Research, 55*(2), 163-177.

Wang, L., Ren, X., Wan, H., & Yan, J. (2020). Managerial responses to online reviews under budget constraints: Whom to target and how. *Information & Management*, *57*(8), 103382.

Whang, J. B., Song, J. H., Lee, J. H., & Choi, B. (2022). Interacting with Chatbots: Message type and consumers' control. *Journal of Business Research*, *153*, 309-318.

Xia, L. (2013). Effects of companies' responses to consumer criticism in social media. *International Journal of Electronic Commerce*, *17*(4), 73-100.

Yhee, Y., Kim, H., Kim, J., & Koo, C. (2023). Trust in managerial response offsets negative review. *Annals of Tourism Research*, *102*, 103641.

Yim, M. C. (2023). Effect of AI Chatbot's Interactivity on Consumers' Negative Word-of-Mouth Intention: Mediating Role of Perceived Empathy and Anger. *International Journal of Human–Computer Interaction*, 1-16.

Zhang, K., Kwon, O., & Xiong, H. (2023). The Impact of Generative Artificial Intelligence. *arXiv preprint arXiv:2311.07071*.

Zhao, H., Jiang, L., & Su, C. (2020). To defend or not to defend? How responses to negative customer review affect prospective customers' distrust and purchase intention. *Journal of Interactive Marketing*, *50*(1), 45-64.

Zhao, Y., Wen, L., Feng, X., Li, R., & Lin, X. (2020). How managerial responses to online reviews affect customer satisfaction: An empirical study based on additional reviews. *Journal of Retailing and Consumer Services*, *57*, 102205.

Chapter 3.

GenAI vs. Human: A Linguistic Battle in Managerial Responses to Customer Complaints

3.1. Introduction

Existing literature has found that managerial responses to customer reviews, especially negative ones, are crucial in shaping consumer perceptions and influencing purchase decisions. Effective responses enhance the perceived trustworthiness of the brand, positively impact the purchase intentions of potential customers, and encourage more customer reviews, which further bolster the brand's online reputation (Moore & Lafreniere, 2020; Darani et al., 2023). These responses are particularly important when addressing negative reviews, where thoughtful engagement can mitigate the impact of negative feedback and even improve customer ratings (Zhao et al., 2020; Gao et al., 2019).

In this context, Generative Artificial Intelligence (GenAI) tools like ChatGPT have already demonstrated significant potential in automating the creation of managerial responses. As noted in Chapter 2, ChatGPT often outperforms human managers in generating responses that are quick, coherent, and contextually appropriate, particularly for straightforward complaints where consistency and speed are critical. This makes GenAI a valuable tool for firms looking to streamline customer service operations. Recent advancements, such as the introduction of ChatGPT's recent models such as ChatGPT 4, 4o, o1, and o3, have further enhanced these tools' capabilities in natural language processing, contextual understanding, reasoning, and the ability to handle more nuanced interactions (Feng et al., 2024; OpenAI, 2024). Moreover, the growing range of GenAI tools, including Gemini, Copilot, and Claude, offers organizations more options to integrate GenAI into their customer service strategies. To be specific, Gemini, developed by Google, emphasizes cross-platform integration and advanced conversational abilities, making it a strong contender for organizations looking to streamline customer interactions (Google, 2024). Copilot, integrated into Microsoft's suite of products, leverages AI to assist with a wide range of tasks, including drafting and refining customer communications (Microsoft, 2024). Claude, from Anthropic, focuses on

safety and ethics in AI, offering a more controlled approach to AI-generated content (Anthropic, 2024). These advancements suggest that GenAI's potential to enhance customer service efficiency and effectiveness is even greater as these technologies continue to evolve.

However, despite these promising developments, limitations persist. In Study 1 of Chapter 2, qualitative analyses revealed that while ChatGPT performs well in generating responses to general complaints, its effectiveness diminishes when addressing procedural complaints. Study 2 further demonstrated that AI-generated responses lack the necessary concreteness to adequately address procedural unfairness where customers expect detailed, fact-based explanations, leading to lower purchase intentions compared to human responses. Study 3 explored potential solutions and found that training ChatGPT with domain-specific knowledge improved its performance, enhancing its ability to provide more concrete and contextually relevant responses. These findings highlight the need for a deeper investigation into the specific contexts in which GenAI may struggle and how these tools can be further refined to address these shortcomings.

As noted at the end of Chapter 2, the small sample size and specific review criteria used in those studies limit the generalizability of the findings, underscoring the need for broader empirical research to validate results and enhance GenAI's performance in more complex scenarios. While concreteness has been identified as a key factor affecting AI-generated responses, additional linguistic differences between human and AI-written responses may also influence their effectiveness. Moreover, existing managers and marketers often rely on templated or inconsistent responses, rather than crafting detailed, thoughtful managerial replies (Deng & Ravichandran, 2023; Wang & Chaudhry, 2018). This study, therefore, examines cases where human managers actively engage in writing high-quality responses, allowing for a direct comparison of their linguistic strategies with those of two leading GenAI models—ChatGPT and Gemini. By identifying which factors contribute most to effective managerial responses, this study provides deeper insights into the evolving role of AI in customer communication.

Expanding beyond Chapter 2, this research contributes to marketing communication and electric word-of-mouth (eWOM) literature by systematically assessing how various linguistic factors impact the effectiveness of AI- and human-

generated responses. It empirically evaluates where GenAI excels and where human managers retain an advantage. This study also provides practical guidance for AI adoption, highlighting key areas for improvement in AI-generated responses and informing businesses on the strategic use of AI-human hybrid approaches in customer service. By deepening the understanding of how linguistic factors shape GenAI's effectiveness in eWOM interactions, this research lays the groundwork for optimizing AI-human collaboration in future customer engagement strategies.

3.2. Literature Review

As discussed in Chapter 2, managerial responses to customer reviews, especially negative ones, are crucial in shaping consumer perceptions and influencing purchase decisions. These responses enhance the perceived trustworthiness of the brand by demonstrating the organization's commitment to addressing customer concerns (Moore & Lafreniere, 2020). Thoughtful engagement with negative reviews can significantly impact the purchase intentions of other customers, who may view the brand more favorably when they see active engagement with feedback (Darani et al., 2023; Lui et al., 2018). Additionally, effective responses encourage more customers to leave reviews, contributing to a more robust online reputation (Ravichandran & Deng, 2023; Wang et al., 2020). Addressing issues in negative reviews can also lead to improved customer ratings and the possibility of customers updating their reviews positively (Zhang et al., 2020; Lopes et al., 2023; Gao et al., 2019). Given managerial responses' significant roles, understanding the factors that influence their effectiveness and how these can be optimized through (Gen)AI tools is crucial for enhancing customer service strategies.

3.2.1. Linguistic Factors Influencing Managerial Response Effectiveness and (Gen)AI Adoption

In examining the effectiveness of managerial responses, most factors discussed in Chapter 2 are related to consumers' subjective perceptions, such as how customers interpret the intent and effort behind a response. While these elements are crucial in shaping customer reactions, there are more linguistic factors that directly influence response effectiveness through the language and word choices themselves. These factors alter the effectiveness of responses by influencing the style, tone, and clarity of

communication, thereby affecting how customers interpret and react to the information provided. Understanding these linguistic elements is essential not only for improving human-generated content but also for optimizing the use of GenAI in customer service.

Chapter 2 already mentioned several linguistic elements, highlighting how the *length*, *formality/structure*, and *concreteness* of a managerial response can significantly impact its effectiveness. *Longer* responses are often perceived as more sincere and attentive, potentially leading to higher customer satisfaction and trust (Sheng et al., 2021; Lopes et al., 2023). *Formality and structure* in responses are also essential for conveying professionalism and credibility. A well-organized and formally worded response demonstrates respect and seriousness, enhancing the brand's image in the eyes of the customer (Gong et al., 2022). *Concreteness*, or the use of specific and tangible language, is another important factor in determining the effectiveness of a managerial response. Concrete responses are necessary for clearly communicating facts and addressing the specific concerns raised by customers (examples provided in Chapter 2), which is particularly important in resolving complaints (Jiménez-Barreto et al., 2023; Packard & Berger, 2021).

In addition to these three factors discussed in Chapter 2, six more linguistic factors will be examined in Chapter 3. First, *mimicking a customer's language* in the responses to online reviews can significantly influence the perceived attentiveness and trustworthiness of a brand. Research has shown that when companies engage in verbal mimicry—by echoing the words and phrases used by customers—it can enhance the customer's perception that the company is attentive and responsive, which in turn fosters greater trust and increases purchase intentions (Kulesza et al., 2014; Darani et al., 2023; Moore & McFerran, 2017; Swaab et al., 2011). In the context of online reviews, mimicry in written communication signals to third-party observers that the firm is attentive and actively engaged with its customers (Darani et al., 2023). This is crucial in building a strong relationship between the brand and its customers, as well as in influencing the perceptions of other potential customers who may read these reviews.

The effectiveness of *distinct managerial responses over time* has been highlighted as another critical factor in managing customer engagement and mitigating the effects of negative online interactions. Research suggests that when companies provide varied responses—particularly in their level of empathy and explanation—rather

than repetitive or identical responses, they are more successful in reducing the virality of negative electronic word-of-mouth and preventing online firestorms (Herhausen et al., 2019). The variation in response helps maintain customer interest and engagement, preventing disengagement that might arise from perceiving the responses as automated or insincere. Moreover, varying the intensity of empathy and explanation in responses has been shown to decrease the likelihood of further negative responses, which is crucial for maintaining a positive brand image over time (Grewal et al., 2008; Herhausen et al., 2019). This approach emphasizes the importance of maintaining a consistent yet varied response strategy.

The *sentiment* expressed in a managerial response is also critical in shaping customer perceptions and influencing their subsequent actions. Positive sentiment in responses is associated with increased customer satisfaction and improved online ratings (Sheng et al., 2021). A positive-valenced response, emphasizing politeness, appreciation, and accountability, fosters greater satisfaction by enhancing customers' perception of the firm’s professionalism (Min et al., 2015; Smith & Rose, 2020). On the other hand, a negative-valenced response, characterized by denial of responsibility, lack of sincerity, or an aggressive tone, can lead to dissatisfaction and harm the brand's reputation (Wang & Chaudhry, 2018). The careful crafting of emotional cues in responses is crucial for maintaining customer trust and engagement, particularly in digital interactions where non-verbal communication is absent (Sheng et al., 2021).

Time orientation plays a crucial role in shaping how brands communicate warmth and competence in response to customer complaints, ultimately influencing consumer attitudes and purchase intentions. Interactive unfairness, where customers feel disrespected or dismissed in their interactions, is best addressed through present-oriented language that emphasizes warmth, such as friendliness, empathy, and immediate acknowledgment of concerns (Roy & Naidoo, 2021). By focusing on the "here and now," these responses help rebuild trust and repair damaged relationships. Distributive unfairness, which arises when customers perceive inequalities in compensation, pricing, or benefits, can be mitigated with future-oriented language that highlights competence by offering clear solutions, outlining planned improvements, and demonstrating long-term commitment to fairness (Zimbardo & Boyd, 2014; Ravichandran & Deng, 2023). In contrast, procedural unfairness, where decision-making processes seem biased, inconsistent, or unclear, is best addressed using past-oriented

language that references established policies, historical precedents, and long-standing company practices to justify decision-making and assure customers of procedural consistency and fairness (Klicperová-Baker et al., 2014). By aligning response strategies with time-oriented messaging, brands can better manage different types of fairness concerns and enhance consumer trust.

Empathy has been shown to enhance the effectiveness of managerial responses to customer complaints because it helps to establish a connection between the firm and the customer, addressing not only the practical concerns but also the emotional needs of the customer (Davis, 2018). Empathy allows managers to acknowledge the customer's feelings, thereby reducing negative emotions and increasing customer satisfaction (Xiao et al., 2020; Yim 2023). When responding to negative reviews, demonstrating empathy by addressing the customer's emotions is vital; it reassures the customer that their concerns are understood and valued, which can help to mitigate the impact of the complaint on the firm's reputation (Bolton & Drew, 1991; Smith et al., 1999). Moreover, research shows that using "I" pronouns in managerial responses conveys a higher level of empathy compared to "we" pronouns (Packard et al., 2018). This is because "I" pronouns suggest a personal, one-on-one engagement, indicating that the manager is personally invested in resolving the customer's issue, which enhances the perception of empathy and leads to greater customer satisfaction and purchase intentions (Packard et al., 2018; Pennebaker, 2011).

Lastly, *processing fluency*, or the ease with which information is processed, significantly impacts how consumers perceive the quality and effectiveness of managerial responses. High processing fluency is associated with positive evaluations because it reduces cognitive effort, making communication seem more reliable and persuasive (Alter & Oppenheimer, 2009; Reber et al., 2004). In the context of managerial responses, clear and easily understood language can enhance customer satisfaction and trust in the brand.

To summarize, the examination of linguistic factors such as length, structure, concreteness, mimicry, distinctiveness, sentiment, time orientation, empathy, and processing fluency sets the stage for a deeper investigation into whether human managers or GenAI tools perform better in crafting effective managerial responses.

3.2.2. Comparing ChatGPT and Gemini

The decision to select ChatGPT-4o and Google's Gemini for this study was driven by the prominence and advanced capabilities of these two models in the current landscape of Generative AI. ChatGPT-4o and Gemini represent two of the most sophisticated and widely used AI tools available (Ticong, 2025), Their advanced capabilities and unique features make them ideal candidates for examining the effectiveness of AI in generating managerial responses.

In terms of the capabilities, as one of the most recent models, ChatGPT-4o, offers improved contextual understanding and the ability to generate more nuanced responses than its predecessors. This version enhances the model's capability to handle complex interactions, making it particularly useful for automating customer service tasks that require quick, coherent, and contextually appropriate responses (OpenAI, 2024; Raffo, 2024).

While a newer player in the AI field, since its launch in December 2023 (Pichai & Hassabis, 2023), Google's Gemini has quickly distinguished itself through its emphasis on cross-platform integration and advanced real-time conversational capabilities. Unlike ChatGPT, which primarily functions as a standalone AI tool, Gemini is integrated into Google's ecosystem, allowing it to leverage a wide array of data sources from across Google's services (Google, 2024). For example, Gemini is integrated into Google Workspace applications like Gmail and Docs, enabling AI-assisted email drafting and document generation (Google, n.d.). This integration enables Gemini to provide more contextually relevant and up-to-date responses, which can be particularly advantageous in scenarios where the latest information is crucial. Additionally, Gemini's integration with Google Maps allows it to provide real-time location-based suggestions and travel recommendations (Cai, 2024). Moreover, Gemini's design focuses on real-time data processing, making it effective in handling inquiries that require immediate and accurate responses (Marr, 2024; Morrison, 2024).

Extensive testing within the technology industry has highlighted key distinctions between ChatGPT-4o and Gemini, as documented by Raffo (2024), McKenzie (2025), and Ticong (2025). These findings reveal critical differences particularly relevant to the study of managerial responses in customer service. ChatGPT tends to provide more

general responses that may lack depth in addressing specific customer concerns, whereas Gemini delivers more detailed and specific replies. Additionally, ChatGPT-4o is highly versatile and effective in generating sophisticated responses, it primarily relies on pre-existing data, while Gemini accesses real-time data, making it more effective for responding to complaints involving emerging issues. ChatGPT may also exhibit biases in its responses, whereas Gemini emphasizes fairness and neutrality, producing more balanced and consistent replies. In terms of personalization, ChatGPT retains user preferences even in its free version, facilitating more consistent engagement, while Gemini offers memory features only in its paid subscription. Lastly, ChatGPT functions as a standalone AI tool, while Gemini's integration into Google's ecosystem, including Gmail, Google Docs, and Maps, allows it to generate responses that incorporate real-time data and contextual insights.

Despite these observed differences, existing studies have not examined how these models perform in generating managerial responses, leaving uncertainty about how these distinctions influence linguistic factors that may affect response effectiveness. Thus, empirical testing is necessary. Understanding these distinctions is critical for this study, as it seeks to evaluate which author—human, ChatGPT-4o, or Gemini—is more effective in crafting managerial responses that meet customer expectations and lead to positive outcomes such as enhanced brand trust and purchase intention.

3.3. Data Collection and Study Methods

3.3.1. Data Collection

Data for this study were collected from Trustpilot.ca, a leading consumer review platform founded in Denmark in 2007 and now widely used in Canada (Littlechild, 2021). Trustpilot allows any customer to leave reviews for firms without needing an invitation, promoting transparency (Trustpilot, 2024). This open nature, combined with its large user base, makes Trustpilot ideal for analyzing negative reviews and managerial responses. Additionally, from my data analysis (evidence provided in the next paragraph), Trustpilot has a higher response rate from firms to negative reviews compared to other platforms. GatherUP (2024) found that 75% of organizations do not respond to reviews, and a 30% response rate should be an industry benchmark to enhance review readers' purchase intentions. Moreover, Trustpilot's framework allows

organizations to create or claim profiles, manage customer feedback, and engage with reviewers, providing rich data for examining customer-organization interactions (Littlechild, 2021; Trustpilot, 2024). To ensure the ethical collection of data, public reviews and managerial responses were scraped using Octoparse software, adhering strictly to Trustpilot's guidelines and ethical standards (Huang, n.d.). Identifiable information, such as user IDs, was removed before data processing to maintain privacy.

The dataset was drawn from 613,461 reviews spanning 132 organizations across 20 categories on Trustpilot.ca. The categories included a diverse range of industries, such as Animals & Pets, Beauty & Well-being, Business Services, Electronics & Technology, and Travel & Vacation, and so on. Only negative reviews (1 & 2 stars) were included in the initial selection, which resulted in 72,431 reviews. These reviews were further filtered to include only those with managerial responses, leaving 48,650 reviews—a 67% response rate, notably higher than the industry benchmark of 30% (GatherUP, 2024). To ensure that the dataset was free from AI-generated content, only reviews posted before 2022 were selected, reducing the number to 17,682 reviews. Given the aim of this study is to compare the full capabilities of human responses with those generated by GenAI, it was important to include only those managerial responses that were thoroughly and carefully crafted, targeting problem-solving without relying on templates. Therefore, to focus on thoughtfully crafted responses, managerial responses with more than 150 words were included, narrowing the dataset to 544 reviews. Finally, to avoid the use of template-based responses, only unique organizational responses were selected, resulting in a final dataset of 500 reviews.

All names, salutations, and signatures were manually removed from both reviews and responses to ensure the privacy of individuals and focus solely on the textual content for analysis. This meticulous filtering process ensured a robust dataset, suitable for analyzing the effectiveness of managerial responses and generating comparative data using Generative AI tools.

For the proposed analyses, I used ChatGPT-4o mini and Gemini, both free versions, to generate responses to the selected negative reviews. The prompt used for both tools was: “Here is a negative online review against a business. Please read this review and write a response to the reviewer as the manager of this business (just the body of the response without salutation or closing):”. Although the tools sometimes

ignored the instruction to omit salutations or closings, AI-generated salutations (such as “dear customer”) and closings (such as “many thanks and best regards”) were manually removed before data analysis to maintain consistency.

The final dataset comprised 500 rows (complaint cases) and four columns (original review, human seller’s response, ChatGPT-generated response, and Gemini-generated response), allowing for a comprehensive comparison of human versus AI-generated responses. Three randomly selected rows are presented in Appendix E to show examples of response comparisons.

3.3.2. Study Methods

The proposed study utilizes several advanced text analysis tools to explore linguistic differences between human-generated and AI-generated managerial responses, focusing on various linguistic factors such as length, structure, concreteness, mimicry, distinctiveness, sentiment, time orientation, empathy, and processing fluency. The three tools that I used to conduct these text analyses are Linguistic Inquiry and Word Count (LIWC) 2022, IBM Watson Natural Language Understanding (NLU), and VOSViewer.

LIWC 2022 (Boyd et al., 2023) categorizes words based on psychological and linguistic dimensions. The 2022 version includes enhancements like the Language Style Matching (LSM) function, which measures stylistic similarity between texts, as well as detailed metrics on word count, linguistic structure, emotional tone, and time orientation. For example, LIWC's sentiment analysis is refined, distinguishing between positive and negative emotions and analyzing specific emotion words like anxiety, anger, and sadness. Additionally, LIWC’s time orientation analysis focuses on past, present, and future word usage, providing insights into the temporal focus of the texts.

IBM Watson NLU (IBM, 2024) extracts metadata from text, such as sentiment, emotion, and keywords. The sentiment analysis feature categorizes the overall emotional tone of the text on a scale from negative to positive, while the emotion analysis detects primary emotions like joy, anger, sadness, fear, and disgust. Additionally, Watson’s keyword extraction function identifies significant terms and phrases within the text, highlighting key themes or topics discussed. This capability is

particularly valuable for comparing the thematic content of human and AI-generated responses, offering insights into how each group addresses core issues in customer reviews.

Next, I will explain how I operationalized each linguistic factor that may affect managerial response effectiveness with these tools. First, I measured the *length* of each response using the word count feature in LIWC, a fundamental metric for comparing the verbosity of human and AI-generated responses. The level of *formal structure* of the responses was also be analyzed using the analytics variable in LIWC, which assesses the level of analytical thinking and formal reasoning within a text. This variable reflects the degree to which responses are logically structured and coherent, capturing the cognitive effort involved in crafting the message.

In terms of *concreteness*, the study utilized a customized dictionary in LIWC, based on the concreteness dictionary developed by Brysbaert et al. (2014). This method measures the extent to which the language used in responses refers to tangible, specific, and imaginable objects, processes, or relationships. Words referring to physical objects, specific events, and direct actions are categorized as concrete, while abstract terms are categorized as less concrete.

Mimicry, or the stylistic alignment between the language of the review and the response, was evaluated using the Language Style Matching (LSM) feature in LIWC. Here, LSM was used to assess the degree of stylistic similarity between each review and its corresponding response. High LSM scores indicate that the response closely mirrors the style of the original review, which can enhance rapport and perceived empathy in customer service interactions. *Distinctiveness* was also be analyzed using LSM but by comparing the linguistic style across all responses within the same author group (human, ChatGPT, or Gemini). High within-group LSM indicates low distinctiveness over time, suggesting that the author tends to produce similar responses, which might be ineffective if the responses become too uniform or formulaic.

Both LIWC's tone variable and IBM Watson’s sentiment analysis feature were employed for comparing the *sentiment* of responses. LIWC’s tone variable measures the positivity or negativity of a text with percentages ranging from 0% to 100% (50% means neutral), while Watson provides a more nuanced sentiment score on a scale from -1 to

+1 (0 means neutral). This dual approach allows for a comprehensive assessment of the emotional tone of the responses, offering insights into how effectively each author group addresses the emotional content of the reviews.

Empathy was be measured through a combination of IBM Watson's emotion analysis and the use of pronouns in LIWC. Watson's analysis can quantify the emotional content of the responses, allowing to see how well they align with the emotions expressed in the original reviews. Additionally, LIWC's analysis of pronoun usage (e.g., "I" vs. "we") can provide insights into the perspective and relational focus of the responses, which are key indicators of empathetic communication. The use of first-person pronouns is often associated with a more personal and empathetic tone, which is vital in customer service contexts.

Time orientation in this study was be assessed using LIWC's time focus variable, which categorizes words based on their temporal references into past, present, or future focus (e.g., "was," "is," "will"). These variables measure the frequency of temporal words to determine the dominant time orientation in the text. Understanding this time focus is crucial, as it influences perceptions of warmth, competence, and responses to various types of unfairnesses, aligning with the conceptual frameworks discussed earlier in this chapter.

Processing fluency, or the ease with which text can be read and understood, was be assessed using the Flesch Reading Ease Score (Flesch, 1948). This metric evaluates readability based on sentence length and the number of syllables per word, calculated as:

$$Reading\ Ease = 206.835 - (1.015 \times ASL) - (84.6 \times ASW)$$

where ASL (average sentence length) is the total number of words divided by the number of sentences, and ASW (average syllables per word) is the total number of syllables divided by the number of words. Higher scores indicate greater readability and ease of comprehension, whereas lower scores suggest denser, more complex language that may hinder communication effectiveness. The Flesch formula is widely used in psycholinguistics and readability research, demonstrating strong predictive power for text processing fluency (Kincaid, 1975). In this study, the *textstat* Python package was

utilized to compute the Flesch Reading Ease Score for each response, ensuring a consistent and replicable assessment of readability (PyPI, n.d.).

Table 3.1 summarizes all linguistic factors influencing managerial response effectiveness that were examined in this study, their definitions, and the operationalizing variables.

Table 3.1. Influencing Factors and Operationalizing Variables

Influencing Factor	Definition	Operationalizing Variable
Length	The total number of words used in a response. Longer responses are perceived as more sincere and attentive, potentially leading to higher customer satisfaction and trust (Sheng et al., 2021; Lopes et al., 2023).	Word count (LIWC)
Structure	The organization and formality of the response, conveying professionalism and credibility. Structured responses demonstrate respect and seriousness, enhancing brand image (Gong et al., 2022).	Analytics (LIWC)
Concreteness	The use of specific, tangible language to clearly communicate facts and address concerns. Concrete language is crucial for resolving complaints effectively (Jiménez-Barreto et al., 2023; Packard & Berger, 2021).	Concreteness (LIWC - Custom Dictionary)
Mimicry	The stylistic alignment between a customer's language and the response, signaling attentiveness and responsiveness. Mimicry can enhance perceived trust and increase purchase intentions (Kulesza et al., 2014; Darani et al., 2023).	Language Style Matching (LSM between review and response) (LIWC)
Distinctiveness	The variability in language use across responses. Low distinctiveness indicates a formulaic approach, while high distinctiveness suggests tailored responses. Effective distinctiveness helps maintain customer engagement (Herhausen et al., 2019).	Language Style Matching (LSM within author group) (LIWC)
Sentiment	The emotional tone of the response, affecting customer satisfaction and brand perception. Positive sentiment in responses is associated with increased satisfaction (Sheng et al., 2021; Wang & Chaudhry, 2018).	Tone (LIWC), Sentiment (IBM Watson NLU)
Empathy	The expression of emotions and the use of pronouns, influencing perceived empathy. Responses that align with customer emotions and use inclusive language can build rapport (IBM Watson NLU documentation; LIWC analysis).	Emotion analysis (IBM Watson NLU), Pronoun usage (LIWC)
Time Orientation	The focus on past, present, or future within a response. Time orientation influences perceptions of warmth, competence, and how unfairness is addressed (Boyd et al., 2022).	Time focus (LIWC)

Processing Fluency	The ease with which a response can be read and understood, sentence length, and the number of syllables per word. Higher processing fluency improves comprehension and effectiveness (Flesch, 1948).	Flesch Reading Ease (Python textstat)

3.4. Results

Building on the prior discussion of linguistic factors influencing managerial response effectiveness, this section compares all the aforementioned factors across three groups: human-generated responses, ChatGPT-4o responses, and Gemini responses. Specifically, these analyses examine length, structure, concreteness, mimicry, distinctiveness, sentiment, time orientation, empathy, and processing fluency to evaluate how these linguistic elements could influence customer perceptions and purchase intentions.

For *length*, a one-way ANOVA was conducted to compare word count (WC) of the three groups. The results indicated a significant difference between the three groups (M_{Gemini} = 240.9, SD_{Gemini} = 62.38, $M_{ChatGPT}$ = 169.4, $SD_{ChatGPT}$ = 66.38, M_{Human} = 200.5, M_{Human} = 80.2, $F(2, 988) = 155.00$, $p < .001$). Post hoc analyses revealed that Gemini-generated responses were significantly longer than both ChatGPT-generated responses (mean difference = 71.5, $p < .001$) and human-generated responses (mean difference = 31.2, $p < .001$). Additionally, human-generated responses were significantly longer than ChatGPT (mean difference = 40.3, $p < .001$). These findings suggest that AI-generated responses, particularly those from Gemini, tend to be lengthier than human-written responses, which may influence perceived effort and thoroughness in customer engagement.

A correlation analysis examined the relationship between the *length* of the original reviews and the length of responses generated by each author group. The results showed a significant positive correlation between review length and response length across all three groups. Specifically, human-generated responses exhibited a weak correlation with review length ($r = 0.259$, $p < .001$), indicating a relatively low dependence on the original review's length when crafting responses. Both ChatGPT-generated ($r = 0.474$, $p < .001$) and Gemini-generated responses ($r = 0.416$, $p < .001$)

exhibited moderate correlations with review length, indicating that these GenAI models tend to mirror the length of the original review more closely than human authors. These findings imply that AI-generated responses, particularly from ChatGPT and Gemini, are more responsive to variations in review length compared to human-written responses, which may reflect AI models' tendency to align response length with the input text.

A one-way ANOVA was carried out to examine differences in the *Analytic* variable across the three groups. As previously introduced, *Analytic* serves as a proxy for structure by assessing the degree of logical organization, formality, and complexity in written text, capturing how systematically and coherently information is conveyed in a response. The results indicated a significant difference among the three groups (M_{Gemini} = 59.1, SD_{Gemini} = 14.74, $M_{ChatGPT}$ = 38.5, $SD_{ChatGPT}$ = 15.29, M_{Human} = 50.9, SD_{Human} = 16.00, $F(2, 983) = 237.1$, $p < .001$). Post hoc analyses revealed that Gemini-generated responses had significantly higher Analytic scores than both ChatGPT-generated responses (mean difference = 20.6, $p < .001$) and human-generated responses (mean difference = 8.24, $p < .001$). Additionally, human-generated responses exhibited significantly higher structure scores than ChatGPT-generated responses (mean difference = 12.39, $p < .001$). These findings suggest that Gemini-generated responses tend to exhibit a more structured and formal organization, while ChatGPT responses appear less structured in comparison.

A one-way ANOVA was performed to assess differences in linguistic *Concreteness* among the three groups of responses. The results indicated a significant difference among the three groups (M_{Human} = 217, SD_{Human} = 12.15, $M_{ChatGPT}$ = 209, $SD_{ChatGPT}$ = 9.07, M_{Gemini} = 203, SD_{Gemini} = 11.97, $F(2, 978) = 146.0$, $p <.001$). Post hoc analyses revealed that human-generated responses were significantly more concrete than both ChatGPT-generated responses (mean difference = 8.16, $p < .001$) and Gemini-generated responses (mean difference = 13.97, $p < .001$). Additionally, ChatGPT-generated responses were significantly more concrete than Gemini-generated responses (mean difference = 5.81, $p < .001$). The current findings align with prior discussions in Chapter 2, where AI-generated responses were noted to struggle with

providing concrete information. The lower concreteness scores of AI-generated responses, especially from Gemini, suggest that while AI models can generate fluent and structured content, they may still lack the necessary specificity to fully address customer concerns.

As previously introduced, Pairwise Language Style Matching (LSM) analysis in LIWC evaluates the degree of *Mimicry* in a response by measuring its alignment with the linguistic style of the original review, specifically in terms of function words and syntactic structures. Higher LSM scores indicate stronger mimicry, reflecting a response's ability to closely replicate the review's linguistic patterns, while lower scores suggest weaker mimicry and less stylistic adaptation. A one-way ANOVA was conducted to compare pairwise LSM scores which represent the similarity between the linguistic style of the original reviews and each response group, revealing a significant difference across the groups (M_{Human} = 0.725, SD_{Human} = 0.125, $M_{ChatGPT}$ = 0.707, $SD_{ChatGPT}$ = 0.121, M_{Gemini} = 0.706, SD_{Gemini} = 0.113, $F(2, 978) = 3.68$, $p = .026$). Post hoc analyses found that human-generated responses exhibited a higher but relatively weak level of linguistic style matching with the original reviews compared to both ChatGPT-generated responses (mean difference = 0.018, $p = .059$) and Gemini-generated responses (mean difference = 0.019, $p = .037$). However, there was no statistically significant difference between ChatGPT and Gemini in their ability to mimic the style of the original reviews ($p = .993$, ns). These findings suggest that while AI-generated responses demonstrate some capacity for adapting to the language style of the original review, they do not match the mimicry level of human responses.

To evaluate the *Distinctiveness* of linguistic patterns within each response group, within-group LSM was analyzed using a one-way ANOVA. Higher within-group LSM scores indicate that responses within a group share more stylistic similarities, suggesting a more formulaic or uniform approach, whereas lower scores imply greater linguistic variability and individuality in responses. The results indicated a significant difference among the three groups (M_{Gemini} = 0.866, SD_{Gemini} = 0.053, $M_{ChatGPT}$ = 0.861, $SD_{ChatGPT}$ = 0.065, M_{Human} = 0.855, SD_{Human} = 0.057, $F(2,992) = 4.98$, $p = .007$). Post hoc analyses

showed that Gemini-generated responses had significantly higher within-group LSM than human-generated responses (mean difference = 0.011, p = .005), while the differences between ChatGPT and Gemini (p = .486) and between ChatGPT and human responses (p = .198) were not significant. These findings suggest that ChatGPT-generated responses exhibit a level of linguistic distinctiveness comparable to human-generated responses, whereas Gemini-generated responses are significantly more uniform in style.

Sentiment analysis was conducted using two methods: IBM Watson's categorical sentiment classification and LIWC's Tone variable, which measures sentiment on a scale from 0 (extremely negative) to 100 (extremely positive).

The chi-square test for IBM Watson's sentiment classification showed a significant association between sentiment and response group ($\chi^2 = 870$, $p < .001$). ChatGPT-generated responses exhibited the highest proportion of positive sentiment (451 out of 500 responses), while Gemini-generated responses had the highest proportion of negative sentiment (149 out of 500 responses). Human-generated responses fell between the two GenAI models, with more positive sentiment (363 out of 500) but also a notable level of negativity (136 out of 500).

For LIWC's Tone analysis, a one-way ANOVA revealed a significant difference among the three groups ($M_{ChatGPT} = 72.9$, $SD_{ChatGPT} = 22.78$, $M_{Human} = 62.2$, $SD_{Human} = 22.39$, $M_{Gemini} = 58.6$, $SD_{Gemini} = 23.46$, $F(2,998) = 52.7$, $p < .001$). Post hoc analyses showed that ChatGPT-generated responses had significantly higher Tone scores than both Gemini-generated (mean difference = 14.3, $p < .001$) and human-generated responses (mean difference = 10.75, $p < .001$). Additionally, human-generated responses had a slightly but significantly higher Tone score than Gemini-generated responses (mean difference = 3.58, p = .037). Both methods produced identical results, confirming the robustness of the findings: ChatGPT-generated responses were the most positive, Gemini-generated responses were the most negative, and human-generated responses fell in between.

Empathy in responses was evaluated using two methods: IBM Watson's emotion analysis and LIWC's pronoun analysis. IBM Watson's emotion analysis assessed how well each response group aligned with the emotional tone of the original reviews, where higher correlation values indicate stronger emotional alignment and, consequently, greater empathy. Additionally, LIWC's analysis of pronoun usage examined the prevalence of first-person singular ('I') and first-person plural ('we') pronouns, with a higher use of 'we' pronouns indicating greater inclusivity and a stronger empathetic connection with customers.

Table 3.2. Descriptives of IBM Watson Emotion Scores

Group Descriptives

	Category	N	Mean	SD
Sadness	ChatGPT	500	0.3132	0.0696
	Gemini	500	0.306	0.0712
	Human	500	0.3015	0.0913
	Review	498	0.372	0.1309
Joy	ChatGPT	500	0.3156	0.0923
	Gemini	500	0.276	0.0927
	Human	500	0.2901	0.1176
	Review	498	0.2033	0.1237
Fear	ChatGPT	500	0.054	0.0217
	Gemini	500	0.0538	0.0211
	Human	500	0.0613	0.0255
	Review	498	0.0754	0.04
Disgust	ChatGPT	500	0.0267	0.0168
	Gemini	500	0.0261	0.0136
	Human	500	0.0322	0.0199
	Review	498	0.0522	0.0401
Anger	ChatGPT	500	0.0826	0.0312
	Gemini	500	0.0935	0.0319
	Human	500	0.0838	0.0342
	Review	498	0.1297	0.0704

Table 3.2 displays the group descriptives of the five IBM Watson emotional scores across the four text groups: review, human, ChatGPT, and Gemini. To assess empathy, correlation analyses measured the alignment of each response group's emotions with those expressed in the original review. Higher correlation values indicate stronger emotional alignment, suggesting greater responsiveness to customer

sentiment. The results showed that Gemini-generated responses had the highest correlations across all five emotions (r_{anger} = 0.320, $r_{disgust}$ = 0.302, r_{fear} = 0.234, r_{joy} = 0.356, $r_{sadness}$ = 0.315, all ps < .001). ChatGPT-generated responses exhibited moderate correlations (r_{anger} = 0.218, $r_{disgust}$ = 0.196, r_{fear} = 0.174, r_{joy} = 0.289, $r_{sadness}$ = 0.241, all ps < .001), suggesting a balanced but less pronounced emotional alignment compared to Gemini-generated responses. Human-generated responses demonstrated the weakest empathy, with non-significant alignment for disgust ($r_{disgust}$ = 0.080, p = .074, ns) and fear (r_{fear} = 0.102, p = .054, ns), while only weak alignment was observed for anger, joy, and sadness (r_{anger} = 0.153, r_{joy} = 0.178, $r_{sadness}$ = .158, all ps < .001). These findings indicate that human responses were the least responsive to the emotional tone of the original reviews, particularly for negative emotions.

For the pronoun usage measured in LIWC, a one-way ANOVA revealed significant differences among the three groups for both 'I' (M_{Gemini} = 1.54, SD_{Gemini} = 0.99, M_{Human} = 0.78, SD_{Human} = 1.11, $M_{ChatGPT}$ = 0.20, $SD_{ChatGPT}$ = 0.64, F(2, 936) = 334.28, p < .001) and 'we' ($M_{ChatGPT}$ = 8.04, $SD_{ChatGPT}$ = 1.85, M_{Human} = 5.38, SD_{Human} = 2.09, M_{Gemini} = 6.99, SD_{Gemini} = 1.37, F(2, 965) = 228.78, p < .001). Post hoc analyses showed that ChatGPT-generated responses used significantly more 'I' pronouns than both human (mean difference = 0.76, p < .001) and Gemini-generated responses (mean difference = 1.35, p < .001), while human responses also contained significantly more 'I' pronouns than Gemini responses (mean difference = 0.59, p < .001). Similarly, ChatGPT-generated responses had significantly higher 'we' pronoun usage than both human (mean difference = 2.66, p < .001) and Gemini-generated responses (mean difference = 1.06, p < .001), with Gemini responses also showing significantly greater use of 'we' than human responses (mean difference = 1.61, p < .001).

The findings from both methods align closely. Gemini-generated responses exhibited the highest emotional alignment with the original reviews and the highest usage of 'I' pronouns. In contrast, human-generated responses displayed the weakest emotional alignment and the least inclusive language, suggesting a lower degree of empathy when measured through both methods.

Time orientation in responses was examined using LIWC's time focus analysis, categorizing words based on their temporal references—past, present, or future. A one-way ANOVA revealed significant differences among the three response groups for past focus (M_{Gemini} = 1.85, SD_{Gemini} = 1.07, $M_{ChatGPT}$ = 2.49, $SD_{ChatGPT}$ = 1.24, M_{Human} = 3.21, SD_{Human} = 2.25, $F(2, 942) = 88.7$, $p < .001$), present focus ($M_{ChatGPT}$ = 4.79, $SD_{ChatGPT}$ = 1.33, M_{Human} = 4.66, SD_{Human} = 1.67, M_{Gemini} = 3.72, SD_{Gemini} = 1.20, $F(2, 982) = 104.6$, $p < .001$), and future focus (M_{Gemini} = 2.09, SD_{Gemini} = 1.01, M_{Human} = 1.57, SD_{Human} = 1.15, $M_{ChatGPT}$ = 1.52, $SD_{ChatGPT}$ = 0.94, $F(2, 991) = 50.0$, $p < .001$). Post hoc analyses showed that human-generated responses contained significantly more past-focused language than both ChatGPT (mean difference = 0.72, $p < .001$) and Gemini (mean difference = 1.36, $p < .001$), while ChatGPT responses had significantly higher past focus than Gemini (mean difference = 0.64, $p < .001$). For present focus, ChatGPT responses contained significantly more present-oriented language than Gemini (mean difference = 1.07, $p < .001$), while human responses were not significantly different from ChatGPT ($p = .337$). Regarding future focus, Gemini responses included significantly more future-oriented language than both ChatGPT (mean difference = 0.58, $p < .001$) and human responses (mean difference = 0.52, $p < .001$), while the difference between ChatGPT and human responses was not significant ($p = .700$).

To assess *processing fluency*, a one-way ANOVA was performed to examine the Flesch Reading Ease Scores among the three groups. The results indicated a significant difference among the three groups (M_{Gemini} = 34.5, SD_{Gemini} = 10.39, $M_{ChatGPT}$ = 49.7, $SD_{ChatGPT}$ = 6.81, M_{Human} = 60.1, SD_{Human} = 8.29, $F(2, 971) = 924$, $p < .001$). Post hoc analyses showed that Gemini-generated responses are more difficult to read and process than ChatGPT-generated responses (mean difference = -15.2, $p < .001$) and human-generated responses (mean difference = -25.6, $p < .001$). ChatGPT-generated responses are also more complex than human-generated responses (mean difference = -10.4, $p < .001$). These results suggest that AI-generated responses tend to use more complicated vocabulary, which may affect readability and perceived fluency in customer communications.

3.5. General Discussion and Future Research Directions

This study aims to explore the linguistic differences between managerial responses crafted by humans and those generated by two leading GenAI models—

ChatGPT and Gemini. By examining various linguistic factors, such as length, structure, concreteness, mimicry, distinctiveness, sentiment, time orientation, empathy, and processing fluency, this research sought to identify which factors critical to managerial response effectiveness are best addressed by each type of author.

To be specific, Gemini-generated responses, being the longest, demonstrate the most effort and comprehensiveness. Humans, while slightly shorter, maintain clarity and engagement, ensuring readability. In terms of structure, Gemini responses are the most structured, followed by human responses, while ChatGPT is the least structured. Greater structure enhances professionalism and credibility, giving Gemini an advantage in formal communication.

Human responses are the most concrete, providing clear details and direct resolutions, whereas AI-generated responses, particularly Gemini's, lack specificity. This aligns with the findings of Chapter 2, which emphasize that AI-generated responses, while efficient, often fail to deliver precise and meaningful resolutions to customer complaints. These results further reinforce the need for human oversight in AI-generated content to ensure that managerial responses address concerns effectively and maintain customer trust.

The stronger mimicry observed in human-generated responses highlights their ability to naturally align with customer language, contributing to greater perceived authenticity and responsiveness. In contrast, AI-generated responses may lack the nuanced stylistic adaptation that makes human responses more engaging and effective in customer service interactions.

The greater variability in human responses is expected, given that multiple individuals authored them, further reinforcing their distinctiveness. ChatGPT responses exhibit variation similar to human authors, making them equally engaging and natural in tone. In contrast, Gemini-generated responses follow a more formulaic approach, potentially making these responses appear less personalized and overly standardized. While structured responses ensure consistency, they may lack the flexibility and nuance needed to foster stronger customer engagement.

ChatGPT responses are the most positive, enhancing customer perceptions of warmth, while Gemini's responses are the most negative. In addition, both GenAI tools

demonstrate higher empathy than human responses, with Gemini showing the strongest emotional alignment with customer sentiment. This greater alignment suggests that AI-generated responses can effectively mirror customer emotions, which may enhance perceived attentiveness and engagement.

In terms of time orientation, ChatGPT responses are the most present-focused, making them particularly effective in addressing interactional unfairness by emphasizing warmth, immediate acknowledgment, and corrective action to rebuild trust. Meanwhile, Gemini's future-oriented approach is better suited for addressing distributive unfairness by outlining planned improvements and procedural changes to ensure long-term fairness. Human responses, which exhibited a stronger past focus, may be more effective in addressing procedural unfairness by referencing historical precedents and company policies to justify decision-making and reinforce reliability. This aligns with the conclusions of Chapter 2, which found that AI-generated responses often lack the depth needed to effectively reassure customers about procedural fairness. Human-written responses, by drawing on historical consistency, may better mitigate concerns regarding biased or inconsistent decision-making, reinforcing customer trust in company policies.

Finally, GenAI responses use longer sentences and more complex vocabulary, which may reduce readability. While professional, excessive complexity can hinder clarity, particularly for frustrated customers needing straightforward resolutions.

Table 3.3 summarizes the comparative performance of human-generated responses, ChatGPT, and Gemini across key linguistic factors relevant to managerial response effectiveness discussed above. This table serves as a practical guide for marketers seeking to leverage GenAI tools to enhance efficiency in crafting managerial responses, helping them understand where AI excels and where human oversight or model training is needed. Since no single tool excels across all linguistic factors, businesses can use this table to develop a hybrid approach that strategically integrates AI tools with human intervention. By balancing automation with human adaptability, companies can optimize response effectiveness, ensuring that each response aligns with customer expectations and specific service needs.

Table 3.3. Performance Comparison Across Linguistic Factors

Linguistic Factor	Best Performer	Worst Performer

Length	Gemini	ChatGPT
Structure	Gemini	ChatGPT
Concreteness	Human	Gemini
Mimicry	Human	Gemini
Distinctiveness	ChatGPT/Human	Gemini
Sentiment	ChatGPT	Gemini
Time Orientation - Past	Human	Gemini
Time Orientation - Present	ChatGPT	Gemini
Time Orientation - Future	Gemini	Human
Empathy	Gemini	Human
Processing Fluency	Human	Gemini

The findings contribute to the broader understanding of how GenAI tools can be utilized in customer service, particularly in crafting responses to negative online reviews. By providing empirical evidence on the linguistic capabilities of both human and AI-generated responses, this study offers valuable insights for marketers and managers on the practical applications and limitations of GenAI in marketing communication.

Despite the contributions of this study, several limitations should be acknowledged. First, the observational nature of the research inherently limits the ability to draw causal inferences. The study relied on naturally occurring data from Trustpilot.ca, which, while extensive, does not allow for controlled experimental conditions. Additionally, the selection of Trustpilot as the sole review platform may limit the generalizability of the findings. Other platforms may have different user bases, review structures, or managerial response patterns that could influence the results. Moreover, while this study focused on two leading GenAI models, the rapidly evolving nature of AI technology means that newer models may exhibit different capabilities, potentially altering the conclusions from the proposed studies.

Building on the findings of this study, future research should explore whether training AI to improve weaker factors, such as concreteness, mimicry, distinctiveness, and processing fluency, can enhance managerial response effectiveness. Experimental studies could assess if refining these aspects through model fine-tuning leads to improved customer satisfaction and trust. Additionally, research should examine the

impact of hybrid AI-human customer service approaches, measuring how human oversight influences AI credibility and customer engagement. As GenAI technology continues to advance, ongoing research will be essential to ensure these tools not only match but also exceed human capabilities in delivering nuanced, personalized, and contextually appropriate responses, ultimately setting new standards for excellence in customer service.

References

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and social psychology review*, *13*(3), 219-235.

Anthropic. (2024). *Meet claude*. \ Anthropic. https://www.anthropic.com/claude

Bolton, R. N., & Drew, J. H. (1991). A multistage model of customers' assessments of service quality and value. *Journal of consumer research*, *17*(4), 375-384.

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*, 10.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods*, *46*, 904-911.

Cai, K. (2024, October 31). *Google brings AI answers to map applications*. Reuters. https://www.reuters.com/technology/artificial-intelligence/google-brings-ai-answers-map-applications-2024-10-31/

Darani, M. M., Mirahmad, H., Raoofpanah, I., & Groening, C. (2023). Managerial responses to online communication: The role of mimicry in affecting third-party observers' purchase intentions. *Journal of Business Research*, *166*, 113979.

Davis, M. H. (2018). *Empathy: A social psychological approach*. Routledge.

Deng, C., & Ravichandran, T. (2024). Managerial response to online positive reviews: Helpful or harmful?. *Information Systems Research*, *35*(4), 1802-1823.

Feng, C. M., Botha, E., & Pitt, L. (2024). From HAL to GenAI: Optimizing chatbot impacts with CARE. *Business Horizons, 67*(5), 537-548.

Flesch, R. (1948). *A new readability yardstick. Journal of Applied Psychology, 32*(3), 221-233.

GatherUp. (2024, July 5). *120+ Online Review Statistics You Need to know in 2020*. GatherUp. https://gatherup.com/resources/100-online-review-statistics/#:~:text=Businesses%20that%20don't%20reply,9%25%20less%20revenue%20than%20average.&text=Businesses%20that%20reply%20to%20their,time%20average%2035%25%20more%20revenue.&text=75%25%20of%20businesses%20don't,to%20any%20of%20their%20reviews.&text=A%2030%25%20reply%20rate%20is,growth%20to%20outstrip%20your%20competitors.

Gao, C., Zeng, J., Xia, X., Lo, D., Lyu, M. R., & King, I. (2019, November). Automating app review response generation. In *2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE)* (pp. 163-175). IEEE.

Gong, C., Liu, J., Law, R., & Ye, Q. (2022). Exploring the effects of official-structured managerial responses on hotel online popularity. *International Journal of Hospitality Management*, *106*, 103293.

Google. (2024). *Gemini - Chat to Supercharge Your Ideas*. https://www.google.com/gemini

Google. (n.d.). *Ai tools for Business | Google Workspace*. https://workspace.google.com/solutions/ai/

Grewal, D., Roggeveen, A. L., & Tsiros, M. (2008). The effect of compensation on repurchase intentions in service recovery. *Journal of Retailing*, *84*(4), 424-434.

Herhausen, D., Ludwig, S., Grewal, D., Wulf, J., & Schoegel, M. (2019). Detecting, preventing, and mitigating online firestorms in brand communities. *Journal of Marketing*, *83*(3), 1-21.

Huang, Y. (n.d.). *Is web scraping via octoparse legal and ethical?*. Help Center. https://helpcenter.octoparse.com/en/articles/6471169-is-web-scraping-via-octoparse-legal-and-ethical

IBM. (2024, May 16). *IBM Watson Natural language understanding*. https://www.ibm.com/products/natural-language-understanding

Jiménez-Barreto, J., Rubio, N., & Molinillo, S. (2023). How chatbot language shapes consumer perceptions: The role of concreteness and shared competence. *Journal of Interactive Marketing*, 10949968231177618.

Kincaid, J. P. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. *Chief of Naval Technical Training*.

Klicperová-Baker, M., Košťál, J., & Vinopal, J. (2014). Time perspective in consumer behavior. *In Time perspective theory; review, research and application: Essays in honor of Philip G. Zimbardo* (pp. 353-369). Cham: Springer International Publishing.

Kulesza, W., Dolinski, D., Huisman, A., & Majewski, R. (2014). The echo effect: The power of verbal mimicry to influence prosocial behavior. *Journal of Language and Social Psychology*, *33*(2), 183-201.

Littlechild, S. (2021). Exploring customer satisfaction in Great Britain's retail energy sector Part I: The comparative use of Trustpilot online reviews in four sectors. *Utilities Policy*, *73*, 101298.

Lopes, A. I., Dens, N., De Pelsmacker, P., & Malthouse, E. C. (2023). Managerial response strategies to eWOM: A framework and research agenda for webcare. *Tourism Management*, *98*, 104739.

Lui, T. W., Bartosiak, M., Piccoli, G., & Sadhya, V. (2018). Online review response strategy and its effects on competitive performance. *Tourism Management*, *67*, 180-190.

McKenzie, L. (2025, January 26). *Google Gemini vs Chatgpt: Which AI assistant wins in 2025?*. Backlinko. https://backlinko.com/gemini-vs-chatgpt

Marr, B. (2024, July 2). *Ai Showdown: Chatgpt vs. Google's gemini – which Reigns Supreme?* Forbes. https://www.forbes.com/sites/bernardmarr/2024/02/13/ai-showdown-chatgpt-vs-googles-gemini--which-reigns-supreme/

Microsoft. (2024). *Microsoft copilot*. Microsoft AI. https://www.microsoft.com/en-us/copilot

Min, H., Lim, Y., & Magnini, V. P. (2015). Factors affecting customer satisfaction in responses to negative online hotel reviews: The impact of empathy, paraphrasing, and speed. *Cornell Hospitality Quarterly*, *56*(2), 223-231.

Moore, S. G., & Lafreniere, K. C. (2020). How online word-of-mouth impacts receivers. *Consumer Psychology Review*, *3*(1), 34-59.

Moore, S. G., & McFerran, B. (2017). She said, she said: Differential interpersonal similarities predict unique linguistic mimicry in online word of mouth. *Journal of the Association for Consumer Research*, *2*(2), 229-245.

Morrison, R. (2024, March 3). *I tested Google Gemini vs OpenAI CHATGPT in a 9-round face-off - here's the winner*. Tom's Guide. https://tomsguide.com/ai/google-gemini-vs-openai-chatgpt

OpenAI. (2024). *Introducing GPT-4O and more tools to chatgpt free users*. https://openai.com/index/gpt-4o-and-more-tools-to-chatgpt-free

Packard, G., & Berger, J. (2021). How concrete language shapes customer satisfaction. *Journal of Consumer Research*, *47*(5), 787-806.

Packard, G., Moore, S. G., & McFerran, B. (2018). (I'm) happy to help (you): The impact of personal pronoun use in customer–firm interactions. *Journal of Marketing Research*, *55*(4), 541-555.

Pennebaker, J. W. (2011). The secret life of pronouns. *New Scientist*, *211*(2828), 42-45.

Pichai, S., & Hassabis, D. (2023, December 6). *Introducing Gemini: Our largest and most capable AI model*. Google. https://blog.google/technology/ai/google-gemini-ai/

PyPI. (n.d.). *textstat 0.7.5*. https://pypi.org/project/textstat

Raffo, D. (2024, June 10). *Gemini vs. CHATGPT: What's the difference?: TechTarget*. Enterprise AI. https://www.techtarget.com/searchenterpriseai/tip/Gemini-vs-ChatGPT-Whats-the-difference

Ravichandran, T., & Deng, C. (2023). Effects of managerial response to negative reviews on future review valence and complaints. *Information Systems Research*, *34*(1), 319-341.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience?. *Personality and social psychology review*, *8*(4), 364-382.

Roy, R., & Naidoo, V. (2021). Enhancing chatbot effectiveness: The role of anthropomorphic conversational styles and time orientation. *Journal of Business Research*, *126*, 23-34.

Sheng, J., Wang, X., & Amankwah-Amoah, J. (2021). The value of firm engagement: How do ratings benefit from managerial responses?. *Decision Support Systems*, *147*, 113578.

Smith, A. K., Bolton, R. N., & Wagner, J. (1999). A model of customer satisfaction with service encounters involving failure and recovery. *Journal of marketing research*, *36*(3), 356-372.

Smith, L. W., & Rose, R. L. (2020). Service with a smiley face: Emojional contagion in digitally mediated relationships. *International Journal of Research in Marketing*, *37*(2), 301-319.

Swaab, R. I., Maddux, W. W., & Sinaceur, M. (2011). Early words that work: When and how virtual linguistic mimicry facilitates negotiation outcomes. *Journal of Experimental Social Psychology*, *47*(3), 616-621.

Ticong, L. (2025, January 6). *Gemini vs ChatGPT 4 (2025): The winning AI revealed*. eWeek. https://www.eweek.com/artificial-intelligence/gemini-vs-chatgpt/

Trustpilot. (2024). *Read reviews. write reviews*. Trustpilot reviews. https://www.trustpilot.com/

Wang, L., Ren, X., Wan, H., & Yan, J. (2020). Managerial responses to online reviews under budget constraints: Whom to target and how. *Information & Management*, *57*(8), 103382.

Wang, Y., & Chaudhry, A. (2018). When and how managers' responses to online reviews affect subsequent reviews. *Journal of Marketing Research*, *55*(2), 163-177.

Xiao, Z., Zhou, M. X., Chen, W., Yang, H., & Chi, C. (2020, April). If I hear you correctly: Building and evaluating interview chatbots with active listening skills. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-14).

Yim, M. C. (2023). Effect of AI Chatbot's Interactivity on Consumers' Negative Word-of-Mouth Intention: Mediating Role of Perceived Empathy and Anger. *International Journal of Human–Computer Interaction*, 1-16.

Zhang, K., Kwon, O., & Xiong, H. (2023). The Impact of Generative Artificial Intelligence. *arXiv preprint arXiv:2311.07071*.

Zhao, Y., Wen, L., Feng, X., Li, R., & Lin, X. (2020). How managerial responses to online reviews affect customer satisfaction: An empirical study based on additional reviews. *Journal of Retailing and Consumer Services*, *57*, 102205.

Zimbardo, P. G., & Boyd, J. N. (2014). Putting time in perspective: A valid, reliable individual-differences metric. In *Time perspective theory; review, research and application: Essays in honor of Philip G. Zimbardo* (pp. 17-55). Cham: Springer International Publishing.

Appendix A

Reviews and Responses Used in Study 1

Table A1. Reviews and Responses Used in Study 1

ID	Review	Respondent	Response
1	Good food and drinks, but the manager literally came to our table and said they “need to turn the table” and asked us to settle. So, we cancelled the food and drinks on the way (to his surprise) and we chose to go to (another restaurant's name) instead. They have a time limit, so, if you want a good night, go somewhere else.	ChatGPT	Subject: Our Sincere Apologies and Thanks for Your Feedback Dear [Reviewer's Name], I hope this message finds you well. My name is [Manager Name], and I am the manager at [Restaurant's Name]. I wanted to personally reach out to you after reading your review, and I truly appreciate you taking the time to share your experience with us. First and foremost, I would like to extend my sincerest apologies for the discomfort and inconvenience you felt during your recent visit to our restaurant. It was never our intention to make you feel rushed or unwelcome, and I am truly sorry that this was your takeaway from our interaction. I understand that our request to settle the bill was untimely and disrupted your dining experience, which is something we never want for any of our guests. Our intention behind implementing a time limit during peak hours is to accommodate as many guests as possible, especially during busy times, but I realize that this policy may have been miscommunicated or applied too strictly in your case. Please know that we value your feedback immensely as it helps us to improve our service and ensure that all our guests leave feeling satisfied and eager to return. I have taken your comments to heart and will be revisiting our policies and staff training to ensure that this situation does not occur again in the future. While I understand that you chose to dine elsewhere following your experience, I hope you might consider giving us another chance in the future to show you the true hospitality and enjoyable dining experience that we strive to provide to all our guests. Once again, I apologize for any frustration we may have caused, and I thank you for bringing this matter to our attention. Your feedback is invaluable to us, and we are committed to making the necessary changes to ensure a better experience for all our guests. Wishing you all the best, [Manager Name] Manager, [Restaurant's Name]

		Human Manager	Hi (customer name), This is extremely disappointing to hear about your experience. I can only imagine how frustrated you were waiting for your food and then being asked to leave. I want to ensure you that is not our policy surrounding our reservations and no guest should be asked to leave prior to there experience being complete. I would love the opportunity to talk further about this over the phone with you so I can collect as much feedback as possible. Please reach out to me directly at (manager email) Thanks, (manager name)
2	We were seated at the very back behind a wall and could not get any service. It was like being in a black hole. We literally had to flag and wave a server down for every single drink and all our food. We ordered a bottle of wine that did not come so we cancelled it and then we were charged on our bill. We left very disappointed with the overall experience.	ChatGPT	Subject: Our Apologies and Commitment to Improvement Dear Guest, I hope this message finds you well. My name is [Manager Name], and I am the manager at [Restaurant's Name]. I recently read your review about your experience at our restaurant, and I want to extend my sincerest apologies for the disappointment you faced. Firstly, I am truly sorry that the seating arrangement and service did not meet your expectations. It is concerning to hear that you felt isolated and neglected, especially since we aim to provide a welcoming and attentive environment for all our guests. Your feedback has been taken very seriously, and I assure you that we are reevaluating our seating arrangements and service protocols to ensure that every guest feels valued and well-served, regardless of where they are seated. Regarding the issue with the wine order and the billing error, I deeply apologize for the oversight and any inconvenience it may have caused. This is certainly not the standard we strive for, and I have addressed this with our team to prevent such occurrences in the future. While I understand that your experience was not up to the mark, I hope you will give us another chance to demonstrate our commitment to excellence in both service and dining. We value your feedback as it is crucial for our improvement, and we are dedicated to making the necessary changes to enhance our guests' experiences. If you wish to discuss this matter further or have any more feedback to share, please feel free to contact me directly at [Manager Contact Information]. Your satisfaction is important to us, and we are committed to regaining your trust. Thank you for taking the time to share your experience. We hope to have the opportunity to welcome you back and provide you with the quality service and dining experience that you rightfully expect from us. Warm regards, [Manager Name] Manager, [Restaurant's Name]

		Human Manager	Hi (customer name), Thank you for taking the time to leave us a review. I'm so sorry to hear that we missed the mark with your experience. I would love the opportunity to connect with you further and get more details on how we can provide a better experience for our guests. Please feel free to reach out at (manager email). I look forward to hearing from you. Kindly, (restaurant name)
3	We were seated beside the open take out door. We were going to ask to change tables but after 15 minutes of sitting there still in our coats due to the cold wind, and no sign of any server, we decided to leave. I know the food in (restaurant name) is generally good, however, it is not good enough for one to put up with the lack of service in this particular location.	ChatGPT	Subject: Addressing Your Recent Visit to [Restaurant's Name] Dear Guest, I hope this message finds you well. My name is [Manager Name], and I am the manager at [Restaurant's Name]. I was disheartened to read about your recent experience at our restaurant, and I want to extend my sincerest apologies for the discomfort and service issues you encountered. Firstly, I am truly sorry that the seating arrangement near the open takeout door caused you discomfort due to the cold wind. We always aim to provide a comfortable and welcoming environment for our guests, and it seems we failed to do so in your case. I appreciate you bringing this to our attention, as it helps us to reassess our seating plan, especially during colder months or busy periods. Furthermore, the lack of prompt service you experienced is equally concerning. It is our standard to greet and attend to our guests promptly, ensuring a pleasant dining experience from the moment you are seated. The delay and lack of attention you faced are not reflective of our service standards, and I assure you that this matter will be addressed with our team to prevent such occurrences in the future. I understand that your experience has impacted your view of our restaurant, but I hope you will consider giving us another opportunity to serve you. We value your feedback and are committed to making the necessary improvements. If you are open to discussing this matter further or have any additional feedback, please do not hesitate to contact me directly at [Manager Contact Information]. Your satisfaction is important to us, and we would like to make things right. Thank you for taking the time to share your experience. We hope to have the chance to welcome you back to [Restaurant's Name] and provide you with the quality service and dining experience that you deserve. Warm regards, [Manager Name] Manager, [Restaurant's Name]

		Human Manager	Hi (customer name), my name is (manager name) and I am the lounge manager here at (restaurant name). I wanted to personally follow up and apologize for your experience at our restaurant. I would love for the opportunity to discuss this further. If you could email me at (manager email) it would be greatly appreciated. Kind Regards, (manager name)
4	Will never come again. Ordered 2 bowls of noodle waited almost an hour to found out they lost the order and had to reorder again. No apologize no nothing only excuses after excuses. Waste of time.	ChatGPT	Subject: Addressing Your Recent Experience at [Restaurant's Name] Dear Guest, I hope this message finds you well. My name is [Manager Name], and I am the manager at [Restaurant's Name]. I was very concerned to read about your recent experience at our restaurant, and I want to extend my sincerest apologies for the frustration and inconvenience you faced. Firstly, I am truly sorry to hear about the significant delay and confusion regarding your order. Waiting an extended period for a meal, only to learn that the order was lost, is understandably disappointing and not the level of service we aim to provide. I assure you that this incident is being taken very seriously, and I am personally investigating the cause of this oversight to ensure it does not happen again. I also want to apologize for the lack of a proper apology and the unsatisfactory handling of the situation when it was brought to our attention. Our guests' time and dining experience are valuable, and it is clear that we failed to acknowledge and address your concerns appropriately. This is not reflective of our customer service standards, and I will be addressing this with our entire team to improve our communication and response in such situations. I understand that this experience may have affected your perception of our restaurant, but I hope you will consider giving us another opportunity to serve you. We value your feedback as it is crucial for our improvement. If you are open to discussing this matter further or have any additional feedback, please do not hesitate to contact me directly at [Manager Contact Information]. Your satisfaction is important to us, and we are committed to making things right. Thank you for taking the time to share your experience. We hope to have the chance to welcome you back to [Restaurant's Name] and provide you with the efficient service and enjoyable dining experience that you deserve. Warm regards, [Manager Name] Manager, [Restaurant's Name]
		Human Manager	Hi (customer name), I want to apologize for the poor service. I'd love to find out more about your visit such as the date and time that you visited us so that we can make sure that we improve our service in the future. Please email us at (restaurant email) with any additional feedback. Thank you again for the review.

5	Human Manager	Human Manager	Dear Valued Guest, I hope this message finds you well. My name is [Manager Name], and I am the manager at [Restaurant's Name]. I recently came across your review, and while I am pleased to hear that you found many aspects of your visit amazing, I am truly concerned about your experience with our service staff. First and foremost, I want to extend my sincerest apologies for any discomfort or pressure you felt regarding tipping. It is our policy to provide professional, courteous service without making our guests feel obligated or coerced in any way, especially when it comes to gratuities. The behavior you described is not in line with our standards, and I assure you that this matter will be addressed promptly and appropriately with the staff member involved. We deeply value the feedback of our guests as it helps us to improve and ensure that we are providing the best possible experience. I am grateful that you brought this issue to our attention, and I want to reassure you that we are taking your comments seriously. If you are open to discussing this matter further, or if there is anything else you would like to share about your visit, please do not hesitate to contact me directly at [Manager Contact Information]. Your satisfaction and comfort are of utmost importance to us, and we are committed to making things right. Thank you for taking the time to share your experience. We hope to have the opportunity to welcome you back to [Restaurant's Name] and provide you with an experience that not only meets but exceeds your expectations. Warm regards, [Manager Name] Manager, [Restaurant's Name]
		Human Manager	Hi (customer name) , Thank you for taking the time to leave us a rating. We take feedback very seriously and want to know where we missed the mark to ensure an outstanding experience for your next visit. If you could email me at (manager email), I would like to discuss your experience with us further. We look forward to hearing from you soon, (Restaurant name)

6	My friend who is pregnant had found a piece of iron scrubber in her mouth. We told the server right away, the first word she told us “That’s a herb” BUT IT WAS NOT. It was an iron scrubber. As a server, I think she should’ve checked if my friend is okay or not. She brought it to the manager. I expected the manager to come to us as front house manager. BUT WE NEVER SEEN THE MANAGER.	ChatGPT	Subject: Addressing Your Recent Concern at [Restaurant's Name] Dear Valued Guest, I hope this message finds you well. My name is [Manager Name], and I am the manager at [Restaurant's Name]. I was deeply concerned to read about the incident you described in your recent visit to our restaurant, and I want to extend my most sincere apologies for the distress and discomfort experienced by you and your friend. First and foremost, the presence of a foreign object, such as a piece of iron scrubber, in any dish is completely unacceptable and goes against our strict standards for food safety and quality. I am truly sorry for the alarm and concern this must have caused, especially considering your friend's pregnancy. Please be assured that we take this matter extremely seriously, and I am personally investigating how this incident occurred to ensure it is never repeated. Regarding the response from our server and the lack of follow-up from a manager, I am equally concerned. It is our policy to prioritize the well-being and comfort of our guests above all else, especially in a situation as serious as this. The reaction you described is not in line with our training or standards of care. I apologize for the lack of appropriate response and empathy at the moment, and I assure you that this will be addressed with our entire staff to improve our procedures and sensitivity in handling such incidents. I understand that this experience has likely affected your trust in our restaurant, but I hope you will allow us the opportunity to make amends. Your and your friend's health and safety are of the utmost importance to us. If you are willing, I would greatly appreciate the opportunity to speak with you directly to discuss this matter further and to check on your friend's well-being. Please feel free to contact me at [Manager Contact Information]. Thank you for bringing this matter to our attention. We are committed to making the necessary improvements and ensuring the safety and satisfaction of all our guests. Warm regards, [Manager Name] Manager, [Restaurant's Name]
		Human Manager	Hello (customer name), I am sorry to hear that this was your experience with us. I will speak with the team as this is not how we handle situations. I do apologize that this happened to you and your friend. I do hope that you give us another chance so we can show you that we can deliver the 5 star experience you deserve. Please feel free to reach out to me directly via email or you can call me at the restaurant and I would be more than happy to make a reservation for you. (Manager name) Assistant General Manager (Manager email) (Manager number)

7	Most disrespectful and racist place we have ever been , we actually felt uncomfortable and left the restaurant. The server looked us up and down as if we are not normal people .	ChatGPT	Dear Guest, I hope this message finds you well. My name is [Manager Name], and I am the manager at [Restaurant's Name]. I am writing to you with great concern after reading your review about your recent visit to our restaurant. Please accept my deepest apologies for the experience you described, which is absolutely unacceptable and against everything we stand for. First and foremost, I want to express our unequivocal stance against any form of discrimination or disrespectful behavior. The treatment you described is deeply troubling and goes against our core values of inclusivity, respect, and hospitality. I am truly sorry that you felt uncomfortable and unwelcome – this is never the experience we want any guest to have. I take your comments very seriously and am committed to addressing this issue immediately. I will be conducting a thorough investigation into your experience, including speaking directly with the staff involved. It is crucial for us to understand what happened and to ensure that such behavior is corrected and does not occur again. I understand that we have significantly failed in providing you with the respectful and enjoyable experience you deserve. While I know that our actions cannot undo your experience, I hope you will allow us the opportunity to make amends. If you are willing, I would greatly appreciate the chance to speak with you directly to offer a personal apology and to discuss how we can make this right. Please feel free to contact me at [Manager Contact Information]. Thank you for bringing this matter to our attention. We are committed to making the necessary changes to ensure our restaurant is a welcoming and safe environment for all guests. Warm regards, [Manager Name] Manager, [Restaurant's Name]
		Human Manager	Dear (customer), Thank you for taking the time to post this review. We are most distraught to hear of your experience at our restaurant. We would like to take this opportunity to extend our sincerest apologies. Please know that this is not in line with our values, and we will be doing everything we can to address the situation. We would very much appreciate the opportunity to speak with you about your experience. We kindly ask you to contact us by email at (manager email) or by phone at (manager number). Warm regards, (manager name) (restaurant name)

8	I have never been disappointed by any dish to this extent. Very carelessly made. Pasta came uncooked dry with little to no sauce. The chicken parm was stale. Not ordering again	ChatGPT	Subject: We Are Truly Sorry: Addressing Your Recent Dining Experience Dear Valued Guest, I hope this message reaches you in good spirits. My name is [Manager Name], and I am the manager at [Restaurant's Name]. I recently came across your review, and I felt compelled to reach out to you directly to address your concerns and extend our sincerest apologies. Firstly, I want to express how truly sorry we are to hear about your disappointing experience with us. It is evident from your feedback that we fell short of the high standards we set for our food quality and overall guest experience, and for that, I am genuinely sorry. The issues you've highlighted regarding the pasta and chicken parm are concerning and certainly not reflective of the quality we strive to deliver. Please be assured that I have personally addressed these issues with our kitchen team to ensure that this was an isolated incident and to prevent such mistakes in the future. We take great pride in our food, and it is disheartening to know that we let you down. I understand that we have lost your trust, but I sincerely hope you will consider giving us another opportunity to serve you and to demonstrate the true quality of our dishes. Your feedback is invaluable to us, and we are committed to making the necessary improvements to ensure that all our guests have a delightful dining experience. If you are open to discussing your experience further, I would greatly appreciate the opportunity to speak with you directly. You can reach me at [Manager Contact Information]. Once again, I apologize for the disappointment we caused, and I thank you for taking the time to share your feedback. We are taking your comments to heart, and we are dedicated to making things right. Warmest regards, [Manager Name] Manager, [Restaurant's Name]
		Human Manager	Hi (customer name) , Thank you for taking the time to leave your review. Customer feedback is extremely important to us and I sincerely apologize that your last experience fell short on the standard we strive for at (restaurant name). I would like to have the opportunity to connect with you further. If you could please email me at (manager email) (manager name)

9C	My Coleslaw was drown in creamy slush. Potato salad was tasteless. My Sangria tasted like Jello with complete with an Aspartame after taste, which quickly went back! What a waste of $29. Make an effort you guys. Do better.	ChatGPT	Subject: Addressing Your Recent Dining Experience at [Restaurant's Name] Dear Guest, I hope this message finds you well. My name is [Manager Name], and I am the manager at [Restaurant's Name]. I recently came across your review, and I want to extend my sincerest apologies for the disappointment you experienced with your meal. Firstly, I am truly sorry to hear that the coleslaw and potato salad did not meet your expectations. We aim to delight our guests with every dish, and it is clear that we missed the mark in your case. Your feedback about the excessive creaminess of the coleslaw and the lack of flavor in the potato salad is concerning, and I assure you that I will personally discuss this with our culinary team to review our recipes and preparation methods. Regarding your experience with the Sangria, I am sorry that it did not provide the enjoyable and refreshing experience we intend. We take pride in our beverage offerings, and it is disheartening to learn that it fell short, especially in terms of taste. I will be sharing your comments with our bar staff to reassess the ingredients and preparation of our Sangria. We value your feedback as it is crucial for our improvement. I would greatly appreciate the opportunity to discuss your experience in more detail and to learn more about how we can make amends. Please feel free to contact me directly at [Manager Contact Information]. Your satisfaction is important to us, and we are committed to making the necessary changes to enhance our guests' dining experiences. Thank you for taking the time to share your experience. We hope to have the opportunity to welcome you back and provide you with the quality dining experience that you rightfully expect from us. Warm regards, [Manager Name] Manager, [Restaurant's Name]
		Human Manager	Hi (customer name), My name is (manager name) and I am the Restaurant Manager at (restaurant name). I wanted to personally reach out and apologize for the lack of food and drink quality you experienced at my restaurant and want to assure you this is not the standard we strive to operate at. I would love for the opportunity to chat more about your experience over the phone so I can follow up with all the appropriate people. If you could email me at (manager email) it would be greatly appreciated. Kindest Regards, (manager name)

10	We ordered: spicy chicken - came cold on a small plate Warm potato salad - straight fridge cold. Tasted like cardboard Spinach and artichoke dip and the butter- they were OUT of these items. However we weren't informed of this until the last of our "food" was brought out Management was apologetic but our overall experience and dinner was ruined. We were super hungry and hoping for a solid meal. Should've gone to another establishment :(	ChatGPT	Subject: Our Apologies for Your Recent Dining Experience at [Restaurant's Name] Dear Guest, I hope this message finds you well. My name is [Manager Name], and I am the manager at [Restaurant's Name]. I am writing to you after reading your review about your recent dining experience with us, and I want to extend my sincerest apologies for the series of disappointments you encountered. Firstly, I am truly sorry to hear that the dishes you ordered, particularly the spicy chicken and warm potato salad, did not arrive at your table as expected. Serving cold food that is meant to be warm is not in line with our standards, and I understand how this significantly detracted from your dining experience. Please be assured that I will be addressing this with our kitchen staff to ensure that all dishes are prepared and served at the appropriate temperatures. I am also concerned to learn about the unavailability of the spinach and artichoke dip and the butter, and that you were not informed of this until later in your meal. It is our responsibility to keep our guests informed about menu changes and availability, and we clearly failed in this regard during your visit. I will be speaking with our team to improve our communication and ensure that our guests are promptly and accurately informed about any menu changes. While I am glad to hear that our management team was apologetic, I understand that this does not change the fact that your overall experience was not what you expected or deserved. We value your feedback as it is crucial for our improvement. I would greatly appreciate the opportunity to discuss your experience in more detail and to learn more about how we can make amends. Please feel free to contact me directly at [Manager Contact Information]. Your satisfaction is important to us, and we are committed to making the necessary changes to enhance our guests' dining experiences. Thank you for taking the time to share your experience. We hope to have the opportunity to welcome you back and provide you with the quality dining experience that you rightfully expect from us. Warm regards, [Manager Name] Manager, [Restaurant's Name]
		Human Manager	Hi (customer name), Thank you for sharing your review. I am sincerely sorry that your experience with us was a negative one. That is definitely not our standard and I will follow up with my kitchen team to find out what happened, and ensure this never happens again. Please feel free to reach out to us anytime at (restaurant email). We would love to hear more of your feedback and make this right. Thanks again, Sincerely, (restaurant name)

Appendix B

Reviews and Responses Used in Study 2

Table B1. Reviews and Responses Used in Study 2

ID	Review	Respondent	Response
1	The restaurant is not very child-friendly at all. When kids go there, they feel discriminated against and stared at. I brought homemade baby food for my one-year-old daughter, while the other adults ordered regular food, and the manager said outside food was not allowed. Excuse me? If children aren't allowed to dine here, they should make that clear when making reservations.	ChatGPT	Dear Guest, Thank you for taking the time to share your experience with us. I'm very sorry to hear that your visit to our restaurant was not as pleasant as you had hoped, especially concerning the treatment of your child and the policy on outside food. We strive to create a welcoming environment for all our guests, including families with young children, and I regret that you felt discriminated against. Your feedback highlights an important area where we need to improve our communication and sensitivity towards families. Regarding the policy on outside food, it is standard for many restaurants to restrict outside food for health and safety reasons. However, we should have handled your situation with more understanding and flexibility, especially considering the specific needs of feeding a young child. I apologize for the lack of clarity and the inconvenience this caused. We value your feedback as it helps us identify where we need to make changes to better accommodate all our guests. I will review this policy with our team to ensure we approach such situations with greater empathy and clearer communication in the future. Thank you again for bringing this to our attention. We hope to have the opportunity to serve you and your family again, providing a more welcoming and inclusive experience. Warm regards, [Manager Name] Restaurant Manager [Restaurant Name]

		Human Manager	To whom it may concern, Thank you for visiting us providing us with the feedback. As there is a word limit to how much I can type here may I have you reach out to me where I can make up to you. As of the incident that happened that afternoon, I can assure you that we do allow kids here and the staff that spoke to you will be reprimanded and re-trained properly. I would also like to re-invite you to another dining experience in the future and take care of you personally to make up to you and your group that night, Please send me an email with a contact number at (manager email) "Attn to (manager name)" and I hope to speak to you soon. Warmest Regards, General Manager (manager name)
2	I always like the food and the vibe at (restaurant name). But the lining up system is a mess. After lining up, the wait time jumped to one hour later. And even when there was no one in front me online and we were the only party waiting at the door, they still wouldn't seated us with a few tables open. Don't recommend this location.	ChatGPT	Dear valued guest, Thank you for taking the time to share your feedback regarding your recent experience at [restaurant name]. We are pleased to hear that you have enjoyed our food and the ambiance we strive to create for our guests. Your satisfaction is our top priority. However, I am genuinely disheartened to learn about your experience with our seating and queue management. It is certainly not our intention to make our guests feel unwelcome or to have them endure unnecessary wait times. I apologize for any inconvenience this may have caused you. We are continuously working to improve our service and your feedback is invaluable in this process. I will personally look into the concerns you've raised regarding our lining up system and the seating protocol. We aim to ensure that our guests receive prompt and courteous service, and your comments will help us to refine our processes to better meet those standards. While I understand your disappointment, I hope you will consider visiting us again. We value your patronage and would welcome the opportunity to serve you again, hopefully under more favorable circumstances. Thank you once again for your feedback. Please feel free to reach out to us directly if you have any further comments or suggestions. We look forward to the opportunity to regain your trust and deliver the quality experience you expect and deserve from [restaurant name]. Warm regards,

			[Manager Name] Restaurant Manager [Restaurant Name]
		Human Manager	Hi (customer name), thank you for your feedback. Apologies on the confusion about our waitlist. We do very much appreciate all of our guests time as we know that we get very busy up to 2 hours on week days and sometimes over 3 hours on the weekend. The idea behind the (review service site) waitlist is so that you can add your name to the waitlist from at home and our hosts will message you once your table is ready so you do not have to wait around. It does happen that some guests decide to cancel thus benefiting others on the waitlist or sometimes not showing up to their scheduled wait list time which can cause for empty tables. However, we do honor a 10 minute grace period for all parties once they have been notified of their table being ready. Should a party be late, we do restore the party to the top of the waitlist which can cause for the waitlist times to vary for others in line unfortunately. If you wish to discuss this furthermore, I would be happy to make the time to explain in more detail about how our (review service site) online waitlist works. Please do not hesitate to contact me directly, (manager email). Best Regards

Appendix C

Example of ChatGPT Training Results

Table C1. Example of ChatGPT Training Results

Text Category	Content
Original Review	we had a very poor experience of [restaurant name] and in fact did not eat there at all. We joined the queue outside last Saturday and patiently waited as slowly it moved towards the door. We assumed that the host would take our name as we got closer. However, we discovered after about 30 minutes that the position in the queue was not what counted and that we should have entered the restaurant at the start to put down our name. When we discovered that, we spoke to the host who told us there was a 90 minute wait. Needless to say we didn't wait but better communication of the system would have really helped ....
Human Response	Firstly, thanks for taking the time to provide feedback. It is very important to us and I am grateful for it. I am sorry you did not inquire first as our waitlist policy is clearly stated on our entrance window and in addition it is our mandate to come outside every 5-10 minutes to ensure if there is a queue it is kept orderly so that neighbours and passers by are not infringed upon and also to make sure that all those waiting actually have their names on our list as sometimes people do actually assume that they should just queue instead of inquiring first either with us or those that are in the queue. We also encourage those who do have their names on our list to meander and return at a realistic time for their table to be ready. I am sorry for what transpired for you that day as it inhibited you from actually have the experience you should have. I hope that if you are still in town you do come back and give us another try. Kindest regards, [manager name]
Policies Extracted from Human Response	The waitlist policy is stated on the entrance window. It is the restaurant's mandate to come outside every 5-10 minutes to ensure if there is a queue it is kept orderly so that neighbours and passers by are not infringed upon and also to make sure that all those waiting have their names on list. The restaurant encourages those who do have their names on wait list to meander and return at a realistic time for their table to be ready.
Untrained ChatGPT Response	Dear [Reviewer's Name], Thank you for sharing your feedback regarding your recent attempt to dine at [Restaurant Name]. I deeply regret to learn that your experience with our waiting system fell short of your expectations, and I apologize for any inconvenience this may have caused. I understand your frustration with the confusion surrounding our queue process. It's certainly not our intention to leave our guests feeling uncertain or overlooked. We strive to create a welcoming environment where every guest feels valued and well-informed from the moment they arrive.

	We are taking your feedback very seriously. Rest assured, we will be reevaluating our queue management procedures to ensure they are clear and effective for all our guests. This includes improving our communication at the entrance to provide better guidance on how to secure a table, particularly during busy periods like Saturdays. As part of our commitment to continuous improvement, we will also discuss implementing additional measures such as improved signage and clearer instructions from our hosts. Our goal is to enhance the overall experience for all our patrons and prevent similar situations from occurring in the future. I sincerely apologize once again for the disappointment you experienced. We value your input, as it helps us identify areas where we can improve. I hope you will consider giving us another opportunity to demonstrate the level of service and hospitality that we are known for. Thank you for bringing this matter to our attention. Should you wish to discuss your experience further or provide additional feedback, please feel free to reach out to me directly at [Your Contact Information]. Warm regards, [Your Name] Manager, [Restaurant Name]
Trained ChatGPT Response	Dear [Reviewer's Name], Thank you for taking the time to share your feedback regarding your recent visit to [Restaurant Name]. We sincerely apologize for the inconvenience and confusion you experienced while waiting to be seated. We understand how frustrating it must have been to wait for an extended period only to discover that your name was not on the waitlist. Our waitlist policy is stated on the entrance window, and it is our mandate to come outside every 5-10 minutes to ensure the queue is kept orderly and to verify that all those waiting have their names on the list. This helps us manage the waitlist efficiently and ensure a fair process for all guests. We regret that this procedure was not communicated to you clearly and that our host did not approach you sooner to confirm your place on the list. Your feedback highlights an area where we need to improve our communication and ensure that all guests are informed about the waitlist process as soon as they arrive. We are taking your comments seriously and will work with our team to enhance our communication with guests waiting outside. Our goal is to provide a seamless and enjoyable experience for everyone who visits us, and we appreciate your input in helping us achieve that. We hope you will consider giving us another opportunity to serve you in the future. If you have any further concerns or suggestions, please do not hesitate to reach out. Kind regards, [Your Name] [Restaurant Name] Management Team

Appendix D

Reviews and Responses Used in Study 3

Table D1. Reviews and Responses Used in Study 3

ID	Review	Respondent	Response
1	we had a very poor experience of [restaurant name] and in fact did not eat there at all. We joined the queue outside last Saturday and patiently waited as slowly it moved towards the door. We assumed that the host would take our name as we got closer. However, we discovered after about 30 minutes that the position in the queue was not what counted and that we should have entered the restaurant at the start to put down our name. When we discovered that, we spoke to the host who told us there was a 90 minute wait. Needless to say we didn't wait but better communication of the system would have really helped	ChatGPT	Dear [Reviewer's Name], Thank you for sharing your feedback regarding your recent attempt to dine at [Restaurant Name]. I deeply regret to learn that your experience with our waiting system fell short of your expectations, and I apologize for any inconvenience this may have caused. I understand your frustration with the confusion surrounding our queue process. It's certainly not our intention to leave our guests feeling uncertain or overlooked. We strive to create a welcoming environment where every guest feels valued and well-informed from the moment they arrive. We are taking your feedback very seriously. Rest assured, we will be reevaluating our queue management procedures to ensure they are clear and effective for all our guests. This includes improving our communication at the entrance to provide better guidance on how to secure a table, particularly during busy periods like Saturdays. As part of our commitment to continuous improvement, we will also discuss implementing additional measures such as improved signage and clearer instructions from our hosts. Our goal is to enhance the overall experience for all our patrons and prevent similar situations from occurring in the future. I sincerely apologize once again for the disappointment you experienced. We value your input, as it helps us identify areas where we can improve. I hope you will consider giving us another opportunity to demonstrate the level of service and hospitality that we are known for.

			Thank you for bringing this matter to our attention. Should you wish to discuss your experience further or provide additional feedback, please feel free to reach out to me directly at [Your Contact Information]. Warm regards, [Your Name] Manager, [Restaurant Name]
		Human Manager	Firstly, thanks for taking the time to provide feedback. It is very important to us and I am grateful for it. I am sorry you did not inquire first as our waitlist policy is clearly stated on our entrance window and in addition it is our mandate to come outside every 5-10 minutes to ensure if there is a queue it is kept orderly so that neighbours and passers by are not infringed upon and also to make sure that all those waiting actually have their names on our list as sometimes people do actually assume that they should just queue instead of inquiring first either with us or those that are in the queue. We also encourage those who do have their names on our list to meander and return at a realistic time for their table to be ready. I am sorry for what transpired for you that day as it inhibited you from actually have the experience you should have. I hope that if you are still in town you do come back and give us another try. Kindest regards, [manager name]
		ChatGPT-Trained	Dear [Reviewer's Name], Thank you for taking the time to share your feedback regarding your recent visit to [Restaurant Name]. We sincerely apologize for the inconvenience and confusion you experienced while waiting to be seated. We understand how frustrating it must have been to wait for an extended period only to discover that your name was not on the waitlist. Our waitlist policy is stated on the entrance window, and it is our mandate to come outside every 5-10 minutes to ensure the queue is kept orderly and to verify that all those waiting have their names on the list. This helps us manage the waitlist efficiently and ensure a fair process for all guests.

			We regret that this procedure was not communicated to you clearly and that our host did not approach you sooner to confirm your place on the list. Your feedback highlights an area where we need to improve our communication and ensure that all guests are informed about the waitlist process as soon as they arrive. We are taking your comments seriously and will work with our team to enhance our communication with guests waiting outside. Our goal is to provide a seamless and enjoyable experience for everyone who visits us, and we appreciate your input in helping us achieve that. We hope you will consider giving us another opportunity to serve you in the future. If you have any further concerns or suggestions, please do not hesitate to reach out. Kind regards, [Your Name] [Restaurant Name] Management Team
2		ChatGPT	Dear [Reviewer's Name], Thank you for sharing your feedback about your recent visit. I am truly sorry to hear about your experience and appreciate you bringing this to our attention. Our goal is to provide exceptional service to all our guests, and it's disappointing to learn that we fell short during your visit. We understand that not having a reservation can sometimes result in limited availability, but it’s important to us that every guest feels valued and accommodated to the best of our abilities. I apologize for any inconvenience you faced and assure you that your comments will be shared with our team to ensure better handling of similar situations in the future. We strive to balance our reservations with walk-in guests, and we will work on improving our approach to make sure all guests feel welcome. We hope to have the opportunity to serve you in the future and provide a more positive experience.

	Make reservations or feel low life customer ? My partner and I did not make reservations. We knew going in we might not get a table. Our intention was to have a quick happy hour experience and be out to finish our shopping. We arrived at 4pm with only 4 tables occupied for entire restaurant. The hostess working that night blew us off by telling us they were fully reserved, starting at 430. I find it hard to believe that all reservations were showing up at once. Most likely staggering 430, 500, 530, and could have offered us something if she were willing to juggle but she made it abundantly clear she couldn't be bothered. We went next door and they managed to fit us in despite reservations (later) and we ate and left. Too bad they couldn't bother trying.		Warm regards, [Your Name] Restaurant Manager [Restaurant Name] [Restaurant Name]
		Human Manager	Dear Guest, Thank you for your review. I'm sorry to hear you were unable to join us Saturday, [review date]. With it being the busy holiday season we are quite often fully committed with reservations most days due to large parties for holiday events reserving multiple tables in advance, especially on weekend evenings. Due to that factor a lot of our tables are reserved from 4pm onwards. Our hostess team is very accommodating and will always try move things around if at all possible. On that particular evening we had been fully reserved for about a week. We do recommend making a reservation for a weekend evening, especially in December or a holiday weekend, as far ahead as possible, but do take walk in's when we are able to. I apologize we were unable to host you on that evening. Sincerely, [manager name]
		ChatGPT-Trained	Dear [Customer Name], Thank you for sharing your experience with us. We apologize for the disappointment and frustration you felt during your recent visit to [Restaurant Name]. During busy seasons, our restaurant is often fully committed with reservations, especially on weekend evenings. Many large parties reserve multiple tables in advance, which can make it challenging to accommodate walk-in guests. While our hostess team always strives to move things around when possible, it appears that we were unable to do so during your visit. We understand your perspective and regret that we couldn't provide a table for your quick happy hour experience. Your feedback is important, and we will use it to improve our guest experience and better manage our reservation system. We recommend making a reservation ahead of time, particularly on weekends and during busy seasons, to ensure a more seamless experience. We hope you will consider giving us another chance in the future.

			Best regards, [Your Name] [Restaurant Name] Management Team

Appendix E

Randomly Selected Samples of Dataset used in Chapter 3

Table D1. Reviews and Responses Used in Study 3

Review	ChatGPT-Generated Response	Gemini-Generated Response	Human Response
I got an arc'teryx jacket and after washing it, the glue came off from the entire jacket. I couldn't believe my eyes. It looks as if it's made of paper. Such an expensive brand with such a low quality. I contacted the customer service and they would not do anything about it. The company only offered 40% off from my next purchase after I send in my old jacket. The entire store is offering 40% off!!! It's not that big of a favor to me. Do not buy from this brand, you never know if you'll get a good jacket or you'll get a bomb. It is like playing with Russian roulette.	Thank you for taking the time to share your feedback with us. I'm deeply sorry to hear about your experience with your Arc'teryx jacket. We strive to maintain the highest quality in our products, and it's clear we fell short in this instance. I understand how frustrating it must have been to find the glue coming off after washing, especially given the expectations associated with our brand. Your experience does not reflect the standards we aim to uphold. Regarding your interaction with our customer service, I apologize if the resolution offered did not meet your expectations. We always aim to provide solutions that fully address our customers' concerns. I would like to personally assist in resolving this issue to your satisfaction. Could you please reach out to me directly at [contact information] so we can discuss this further and find a more suitable resolution for you? We value your feedback and would like to learn from this to ensure better experiences for all our customers in the future. Thank you for bringing this to our attention and giving us the opportunity to make things right.	Thank you for taking the time to share your experience with your Arc'teryx jacket. I'm incredibly sorry to hear that the glue came off after washing and that you're disappointed with the quality. Arc'teryx is known for its high-performance garments, and this clearly falls short of those standards. I understand your frustration with the initial customer service response. A 40% discount on a new jacket wouldn't be a fair resolution for a defective product. I'd like to offer you a full refund for your jacket. We can arrange for a prepaid return label for your convenience. Additionally, I'd like to personally investigate what happened with your jacket. This will help us ensure such an issue doesn't occur again. If you're open to it, could you please send me a direct message with your order details or a copy of your receipt? This will expedite the refund process. We value your business and hope this experience doesn't deter you from Arc'teryx in the future.	Thank you for taking the time to provide us with your review. We're sorry to hear about the issue you're experiencing and we hope we can provide further details regarding your case. Gore-Tex is a very durable and functional material, however, it does require regular care and maintenance to ensure it will perform well and last a long time. If a Gore-Tex jacket is not washed regularly with technical detergent, then sweat and body oils can build up over time and contaminate the material. This contamination can lead to issues like the jacket not repelling water like it once did, and internal seams coming loose. We stand behind the quality of our products and our team of experts who assess the condition of items experiencing issues. We do our best to keep products functional for many years, however, in some cases there are no repair options available. We hope that information helps clarify this situation, and please feel free to contact our Customer Service team with any questions or concerns.

Unfortunately I have very bad experience with this company. I bought bus tickets from NY to Washington DC, using Busbud.Bus company-Greyhound Inc. Arriving at bus terminal in NY, I was announced by Greyhound representatives at front desk, that my tickets are fake, with false reference No.and that Busbud is not their ticket agent/partner. I had to buy new tickets. After contacting Busbud, to understand what was what, I received long explanation letters, very caring attitude, like some mistake has happened in Greyhound side. Greyhound does not comment it anyhow, after 2 information requests there is zero feedback. I don't want to spend by time, writing or calling several times, investigating, who is liar or not. So just be careful-always buy tickets from direct bus company sites, not using third parties.	We sincerely apologize for the inconvenience you experienced with your recent ticket purchase through Busbud for travel with Greyhound Inc. It's disheartening to hear about your negative experience, and we understand the frustration caused by the unexpected situation at the bus terminal in New York. We take matters like these very seriously, and we have investigated the issue further. While we cannot speak on behalf of Greyhound Inc., we can assure you that we are committed to resolving any discrepancies that may have occurred with your ticket purchase. We strive to maintain a high standard of service and accountability, and we regret any confusion or inconvenience this may have caused you. We appreciate your feedback and understand your concerns about purchasing tickets through third-party platforms. Rest assured, we are working diligently to address this issue and prevent similar incidents from happening in the future. If you have any further questions or concerns, please don't hesitate to reach out to us directly. We value your business and hope to have the opportunity to restore your faith in our services.	Thank you for bringing this unfortunate situation to our attention. We sincerely apologize for the incredibly frustrating experience you had with your bus tickets from NY to Washington DC. We understand how stressful it must have been to be told your tickets were fake at the terminal, especially after purchasing them through Busbud. While we can't speak for Greyhound, we take full responsibility for ensuring our customers receive valid tickets. Here's what we can do to help: * **Investigate the Issue:** We'll immediately launch an investigation into what happened with your tickets. We take such matters very seriously and will work to prevent similar issues from occurring in the future. * **Reimbursement:** We'll ensure you're fully reimbursed for the new tickets you had to purchase at the Greyhound terminal. Please reach out to us directly at [phone number] or [email address] with a copy of your new ticket receipt, and we'll process your refund promptly. We understand your frustration with having to deal with multiple parties. We hope by taking these steps, we can regain your trust and demonstrate our commitment to providing a smooth travel experience.	we are really sorry about this issue with your tickets, thank you for letting us know that you are still waiting for a refund. We are officially working with Greyhound, and we are selling tickets for customers traveling across the US and Canada every day. All our tickets are valid and generated by Greyhound, they are exactly the same as if you buy them on their site. As you can see on your credit card statement, you were charged by Greyhound directly. Could it be that you showed the email instead of the ticket? Greyhound is very strict with their boarding requirements. I'm really sorry that they were not able to help you at the station when you asked for an exchange of your tickets. I understand that this situation is really frustrating for you and this is not the experience of travel we want to bring to our customers. I sent you an email this morning with more information to solve this issue with you as we want to make it right. If you have any questions don't hesitate to contact me back.

Be very careful of the "extras" including of pre-purchase of miles/kilometers. Make sure you check that you have everything you need, in our case, two adults received one pillow for our journey. If you have no experience with RV rental or use of an RV or boat do not expect proper orientation/instruction. The Halifax, Nova Scotia office is a complete mess, confusion and lack of organization specifically attributable to the manager(?) Willie who is also rude/aggressive/nasty. He has the very lovely ladies working there running in circles rather than focusing in on the task at hand. Terrible management.	Thank you for bringing your concerns to our attention. I sincerely apologize for the issues you encountered with our services. We take such feedback seriously and are committed to addressing these concerns promptly. Regarding the pre-purchase of miles/kilometers and the amenities provided, we understand how important these are for a comfortable and hassle-free journey. We will review our inventory procedures to ensure that all necessary items, like pillows and other essentials, are adequately supplied and provided to our customers. We also recognize the importance of proper orientation and instruction for those new to RV rentals. We will enhance our orientation process to ensure all customers, regardless of their experience level, receive comprehensive and clear instructions to make their rental experience smoother and more enjoyable. Your feedback about the Halifax, Nova Scotia office is particularly concerning. I apologize for the confusion and disorganization you experienced. We are currently reviewing our management practices and staff training to improve efficiency and customer service. Specific feedback about individuals is valuable, and we will address the behavior and management style you described to prevent such occurrences in the future. Thank you for your honest feedback. We are committed to making the necessary improvements to ensure a better experience for all our customers.	Thank you for your feedback regarding your recent RV rental experience at our Halifax, Nova Scotia location. We sincerely apologize for the issues you encountered, particularly with: * **Limited "Extras" Explanation:** We understand your concern about transparency regarding add-on options like pre-purchased miles/kilometers. We'll be reviewing our communication materials and staff training to ensure clear explanations of all additional charges and inclusions. * **Missing Supplies:** We apologize that you only received one pillow for two adults. Proper inventory checks should have prevented this oversight. We'll be revising our pre-rental inspection procedures to ensure all necessary items are included in each RV. * **Lack of Orientation for New RV Users:** We acknowledge that renters unfamiliar with RVs may require additional guidance. We'll be reevaluating our orientation process at the Halifax location to ensure all customers, regardless of experience level, receive comprehensive instructions on RV operation and safety protocols. * **Halifax Staff Performance and Management:** We take your comments about confusion, disorganization, and a rude/aggressive manager (Willie) very seriously. This falls far short of our customer service standards. We'll be investigating this matter thoroughly, including speaking directly with Willie and the staff at the Halifax location. We're committed to addressing any identified issues and ensuring a professional and helpful environment for all our customers.	We appreciate all reviews, both positive and negative as it enables us to improve the service we offer to future Guests. In your case, we are a little confused, as the drop-off satisfaction survey you or your wife completed when you brought the RV back to Halifax shows your overall impression as Excellent confirms you would recommend us to family and friends. The notes in your booking do indicate that we were unable to refund you for the 300 km you prepaid for and didn't use, however it is clearly outlined in all correspondence we have with Guests (and on our website) that we do not refund for unused kilometres. Our Manager was simply adhering to this policy when he explained he was unable to provide you with a refund.You received only one pillow because your booking indicated only one person was travelling. Kits were included in your package at no additional cost and we would have been happy for you to purchase an additional pillow and reimburse you for it at drop-off with the receipt.Our Halifax manager is regularly complimented for his personality and management style so we are sorry if this wasn't your experience. One has only to read the positive online reviews on our Halifax location to conclude that your experience was not shared by others.We thank you for taking advantage of one of our special deals to experience Nova Scotia at your own pace and hope that, despite your dissatisfaction at drop-off, you enjoyed exploring the Canadian Maritimes. It was a pleasure having you as our Guests.

		We appreciate your feedback about the "lovely ladies" working there. We value our staff and will ensure they have the proper support and direction to effectively assist renters. We strive to provide a positive and well-equipped RV rental experience. We clearly fell short in your case. We would like to regain your trust. Please contact us directly at [insert contact information] to discuss how we can compensate you for the inconvenience you experienced.	