



UNIVERSITY OF TWENTE.

Faculty of Electrical Engineering,
Mathematics & Computer Science

Empathic chatbot for complaint handling in customer service

Babiche L. Pompe
Master Thesis
June 2023

Supervisor:
dr. M. Theune
dr. S. Borsci

Research group Human Media Interaction (HMI)
Faculty of Electrical Engineering,
Mathematics and Computer Science
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Abstract

Empathic chatbots for complaint handling in customer service

by Babiche POMPE

This research investigates the impact of empathy in chatbots on customers' perceived interactional justice during service recovery. Interactional justice refers to the perceived fairness of inter-personal behaviour, influenced by aspects such as the level of empathy of an employee, during service recovery. To this end, a complaint handling chatbot was developed based upon a combination of existing models of complaint handling and empathy for robots. A mixed-methods between-participant study was conducted with 25 participants, divided into two conditions: one interacting with an empathic chatbot (n=13) and the other with a non-empathic chatbot (n=12). The empathic chatbot includes several different empathic strategies, including use of empathic language, empathic intent mapping, emotion validation, and empathic questions. Using sentiment analysis to extract the user's dominant emotion from text, the chatbot is able to map its response to an appropriate empathic intent. It can also provide personalized responses based upon knowledge obtained through entity extraction from the conversation, increasing the overall feeling of empathy. The overall flow of the complaint conversation was designed using a decision tree, and a custom intent classifier was incorporated to optimize the chatbot's functionality. The findings showed that customers interacting with an empathic chatbot reported higher levels of interactional justice compared to those interacting with the non-empathic chatbot. Participants also expressed feeling more understood and helped by the empathic chatbot. Additionally, participants reported a more positive overall perception of the chatbot in the empathic condition, highlighting the possible positive effect of empathy on user experience. This research contributes to the understanding of the impact of empathy in chatbots for successful service recovery. By implementing empathic strategies, companies can enhance customer's perceived interactional justice, leading to increased customer satisfaction, loyalty, and overall user experience. Also, the proposed frameworks derived from the literature review could provide a foundation for future research in the field. However, this research also has its limitations. The study sample size is relatively small, which may impact the generalizability of the findings. In the future, studies may repeat the research with a larger sample size and in an offline setting, explore the relationship between empathy and user engagement, and investigate the direct link between chatbot empathy and emotion regulation. In conclusion, this study has revealed that empathy in chatbots positively influences customers' perceived interactional justice.

Acknowledgements

First and foremost I would like to thank my supervisors for their constant support and expertise. I would like to express my heartfelt gratitude to dr. Mariët Theune, who not only supported me through academic guidance and insightful comments on my work, but also served as a mentor with immense kindness and consideration. Without your constant support and mentorship this research would not have been the same. Also, I want to thank you for your great attention to detail when providing feedback on my research, and consistently pointing out my misuse of semicolons. I think all of this is a testament to your exceptional guidance and expertise, and I am very grateful to have had you as my supervisor for my Research Topics and final thesis.

I am also extremely grateful to dr. Simone Borsci, whose contributions, feedback and expertise have enriched my research to what it is now. Your new insights helped me better form my research and take into consideration new angles which I had not thought about before. Also, your support, positive comments and enthusiasm during meetings did not go unnoticed and helped me to stay motivated and confident about my research. I am thankful for your willingness to offer so much guidance to a student outside of your own faculty, and am very grateful for all of the support and expertise that you have provided me with during this process.

I also want to thank my family for their encouragement and belief in my abilities, which have been a constant source of motivation for me. I am especially grateful that they occasionally forced me to take a break or come home to relax which helped me renew my energy and motivation to continue. I am truly fortunate to have such a caring and supportive family by my side. Also, many thanks to my close friends, who have constantly helped me stay motivated and supported me when I had doubts or questions. Not only did they constantly remind me of their belief in my abilities, but they also reminded me that setbacks during a thesis were normal and convinced me to go for walks or coffee to help me take a break.

I also want to thank all of the participants who generously volunteered their time and provided me with their insights to this study. Their involvement has provided me with a substantial amount of interesting findings and improved the overall quality and relevance of this research. Also, the fun interactions and comments energized me and reassured me of the relevance of my research. Lastly, I want to thank dr. Alma Schaafstal for advising me to contact Mariët about a possible thesis research. Without this tip this research would likely not have happened.

Contents

Abstract	i
Acknowledgements	ii
List of Figures	vi
List of Tables	viii
1 Introduction	1
2 Context	3
2.1 History	4
2.2 Benefits and limitations	5
2.2.1 Benefits	5
2.2.2 Limitations	7
2.3 Chatbots in customer service	8
2.3.1 Uses of chatbots	8
2.3.2 Customer service	9
2.4 Conclusion	10
3 Theory	11
3.1 Complaint handling	11
3.1.1 Importance of satisfied customers	11
3.1.2 Definitions of complaint handling	12
3.1.3 Benefits of successful complaint handling	12
3.1.4 Effective complaint handling	12
3.2 Empathy in complaint handling	14
3.2.1 Empathy	14
3.2.2 Empathy related to complaint handling	15
3.3 Measurement methods	17
3.3.1 Empathy	17
3.3.2 Perceived interactional justice	18
3.4 Conclusion	19
4 Related work	21
4.1 Empathic conversational agents	21
4.1.1 Neural empathic agents	22
4.1.2 Framework-based empathic agents	24
4.2 Complaint handling	30
4.2.1 Complaint handling process	30
4.2.2 Internal complaint mechanisms	31
4.2.3 Concluding remarks	33
4.3 Conclusion	34

5	Method	35
5.1	User tests	36
5.2	Research Design	36
5.2.1	Scenario	36
5.2.2	Variables	37
5.3	Procedure	37
5.3.1	Step-by-step	37
5.3.2	Data collection	38
5.4	Measures	39
5.4.1	Quantitative Measures	39
5.4.2	Qualitative Measures	39
5.5	Participants	40
5.5.1	Sample group distribution	40
5.5.2	Ethical considerations	40
5.6	Limitations	41
5.7	Conclusion	41
6	Realisation	42
6.1	Design	42
6.1.1	Empathy framework	42
6.1.2	Complaint handling framework	45
6.1.3	Requirements List	46
6.2	Implementation	47
6.2.1	Non-empathic chatbot	48
6.2.2	Empathic chatbot	49
6.2.3	Response dataset	49
6.2.4	Dialogue structure	50
6.2.5	Models	52
6.2.6	User interface	53
6.3	User Testing	53
6.3.1	Prototype 1	54
6.3.2	Prototype 2	54
6.4	Conclusion	55
7	Results	56
7.1	Quantitative results	57
7.1.1	Perceived interactional justice	57
7.1.2	Perceived empathy	57
7.1.3	Perceived usability	58
7.1.4	Interaction dynamics	59
7.2	Qualitative results	62
7.2.1	Perception of interaction	62
7.2.2	Realism of scenario and chatbot behaviour	64
7.2.3	Comparison to human customer service	65
7.3	Relations qualitative and quantitative results	66
7.4	Conclusion	67
8	Discussion	68
8.1	Analysis of results	68
8.1.1	Effect empathy on user experience	68
8.1.2	Effect empathy on user behaviour	69

8.1.3	Effect empathy on user engagement	70
8.1.4	Implications of results	70
8.2	Limitations	71
8.3	Future Work	72
9	Conclusion	74
	References	75
A	Implementation details	84
A.1	Models	84
A.1.1	Emotion recognizer	84
A.1.2	Intent classifier	84
A.1.3	Entity extraction	85
A.2	User interface	85
B	Complaint handling framework	86
C	Complete decision tree	87
D	Experiment Instructions	90
D.1	Informatie voor deelnemers (NL)	90
D.2	Information for participants (ENG)	91
E	Extra (Fallback) Response Intents	92

List of Figures

2.1	General chatbot architecture based on design by Adampoulou and Moussiades [5]	3
2.2	Number of published documents related to "chatbot", "conversation agent", "conversational agent", and "dialogue system" on Scopus from 2000 to 2022.	5
3.1	Overview of complaint handling concepts	17
3.2	Items of RoPE Scale as proposed by Charrier et al. [76]	18
3.3	Items for measuring perceived interactional justice in complaint handling as proposed by Homburg et al. [52]	19
3.4	Items for measuring perceived interactional justice as proposed by Waqas et al. [77]	19
4.1	Models for an empathic conversational agent as designed by Yalçın and DiPaula [84]	26
	(a) Information flow	26
	(b) Framework of empathy mechanism	26
4.2	Empathy model diagram as proposed by Rodrigues et al. [87]	28
4.3	Top-level model of the complaint handling process as proposed by Trappey et al. [12]	31
4.4	Proposed framework on complaint management by Zairi [101]	32
4.5	Framework by Razali and Jaafar [99]	32
4.6	Model of diffusing customer anger strategies as proposed by Nguyen and McColl-Kenney [51]	33
5.1	Screenshot of example setup	38
6.1	General empathy framework for dialogue agent	43
6.2	Avatar of chatbot	45
6.3	Architecture of Chatbot	47
6.4	Example of Flow's fallback response	48
6.5	Info decision tree dialogue	51
6.6	Interface Flow	54
7.1	Perceived Interactional Justice per condition	57
7.2	Perceived Empathy per condition	58
7.3	Perceived Usability per condition	59
7.4	User turns per condition	60
7.5	Words used by user per condition	60
7.6	Interaction times per condition	61
B.1	Model of dealing with complaints based upon Stauss and Seidel [100]	86
C.1	Greeting phase	88

C.2	Aggression reduction + conflict settlement phase	88
C.3	Problem solution phase	89
C.4	Conclusive phase	89

List of Tables

3.1	Key strategies for complaint handling per category of justice theory (<i>italics</i> show most frequently mentioned strategies)	13
3.2	Benefits of empathy as proposed by Bove [65]	15
4.1	Most common emotion/intent pairs as found by Welivita and Pu [88] .	29
6.1	Overview of mapping between emotions and empathic response intents	44
6.2	Empathic strategies per phase	46
6.3	Example of different phrasings per empathic intent	50
7.1	Responses on chatbot perception (positive) <i>*Only relevant for NonE condition</i>	63
7.2	Responses on chatbot perception (negative)	64
7.3	Responses on realism chatbot behaviour	65
7.4	Responses on human agent behaviour compared to chatbot	65
7.5	Responses on user behaviour compared to interaction with human agent	66

Chapter 1

Introduction

From Research Topics (including adaptations)

Over the past few years large technological developments have taken place in the field of communication and people have become more dependent on modern technology [1]. Keeping in touch with friends and family is no longer a time-consuming process, as most people have applications which allow them to quickly text or call the other person. Similarly, people no longer have to travel to a shop to check if they have their item on stock, but can simply chat or call with an employee from home. Technology is seen as one of the key factors to bring revolution in social communication [2].

Technologies are no longer only used to communicate via, but also to communicate with. A popular example is the chatbot, described by the Oxford lexicon as “a computer program designed to simulate conversation with human users, especially over the internet.” [3]. In short, they are computer programs to chat with. Today, chatbots are not only used for task-oriented jobs, as the development of social chatbots means that chatbots are increasingly used for social interactions [4]. An example of a mix between a task-oriented and a social bot is can be found in customer service. Here, chatbots are widely implemented [5–8], due to their cost efficiency, fast responses, and multiple other benefits. Complaint handling is one of the areas where chatbots are increasingly employed in customer service [5, 6, 9].

Complaint handling is a part of service recovery, a process where the service provider actively attempts to recover the relationship with disgruntled customers [10]. Successful complaint handling is crucial for service recovery, as it helps companies to keep and strengthen customer relationships [10–12]. Research suggests a plethora of different strategies of successful complaint handling [13–15]. However, most researchers agree upon using *justice theory* as a way to measure satisfaction in service recovery [11, 16, 17]. This includes three types of justice, distributive, procedural, and interactional [13]. Interactional justice is focused on the perceived fairness of the inter-personal behaviour during the service recovery process [17]. When this is high, research found that the need for the other ‘justices’ is decreased [17]. An important strategy to improve interactional justice is the portrayal of empathy by the employee during the service recovery process [17, 18].

This raises the question of the capability of chatbots to successfully handle complaints, as they are often seen as ‘cold, rational’ machines which lack the ability to handle emotional tasks [19]. Given the increasing implementation of chatbots for complaint handling in customer service, this is a relevant topic for more research. Therefore, my research will focus upon the following research question:

What is the effect of empathy in chatbots on perceived interactional justice during service recovery?

An experiment with two chatbots was designed to find an answer to this question. The study used a between-subject design and each participant interacted with either an empathic or a non-empathic chatbot and filled out a questionnaire about perceived interactional justice.

In the upcoming chapters the process of this research will be explained. In chapter 2 more information on the concept of chatbots will be provided, including history, benefits and limitations, and use cases. In chapter 3 more information will be given on core principles in this research, such as complaint handling and empathy, including the relation between the two. Then, chapter 4 will discuss the related work in the field of empathic chatbots, and identify key components of empathy in chatbots. The method used to answer the research question will be discussed in chapter 5, which describes all necessary steps to replicate this study. The details on the realisation of the two prototype chatbots will be discussed in chapter 6. This will be followed by an examination of the results (chapter 7). Then, significant discoveries and suggestions for future work will be discussed in chapter 8. Moreover, chapter 9 will present the conclusion to this research.

End Research Topics

Chapter 2

Context

From Research Topics (including adaptations)

Chatbots, also known as conversational agents, are disembodied conversational agents designed to converse with humans [20]. A chatbot is an interactive computer program, which takes natural language as input, and generates natural language as output [21]. This way a chatbot is able to interact with human users through text or speech [5, 20]. In order to do so successfully, the chatbot requires natural language processing (NLP), including natural language understanding (NLU) and natural language generation (NLG) [5]. NLP is the area of artificial intelligence which focuses on the use of natural language by machines [5]. At the core of NLP is NLU, focused on extracting context and meaning from natural language, as to create an abstract representation of the natural language [5]. Lastly, NLG allows the machine to generate natural language, to provide human-like responses [5]. A basic overview of chatbot architecture is shown in Figure 2.1.

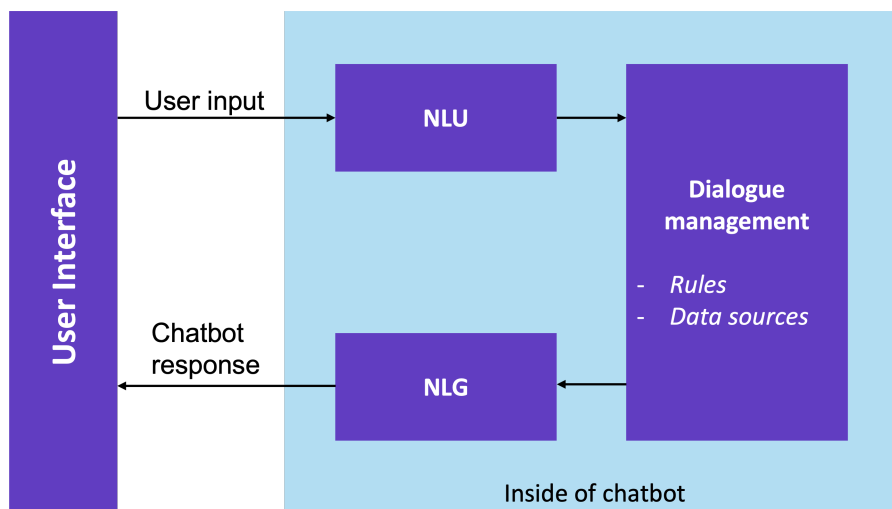


FIGURE 2.1: General chatbot architecture based on design by Adampoulou and Moussiades [5]

Chatbots are used in a variety of fields, including the customer service [6, 21]. The popularity of chatbots in customer service is best explained when looking at the benefits of using a chatbot for customer support. For example chatbots are able to reply quickly, they are usually cheaper than human workforce, and look up items or stock more efficiently than most of their human counterparts [9]. However, chatbots also come with limitations compared to human workers, such as their lack of inherent natural language understanding [22] or their lack of emotional intelligence [7, 19].

The following chapter will focus on the area of chatbots in more detail. Specifically, the concept of chatbots will be discussed through history, benefits and limitations, and use in customer service.

2.1 History

Even though the popularity of chatbots has significantly risen the past few years [5] (see Figure 2.2), the idea of a chatbot is not completely novel. In 1950 Alan Turing introduced the 'Turing Test' [23], a test in which a participant asks questions to two agents, one human and one robot, and has to distinguish which one is human and which one is a robot. The goal of this test is to design a robot which can trick the participant into thinking it is human. This introduced and popularized the first idea of today's chatbot.

One of the first, and best known chatbots is the ELIZA chatbot. Developed by Weizenbaum in 1966 [24], ELIZA was able to converse with humans through pattern matching: looking for keywords in the input text and selecting a response from a set of predefined responses [24]. This approach proved to be moderately successful as many people believed that they were conversing with a human [25]. Inspired by ELIZA, the ALICE chatbot was developed by Wallace in 1995 [26]. Similar to ELIZA, ALICE uses pattern templates to recognize user input. However, it is extended with AIML¹ to allow more custom conversations. ALICE is an example of a rule-based chatbot, meaning that the responses are based upon a predefined set of rules.

Whereas these chatbots were already able to fool some users into thinking they were human, large developments have been made since. This was possible due to the recent developments in artificial intelligence and wireless communication [25]. This resulted in intelligent personal assistants such as Amazon's Alexa², Apple's Siri³, and Google Duplex which is able to book hair appointments without the hairdresser realizing they speak to a robot [28].

Following these personal assistant chatbots a new type of chatbot has emerged, the social chatbot. Instead of task oriented chatting, these chatbots are built to have natural conversations while showing social cues and emotional behaviour [25]. This makes them a social conversational partner which is able to converse about a variety of topics, rather than a means to get to an end. An example of this is the XiaoIce chatbot, developed by Microsoft⁴.

¹"AIML contains a collection of rules which define the conversational capabilities of the chatbot. it's used with a linguistic communication Understanding (NLU) processor which takes AIML rules to investigate and reply to the text queries asked via the chatbot." [27]

²<https://developer.amazon.com/alexalive>

³<https://www.apple.com/siri/>

⁴<https://www.xiaoice.com/>

Currently the uses of chatbots are widespread. But over history their uses have changed, starting as a test to see 'if machines can think' [23], they were developed into human-mimicking speech entities, then into task-oriented smart assistants, and eventually developed to participate in full social interactions. The growing interest in chatbots is reflected in research, which shows a steep increase in published papers related to chatbots in the last few years, as can be seen in Figure 2.2.

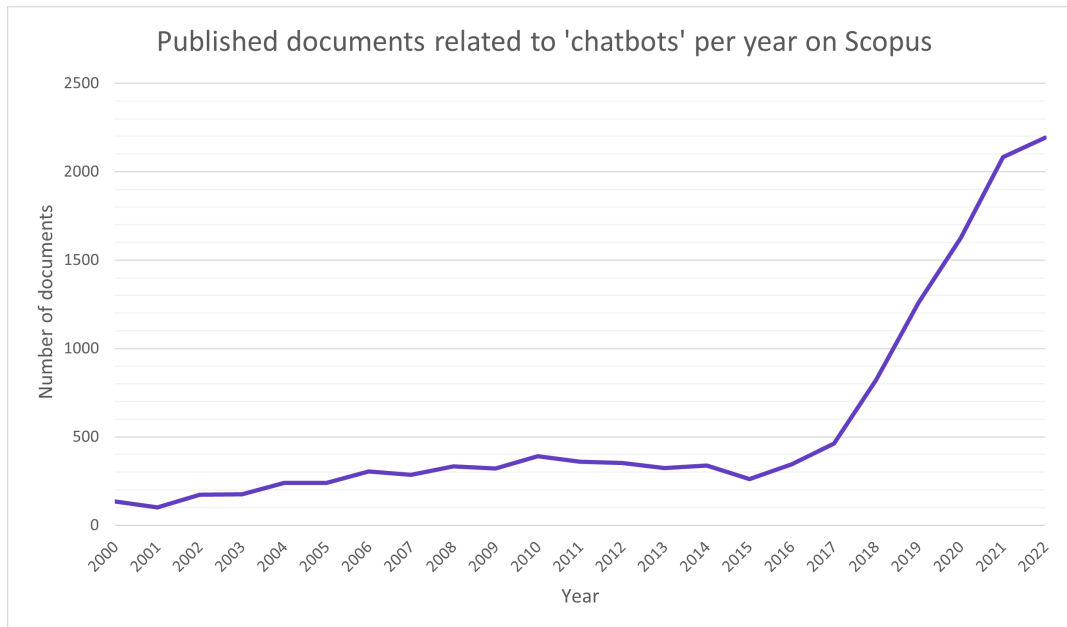


FIGURE 2.2: Number of published documents related to "chatbot", "conversation agent", "conversational agent", and "dialogue system" on Scopus from 2000 to 2022.

2.2 Benefits and limitations

The growing interest in chatbots is due to their variety of benefits. However, the increase in use of chatbots also shines a light on their limitations. The following sections will discuss both in more detail, to provide a thorough overview of the benefits and limitations of chatbots.

2.2.1 Benefits

Klopfenstein et al. [29] discuss the benefits of chatbots in terms of user benefits and company benefits. Whereas the first focuses on the benefits for end-users, the latter highlights the benefits for companies which deploy them. The benefits of chatbots can be seen in relation to their human counterparts (e.g. chatbots don't require rest whereas humans do) or in relation to other technological solutions, such as an email or form on a website (e.g. a chatbot can be more personal than a website).

User benefits

An important user benefit is the *instant availability* of chatbots [29], compared to both human and technological competitors. This benefit is most commonly mentioned among researchers, for example by Adamopoulou et al. [5], Devaram [30], Brandtzaeg and Følstad [6]. The importance of instant availability is also highlighted

by Xu et al. [31], whose research showed that the average response time of customer service teams on social media is 6.5 hours, whereas 53% of Twitter users expect a response within an hour [32]. These expectations are clearly not met, and this is where chatbots could boost productivity and user satisfaction. After all, chatbots do not require rest, while human agents do. Brandtzaeg and Følstad [6] state that in some scenarios, such as for booking a taxi or looking for medical advice, this means that chatbots are preferable over other means of assistance (phone call or online search), as they are easier to reach and respond fast. Jenneboer [9] adds that fast support and effective assistance also influences customer loyalty.

Another frequently mentioned benefit is the *efficiency* of chatbots, which is closely related to the previous benefit. However, whereas instant availability has to do with the fact that chatbots do not need sleep, food or breaks, the efficiency has to do with the fact that chatbots can multitask and search data more efficiently. The efficiency of chatbots is a benefit in relation to humans [5, 6]. Closely connected to their ability to instantly reply, chatbots are able to look up information in databases faster than most humans. This makes that chatbots need significantly less time to draft a response than their human counterparts, hence increasing their efficiency. Additionally, chatbots can do a multitude of tasks at the same time, whereas humans cannot.

Besides their efficiency, chatbots have an additional advantage over humans, in terms of *threshold of communication* [9, 33]. Vaidyam et al. [33] explain that chatbots may be easier to open up to than to humans, specifically in the mental-health domain. People may be less keen to disclose sensitive information with humans, but the easy accessibility of chatbots combined with their machine-nature, makes them a easier alternative [30].

Whereas these three benefits are most commonly referenced in literature, more benefits are mentioned. Other benefits of chatbots compared to humans include, *entertainment value* and the *novelty effect*, which makes some users more likely to interact with chatbots [5]. Furthermore, compared to other technologies, chatbots are usually deemed as more *adaptive* and *personal* [21], and able to keep calm in scenarios where humans might feel stressful [19]. Additionally, Brandtzaeg and Følstad [6] suggest that chatbots can provide more engaging and natural feeling conversations compared to searching answers on a mobile application.

Company benefits

First of all, user satisfaction is crucial for companies. Since the aforementioned user benefits of chatbots lead to higher user satisfaction [9], they can be seen as indirectly beneficial for companies which use the chatbots. However, there are more specific benefits as well.

The most obvious advantage of chatbots over humans is that they are *cost efficient* due to a variety of reasons. First, using chatbots reduces the cost of human workers, as machines do not get paid [19]. Second, chatbots are able to handle multiple messages at the same time, meaning that they can be more efficient in a shorter amount of time than humans. So, even in the case where employing chatbots is equally costly as employing people, chatbots would still have an advantage over humans.

In line with this benefit is that chatbots increase *productivity* [6, 9]. An illustration of this benefit can be seen in a customer service setting where a chatbot handles simple requests for a customer service company and transfers difficult requests to a human

agent. This relieves the human agent of routine tasks and allows them to focus on complicated requests. Furthermore, since chatbots can handle multiple requests at the same time, overall productivity is increased.

Another advantage is highlighted by Miklosik et al. [21], who states that chatbots which employ machine learning algorithms are able to improve with every user interaction. This way the chatbots can develop *without human interference*, and constantly adapt to user needs. Whereas this is no advantage over human agents, it is distinct benefit compared to other technologies. In addition to this, the chatbot can be easily expanded to *collect user data* [34], which the company can use for future reference and training purposes. Lastly, Tsai et al. [19] add the benefit that chatbots can be *easily embedded* in other communication channels, such as websites and mobile apps. This makes them an accessible solution to implement for companies.

2.2.2 Limitations

Besides advantages, chatbots also come with limitations. The main limitations will be discussed and explained in the following section.

One large disadvantage is that chatbots are *not naturally able to understand natural language* the way humans are. This means that slang, typos, and sarcasm are often not understood by chatbots [22]. Since these types of input are inescapable when interacting with people, this negatively affects the interaction. A way to overcome this is by expanding the capacities of the chatbot. Some dynamic chatbots are equipped with machine learning algorithms which allow them to learn from interactions [21]. This way, the chatbot could learn more about these types of inputs and be better prepared for the future.

Another limitation is the *limited domain of knowledge* of the chatbot, which restricts its flexibility. Whereas humans can effortlessly switch between conversational topics, this is a challenge for chatbots, as they only 'know' of the data that they have access to. This means that open-ended conversations are complex to design [35]. When conversations are able to run freely, the chatbot requires all types of knowledge, including domain specific terms and expressions. This is often hard to obtain. For open-ended conversations the chatbot needs to be able to not only recognize the topic and respond accordingly, but understand context, ontologies and semantic spaces as well [7]. If done incorrectly, the chatbot will come up with non-sensible responses, and fail to fulfill the user's needs [6]. This constrains most chatbots to a specific domain, whereas humans are more flexible. This limitation is at the core of an important trade-off for chatbots; choosing between highly human-like interactions and efficiency [29]. When provided with a broad knowledge domain, chatbots will be able to have more fluent conversations. However, this also means that the information is accessed less quickly, due to the large databases that the chatbot will have to go through.

Furthermore, a big limitation is the *lack of emotional intelligence* of chatbots. This is a limitation for two reasons. First, Tsai et al. [19] explain that a chatbot is an "inherently logic-driven, mechanical, and depersonalized computer program", making it "inferior in handling emotional tasks compared to humans". So a chatbot might be efficient and easy to talk to, but lacks the emotional intelligence of a human worker [5]. Emotional intelligence is important as it allows the chatbot to correctly identify user's emotions and come up with an appropriate response [36]. This leads to better interaction experience and higher user satisfaction [36]. Secondly, according to

Augello et al. [7] chatbots are unable to correctly manage social interactions due to the complex nature of human dialogue. One of the challenges lies within interpreting the social context, which is naturally more difficult for chatbots than for humans, as chatbots lack emotional intelligence [7]. Overall, emotional intelligence is crucial in improving user experience, and managing social interactions.

Lastly, researchers mention *impersonality* as a limitation of chatbots. As described in the previous paragraph, chatbots are often perceived as 'impersonal' and 'cold' [9, 19]. Research has shown that users often prefer interacting with a human-like and personal chatbot [20], and that this may even increase the trust in the chatbot [9]. Possible ways to overcome the impersonal feel of chatbots would be by adapting the use of language [9], and adding visual or identity cues [37].

Concluding remarks

There are a variety of benefits and limitations when using chatbots. The benefits can be categorized in two groups; benefits for the end-users, and benefits for the company that deploys them. Important benefits include chatbots' 24/7 *availability* and *efficiency*. However, chatbots also come with limitations when compared to human interactions. An important limitation of chatbots is their lack of emotional intelligence, which is found to be important for improving user experience. When designing a chatbot it is important to understand the necessary trade-offs and prioritize certain design decisions accordingly.

2.3 Chatbots in customer service

There are a variety of current uses for chatbots. Due to their benefits, one of the main implementations is in the field of customer service. This section will describe the most common applications, and give a more in-depth description of chatbots in the customer service.

2.3.1 Uses of chatbots

The popularity of chatbots is reflected in their multitude of applications. The most common ones include customer service, information retrieval, education, emotional support, and virtual assistance [5–7]. Chaves and Gerosa [8] add the use of chatbots for entertainment. These areas will be shortly discussed.

Customer service - The area of customer service is widely regarded as an important area for chatbots [5–8]. Mainly due to the benefits of chatbots; instant availability, productivity, and cost efficiency, chatbots are a practical solution for customer service. Given the relevance of this area for this research, it will be discussed in more detail in subsection 2.3.2.

Information retrieval - Chatbots can be used to do information retrieval. Given their efficiency, access to databases and computing power, they are usually more efficient than humans. For this application chatbots are asked questions by the user and then try to find pieces of information from a database [38]. Additionally, natural language understanding can make the queries feel more natural, which was found to be desirable by end-users [20].

Education - Chatbots can be implemented for educational purposes, and are the topic of a variety of research [8, 34, 39]. An example of an educational chatbot is a chatbot which tutors children [8]. Research by Chaves and Gerosa [8] found that chatbot tutors are able to increase the students' engagement and learning outcomes.

Emotional support - Chatbots are also employed to provide emotional support, for example with mental health issues. Chatbots are a cost-efficient way to help relieve some symptoms of mental conditions, such as stress [30]. Additionally, chatbots may be helpful for people who are uncomfortable talking to another person, but still need to express their feelings [33]. The instant availability of chatbots also helps in this domain; people are always able to talk to them to find relief [30]. Additional uses of emotional support chatbots include; providing reminders of medication and exercise, identify emergencies, and helping to regulate emotions [30]. For emotional support chatbots it is important that they are able to recognize emotions [30], so they can give appropriate responses or suggestions. Sentiment analysis is commonly implemented for this purpose [30].

Virtual assistance - An example of a virtual assistant is Apple's Siri⁵. Rossmann et al. [40] describe virtual assistants as software agents which do tasks and provide services for their user. Tasks include; responding to queries, acting as tutor, leading people to relevant sites, helping with shopping decisions [40].

Entertainment - Research has shown that many users also resort to chatbots for entertainment [6, 8]. Brandtzaeg and Følstad [6] found that some users like talking to chatbots because they "like chatbots that have funny things to say", or when they feel bored. Lastly, the novelty effect adds to the entertainment value [5].

2.3.2 Customer service

Due to their ability to work faster and more effectively than most humans, chatbots are frequently used in the customer service sector. Tsai et al. [19] even state that some estimate that chatbots will handle 85% of all customer service interactions in a few years. Chatbots positively affect key concepts in customer service, such as customer loyalty and customer experience [9]. Brandtzaeg and Følstad [6] add that customer engagement is also positively influenced by use of chatbots. They allow for direct user interaction, which static apps or webpages are unable to do, and this opens new possibilities for connecting to customers and marketing [31].

Possible uses of chatbots in customer service include helping with queries, providing personal advice, and complaint handling [40]. Complaint handling is one of the most common uses for chatbots in customer service [9, 19]. Whereas humans might feel uncomfortable in stressful situations, the advantage of chatbots is that they will always remain calm [19]. Furthermore, in complaint handling customers are often looking to vent to alleviate their anger [19], which can be done to either human employees or chatbots. Some people prefer a human service employee as it makes them feel understood and heard, but there are also instances in which people prefer to interact with a chatbot as they feel too embarrassed to speak to a human employee [19].

An interesting dilemma for chatbots in customer service, specifically complaint handling, is the trade-off between showing compassion and being task-oriented. Whereas users claim to prefer personal interactions [20], they also want to be helped quickly

⁵<https://www.apple.com/siri/>

[5, 6, 29]. Since chatbots are often deemed impersonal [9, 19], the users may question the authenticity of the personal interaction. As a result, they may become even more frustrated when the chatbot is displaying compassion instead of being task-oriented.

Concluding remarks

Overall, chatbots are widely researched a variety of domains, but particularly in customer service [5, 6, 9]. Chatbots' benefits make them a good alternative for human employees. One of the most fruitful aspects of chatbots in customer service is in complaint handling. The following chapter will focus on that in more detail, and describe the importance of emotional intelligence in relation to it.

2.4 Conclusion

This chapter provided a general overview of chatbots, including history, benefits and limitations, and current uses, specifically in customer service. The popularity of chatbots has increased significantly over the past decade, and so have their amount of different uses. This is largely due to their vast amount of benefits. The literature describes a variety of benefits of chatbots, both for users and companies who employ them. User benefits include *instant availability*, *efficiency*, and lowering the *threshold of communication* for some people. Extra benefits for companies are chatbots' *cost efficiency*, *productivity*, and the possibility for development *without human interference*. Besides benefits, chatbots also have their limitations. In this chapter, the following four main limitations were identified; the inability to *naturally understand natural language*, their *limited domain of knowledge*, *lack of emotional intelligence*, and perceived *impersonality*.

Despite their limitations, chatbots are employed in a variety of domains, such as customer service, information retrieval, education, emotional support, virtual assistance, and entertainment. Especially in customer service, chatbots have a lot of potential, and are used frequently. Due to their benefits, they have a positive effect on key concepts in customer service, increasing their employment. One of the main uses for chatbots in customer service is for complaint handling, a type of service recovery. Chatbots are able to remain calm, and allow angry customers to vent their anger. However, an important challenge for chatbots is the trade-off between showing personal behaviour and being task-oriented.

This chapter has shown how chatbots can be used and what their strengths and limitations are. Knowing this enables accurately addressing limitations, such as the lack of emotional intelligence. It also highlights the relevance of this research, since chatbots are increasingly used in customer service, especially for complaint handling. Given the finding that chatbots are useful in complaint handling, the importance of complaint handling will be researched in the following chapter.

End Research Topics

Chapter 3

Theory

From Research Topics (including adaptations)

This chapter will focus on the theory behind this topic. First, the concept and importance of complaint handling will be further explained, as the previous chapter found that this is a relevant area for chatbots. The importance of justice theory in complaint handling is explained, highlighting the importance of interactional justice. Since literature shows that empathy plays an important role in interactional justice, the concept of empathy, benefits, and the relation to complaint handling will also be described. Lastly, common measurement methods of important concepts such as empathy and interactional justice for this research will be discussed.

3.1 Complaint handling

In order to be profitable and successful, companies are constantly aiming to build, improve and maintain customer relations [11, 16, 41]. In cases where customers are dissatisfied, complaints are valuable to the company to identify areas of improvement [12, 16, 42]. Handling complaints efficiently is crucial in rebuilding the customer's relationship with the company, and results in measurable improvement of revenue and profitability [16]. The following sections will address the importance of complaint handling and the key points of effective complaint handling.

3.1.1 Importance of satisfied customers

That upholding customer relationships is positive for companies has been shown in a plethora of research [11, 16, 17, 41, 43]. According to Tax and Brown [16] a good relationships between the customer and the company improves customer satisfaction and loyalty. Subsequently, both these concepts influence whether a customer remains a satisfied customer or leaves for a competitor [41]. Maintaining relations with existing customers is crucial for companies, as it is cheaper than the time-consuming and expensive process of recruiting new customers [9, 42]. Additionally, losing a customer results in decrease in current buying behaviour, meaning a decrease in sales, and thus the profits for the company [11].

An additional advantage of satisfied customers is that they can participate in word-of-mouth marketing, recruiting more customers for the company [42]. Conversely, dissatisfied customers may engage in negative word-of-mouth, and in turn this may result in less current and future customers [9, 11, 42].

3.1.2 Definitions of complaint handling

Unfortunately, every company is bound to make mistakes, resulting in dissatisfied customers. Given the aforementioned importance of satisfied customers, this makes complaint handling a crucial aspect of customer relationship management (CRM).

When there is a service failure, customers have the option to complain. Complaining is a behaviour where a customer shows dissatisfaction with the provided service [12]. Complaint handling is the process to solve this dissatisfaction and compensate the consumer, attempting to improve customer satisfaction [12]. Fornell and Westbrook [44] add to this the importance of identifying the cause of the consumer dissatisfaction in the process.

According to Trappey et al. [12] complaint handling has both a physical and an emotional aspect. The physical aspect has to do with economic compensation, whereas the emotional aspect includes providing an apology and explanation. Simon [43] also discusses an affective component in complaint handling, which appears similar to the emotional aspect mentioned by Trappey et al. [12]. The notion of an affective component in complaint handling is further supported by Iyiola and Ibidunni [45]. Moreover, other studies imply that the emotions of the customer during service encounters also play a role in the satisfaction, loyalty and future buying behaviour [46, 47].

3.1.3 Benefits of successful complaint handling

Successful complaint handling has a number of advantages for companies [12, 41, 42, 48]. First, customers who complain provide valuable information to the company [12, 42]. They give insight into the weaknesses, and allow the company to learn from their mistakes [42]. Both Trappey et al. [12] and Filip [42] agree that complaints are key indicators of organizational performance.

Second, complaints are crucial to uphold positive customer relations and avoid losing customers [11, 12, 42]. When consumers feel dissatisfied by a service, they have the option to look for a competitor or spread negative word-of-mouth. By complaining, they give the original company a chance to keep them as a customer [42]. When their complaint is handled effectively, their satisfaction and loyalty to the company will be restored, and they will stay [11, 42]. In addition to just restoring satisfaction and loyalty, Iyiola & Ibidunni [45] found that properly resolved complaints result in even more trust, commitment and long-term relationships with the customer than before the service failure. Whereas unresolved complaints lead to the opposite [12, 42, 45]. In line with this, research has shown that customers who experienced successful complaint handling show higher levels of brand loyalty than customers who did not experience any service failure [41].

In conclusion, successful complaint handling does not only give insight in the company's weaknesses and performances. It also restores customer satisfaction after a service failure, and research has shown that effective complaint handling could improve customer satisfaction and loyalty to higher levels than before the failure.

3.1.4 Effective complaint handling

All of the aforementioned benefits are dependent on the quality of the complaint handling. Luckily, there has been plenty of research into recovery strategies for service failure, for example by [13–15, 42]. Given the variety of strategies, it is fruitful

to investigate what concepts are commonly identified by research as 'key for successful complaint handling'. Michel et al. [49] found that research primarily focuses on three different types of service recovery. Namely, *employee recovery* - focusing on preparing employees on service recovery, *process recovery* - focusing on how to improve processes and learn from failures for the future, and *customer recovery* - focusing on customer experience and satisfaction. This research focuses on customer recovery.

Whereas a variety of strategies have been proposed ([13, 14]), researchers appear to widely accept social exchange and justice theory as an adequate way of evaluating customer recovery (e.g. [11, 13, 16, 50, 51]). Justice theory is the dominant theory used to study consumer's assessment of the service recovery, and focuses on the fairness of compensation as perceived by the involved parties [52]. Justice theory includes three different types of justice; (i) distributive (perceived fairness of assignment of tangible resources to compensate for failure); (ii) procedural (perceived fairness of methods used during the service delivery); and (iii) interactional justice (perceived fairness of inter-personal behaviour during the service delivery) [13, 50].

Distributive justice has the greatest influence on the overall perceived justice compared to procedural and interactional justice [15, 17]. However, when little can be done in terms of monetary compensation, procedural and interactional justice play an important role in customer satisfaction [17]. Van Herck et al. [17] found that successful service encounter (high interactional justice) can decrease the need for high compensation (high distributive justice) after service failure. Furthermore, procedural justice is crucial to lower the barrier of complaining so more people will be inclined to complain [42]. This is important, as research suggests that only a minority of customers complain [11, 12, 15], while complaints can only be resolved if they are reported. Given the importance of customer's willingness to complain, Filip et al. [42] highlight the importance of accessible and worthwhile complaint handling procedures.

An overview of different strategies for each category of justice theory can be found in Table 3.1. This table provides a framework of key strategies to focus on for successful service recovery, as put forth by different sources of literature. The most frequently mentioned key strategies are *Easy access*, *Fast resolution*, *Friendly employees*, and *Empathetic employees*.

Distributive	Procedural	Interactional
Product repairs [53]	Flexible [16, 40, 50]	Provide explanation [54]
Product replacement [50, 53]	<i>Easy access</i> [16, 40, 42, 50]	<i>Friendly employees</i> [17, 40, 50, 55]
Free service [53]	<i>Fast resolution</i> [9, 12, 15]	<i>Empathetic employees</i> [16, 17, 40, 42]
Discount [50, 53]	[16, 17, 40, 42, 50, 56, 57]	[49, 50]
Refund [50, 53]	Keep informed [9, 40, 42]	Trustworthy employees [55]
Voucher [17]	Customer in control [40, 50]	Personal approach [49]

TABLE 3.1: Key strategies for complaint handling per category of justice theory (*italics* show most frequently mentioned strategies)

Concluding remarks

Customer satisfaction is crucial for a company as satisfied customers can engage in positive word-of-mouth and stay loyal. Finding new customers is more cumbersome than maintaining current customers. However, when customers become disgruntled, complaint handling can be used to restore and even improve customer loyalty and satisfaction. Customer recovery can be evaluated with justice theory, including distributive, procedural, and interactional justice. Research found that interactional justice can reduce the need for compensation. Some key strategies for service recovery are easy access, fast resolution, friendly employees, and empathic employees. The following section will evaluate what role empathy plays exactly.

3.2 Empathy in complaint handling

Empathy plays an important role in complaint handling. To illustrate, the concept of empathy is first defined, including the notion of known benefits of empathy. Then, the relation between empathy and complaint is explained. This section does not specifically focus on chatbots.

3.2.1 Empathy

Empathy is a well-known concept which has been researched for years. Kickstarted with the translation of the German word 'Einfühlung', thousands of researchers explored the concept of empathy [58], in multiple different disciplines [59]. Unfortunately, no clear definition for empathy is unanimously agreed upon by researchers. However, most researchers agree that empathy is a multidimensional construct, which includes both cognitive and affective dimensions [43, 58, 60–62]. The cognitive component covers the tendency to understand what others are feeling, also called empathic concern [43]. The affective component refers to the phenomenon of feeling the same affective state as another person [58, 61], also called emotion contagion [43].

As social interaction inherently involves understanding each other's affective states [58], empathy is an important concept in social interactions [59]. According to Bošnjaković and Radionov [59], empathy is the most "vital and flexible human ability". Clark et al. [58] agree with the importance of empathy, and add that empathy is a critical construct in social crisis management and forgiveness. Researchers agree that empathy has two main effects, facilitating (1) adaptive¹ and (2) prosocial behaviours [63]. Dovidio [64] defines prosocial behaviour as "acts that are socially defined as generally beneficial to other people". In short, people better understand their environment and navigate social interactions because of empathy [63]. Bove [65] created an overview of the different benefits of empathy in social interactions, for different social groups, which can be found in Table 3.2.

¹"The level of everyday performance of tasks that is required for a person to fulfill typical roles in society, including maintaining independence and meeting cultural expectations regarding personal and social responsibility." source: <https://dictionary.apa.org/adaptive-behavior>

Individual or in-group	Service organization	Society
Motivates helping behaviour Facilitates social bonding Enhances social support	Contributes to service quality, customer compliance and sales performance Promotes forgiveness Protects brand reputation Enables design thinking	Improves moral decision making Reduces prejudice Discourages anti-social behaviour

TABLE 3.2: Benefits of empathy as proposed by Bove [65]

3.2.2 Empathy related to complaint handling

A multitude of research has highlighted the importance of empathy in complaint handling [17, 18, 43, 61, 63]. Whereas the use of affective empathy is not always necessary, cognitive empathy is a must to identify the needs of the user, and act accordingly [61]. The use of empathy in complaint handling is two-fold: (1) to regulate the emotions of complainant, (2) to increase perceived interactional justice. The following sections will explain each of the concepts in more detail.

Emotion regulation

First of all, research has shown that complainants usually experience some sort of negative emotion, such as anger or sadness [45]. However, these strong affective states overwhelm the complainant and interfere with the ability to accept solutions [14].

Venting is recognized as an effective way to regulate these emotions [66, 67]. Tsai et al. [19] state that the first thing people want when they contact customer service is to vent. This allows them to decrease the dominance of the negative emotions, and transform these into a cognitive mechanism [19], allowing for more fruitful social interaction. Expressions of remorse are a key strategy to alleviate intense emotions experienced by the complainant [68]. Fjelstad [69] explains that "remorse comes from true empathy for the pain the other person is feeling because of your actions". This shows that empathy is crucial for true remorse, which is in turn necessary to decrease negative affective states of the complainant.

Another interesting concept related to emotion regulation is *emotion validation*. Emotion validation is "the process of learning about, understanding, and expressing acceptance of another person's emotional experience" [70]. The key aspects of emotional validation are recognizing the emotional state (1) accurately, and (2) non-judgementally [71]. Research by Lambie et al. [72] explains that emotion validation supports emotion regulation. In their research they found that children calmed down faster and showed higher levels of emotion regulation when their parents would engage in emotion validation. This suggests that emotional validation is an effective method for emotion regulation, and could be adopted in service recovery as well.

Whereas researchers are not decided on the specific relationship between empathy and emotional validation, they do agree that there is a relation [71, 72]. Given that emotional validation is based upon the importance of understanding emotions, and cognitive empathy encompasses this understanding, some researchers say that empathy is the underlying concept of emotional validation [72]. The main difference between the concepts is that empathy is a knowledge state or feeling, whereas emotional validation is an act [72].

Perceived interactional justice

After regulating the emotion of the complainant, it is time to start solving the dissatisfaction. As described before, successful complaint handling is dependent on the perceived interactional justice. Table 3.1 shows that many researchers agree that empathy is a key concept to increase perceived interactional justice. The relation between increasing customer satisfaction with the service encounter and empathy has also been proven by a variety of research [18, 43, 63]. An important construct for empathy in complaint handling is *employee empathy*, the empathy shown by the employee [17, 63]. Wieseke et al. [63] define employee empathy as the ability of the employee to sense and react to the consumer's affective states, thoughts, and experiences during service recovery. Van Herck et al. [17] explain that employee empathy means that (1) the employee understands the state of the complainant - *cognitive empathy* (2) the employee shows concern for the situation of the complainant - *affective empathy* (3) the employee experiences similar affective states as the complainant, such as frustration - *affective empathy*. When executed correctly, these three pointers positively influence the perceived empathy, and in turn the perceived interactional justice.

First, understanding the state of the complainant allows the employee to respond appropriately and make them feel recognized [43]. This increases the service quality [65], and thus the customer's satisfaction with the process. It is worth mentioning that accurately addressing the user's needs not only positively influences interactional justice, but also procedural justice, as it facilitates a fast resolution of the problem. In short, adding cognitive empathy of the employee results in the benefits shown in the column 'Service organization' of Table 3.2. Second, showing concern for the situation of the complainant positively affects the complainant's attitude towards the employee and company [43, 63]. Simon [43] found that customers may interpret employee empathy as an act of benevolence, meaning that they feel as if the employee puts in extra effort to help them. Showing concern also makes the employee seem more sympathetic, which increases perceived interactional justice, as shown in Table 3.1. When the employee shows affective empathy this results in the benefits shown in the column 'Individual or in-group' of Table 3.2. Third, experiencing similar affective states as the complainant also results in the benefits as shown in column 'Individual or in-group' of Table 3.2. When the employee experiences similar emotions as the complainant, this increases their inclination to help [65]. For the complainant this shared emotion leads to more social bonding, which makes them feel solidarity from the employee [65]. Altogether, these benefits of employee empathy increase the gratitude of the complainant, increasing the overall customer satisfaction [43].

Concluding remarks

Empathy is a multidimensional construct, including a cognitive and an affective aspect. Researchers agree that empathy is crucial for successful social interaction, and identified several benefits for different social groups. Furthermore, there are a variety of ways that empathy helps with solving complaints successfully. Each of them contribute in their own way to the customer's satisfaction with the service encounter. Especially in making the complainant feel heard, and by increasing goodwill towards the employee. An overview of the interactions between all concepts is shown in Figure 3.1.

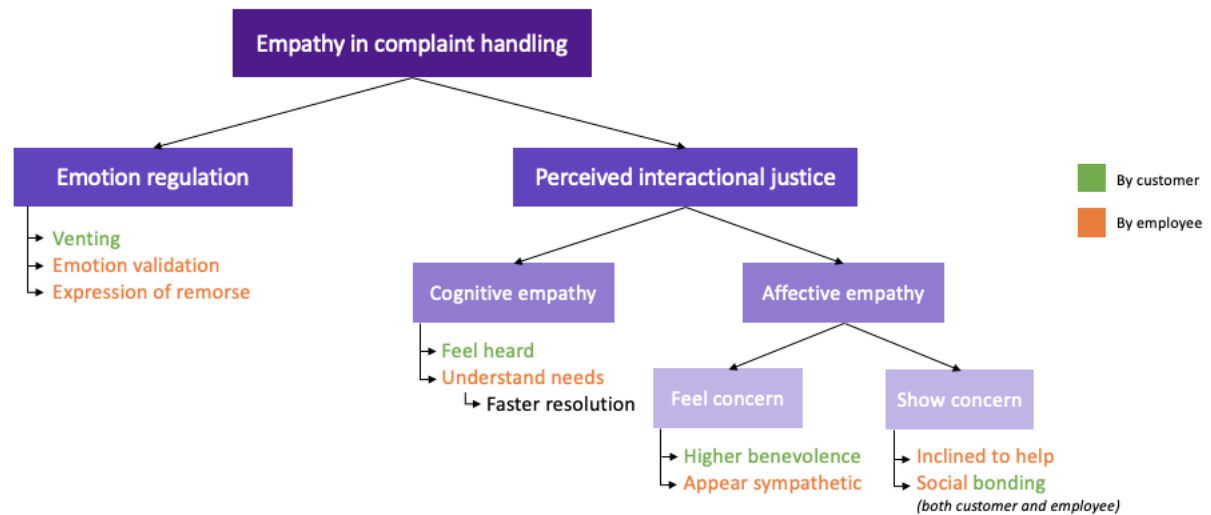


FIGURE 3.1: Overview of complaint handling concepts

3.3 Measurement methods

This section will highlight some common measures for empathy and perceived interactional justice, as previous sections found that these are important concepts for complaint handling. These measures can be used during the design of an experiment to address the research question.

3.3.1 Empathy

In the following section the most common measurements of empathy are described. There are two main measures of empathy in the literature; someone's own empathy, and perceived empathy of others.

Measures of someone's own empathy are grouped into three categories: self-reported measures, behavioural observational measures, and neuroscientific approaches [59]. Most common measures for empathy are self-reported measures [73, 74]. These measures usually consist of a number of questions asking the participant up to what extent they feel empathy. A limitation of this measure is their subjectivity, response bias, and the fact that everyone may have a different internal definition of empathy [59]. Behavioural observational measures usually ask participants to assess experimental stimuli (usually an image/video of a person in a negative situation), and rate up to what extent they can imagine feeling what the person in the stimuli is feeling [59]. The neuroscientific approach focuses on brain imaging techniques, such as MRI's, EEG's, and EMG's [59].

Whereas these types of measurements are often used in empathy research, the current research focuses on perceived empathy. Different measures were found in literature, focusing on perceived empathy between two humans and perceived empathy between a human and robot. Some common measures for both will be briefly discussed in the following sections.

Human-human

Plank et al. [73] propose a measure of perceived empathy, based upon research on sales performance. Whereas they found promising results, their measure did not distinguish between cognitive and affective empathy. No recommendation on the used Likert scale was presented. Similarly, Delpechitre et al. [75] designed a seven-point Likert scale for perceived empathy, but they did include a distinction between cognitive and affective empathy. Similarities can be found between these items, and the items proposed by Plank et al. [73]. This measure by Delpechitre et al. [75] was designed to measure the 'Customer's perception of salesperson's empathy', but as shown by Van Herck et al. [17], the items can easily be adapted to fit the customer service context.

Human-robot

Charrier et al. [76] designed a measure specifically for perceived empathy of robots in human-robot interaction, the RoPE Scale. This scale includes both a measure of empathic understanding (cognitive empathy) and empathic response (affective empathy). The list consists of 16 questions, and 4 filler questions, as shown in Figure 3.2. Again, no recommendation on the used Likert scale was provided.

id	Empathic Understanding subscale items (EU)
EU1	The robot appreciates exactly how the things I experience feel to me.
EU2	The robot knows me and my needs.
EU3	The robot cares about my feelings.
EU4	(−) The robot does not understand me.
EU5	The robot perceives and accepts my individual characteristics.
EU6	The robot usually understands the whole of what I mean.
EU7	(−) The robot reacts to my words but does not see the way I feel.
EU8	The robot seems to feel bad when I am sad or disappointed.
id	Empathic Response subscale items (ER)
ER1	(−) Whether thoughts or feelings I express are "good" or "bad" makes no difference to the robot's actions toward me.
ER2	(−) No matter what I tell about myself, the robot acts just the same.
ER3	The robot comforts me when I am upset.
ER4	The robot encourages me.
ER5	The robot praises me when I have done something well.
ER6	The robot helps me when I need it.
ER7	The robot knows when I want to talk and lets me do so.
ER8	(−) The robot's response to me is so fixed and automatic that I do not get through to it.
id	Filler items (FI)
FI1	The way the robot acts feels natural.
FI2	The robot knows what it is doing.
FI3	The robot is responsible for its actions.
FI4	When I interact with the robot, I feel anxious.

FIGURE 3.2: Items of RoPE Scale as proposed by Charrier et al. [76]

3.3.2 Perceived interactional justice

A variety of measures for perceived interactional justice have been proposed by literature [52][77]. For example, Homburg et al. [52] created items on a 5-point Likert scale for perceived interactional justice. The items are shown in Figure 3.3.

To what extent do you agree with the following statements?
 The employees seemed to be very interested in my problem.
 The employees understood exactly my problem.
 I felt treated rudely by the employees. (r)
 The employees were very keen to solve my problem.
 Overall, the employees' behavior during complaint handling was fair.

FIGURE 3.3: Items for measuring perceived interactional justice in complaint handling as proposed by Homburg et al. [52]

My university told me why the service had failed in the first place.
 The university employees seemed very interested in helping me.
 The university employees tried hard to resolve the problem.
 The university employees were attentive in providing good service.
 I felt that I was treated rudely by university staff while lodging my complaint. (-)
 The employees listened politely to what I had to say.
 They seemed to be very concerned about my problem.

FIGURE 3.4: Items for measuring perceived interactional justice as proposed by Waqas et al. [77]

Another measure was proposed by Waqas et al. [77], which shows some similarities to the items by Homburg et al. [52]. The items are shown in Figure 3.4.

Besides specific measures for interactional justice, literature also shows measures for quality of complaint handling in general. For completeness, some measures will be briefly discussed as well. Van Herck et al. [17] made a measure of complaint handling by combining existing scales. The perceived quality of service is rated by 5 bipolar items. Lastly, Boshoff [78] designed a general measure of customer satisfaction with complaint handling, called the RECOVSAT. Including 17 items covering satisfaction with communication, empowerment, feedback, atonement, explanation, and tangibles.

Concluding remarks

There are a variety of ways to measure empathy and perceived interactional justice. Measures of empathy can be divided into human-human and human-robot measures. However, no robot-oriented measure for perceived interactional justice was found. The items as proposed by literature can be adapted to fit a study's needs, and provide more insight in how a complaint handling process functions.

3.4 Conclusion

This chapter focused on related work to this research. First, the importance and workings of complaint handling was discussed. Literature showed that good complaint handling is beneficial for a company for a variety of reasons. Additionally, literature showed that effective complaint handling focuses on justice theory. Justice theory entails that customers will feel more satisfied when they feel like they are

treated fairly, in terms of distributive, procedural, and interactional justice. Interactional justice was found to be important for successful complaint handling, and in turn highlighted the need for empathy in the complaint handling process.

Empathy is a multidimensional construct, including a cognitive and affective dimension. Both dimensions are important in complaint handling. For example, the cognitive dimension allows the employee to understand what the complainant needs, and the affective dimension shows the complainant that the participant cares. Researching the relation of empathy and complaint handling further showed that the use of empathy helps to (1) regulate the emotions of the complainant, and (2) increases perceived interactional justice. Lastly, some measures of empathy and perceived interactional justice were discussed, which can be used in the later stages of this study.

This chapter has shown the importance of complaint handling for companies, thus highlighting the relevance of this context in our research. In addition, it has shown that there is a relation between empathy and complaint handling, as empathy influences perceived interactional justice, and interactional justice in turn influences how successful the complaint handling is. This emphasizes the importance of this research, especially given the previous knowledge from chapter 2 which showed that chatbots are frequently used in complaint handling, but also lack emotional intelligence, thus are unable to show empathy. Given the theory, the following hypotheses were formed:

1. A chatbot showing empathic behavior will result in significantly higher perceived levels of interactional justice compared to a non-empathic chatbot.
2. Participants will report more positively on an interaction with an empathic chatbot compared to a non-empathic chatbot.

End Research Topics

Chapter 4

Related work

This chapter will focus on the related work on chatbots in complaint handling. This means that this chapter will be split into two parts, first the related work on empathic dialogue agents will be discussed, and second the different ways and strategies of complaint handling using chatbots will be reviewed.

4.1 Empathic conversational agents

This section will discuss and compare a variety of different empathic conversational agents (including embodied agents) to present the related work in this domain. There are roughly two groups of empathic agents: neural empathic agents and framework-based empathic agents. First, some theory on affective dialogue agents will be discussed, followed by some randomly picked examples of neural empathic agents and framework-based empathic agents in terms of the pre-mentioned theory.

Ma et al. [79] describe three key components which enable empathic behaviour in a dialogue agent based upon a literature review on the topic; *affective dialogue system*, *personalized dialogue system*, and *knowledge based system*. The affective dialogue system focuses on emotions during the interaction, including the recognition of the user's emotion and expressiveness of the agent's emotion. The personalized dialogue system contains a user modelling and a personalized response generation part. User modelling aims to make a representation of the user. The personalized response generation is the next step and focuses on the development of user appropriate language. Ma et al. [79] found two different ways of personalizing a response in the literature, through personality-awareness (the agent is aware of the personality of the user), or personality-infusion (the agent has their own personality). Lastly, a knowledge system is a critical component for empathic dialogue agents, as it enables them to keep track of context and use general knowledge to interpret and generate responses accurately [79]. All three components work are important to create an empathic dialogue agent [79]. Therefore, the next sections will focus on whether and how other researchers have implemented them.

Another important concept in empathic dialogue agents is related to the affective system; the mapping between emotion recognition and emotion expression of the agent [80]. An often used approach is mimicry [81–86] where the agent mimics the user's emotion back to them, sometimes in combination with a modulating factor [84–86]. An example is shown below. This mechanism is also referred to as one-to-one mapping.

1. User: (*sad*) I am very sad because my dog died last week.
2. Agent: (*sad*) I'm sorry to hear that, that must be awful :(

Whereas mimicry is seen as one of the essential and most basic aspects of empathy [84, 87], there are other approaches where researchers aim to find an alternative connection between the user's emotion and the response emotion. This can be possible done through careful analysis of dialogue flow or emotion-intent relations [80, 88].

Whereas neural empathic agents usually have an implicit empathy mechanism, their chosen mapping is often hidden. Therefore, the examples of the neural empathic agents will not be analyzed in terms of mapping. On the other hand, the section on framework-based empathic agents will be highlighting the commonalities and differences between the different types of mapping.

4.1.1 Neural empathic agents

Given the rise of deep learning in the field of NLP, it is no surprise that there are a variety of neural empathic agents. In this section, neural empathic agents are defined as fully data-driven models without including expert knowledge [89]. Now, this section will highlight a few examples of neural existing empathic agents.

First of all, Lin et al. [90] developed the CAiRE chatbot. They fine-tuned the Generative Pretrained Transformer (GPT) by Radford et al. [91] on two datasets, the PersonaChat dataset by Zhang et al. [92] and the EmpathicDialogues dataset by Rashkin et al. [93]. The latter is an often used dataset for empathic dialogue agents and consists of 25k labeled dialogues in emotional contexts. Through training on these datasets, the empathic behaviour of the chatbot is enabled by the inclusion of an implicit affective system and the inclusion of personalisation, specifically through personality-infusion, as described by Ma et al. [79]. In other words, CAiRE's language model was fine-tuned to be empathic and fit a persona. Since the empathic behaviour is completely trained upon human data, it is not one-to-one mapping, as humans show more variability in their empathic responses [88]. The researchers neural approach resulted in an end-to-end generative chatbot, which is able to respond to a wide variety of prompts. Lastly, the researchers enabled CAiRE to learn from user feedback, to increase the quality of its answers.

Hu et al. [94] developed a tone-aware chatbot to respond to user requests on social media. To get their data, they collected conversations between customer service and customers on Twitter. Then, they identified common tones in user requests on social media through a formative study. They distinguish between eight types of tones in customer service; empathic, passionate, satisfied, polite, impolite, sad, frustrated, and anxious. Empathic and passionate were found to be most beneficial for increasing user satisfaction. [94]. To build the chatbot, the researchers first identified keywords for each tone from the dataset, and trained the chatbot on these keywords to generate its own tonal responses. They used a sequence-to-sequence model to generate the responses, and included the keywords as tone indicators. In their work, they did not focus on matching tones to user requests or tones, but generated different tonal responses (empathic or passionate) to each request. After human evaluation Hu et al. [94] found significant positive effects in appropriateness and helpfulness for the tones empathic and passionate. They even discovered that their generated responses were perceived as more empathic than human agents. In terms of the key systems by Ma et al.[79], this chatbot includes part of an affective system (emotion

expressiveness but no recognition) and a knowledge system to understand requests, but lacks a personalized system.

The chatbot by Hu et al. [94] was inspired by the work of Zhou et al. [95] who also used the neural approach for their dialogue agent. However, where Hu et al. [94] focus on tones in customer service, Zhou et al. [95] focus on the general emotions anger, disgust, happiness, liking, and sadness. First, the manually labeled training data is classified into these emotions, and then for each prompt the system generates all the different possible emotional responses. Similar to Hu et al. [94], they used a sequence-to-sequence model to generate text. To include emotions in this model, they created embeddings for each emotion category. They also included a mechanism to keep track of the chatbot's internal emotional state, and a mechanism to generate emotional expressions. Again, they did not include a personalized system, but only focused on the emotion expressiveness and knowledge system of an empathic agent.

Another interesting chatbot was designed by Hoegen et al. [96]. Whereas this chatbot does not specifically focus on the affective system, it is centered around the personalised system. More precisely, the researchers focused on style matching between user and chatbot. They designed a voice-based conversational agent which was able to detect the language style of the users and adapt its own style to match. They distinguished between two styles; high consideration (long pauses, hesitations) and high involvement (overlapping speech, louder speech). After asking their participants to interact with the agent for 15 minutes, they found that participants with the high consideration style rated the agent as more trustworthy, whereas the high involvement style participants were indifferent. This study shows that style matching may improve perceived trust. However, it still needs to be tested in textual dialogues.

Wang et al. [80] think that the approach of Zhou et al. [95] and Hu et al. [94] lacks a fundamental part of empathic dialogue generation; the understanding and appropriate generation of empathic emotions. Similar to Lin et al. [90], they used the EmpathicDialogues dataset [93]. In their work they focus on the emotion flow of a dialogue, and how this can be used to predict the emotion of the agent's response to the user. To do so, they propose their SEEK (Serial Encoding and Emotion-Knowledge) method, which includes a model of how emotion and intent flow during the dialogue, which is used to select appropriate emotion to respond to the user. Additionally, they allow emotions and knowledge to interact, so emotions can play a role in selecting relevant knowledge for the situation. This harmony between the affective and knowledge mechanism was not seen in previous examples. Similar to Hu et al. [94] and Zhou et al. [95], they focus on the affective and knowledge system, but not on the personalization of the dialogue agent. When they compared their empathic agent to other baseline models, they found that SEEK outperformed them in terms of coherence, perceived empathy, and fluency.

Concluding remarks

A large variety of empathic dialogue agents have been developed based upon the deep learning approach. Whereas most of them are based upon one-to-one mapping, there is research into alternative mappings which show potential for increasing perceived empathy. The discussed agents all include an affective dialogue and knowledge system, however only some of them include a personalized system. Most agents show promising results in their tests with regards to perceived empathy. A benefit of the discussed neural approaches is that they can handle a variety of prompts and different users since they are trained on large amounts of data. Due to the nature of neural chatbots, the empathy mechanism is implicit in the system, and remains hidden after training. However, an explicit model may be useful for better understanding of the mechanisms of empathy. Therefore, the next section will focus on dialogue agents which were designed based upon frameworks rather than the data-driven approach of neural chatbots.

4.1.2 Framework-based empathic agents

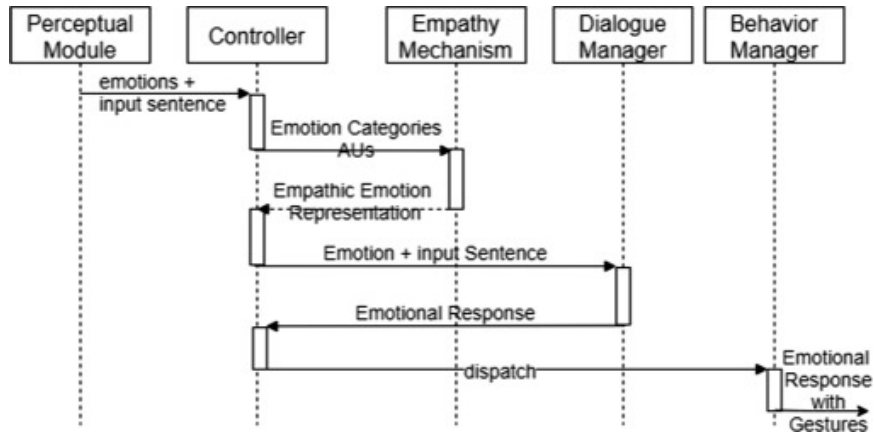
Whereas the neural approach shows promising results on perceived empathy, other researchers choose to focus on the design of an explicit empathy model. Therefore, this section will first shed some light on theoretical concepts related to empathic dialogue agent frameworks, and then highlight some proposed models of empathy for dialogue agents while keeping in mind Ma et al.'s [79] three key components: affective, personality, and knowledge.

First of all, emotions play an important role in empathic dialogue agents, since empathy is related to the perceiving, understanding, and experiencing emotions of another being [87]. It is specifically important for empathic agents with regards to the first feature (affective mechanism), so when designing a framework emotions should be defined properly. Ma et al. [79] found that the computational approach to defining emotions in empathic dialogue systems commonly fall into three categories, the appraisal approach, the dimensional approach, and the discrete approach. The appraisal approach is built upon the idea that emotions are responses to our evaluations of a stimulus, rather than reflexive reactions [97]. So this approach takes into account the context of the dialogue or situation. Another option is the use of a dimensional approach, in which emotions are represented as vectors in the valence-arousal-dominance (VAD) space [79][83]. Lastly, emotions can be represented with the discrete approach, where they are classified in categories [79]. Often a combination between the categories is used. Some researchers claim that the appraisal approach always needs to be combined with either the dimensional or discrete approach [87]. This way, emotions still get a clear label which enables straightforward mapping between user emotions and empathic response emotions.

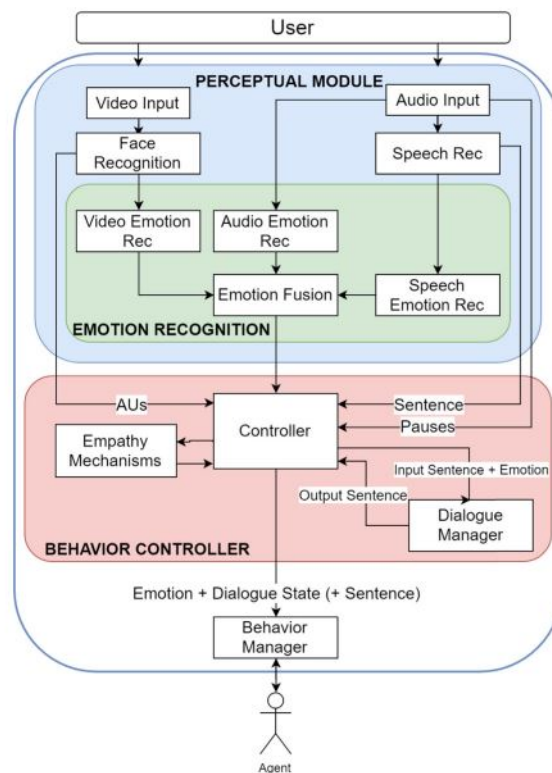
One-to-one mapping

Yalçın and DiPaula [84–86] created a framework to develop M-Path, a conversational system for an empathic embodied agent. The agent uses empathy mechanisms to generate appropriate emotional responses. Interestingly, in their design they make a distinction between low-level (e.g. mimicry), mid-level (affect regulation) and high-level (context reasoning and knowledge) empathic behaviour. This

resonates with the three key concepts by Ma et al. [79], where the low-level focuses on affective mechanism, both on recognition and generation of emotions in the VAD space. They used mimicry to map between the user emotion and response, making it an example of one-to-one mapping. Furthermore, the personality mechanism in the mid-level regulates the emotion found in the low-level based upon mood, liking and personality (i.e. changes the intensity). This shows that this level includes personality-infused characteristics. Lastly, the high-level focuses on the context and appraisal of a situation, like the knowledge mechanism in combination with the affective mechanism. In addition to this, the high-level includes a user model, which indicates the presence of personality-awareness. After evaluation Yalçin and Di-Paula [84–86] found that their approach resulted in higher perception of empathy, usefulness, human-likeness and believability compared to their non-empathic chatbot. The general mechanism of the M-Path chatbot is shown in Figure 4.1a, and a schematic of the empathy model of the chatbot is shown in Figure 4.1b.



(A) Information flow



(B) Framework of empathy mechanism

FIGURE 4.1: Models for an empathic conversational agent as designed by Yalçin and DiPaula [84]

Another empathic conversational agent is ERDAMS, by Ochs et al. [82]. They propose a formalisation of appraisal theory which allows an agent to deduce the appropriate empathic response in a dialogue. They use the appraisal theory to make predictions about the user’s emotions, and again use mimicry as empathy mechanism. After using the appraisal theory for predictions, they make a distinction between the discrete emotions *satisfied*, *frustrated*, *irritated*, *sad*, and *angry*. According to their formalisation, the intensity of an empathic response is defined by the degree of liking of the agent (i) towards the user (u), the degree to which the agent thinks that the user deserves to fulfil their intent (ϕ), and the intensity of the emotion of the user

according to the agent (c). These modulation factors show that a personality system is included in this framework. This is formalised by Ochs et al. [82] as follows:

$$intensity_emp_i(u, c, \phi) = like_i(u) * deserve_i(u, \phi) * c \quad (4.1)$$

ERDAMS is designed to help with using a mail system, meaning that background knowledge about the system and the requests is required. Therefore, we can conclude that this agent includes all three mechanisms of an empathic dialogue agent as described by Ma et al. [79].

Boukricha et al. [83] use the VAD space to model emotions in their empathic chatbot EMMA. Similar to Ochs et al. [82] the intensity of the empathic emotion is flexible and dependent on modulation factors (p_i) and their weights (w_i), in this case liking and familiarity. This shows the inclusion of the personality system. A difference with aforementioned dialogue agents is that EMMA's empathic response is only triggered when a significant change (change > threshold) in the user's emotion is detected. The use of thresholds allow the researchers to manipulate the responsiveness of the agent. Again, they use mimicry as empathy mechanism and focus on manipulating the intensity of the response emotion. The intensity is manipulated by different factors, one of them being the context of the situation, indicating the presence of the knowledge system. Another factor is that Boukricha et al. [83] use the distance between the agent's current emotion and the user's emotion to determine the intensity of the agent's empathic emotion. This inclusion was not seen with the previous dialogue agents. By including this, they allow the chatbot to experience and show more dynamic behaviour during the course of the dialogue. They formalise this as follows:

$$empEmo_{mod} = ownEmo + \frac{(empEmo - ownEmo) * (\sum_{i=1}^n p_i * w_i)}{\sum_{i=1}^n w_i} \quad (4.2)$$

Rodrigues et al. [87] use a combination of the appraisal approach and the dimensional approach. They use the appraisal approach to predict self-oriented emotions (emotions experienced by the agent), and combine these with the results of emotion recognition. Then, they define the emotion as a tuple <type, valence, intensity, cause>, where the *type* is a discrete emotion (joy, anger). Their proposed diagram of empathy is shown in Figure 4.2. Their framework highlights the importance of modulation factors, similar to Boukricha et al. [83], and includes the affective link, similarity, personality and mood as factors (inclusion of personality system). At the end of the *Empathic Appraisal* block, an empathic emotion is determined, and the correct behaviour is selected. Rodrigues et al. [87] explain that their model is based upon the assumption that the emotion recognition engine can return candidate empathic emotions based upon the recognized emotions. However, they do not give a detailed description on how this is achieved. Interestingly, in their evaluation they found that agents which were designed to be more empathic, were generally liked better by the participants.

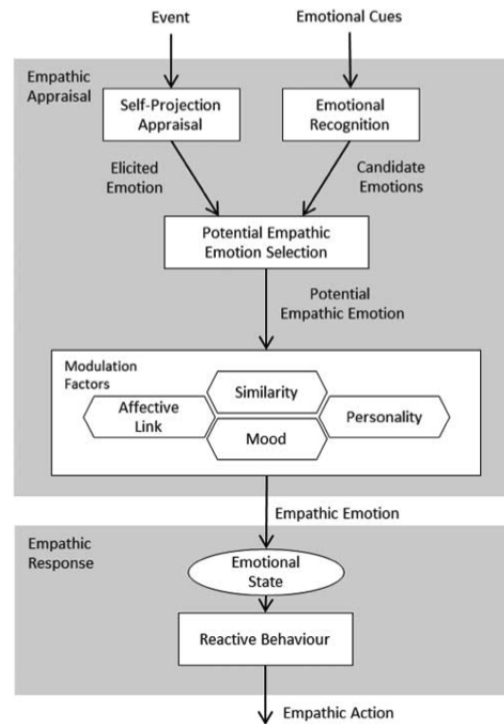


FIGURE 4.2: Empathy model diagram as proposed by Rodrigues et al. [87]

Alternative mapping

Another interesting work is by Lisetti et al. [81]. They created an empathic health intervention virtual agent. For emotion recognition they classify the results into the discrete emotions *happy*, *sad*, *angry*, *surprised*, and *neutral*. Then, they use a decision tree to select and generate the appropriate empathic response. An important aspect of their model is the inclusion of appraisal theory to predict the user's valence towards the agent's question, which is used in the decision tree. This shows the inclusion of knowledge and context. Furthermore, they include a personalized system to make a *user model*, and provide *tailored feedback* to the user's answers. When this virtual agent was compared to a text-only variant which did the same intervention, Lisetti et al. [81] found that 30% of users were more motivated to interact with the virtual empathic agent for future interventions.

Lastly, Welivita and Pu [88] created a taxonomy of empathic response intents. They used the EmpathicDialogues dataset [93] to detect the most used response intents in empathic dialogues. They found the following eight intents: *agreeing*, *acknowledging*, *encouraging*, *consoling*, *sympathizing*, *suggesting* and *wishing*. In addition, they identified the most common combinations of emotions and response intent, this can be seen in Table 4.1. This table highlights the importance of questioning to show empathy, which is backed by other research by Svikhnushina, Welivita and Pu [98] into different types of empathic questions.

Speaker emotion	Listener intent
Anticipating	Questioning Acknowledging
Joyful	Questioning Acknowledging
Trusting	Questioning Acknowledging
Surprised	Questioning Acknowledging
Afraid	Questioning Acknowledging
Sad	Questioning Sympathizing Acknowledging Agreeing
Disgusted	Questioning Acknowledging Agreeing
Angry	Questioning Acknowledging

TABLE 4.1: Most common emotion/intent pairs as found by Welivita and Pu [88]

Concluding remarks

There are many different types of frameworks for empathy in dialogue agents. Most researchers use mimicry as their empathy mechanism [81–86], but alternative methods are also used [81, 88]. Furthermore, the appraisal approach is commonly used among researchers to predict user emotions through evaluation of a situation [79, 81, 82, 84–86], usually in combination with more concrete definitions of emotions, either dimensional or discrete. All frameworks focus on the importance of context for emotion recognition, and show how personal factors (personality, mood, liking etc.) influence the empathic emotion that an agent can show.

4.2 Complaint handling

Since complaint handling is important for customer relations, a collection of research has focused on possible frameworks for smooth complaint handling [12, 51, 99, 100]. Some of this research focus on the complaint handling process, whereas other research focuses on the internal mechanisms of complaint handling, such as relations between the involved parties or strategies for complaint conversations. This section will highlight some of the most referenced frameworks on the complaint handling process and the internal mechanisms of complaint handling.

4.2.1 Complaint handling process

A commonly referenced complaint handling framework is designed by Stauss and Seidel [100]. They divide the process of complaint handling into 5 phases; (1) greeting phase, (2) aggression-reduction phase, (3) conflict settlement phase, (4) problem-solution phase, and (5) conclusive phase. They claim that the first phase determines whether the conversation will result in an argument or in a solution. During this phase, employees should have a friendly tone, indicate their willingness to make time for the complaint, and take the complaint seriously. In the second phase, the complainant should be able to freely voice their complaint and let off steam. According to Stauss and Seidel [100] it is important that the employee apologizes in this phase to show understanding and built a relationship with the complainant. A common pitfall in this stage is interrupting the complainant, which may eventually result in aggravating the complainant. After that, the conflict-settlement phase focuses on the objective facts of the complaint. Whereas employees should show empathy in this phase, they should not admit to guilt on their own or colleagues' behalf according to Stauss and Seidel [100]. The fourth phase is centered around finding a solution. Since the previous phases allow the complainant to tell their story, a solution in this phase is perceived to be based upon the conversation and helps them feel understood, making it more likely for the solution to be accepted. In the scenario where the demands of the complainant cannot be met, the employee should give a detailed justification. Lastly, the conclusive phase wraps up the conversation, checks whether the solution is understood and accepted, and allows the employee to end the conversation on a positive note (e.g. "I'm glad we could solve this problem together"). Whereas they do not explicitly include the post-analysis of the complaint in their framework, they do mention this as an important part of the complaint handling process.

Another often used framework is designed by Trappey et al. [12], shown in Figure 4.3. They include the reception of the complaint and the analysis of the complaint in their framework, which differs from Stauss and Seidel's [100] approach. Again, this framework shows that an apology and comforting of the customer should be done early in the complaint handling process. Then, the reason of the complaint needs to be understood and a way to remedy the situation is made. Also, Trappey et al. [12] include a circular route where they account for possible customer dissatisfaction after the remedy. We clearly see similarities between this framework and Stauss and Seidel's [100]. After all, it follows the same general framework, using the 5 phases of a complaint conversation: (1) greeting phase, (2) aggression-reduction phase, (3) conflict settlement phase, (4) problem-solution phase, and (5) conclusive phase. A difference between the two frameworks is that Trappey et al. [12] include

a circular route whereas Stauss and Seidel propose the framework as if it were linear. Lastly, they include the different types of complaints and the final steps of the complaint handling, where the complaint is stored for later use.

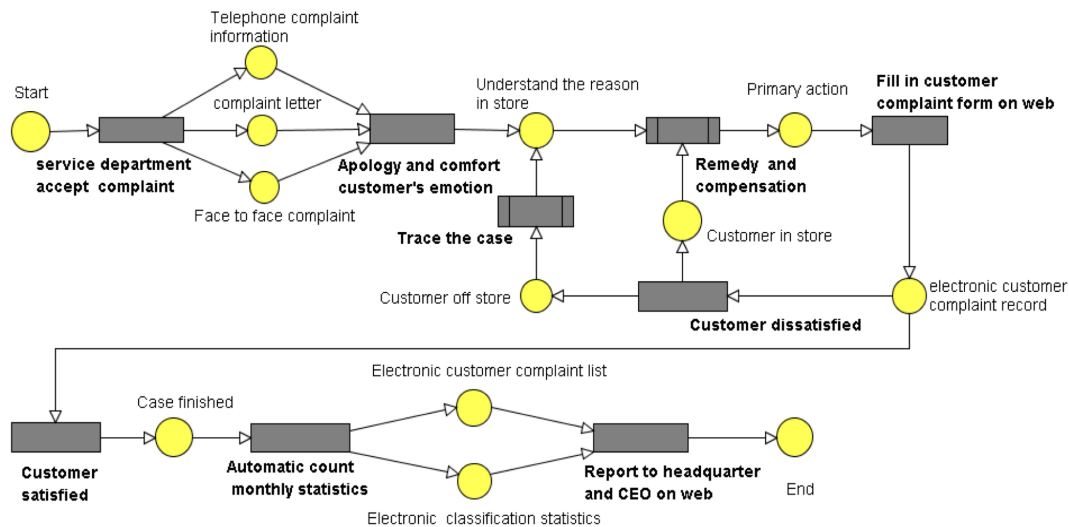


FIGURE 4.3: Top-level model of the complaint handling process as proposed by Trappey et al. [12]

Zairi [101] also researched effective complaint management systems, and designed their own framework, shown in Figure 4.4. The general flow of their framework is similar to the previous frameworks ([12, 100]), but they include a focus on selecting the appropriate team for the complaint handling, which was not seen before. In their first step, they receive the complaint, similar to Trappey et al. [12], and select the correct team to handle the complaint. The second step encompasses steps 2, 3, and 4 from Stauss and Seidel's framework, after which the problem can be solved and analysed for future use. What is interesting about their framework is that they include the 'assignment of responsibility' for the complaint, which was not present in the previous frameworks. In addition to this, they allow for a circular flow of the process, in which a new complaint can arise after analysing, or even solving the problem. A point of critique on this framework is the arrow to a new complaint after closing the problem, which seems counter intuitive and not supported by the frameworks of Stauss and Seidel [100] and Trappey et al. [12], which both use the closing of the complaint as finite stages. Unfortunately, no further explanation about this was provided by the author.

4.2.2 Internal complaint mechanisms

Razali and Jaafar [99] also designed a theoretical framework for complaint handling, focusing on how the different aspects relate to each other rather than the process. They identify four aspects necessary for complaint handling; *complaint service recovery*, *complaint*, *services*, and *customer*, which can be seen in Figure 4.5. The *complaint service recovery* attribute deals with the management of complaints, using the other three components for successful completion of complaints. This attribute can solve or reduce a *complaint*, which can vary in seriousness. Additionally, it can improve the quality of a *service*, which can be low or high quality to begin with. The characteristics of each block (*complaint* and *service*) influence what the type of complaint service

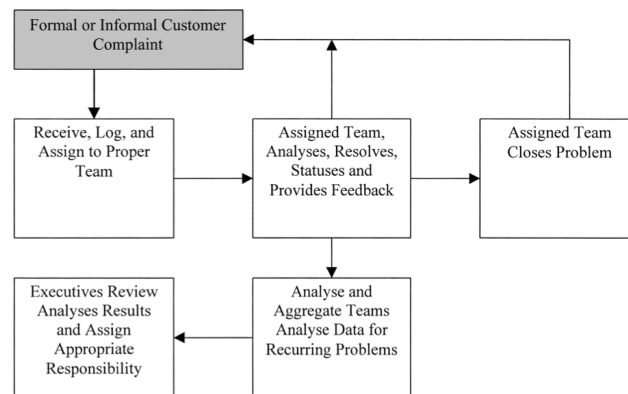


FIGURE 4.4: Proposed framework on complaint management by Zairi [101]

recovery is necessary. Lastly, the complaint service recovery can either satisfy or dissatisfy the *customer*. These relations are important to understand as they give insight into the internal mechanisms of a complaint handling system. Razali and Jaafar [99] also highlight key concepts in the complaint service recovery; apology, explanation, timely response, speedy recovery, easy access, and quality of the solution. Again, we see the importance of an apology and explanation of the problem, similar to Trappey et al. [12]. Razali and Jaafar [99] complement the frameworks in the previous section by providing more detailed information about the relations between different parts of the complaint service recovery, and showing feasible strategies to achieve goals.

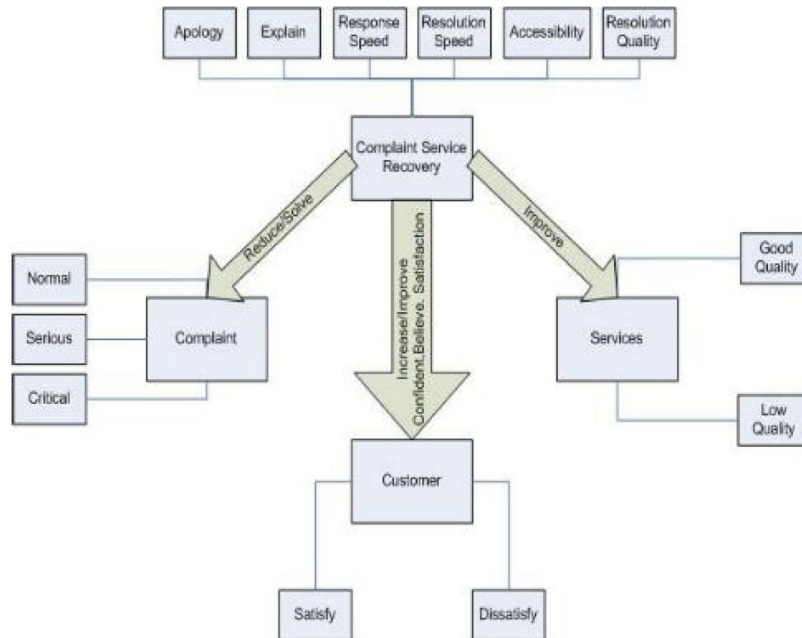


FIGURE 4.5: Framework by Razali and Jaafar [99]

Another research which provides clear strategies for complaint handling is proposed by Nguyen and McColl-Kenney's [51]. They specifically focused on diffusing anger in service recovery with empathy. Therefore, their work can be considered a detailed description of the *aggression-reduction* and *comfort customer's emotion* phases by Stauss and Seidel [100] and Trappey et al. [12] respectively. Interestingly, they use

the appraisal theory in their framework, similar to most researchers in section 4.1. Their proposed model can be seen in Figure 4.6. In their framework they make a distinction between the time before the service recovery (phase 1), and during the service recovery (phase 2). In the first phase, the amount of anger is determined by the extent to which the well-being of the customer is influenced by the problem (goal relevance), the extent to which an event meets the customers expectations (goal incongruence), and the extent to which a person's ego-identity is touched [51]. Then, they propose three strategies for reducing customer anger: (1) Listening to the customer, (2) Blame displacement, (3) Providing an apology. Listening to the customer as well as providing an apology was also mentioned in the other frameworks. Interesting about their work is their idea that displacement of blame may be useful in order to diffuse customer anger, which is based upon causal attribution theory [51]. Furthermore, they explain that blame displacement is successful when the blame is successfully moved away from the service provider (external cause to non-external cause).

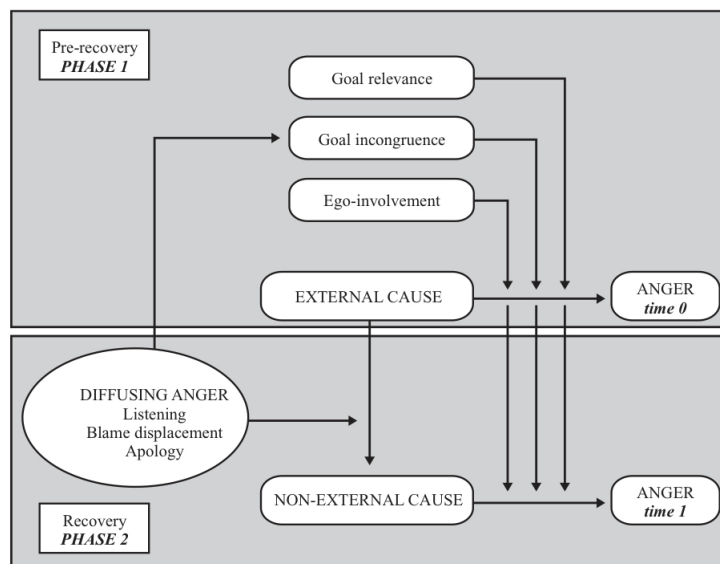


FIGURE 4.6: Model of diffusing customer anger strategies as proposed by Nguyen and McColl-Kenney [51]

4.2.3 Concluding remarks

Overall, there are different types of frameworks for complaint handling. Frameworks focusing on the route from complaint to satisfaction [12, 100], and frameworks focusing on the relations between different aspects of complaint handling or more detailed strategies [51, 99]. Most frameworks follow a similar procedural approach, which can be reduced or expanded to Stauss and Seidel's [100] five phases of service recovery. Lastly, all of the related work in this section highlights the importance of (1) listening to the customer (to let them vent), (2) providing an apology, and (3) post-analysis of the complaint.

4.3 Conclusion

This chapter discussed the different related work in the field of empathic dialogue agents and complaint handling. Both fields showed interesting research and developments.

The wide variety of empathic dialogue agents can be roughly divided into neural and framework-based empathic agents, which both have their strengths and limitations. Neural chatbots can usually handle a larger variety of prompts and use cases, whereas framework-based empathic agents provide more controllable empathic behaviour, and help in understanding the underlying principles. The significance of an affective dialogue system and a knowledge-based system was emphasized by most of the chatbots discussed, highlighting their essential role for empathic chatbots. Additionally, some researchers include a personalised system, which also increased the perceived empathy of the chatbot. Furthermore, researchers showed different types of mapping between recognizing and expressing emotions. Whereas most were based upon mimicry (sometimes including modulation factors), some researchers used alternative ways of mapping based upon theory. The main takeaways from this review are:

- The minimum requirement for an empathic dialogue agent includes an affective and knowledge system.
- An interesting and promising addition to these systems is the personalised system.
- Matching the agent to the personality/style of the user can improve liking
- Mimicry is an often and functional method to map the user's emotions to the agents empathic emotion.
- Emotions can also be matched to intents, which is a feasible strategy for a dialogue agent based upon intents.

Furthermore, the related work on complaint handling provided insight into the general process of the complaint handling as well as more detailed strategies. Research showed that listening to the complainant and providing an apology are key in the beginning of the complaint process. Then, a solution can be determined, which can be accepted or denied by the customer. As a result, the complaint handling process is not strictly linear, but can include loops between different phases. Lastly, the complaint should be analysed and used as feedback to improve the service. The main takeaways from the research in this field are:

- The 5 phases can be used as a base for a complaint process, but need to allow for loops.
- Assigning the correct person or team with knowledge to handle the complaint is important.
- Listening, apologising, and blame displacement may help diffuse customer anger.

These takeaways will be used in the next chapters for the design of my own empathic dialogue agent for complaint handling.

Chapter 5

Method

The purpose of this study is to compare the perceived interactional justice of participants interacting with an empathic chatbot versus a non-empathic chatbot. The study aimed to address the following research question and hypothesis:

What is the effect of empathy in chatbots on perceived interactional justice during service recovery?

H1: A chatbot showing empathic behavior will result in significantly higher perceived levels of interactional justice compared to a non-empathic chatbot.

To provide a more comprehensive understanding of the participants' experiences and behaviors during the interaction, an additional hypothesis about their overall impression of the interaction was formulated and answered using qualitative measures.

What is the effect of empathy in chatbots on the overall interaction experience of the participants?

H2: Participants will report more positively on their interaction with the empathic chatbot compared to the non-empathic chatbot.

An empathic chatbot prototype was developed as the initial step in the research process. Then, two types of studies were conducted to answer the research question. First, user tests were done to assess and if necessary, improve the level of perceived empathy for the empathic chatbot prototype. This prototype and the non-empathic version were then used to conduct the experiments to test interactional justice. Both experiments were between-subjects to avoid order effects and the level of perceived interactional justice was measured with a questionnaire.

To validate the design of the prototypes and account for the reliability and manipulation of variables, additional measures of perceived empathy and usability were added to the questionnaire in the experiments, focused on testing these additional questions and their hypotheses:

Is the empathic chatbot prototype perceived as more empathic than the non-empathic chatbot prototype?

H3: The empathic chatbot prototype will be perceived as significantly more empathic compared to the non-empathic chatbot prototype.

Is the perceived usability similar for both the empathic and non-empathic chatbot prototype?

H4: There will be no significant difference in perceived usability between the empathic and non-empathic chatbot prototypes.

This chapter will provide detailed information on the user tests, research design, participants, procedure, measures, analysis, and limitations of the experiment.

5.1 User tests

The user tests were used to assess and improve the empathy of the chatbot. During the user tests the participants were asked to interact with the chatbot according to the same scenario as during the real experiments. Then, they filled out a questionnaire and answered interview questions about the perceived empathy. Additionally, they were asked questions about the overall interaction, so new conversation topics could be added to the chatbot. A total of 4 user tests were done, with different stages of the chatbot.

5.2 Research Design

During the experiments the participants chatted with a customer service chatbot to complain according to a pre-specified scenario. Each participant would interact with the chatbot and answer the questionnaire which measured (1) perceived empathy, (2) perceived interactional justice, (3) usability. Details on the measurements will be provided in section 5.4. At the end, the participants were asked some additional questions about the interaction.

In the instructions the participants were told that their task was to complain about the late delivery of a package and try to receive some compensation. More details on the scenario will be provided in the next subsection. When the participants decided that they were done with the conversation (this could be at any point in the conversation), they were sent the questionnaire and asked the interview questions after.

5.2.1 Scenario

To minimize differences between interactions and thereby mitigating their effect on perceived interactional justice, a highly detailed and specific scenario was deliberately created. By providing participants with a detailed described scenario, the aim was to ensure that the interactions across different conditions would be comparable. The specific scenario was chosen to trigger an emotional response, as well as the fact that delayed packages are a likely reason for complaint [102].

In order to have a realistic and emotionally engaging user experience, a scenario was designed to elicit frustration by the participants. In this scenario, they were told that their order, intended as a birthday gift for their favorite cousin, had been delayed. As a result they had to show up empty-handed to the birthday party. The participants were specifically instructed to ask compensation for the inconvenience. This carefully designed situation aimed to increase the participants levels of frustration, making it possible for the empathic chatbot prototype to effectively showcase its empathic features. Importantly, the chatbots were unable to provide any compensation, further intensifying the participants' frustration throughout the interaction. The exact description of the scenario provided to the participants is shown below:

Imagine you are in the following scenario:

You are a customer chatting to customer support. You are frustrated because you ordered a package with express delivery but it is too late, it was supposed to be a birthday present for your favourite cousin. You even paid express delivery so it would be in time. Unfortunately it was not, the birthday was yesterday and the package has still not arrived. You want to receive some sort of compensation, but still want to get the package delivered to gift to your cousin.

The full instructions and scenario can be found in Appendix D, which also includes information on how to continue the conversation when the chatbot makes a mistake and what fake personal information to use.

5.2.2 Variables

In the study the effect of the presence of empathy in a chatbot (independent variable) on perceived interactional justice (dependent variable) was measured. To evaluate the manipulation of the independent variable, a measure of perceived empathy was added to the questionnaire.

While the presence of empathy was manipulated in the experiment, the overall functionality (confounding variable) of the chatbot had to be controlled for it to be similar across groups. Therefore, the empathic chatbot was duplicated to become the non-empathic chatbot, and only the empathic additions and framework were deleted from the code. Additionally, a measure for usability was included to test whether both chatbots had the same functionality.

5.3 Procedure

The experiments were conducted online via Microsoft Teams or Zoom to allow the experimenter to keep some control over the chatbot's intent classification (i.e. in case of misclassification, the experimenter could correct the chatbot). Additionally, this made the experiment accessible to a larger pool of potential participants which in turn increased diversity and generalizability of the sample. During the experiment the experimenter would turn off their video and audio to make the participant feel less aware of the experimenter's presence. After the instructions they would share their screen of the chatbot application with the participant. The participants were asked to type their response in the Teams or Zoom chat and the experimenter would copy and paste them to the chatbot. An example of this setup is shown in the screenshot in Figure 5.1. The participants were free to interact as long as they wanted. When they signalled to be done with the conversation they were sent the questionnaire, and then asked the interview questions later. The experiment ended with a short debriefing and time for questions from the participants.

5.3.1 Step-by-step

1. The experimenter sends the participant the information brochure and consent form before the experiment.
2. The experimenter sends the participant a link to the Zoom (or Teams) meeting.

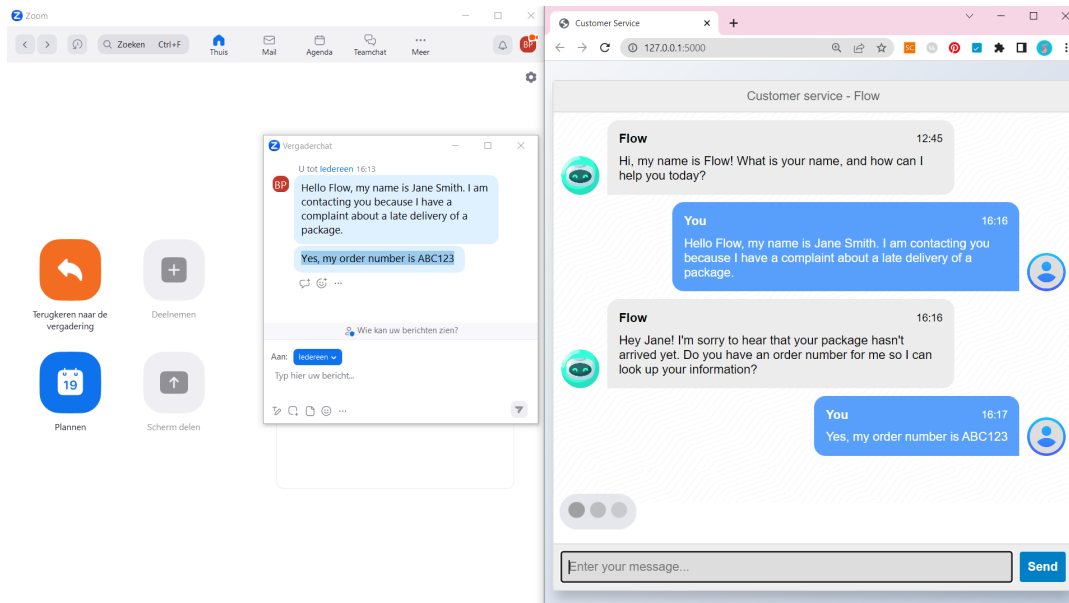


FIGURE 5.1: Screenshot of example setup

3. The experimenter goes over the information in the brochure, and explains the procedure and the consent form. The participant signs the consent form when they agree.
4. The experimenter sends the instructions of the scenario to the participant to read.
5. The chatbot application is started, the experimenter shares their screen and explains where the participant has to type their responses.
6. The participant interacts with the chatbot until they decide that they are done.
7. The experimenter sends the questionnaire to the participant and asks to fill it out and let them know when they are done.
8. The experimenter asks the three interview questions.
9. The experimenter gives the debriefing in which the true purpose of the study is revealed and the participant is told whether they interacted with the empathic or non-empathic chatbot. In this step the participant is asked if they have any questions.
10. The experimenter thanks the participant for their time and lets them know to feel free to ask any questions that may arise later.

5.3.2 Data collection

All data collected during the experiments were given a unique ID per participant, so the results could be compared later. No personal identifiable information was collected (except on the consent forms), and the participants were told to use a fake name while interacting with the chatbot.

During the experiments four different types of data were collected:

1. The log of the conversation (including timestamps, detected intents, emotions etc.)

2. The answers to the questionnaire
3. The answers to the interview
4. Notes taken by the experimenter during the experiment and interview

5.4 Measures

During the experiments different types of measures were done. These can be divided into quantitative measures and qualitative measures. The quantitative measures include the questionnaire on perceived empathy, perceived interactional justice, and usability. Additionally, some qualitative information was stored in the logs. The qualitative measures include the notes taken by the experimenter and the answers to the interview questions.

5.4.1 Quantitative Measures

A measure of *perceived empathy* was introduced to test the validity of the manipulation of empathy presence. For this the Robot's Perceived Empathy (RoPE) scale by Charrier et al. [76] was used. This scale is based upon human empathy metrics and validated with experts from cognitive sciences and robotics. For this study items ER4 ('The robot encourages me') and ER5 ('The robot praises me when I have done something well') were removed as they do not fit the context of the study. All the items were transformed to past tense to make it more intuitive for participants, since the questionnaire was filled in after the interaction. Additionally, the word 'robot' was replaced with 'chatbot' to fit the study better. A 5-point Likert scale was used for this measure.

A measure of *perceived interactional justice* was used to measure the dependent variable. The questionnaire to measure this was adapted from an existing scale to measure perceived interactional justice by Waqas et al. [77]. To fit the context the words 'employees' and 'university' were replaced with 'chatbot'. Again, a 5-point Likert scale was used to keep the whole questionnaire cohesive.

A measure of *usability* was included to check whether the functionality of the chatbot was similar across groups. For this an adaptation of the Bot Usability Scale (BUS) by Borsci et al. [103] was used. The perceived accessibility measures and items 4, 5, 7, and 8 were removed, as they were irrelevant during the experiments. Additionally, the item about response time of the chatbot was changed from 'short' to 'satisfactory', since user tests showed that participants disliked too fast responses in the initial bot. The scale was implemented with a 5-point Likert scale.

In short, the questionnaire consisted of a total of 34 items, measuring perceived empathy, perceived interactional justice, and usability. The survey was split into two parts in a Google Forms, to avoid lengthy scrolling for participants.

5.4.2 Qualitative Measures

After filling out the survey, the participants were asked 3 questions with some follow-up questions. These questions were added to gather qualitative info on the interaction, as well as being a measure of ecological validity.

1. What are your initial thoughts after this interaction?

- (a) Did anything stand out for you during the interaction? Why?
2. Do you think this interaction is realistic?
3. Do you think the interaction would have been the same with a human customer service representative?
 - (a) Do you think you would have acted the same with a human customer service representative?

During the interaction the experimenter would also take notes on interesting behaviour or comments of the participants. Lastly, a log of the conversation was saved, including the time of interaction, detected user intent and emotion by the chatbot, and the actual conversation. These were collected to gain more insight in the conversation, and to check for other correlations during the analysis of the results.

5.5 Participants

A total of 25 participants took part in the experiment, ages between 19-57 (m=10, f=15). No elderly participants were asked to participate to control for too much difference in technological experience. Since the chatbot spoke English, they were asked beforehand whether their English was good enough to hold a conversation. Besides this, all people without previous knowledge of the project were allowed to join the experiments. This was possible since all types of people can interact with a customer service chatbot in real life. Additionally, this broadened the range of possible participants.

Participants were approached via group chats or via other participants and received no compensation. Before the experiments they were informed of this, and general information on the experiment was provided.

5.5.1 Sample group distribution

Group 1 interacted with the empathic chatbot, and consisted of 13 participants, aged 19-53. Group 2 interacted with the non-empathic chatbot, and consisted of 12 participants, aged 19-57. The experimenter strived for an equal distribution of male and female participants between groups. The same was done with regards to experience with chatbots / AI, which participants were asked informally before the experiments.

5.5.2 Ethical considerations

Several considerations were taken into account while recruiting participants. First, no vulnerable groups were approached, and all participants were informed that their participation was completely voluntary. Before participation they were sent the information brochure and consent form, so they had the option to read through them. Both documents were available in English and Dutch. When starting the experiment, the experimenter would explain the information brochure (regardless of whether they read it beforehand), and ask them to sign the consent form if they wanted. During the experiment, they were able to ask questions, which were directly answered by the experimenter.

5.6 Limitations

A limitation of the used procedure is that participants couldn't directly interact with the chatbot, but had to type their responses in a different chat. This may influence the results, but this choice was made to make the functionality of the chatbot more reliable. During user testing it appeared that the chatbot would often lack confidence in classifying intents, or choose the wrong intent. In order to control for too many differences between participants, the possibility for the experimenter to intervene was added. This way, the intent classification would almost always go correctly for all participants.

5.7 Conclusion

The method described above will be used to find out whether there is any difference in perceived interactional justice for an empathic and non-empathic chatbot. The between-subject design and test of usability provide a degree of control over the confounding variables. Additionally, the design of the scenario was chosen to trigger an emotional response from the participants, while also being a realistic scenario. During the experiments both quantitative and qualitative data was collected. Including a questionnaire measuring perceived interactional justice, empathy and usability, and open-ended interview questions about the overall interaction and ecological validity. Whereas a big limitation was the indirect interaction of the participants with the chatbot (via Zoom or Teams), this allowed for more control by the experimenter.

Chapter 6

Realisation

To conduct the experiments, a prototype for the chatbot was realised in a systematic realisation phase. This started with a design phase, in which decisions about the empathy module and complaint handling framework were made. Then, the design was implemented into two separate prototypes - one with empathic behaviour and one without. Lastly, pilot tests were conducted to test different stages of the prototype and refine the design further.

6.1 Design

The design phase consists of two main components: the design of the empathy module for the chatbot and the design of the complaint handling framework. Also, a requirements list for the chatbots was established. This section will explain significant design choices made during the development of the empathic prototype.

6.1.1 Empathy framework

The first step was creating a general framework for empathy inspired by the aforementioned empathy models by Rodrigues et al. [87] (see Figure 4.2) and Yalçın and DiPaula [84–86]. Particularly, the new framework was largely based upon Yalçın and DiPaula's work, which provided the foundation for its construction. Similar to their framework (see Figure 4.1b), the new framework includes a perceptual module and a behavior controller. The resulting framework, shown in Figure 6.1, uses the perceptual module for analyzing input text, the behavior controller for generating appropriate responses based upon the user intent, emotion, and available knowledge, and includes an empathy mechanism responsible for mapping user emotions to the chatbot's empathic intents.

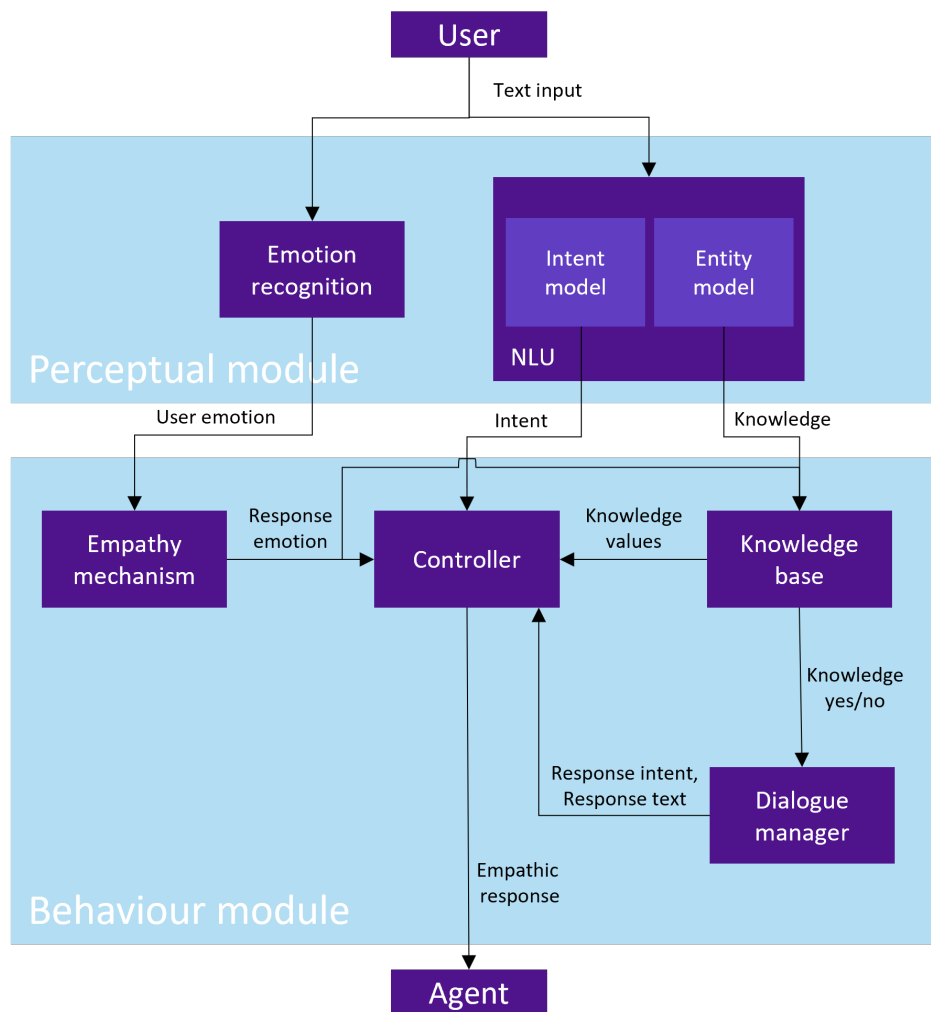


FIGURE 6.1: General empathy framework for dialogue agent

Ma et al.'s. [79] three components for empathic behaviour were implemented to make the chatbot empathic. Each of their implementations will now be discussed in more detail.

Affective dialogue system

In the affective dialogue system, the recognition and representation of emotion is key. Plutchik's eight primary emotions (joy, trust, fear, surprise, sadness, disgust, anger, and anticipation) were selected as discrete emotion categories. This choice was based on Bulgang et al.'s [104] findings, which suggested that these emotions were the most widely used and recognized in previous research. Additionally, this representation allowed for more emotion recognition possibilities.

The output of the emotion recognizer served two purposes: (1) to map the users emotion to an empathic response, and (2) to enable emotion validation responses. Both are examples of cognitive empathy. Whereas affective empathy may increase the perceived sincerity of human employees in customer service, research has also shown that it may have the opposite effect and annoy customers [61]. Given that a chatbot is usually perceived as 'unfeeling' [9, 19], the feeling of insincerity may be amplified, diminishing the only positive effect. Therefore, affective empathy was

Emotion	Empathic intents
Anticipation	Questioning, acknowledging
Joy	Questioning, acknowledging, wishing
Trust	Questioning, acknowledging
Surprise	Questioning, acknowledging, neutral
Fear	Questioning, acknowledging
Sadness	Questioning, sympathizing, acknowledging, agreeing
Disgust	Questioning, acknowledging, agreeing
Anger	Questioning, acknowledging, suggesting
Neutral	Neutral

TABLE 6.1: Overview of mapping between emotions and empathic response intents

mainly shown through expressions to offer emotional support such as 'I understand' or apologies to portray affective empathy [61], and cognitive empathy was prioritized through empathic response mapping and emotion validation.

An essential part of the affective dialogue system is the mapping between the recognized emotion and empathic responses. The chosen mapping is based upon the research by Welivita and Pu [88]. Their study on a substantial amount of empathic conversational data resulted in a mapping between emotions and empathic response intents. Additionally, they made use of Plutchik's primary emotions, making it straightforward to map with the emotion recognizer. An overview of the mapping is shown in Table 6.1. In addition to empathic mapping, an implementation of emotion validation was added to help customers regulate their emotions. Given the aforementioned benefits of emotion validation in complaint handling scenarios [66, 67, 72], it is a logical addition to the affective empathy system.

Overall, the affective system consists of two components: (1) mapping the user's emotion to an empathic response intent and (2) enabling emotion validation responses when customers appear angry.¹

Personalized dialogue system

Given the limited interaction time (a conversation was estimated to take 10 minutes), it was not possible to make an extensive personalized dialogue system. Nevertheless, this empathic component includes some basic user modelling and uses personality-infusion [79] for the personalized response.

In combination with the knowledge base system, the agent is able to store data about the user, resulting in some user modelling. Strategies such as personalizing responses based upon the knowledge and addressing customers by their names, should increase the feeling of personalized responses [105]. To infuse a basic personality into the chatbot, the name "Flow" and a minimalistic, friendly avatar (see Figure 6.2) was selected from Flaticon². The name was chosen to be gender-neutral, to control for any gender biases. The avatar was selected to have a friendly but

¹As discussed in chapter 4 some research also includes modulation factors to determine the intensity of the empathic response [82, 83]. However, this aspect was not included in the prototype for two reasons. (1) given the context of customer service, it is not beneficial to create an empathic agent which can get angry or dislike the customer, (2) these modulation effects focus on an empathic emotional response, whereas the decision was made to use an empathic intent response for this prototype.

²https://www.flaticon.com/free-icon/chatbot_5226034

robot-like look, balancing between customer service friendliness and transparency regarding the chatbot's artificial nature.



FIGURE 6.2: Avatar of chatbot

Knowledge base system

The knowledge base system was designed to improve the customization of responses with knowledge and to identify opportunities for empathic questions [98]. Relevant information, such as the user's name, order number, birthday, and recognized dominant emotion, was stored during the conversation. This information was then used to set rules and personalize responses, and to determine the appropriateness of empathic questions. Based on the research by Svikhnushina et al. [98], which identified 'requesting information' as the most common empathic question act, this study will also use 'requesting information' as empathic question act.

Storing the user's name helped personalizing the responses, and repeating the order number back to the customer enhanced their feeling of being understood. The information on the birthday was used for two purposes. First, when the customer mentions a birthday, the chatbot starts using the word 'present' instead of 'package' to make the conversation more personal. Second, when there is no mention of the birthday, the chatbot can ask whether there was a special occasion, as an empathic question [98]. Lastly, storing the emotion of the user enabled accurate emotion validation.

The combination of the affective dialogue system, personalized dialogue system, and knowledge base system aimed to increase the perceived empathy of the chatbot. The affective dialogue system recognizes and maps user emotions to empathic response intents, and allows for emotion validation responses. The personalized dialogue system includes limited user modelling by gathering some personal information (name, issue), to use during the conversation to increase the feeling of personalized responses, and infuses the chatbot with a personality through a name and avatar. Lastly, through the knowledge system the chatbot can personalize its utterances, include the user's name, and identify opportunities for empathic questions.

6.1.2 Complaint handling framework

The complaint handling framework designed for this research was largely based upon the proposed guidelines for a typical complaint handling conversation by Stauss and Seidel [100]. They put forth the following 5 phases for a typical complaint conversation; *greeting*, *aggression - reduction*, *conflict settlement*, *problem - solution*, and *conclusive*. For the purpose of this study the researcher added another relation between the problem-solution and the aggression-reduction phase to Stauss and Seidel's framework [100]. This was done to address situations in which no satisfactory

Phase	Empathy Strategies
Greeting	- Use of user name - Empathic responses
Aggression-reduction	- Personification using knowledge - Emotion validation - Apology - Empathic responses
Conflict settlement	- Personification using knowledge - Empathic questions - Empathic responses
Problem-solution	- Emotion validation - Empathic responses
Conclusive	- Use of user name - Empathic responses

TABLE 6.2: Empathic strategies per phase

solution could be found, leading to new/more customer frustration. This addition was based upon Trappey et al.'s circular framework [12]. The resulting visualisation of the different phases and their key components are shown in Appendix B.

The aggression-reduction and conflict settlement phase will get the highest focus on empathy. In these phases there are plenty of opportunities for apologies, empathic questions, and emotion validation. Additionally, in the problem-solution phase, the chatbot will be denying user's requests for compensation, giving the opportunity for more apologies and emotion validation. In the other phases, empathy is also included, but only through phrasing of responses, according to the pre-described emotion-intent mapping.

A complete overview of which general strategies are employed in which phase can be found in Table 6.2. In this table *Empathic responses* indicate normal responses which have been slightly adapted according to the empathic intent mapping [88]. *Emotion validation* refers to responses which validate and point out the user's feeling, such as: "I understand that delays in \$PACKAGE delivery can be frustrating and I apologize for the inconvenience caused to you."

6.1.3 Requirements List

During the design phase, a requirements list for the chatbot was created based upon the aforementioned empathy mechanisms and complaint handling framework. Together, these requirements should ensure that the chatbot exhibits empathic behaviour while also adhering to a complaint handling conversation flow. These requirements can be regarded as logical conclusions to the previous sections.

1. **Emotion recognition:** the chatbot should be able to recognize and classify user emotions based on their input text into Plutchik's 8 primary emotions [106].
2. **Intent classification:** an intent classification system should be implemented.
3. **Empathic response generation:** the chatbot should generate empathic responses based on recognized emotions and the mapping between emotions and empathic intents.

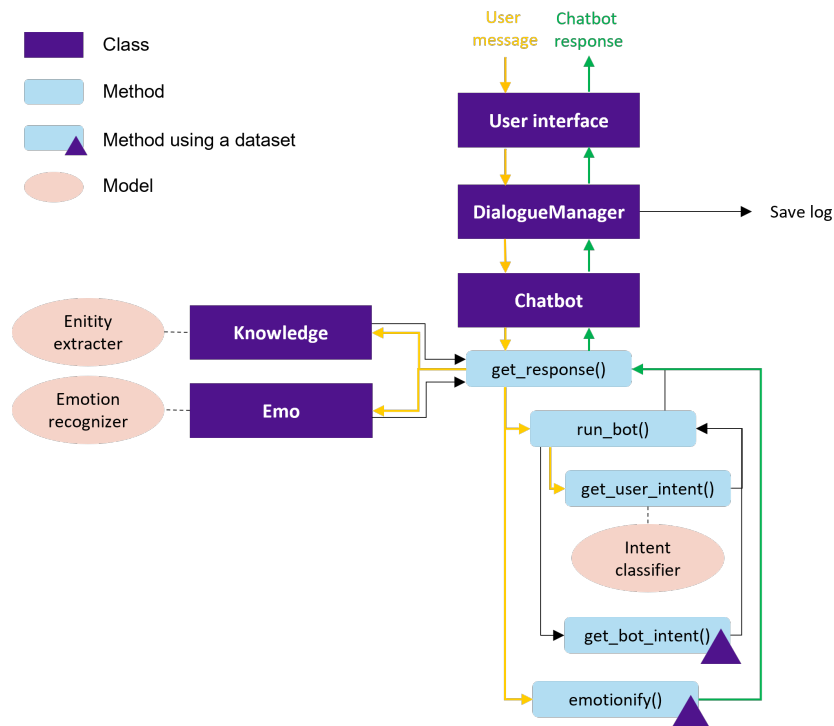


FIGURE 6.3: Architecture of Chatbot

4. **Personalization:** the chatbot should be able to personalize sentences and address users by their names when known.
5. **Knowledge base system:** the chatbot should store relevant user information (name, order number, birthday, emotion).
6. **Complaint handling framework:** the conversation should follow the complaint handling framework, emphasizing empathy in the aggression reduction and conflict settlement phases.
7. **User interface:** the chatbot should have a user-friendly interface showing its avatar and name.

6.2 Implementation

After defining the empathy and complaint handling framework, they were implemented into two prototypes: a non-empathic chatbot and an empathic chatbot. The conversational flow and complaint handling framework of the empathic and non-empathic chatbots was kept the same, but the empathic strategies were only implemented in the empathic chatbot. Both prototypes were built using Python, with additional use of HTML, CSS, and JavaScript for the user interface. The high-level architecture of the system is shown in Figure 6.3.

In the next sections more details will be provided on the basic (non-empathic) chatbot, including relevant classes and important functions. Then, the additional empathy modules for the empathic chatbot will be discussed in more detail. After this, the dialogue structure, selection of response messages and models used in the implementation are briefly discussed.

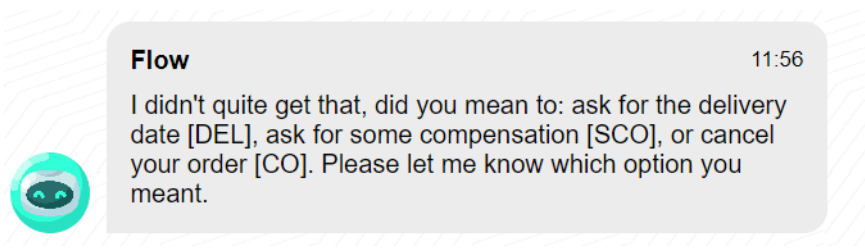


FIGURE 6.4: Example of Flow's fallback response

6.2.1 Non-empathic chatbot

The functionality of the chatbot was implemented using two main classes: the *DialogueManager* and *Chatbot* class. These classes were responsible for managing the conversation and handling user interactions.

First, the *DialogueManager* class manages the conversation. It saves a log containing all relevant information about the conversation and instantiates the *Chatbot*. The most important method in this class is the *query_bot()* method. This method queries the *Chatbot* for a response to the input message. Additionally, it checks if this is the user's first message, and adds a greeting to the response if necessary.

The *Chatbot* class uses a decision tree structure to handle different bot intents and user interactions. It includes the chatbot's knowledge, response options, previous responses and fallback responses. Important methods in this class are *get_response()*, *run_bot()*, *get_user_intent()*, and *get_bot_intent()*. The *get_response()* function is responsible for returning the appropriate response. This overarching method extracts and updates the knowledge from the input message and then calls the *run_bot* method to get the appropriate bot's response intent and corresponding response message using a custom dataset (subsection 6.2.3). Then, the conversation history is updated and, if necessary, relevant knowledge is added to the response message. In special cases (identified by the *run_bot()* method), such as asking for a manager or misunderstanding user input, a response is directly returned to the *DialogueManager* without adding knowledge to the message.

As mentioned previously, the *run_bot()* method selects the appropriate bot intent based on user input. It also handles special cases and returns their appropriate response message. This method determines and returns the bot's response intent.

The *get_user_intent()* method in the *Chatbot* class determines the user's intent based on the user message. It uses a custom classifier model to predict intent probabilities and selects the intent with the highest probability. If the highest probability is below a threshold (determined during testing phases), a fallback response is returned. Depending on the previous conversation, the chatbot then provides the three intents with the highest probabilities as options for the user to choose from. An example is shown in Figure 6.4.

Lastly, the *get_bot_intent()* is responsible for picking the correct bot intent based on the user intent using a decision tree mechanism. Additionally, this method includes functionality to make the conversation more natural and dynamic. For example, it checks if a user requested a refund and how many times this request has been denied. Then, after denying the request twice, the chatbot will offer a more detailed explanation why a refund could not be given. This feature allows for a more realistic behaviour and context-awareness by the chatbot.

6.2.2 Empathic chatbot

The empathic chatbot follows the same structure but includes additional empathy mechanisms. In the Chatbot class, an emotion recognizer is added to extract the user's emotion from the input message. If necessary, a lengthy input text is split into separate sentences to get a more accurate representation of the overall emotion. The extracted sub-emotions are then combined to find the *dominant emotion*.

Another addition is empathic mapping, possibility for emotion validation, and further utilization of the knowledge base. Again, the appropriate response is selected from a dataset (subsection 6.2.3) with different response sentences, but now the selection is based on both the appropriate response intent and the user's dominant emotion. To implement this, for each response intent 6 different phrasings are included, based upon research by Welivita and Pi [88]: questioning, acknowledging, wishing, neutral, sympathizing, and suggesting. It should be noted that an additional implementation in the chatbot prevents the consecutive use of the same empathic intent, to avoid repetitiveness. The response dataset also includes emotion validation sentences and an implementation for the knowledge base. This is done by adding variables in the response sentences ('PACKAGE', 'ORDER_NUMBER', 'EMOTION', 'PERSON') which is replaced with the appropriate knowledge when outputted. The 'EMOTION' variable is included for emotion validation. For example in the sentence "I completely understand that you can feel \$EMOTION when a \$PACKAGE is delayed. (...)", the 'EMOTION' variable is replaced with the current detected emotion³. This functionality is implemented in the Chatbot's `get_response()` method. Furthermore, the `get_bot_intent` method is elaborated with a function to check whether there is information about a birthday, so a follow-up question about a special occasion can be asked. This enables the empathic questions strategy.

6.2.3 Response dataset

The initial response messages for each bot intent are mainly based upon example sentences from Stauss et al.'s Effective Complaint Handling book [100], but also inspired by Jenneboer et al. [9], the TweetSumm dataset [107] and additional online sources⁴⁵⁶. Each bot intent has three slightly different response phrasings (to avoid repetition) or 6 empathic response messages, for the non-empathic and empathic chatbot respectively.

For the empathic chatbot, the phrasings of each empathic intent is carefully adapted from a neutral message based on the most common phrases used for each intent, as identified by Welivita and Pu [88]. An example is shown in Table 6.3, which shows that the general intent of each response was the same, but conveyed different empathic intents. For some intents (e.g. informing about expected delivery date) no empathic response was deemed necessary and the phrasing was kept the same. In some exceptions, the phrasing for a specific empathic intent was illogical with the content of the response, and a neutral response was used. A complete overview of the response dataset for the empathic agent can be found in Appendix F. Whereas

³For phrasing purposes each recognized emotion is translated into a more suitable conjugation. E.g. 'anger' becomes 'angry'.

⁴<https://www.wpdownloadmanager.com/live-chat-scripts-and-customer-service-phrases/>

⁵https://www.comm100.com/wp-content/uploads/2019/07/application/pdf/Comm100_ebook_101Scripts.pdf

⁶<https://www.ltvplus.com/wp-content/uploads/2019/06/LTVplus-100-tried-and-tested-cs-phrases.pdf>

this dataset is similar to the one used for the non-empathic agent, the non-empathic dataset only includes the neutral responses with a few variations.

Neutral	<i>I understand that you experienced issues with our service, what happened?</i>
Questioning	I understand that you experienced issues with our service. Would you mind sharing what happened?
Acknowledging	I bet it must be frustrating to experience issues with our service. What happened?
Wishing	I completely understand how frustrating it can be to experience issues with our service. I hope we can find a solution. What happened?
Sympathizing	I'm sorry to hear that you experienced issues with our service. What happened?
Agreeing	I completely understand how frustrating it can be to experience issues with our service. What happened?
Suggesting	I understand how frustrating it can be to experience issues with our service. Perhaps we can work it out together. What happened?

TABLE 6.3: Example of different phrasings per empathic intent

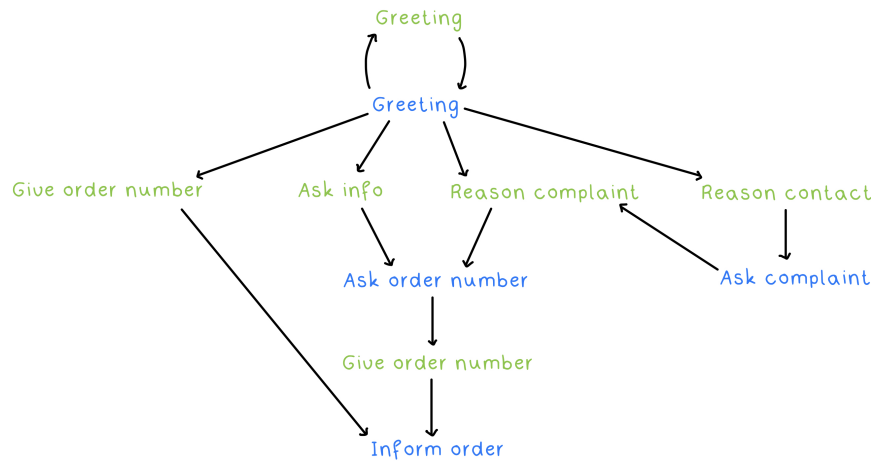
6.2.4 Dialogue structure

The flow of the conversation is represented by a decision tree. This structure is also used to determine possible response intents for each user intent. The decision tree is constructed based upon the complaint handling phases [100]. First a draft of the decision tree was sketched (Figure 6.5a), including the different intents and their connections. Then, this information was added to a JSON file (Figure 6.5b).

The decision tree provides a controllable structure and ensures that all users go through similar steps of the complaint handling process. As a result the results are more cohesive and comparable. Additionally, this allows for more functionality in terms of follow-up responses and less dependence on the variability of humans. The intent options are determined through common phrases used in customer service. Similar to the response messages, this is based upon previous research [9, 100], the TweetSumm dataset [107] and online sources³⁴⁵. Additionally, dialogues were simulated with ChatGPT⁷ to identify unforeseen intents.

The final implementation of the decision tree is based on the 5 complaint handling phases discussed earlier. It was decided to avoid a direct link between the aggression-reduction and problem-solution phases in the implementation in order to keep control over the tree's complexity. Instead, aggression-reduction methods, such as providing additional explanations, offering apologies, and validating emotions, are added as options within the problem-solution phase. Furthermore, in response to user feedback indicating their want to ask further questions even after indicating being finished, the option to try for compensation is added in the conclusive phase. A comprehensive overview of the decision tree can be found in Appendix C. It should be noted that an additional 10 response intents are included as special cases (in case of fallbacks), which can be manually selected at any time in the conversation regardless of, and without any effect on, the stage in the decision tree. The additional special intents to the chatbot prototypes serves to streamline and control the conversation,

⁷<https://openai.com/blog/chatgpt>



(A) Sketch of Greeting phase of decision tree
(green = user intent, blue = chatbot intent)

```

"intents": [
  {
    "bot_intent": "greeting",
    "user_options": {
      "1": "greeting",
      "2": "give_order_number",
      "3": "ask_info",
      "4": "reason_complaint",
      "5": "reason_contact"
    },
    "parents": {
      "1": "greeting"
    }
  },
]

```

(B) Example json entry of decision tree

FIGURE 6.5: Info decision tree dialogue

avoiding the need to account for every possible user prompt. An overview of these intents can be found in Appendix E.

6.2.5 Models

This section describes the key models used in the prototypes, including an Emotion Recognizer, an Intent Classifier, and an Entity Extraction model. These models play essential roles in the chatbot's functionality, such as understanding user inputs, recognizing emotions, and extracting relevant knowledge from user input.

Emotion recognizer

The Emotion Recognizer uses the NRCLex library⁸ for emotion recognition to help the chatbot understand the emotional context of the user interaction. By analyzing the textual content of the user message, it determines the dominant emotion expressed within the input. To fit the purpose of the study, the lexicon was adapted slightly. For example, words as 'birthday' would usually have a positive connotation, but in this case rather express a negative emotion.

```
1 def emotion_recognizer(msg):
2     # Take message as input and return highest ranking emotion
3
4     emotion = NRCLex(msg)
5     aff_freq = emotion.affect_frequencies
6
7     # Remove the unnecessary dictionary items
8     del aff_freq['negative'], aff_freq['positive']
9
10    # Necessary debug
11    if 'anticipation' in aff_freq:
12        del aff_freq['anticipation']
13
14    max_value = max(aff_freq.values())
15    top_emo = {k: v for k, v in aff_freq.items() if v == max_value}
16
17    # Replace 'anticip' with 'anticipation'
18    if 'anticip' in top_emo.keys():
19        top_emo['anticipation'] = top_emo.pop('anticip')
20
21    return top_emo
```

LISTING 6.1: Python code for emotion recognizer function

Intent classifier

The Intent Classifier is a custom model developed specifically for this research. The model takes user messages as input and classifies them into predefined intent categories. The library 'sklearn'⁹ was used to train a classification model on custom data. First, 'CountVectorizer' converted the text into a numerical representation. Then, 'MultinomialNB' was used as a classifier based upon the multinomial naive Bayes algorithm. Additionally, GridSearchCV was used for hyperparameter tuning to find the best set of hyperparameters for the classifier.

The data used for training was custom made. All necessary intents were identified, and example sentences for each intent were added. ChatGPT was used to expand the list of example sentences, by giving commands as 'Give me 15 sentences similar to 'neutral sentence example'.' or 'Give me 15 sentences expressing the intent 'bot intent'.'. Since ChatGPT is trained upon an enormous dataset based upon human data, this

⁸<https://pypi.org/project/NRCLex/>

⁹<https://scikit-learn.org/stable/>

could be used as inspiration to provide more data samples. As a result, each intent had 15-30 (depending on the complexity of the intent) sample sentences for training. This expansion process helped to create a broader and more diverse training dataset, which enhanced the accuracy of intent classification.

```

1 def train_model(data_file):
2     # Function to train an intents classifier model using the data in the
3     # JSON file
4     with open(data_file) as f:
5         data = json.load(f)
6
7     training_data = []
8     labels = []
9
10    for intent in data['intents']:
11        for phrase in intent['training_phrases']:
12            training_data.append(phrase)
13            labels.append(intent['intent_name'])
14
15    pipeline = Pipeline([
16        ('vect', CountVectorizer(ngram_range=(1,2), min_df=2)),
17        ('clf', MultinomialNB(alpha=0.1))
18    ])
19
20    model = pipeline.fit(training_data, labels)
21
22    joblib.dump(model, CLF_DIR)
23
24    return model

```

LISTING 6.2: Python code for intent classification function

Entity extraction

For entity extraction the library 'spacy'¹⁰ was used, which uses natural language processing techniques for identifying and extracting relevant entities from user messages. The code was based upon the tutorial from Medium¹¹. The model was trained on custom training data, which was initially created manually and later expanded using ChatGPT-generated samples. The Entity Extraction model enables the chatbot to identify and extract specific knowledge from user inputs,

6.2.6 User interface

For the user interface (see Figure 6.6), the template from Studygyaan¹² was used and adapted (e.g. colors and background). To this template the chatbot and user avatar were added, as well as a delay in response and 'typing-indicators' to make it seem as if the chatbot was typing a response. This was done after user tests showed that users disliked an instant response.

6.3 User Testing

To improve the interaction, add options for intents/responses, and check perceived empathy, two iterations of prototypes were tested with participants. Additionally,

¹⁰<https://spacy.io/>

¹¹<https://medium.com/mlearning-ai/named-entity-recognition-with-spacy-fd834ff84b86>

¹²https://studygyaan.com/python/create-web-based-chatbot-in-python-django-flask?utm_content=cmp=true

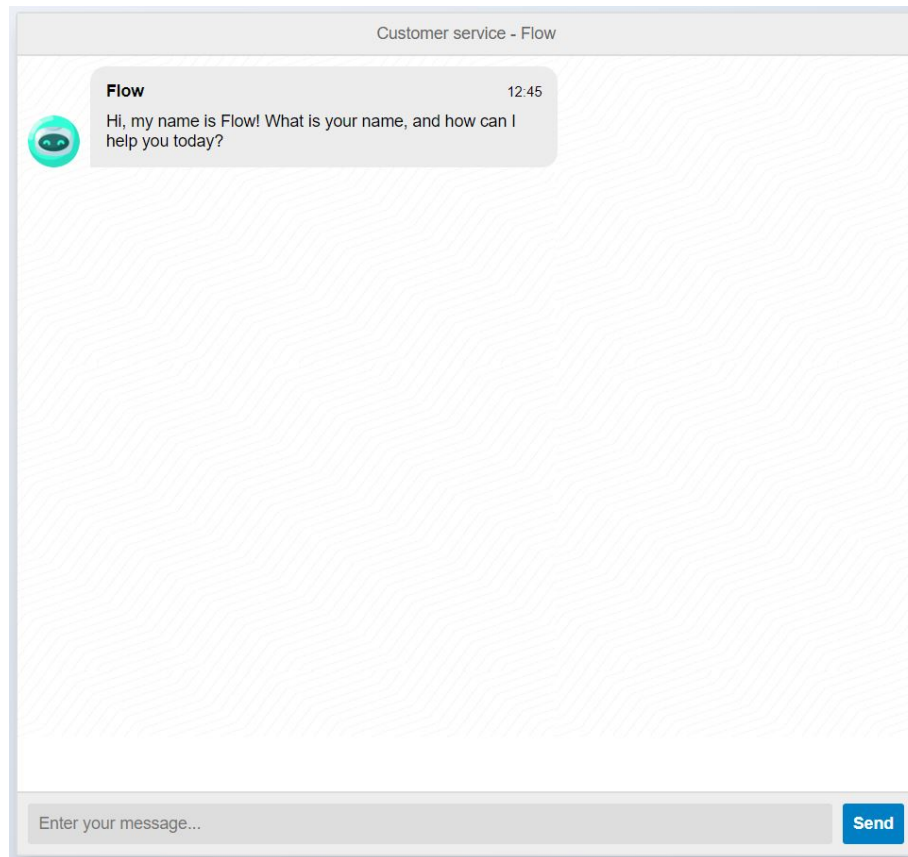


FIGURE 6.6: Interface Flow

these conversations were used for debugging and adding new phrases to the training data for the intent classification and emotion recognition models.

6.3.1 Prototype 1

The initial prototype was a basic chatbot with minimal user interface, using Google's DialogFlow¹³ for intent classification. During testing it became clear that the existing emotion recognizer was not functioning sufficiently and the unpredictability and variability of user behaviour was emphasized. Consequently, improvements were made to the emotion recognizer, and a broader range of user options were added to the chatbot. On a positive note, the participants noticed the expressed empathy of the chatbot.

Changes: improved emotion recognizer, additional user and response intents

6.3.2 Prototype 2

The second round of user testing was conducted with an improved chatbot that featured the final user interface and an enhanced emotion recognizer. The results revealed that using DialogFlow's intent classification was inadequate for the purpose of this study. Consequently, a custom intent classifier was implemented to allow more control in the classification process (e.g. identifying the top 3 possible intents). Additionally, a fallback system was developed incorporating the decision

¹³<https://cloud.google.com/dialogflow>

tree and fallback responses, which was based on commonly used intents and conversational routes taken by users. Another system was implemented to allow the experimenter to intervene in cases where the chatbot made incorrect intent classifications. This was done to account for the unpredictability of the participants and to address the increased self-reported levels of irritation by participants when they encountered intent classification errors during user tests, which could potentially influence the test results. As participants noted, "A full error is way more annoying than a chatbot saying 'I misunderstood'." Lastly, the response time was increased as participants explained that the fast responses made it feel like the chatbot gave them pre-programmed responses instead of listening to their complaint.

Changes: new intent classifier, fallback system, decision tree, experimenter intervention option, increased response time

6.4 Conclusion

This chapter discussed the design and implementation details of the empathic and non-empathic prototypes. The design phase focused on two main components: the design of an empathy framework and the design of a complaint handling framework. Different empathic strategies were identified based on Ma et al's. [79] three components for empathic behaviour: affective dialogue systems, personalized dialogue systems, and a knowledge base system. This resulted in the incorporation of the following strategies for the empathic prototype: empathic intent mapping, emotion validation, empathic questions, and personification using knowledge. The final empathy framework included a perceptual module for analysing input text, a behavior controller for generating appropriate responses based on user intent, emotion, and available knowledge, and an empathy mechanism for mapping user emotions to empathic intents. The complaint handling framework was based upon five phases: greeting, aggression reduction, conflict settlement, problem solution, and conclusion. In the empathic prototype, empathy was mainly emphasized in the aggression-reduction, conflict settlement, and problem-solution phase.

In the implementation phase, the frameworks are incorporated into an empathic and non-empathic prototype with similar functionality. Sentiment analysis is used to select the appropriate empathic response intent from a response dataset based upon the user's dominant emotion. The response dataset includes different phrasings for each empathic intent, such as questioning, acknowledging, wishing, neutral, sympathizing, and suggesting. Additionally, entity extraction enables the detection and use of knowledge in the conversation. The flow of the dialogue is represented in a decision tree, to make the prototypes more controllable and cohesive. To enhance the researcher's control over the prototypes, manual intervention options are included. This allowed the researcher to intervene in instances where the chatbot misclassified a user intent. Lastly, this chapter described some user tests and their resulting insights and adaptations to the prototypes.

Chapter 7

Results

This research focused on the effect of empathy in chatbots on customer service satisfaction. Specifically, whether empathy could increase the perceived interactional justice. Using a mixed-methods approach, both quantitative and qualitative data was collected. Quantitative measures included a questionnaire on self-reported levels of perceived empathy, usability and interactional justice, where qualitative measures focused on the user's overall impressions, emotional responses and thoughts.

To recap, the following research question was answered through this approach, including the consecutive formed hypothesis:

What is the effect of empathy in chatbots on perceived interactional justice during service recovery?

H1: A chatbot showing empathic behavior will result in significantly higher perceived levels of interactional justice compared to a non-empathic chatbot.

The following additional hypotheses were tested and will be discussed in this chapter:

H2: Participants will report more positively on their interaction with the empathic chatbot compared to the non-empathic chatbot.

H3: The empathic chatbot prototype will be perceived as significantly more empathic compared to the non-empathic chatbot prototype.

H4: There will be no significant difference in perceived usability between the empathic and non-empathic chatbot prototypes.

The study included a total of 25 participants, consisting of 15 female and 10 male participants. The age range of the participants varied from 19 to 57 years, with different occupations and levels of education. It was attempted to maintain a balanced distribution of male and female participants across the empathic and non-empathic chatbot conditions. Eventually, 13 participants interacted with the empathic chatbot and 12 participants interacted with the non-empathic chatbot.

This chapter starts by discussing the quantitative results. Then, the qualitative results will provide further insight into the interactions and user experiences. Lastly, both type of results will be compared to identify any similarities or differences. For clarity, the empathic chatbot will be referred to as condition E, and the non-empathic chatbot will be referred to as condition NonE.

7.1 Quantitative results

The quantitative results were interpreted from the scores from the questionnaires on the topics interactional justice, perceived empathy, and usability. During pre-processing all the scores were accumulated per topic and transformed to percentages. According to the outcomes of the Kolmogorov-Smirnov test and Levene's Test, normality and equal variances can be assumed for all three datasets. The results were compared by the use of an independent samples T-Test with a 95% confidence interval. In the next section the results of these measures will be discussed per topic. At the end some additional results will be discussed: amount of user turns, number of words used, and total interaction time.

7.1.1 Perceived interactional justice

No outliers were identified in the data using SPSS. Outliers were defined as data points that fell outside the range of 1.5 times the interquartile range (IQR) above the third quartile and below the first quartile. A visualisation of the results in a boxplot is visible in Figure 7.1. Upon visual inspection there appear large differences between the two conditions, showing a much higher level of perceived interactional justice for the empathic chatbot. Statistically, the participants in the E condition ($M = 73.8, SD = 14.2$) compared to the participants in the NonE condition ($M = 58.6, SD = 12.3$) demonstrated significantly higher perceived interactional justice scores, $t(23) = 2.9, p = .009$. Again, the effect size was large, with a Hedges' g of 1.1. These findings support H1.

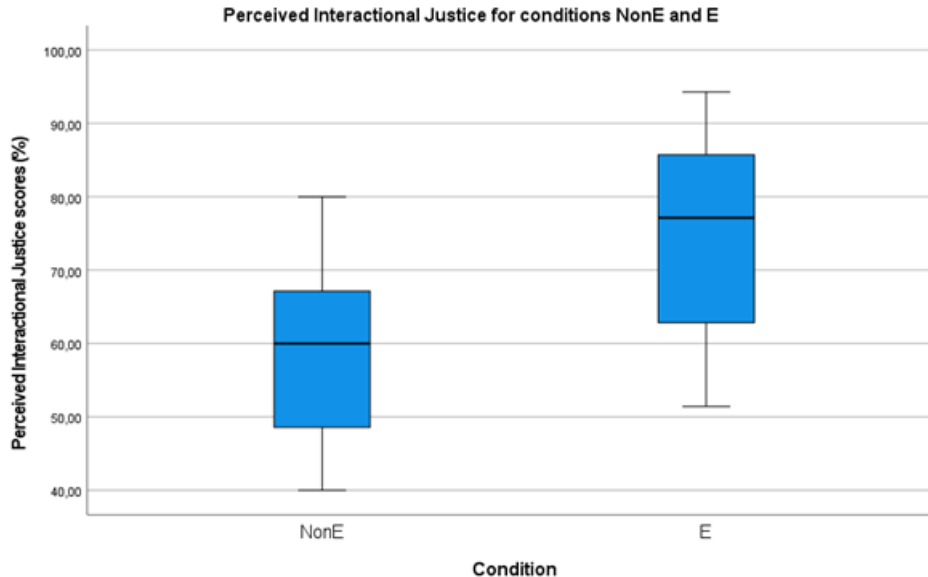


FIGURE 7.1: Perceived Interactional Justice per condition

7.1.2 Perceived empathy

Again, no outliers were identified in the data using SPSS. Figure 7.2 shows a boxplot of the results. Upon visual inspection there appear large differences between the two conditions, showing a higher level of perceived empathy for the empathic chatbot. Statistically, the participants in the E condition ($M = 44.3, SD = 11.3$) compared to

the participants in the NonE condition ($M = 32.5, SD = 10.4$) demonstrated significantly higher perceived empathy scores, $t(23) = 2.7, p = .013$. The effect size was measured using Hedges' g , as this measure is recommended when working with small sample sizes. Generally, a Hedges' g around 0.2 indicates a small effect, around 0.5 a medium effect, and around 0.8 or higher a large effect. For this measure the effect size was large, with a Hedges' g of 1.0. These findings support H3.

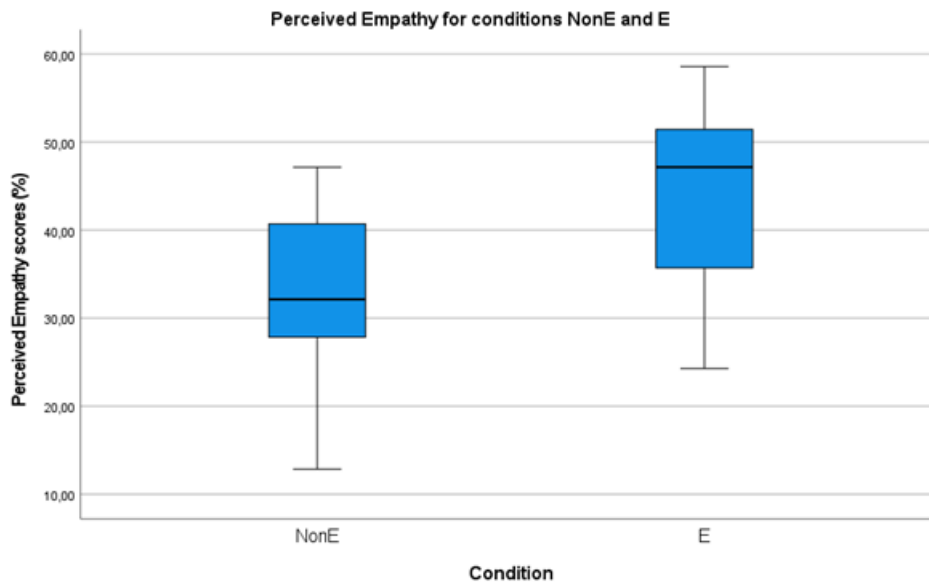


FIGURE 7.2: Perceived Empathy per condition

7.1.3 Perceived usability

One outlier was identified using SPSS, meaning that it fell outside the range of $1.5 \times \text{IQR}$ above the third quartile and below the first quartile. Due to the small sample size, it was decided to still include it in the analysis. As an additional check, the analysis was also carried out without the outlier: the conclusions are not different than when the outlier is included. A visualisation of the results in a boxplot is visible in Figure 7.3. Upon visual inspection there are no large differences between the two conditions, only a slightly higher score for the E condition. Statistically, the participants in the E condition ($M = 74.7, SD = 15.7$) compared to the participants in the NonE condition ($M = 71.9, SD = 10.0$) did not demonstrate significantly different usability scores, $t(23) = 0.54, p = .597$. These findings support H4.

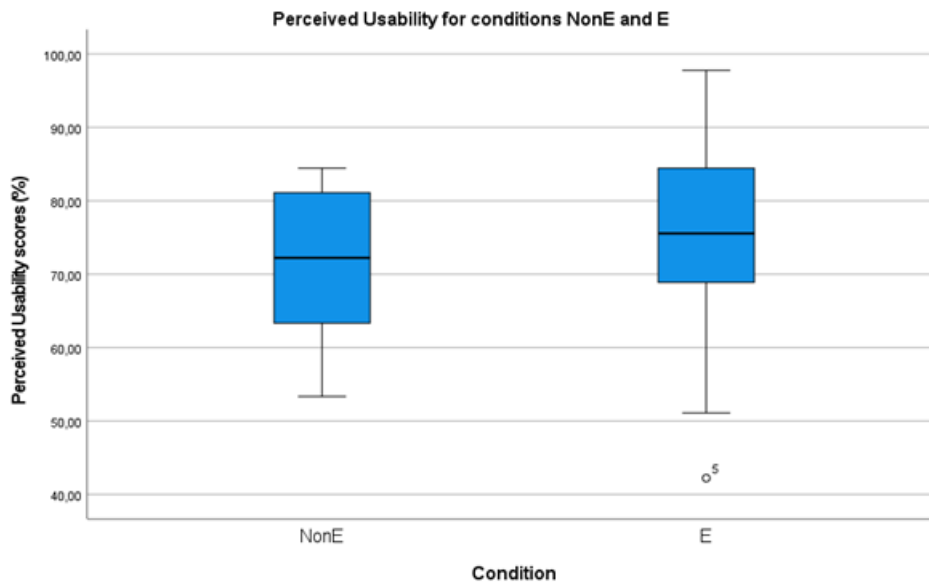


FIGURE 7.3: Perceived Usability per condition

7.1.4 Interaction dynamics

In order to gain further insights and increase the overall understanding of the interaction, three additional quantitative results were analysed. These results offer new perspectives on the overall dynamics of the interaction process, and could possibly show new research opportunities. The following results were analysed from the logs: amount of user turns, total amount of words used by user, and total interaction time. A Mann-Whitney U test was performed to evaluate whether the **amount of user turns** and **amount of words** differed between the E and NonE condition. Mann-Whitney tests were chosen due to the small sample size and violation of the normality test for these variables. An independent samples T-Test was used to compare the total **interaction time**, since the Kolmogorov-Smirnov and Levene's test showed that normality and equal variances could be assumed for this dataset.

Using SPSS some outliers in the amount of user turns and the amount of words were identified (indicated by the small circles and asterisks in the boxplots). Due to the small sample size they were not excluded from the analysis. Again, an additional check was performed: the analyses have been carried out without outliers and the conclusions are not different than when the outliers are included.

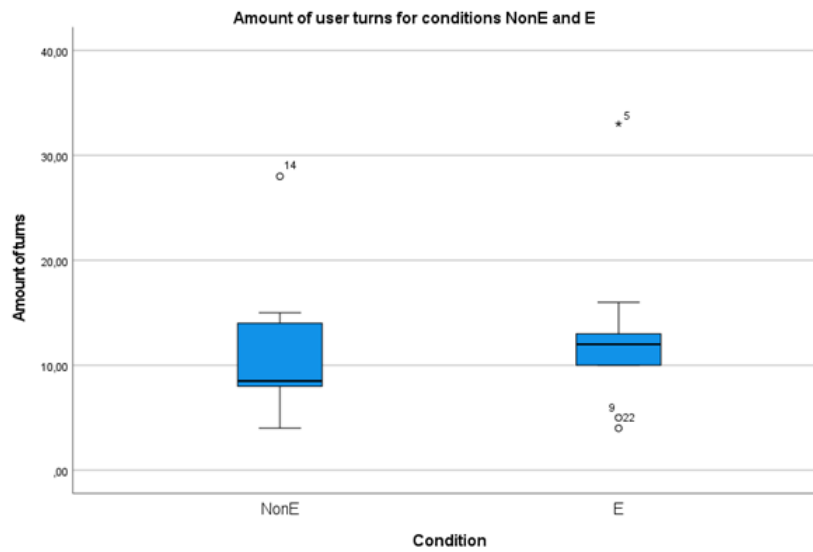


FIGURE 7.4: User turns per condition

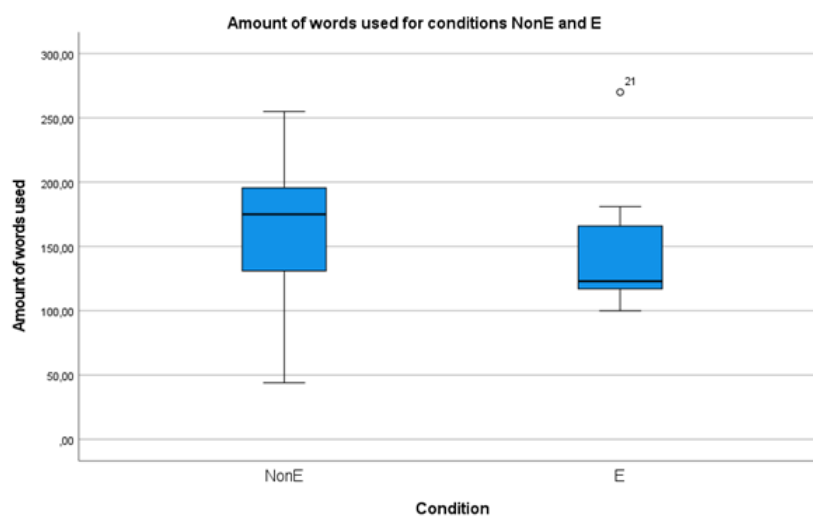


FIGURE 7.5: Words used by user per condition

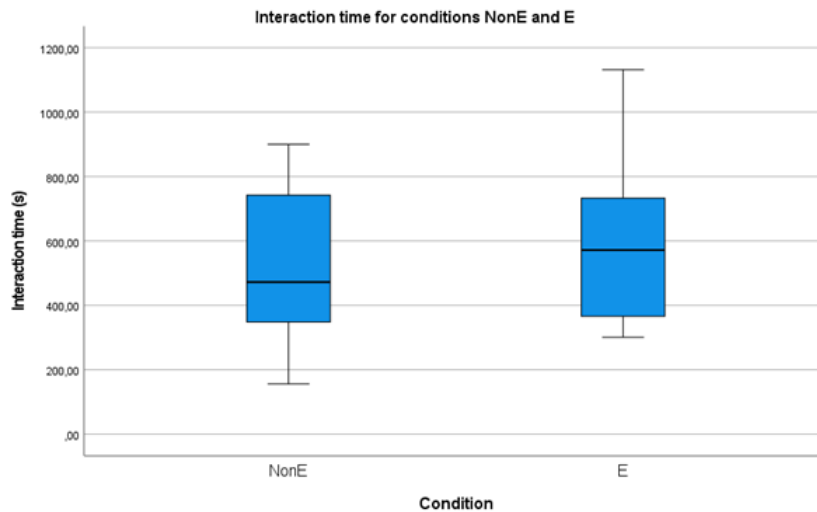


FIGURE 7.6: Interaction times per condition

Visually (see Figure 7.4) it appears as if the participants in the E condition ($M = 12.5, SD = 7.0$) took slightly more turns talking to the chatbot than participants in condition NonE ($M = 11.3, SD = 6.3$). Similarly, the interaction time (see Figure 7.6) appears slightly higher for the E condition ($M = 577.8, SD = 264.9$) than for the NonE condition ($M = 531.7, SD = 237.4$). Contrarily, the amount of words (see Figure 7.5) used is less for the E condition ($M = 145.4, SD = 45.4$) than for the NonE condition ($M = 162.2, SD = 64.3$). A statistic analysis of the results indicated that there was no significant difference between the **amount of user turns** ($U = 90.5, p = .503$), **amount of words** used by participants ($U = 56.0, p = .247$), and total **interaction time** ($t(23) = -0.46, p = .652$) of participants in condition E and condition NonE.

Concluding remarks

The analysis of the quantitative results showed that the scores of perceived interactional justice for participants interacting with the empathic chatbot are significantly higher than those interacting with the non-empathic chatbot. The same is true for the scores of perceived empathy. Both also showed large effect sizes according to Hedges' g . The perceived usability did not differ significantly between conditions, but visual inspection showed a small increase for participants in the E condition. Lastly, additional measures exploring interaction dynamics were not found statistically significant, but showed small observable differences between conditions.

7.2 Qualitative results

After analyzing the quantitative results, the qualitative insights provide a nuanced understanding of how users perceived and experienced interactional justice, empathy and usability in the context of chatbot interactions. During the interview following the interaction, participants were asked open-ended questions about the interaction to gain richer insights into their perception of the chatbot, ecological validity, and possibilities for future implementations. These questions focused on:

1. Perception of the interaction (*What are your initial thoughts after this interaction?*)
2. Realism of scenario and chatbot behaviour (*Do you think this interaction is realistic?*)
3. Comparison to human customer service (*Do you think the interaction would have been the same with a human customer service representative?*)

To analyse the results, common response themes were identified from the responses. Themes which were mentioned by more than one participant were included, given the small sample size. Then, the amount of times each theme was mentioned was recorded. The presence of overlapping responses among participants is interesting because all the answers were given without any specific prompting. The fact that multiple participants independently provided similar comments or observations indicates the importance and relevance of those particular topics or themes to the users. In the subsequent sections the results of each question will be discussed in more detail, identifying common themes and elaborating on interesting comments from participants.

7.2.1 Perception of interaction

For clarity, the positive and negative responses are discussed separately. The positive remarks can be found in Table 7.1. Interestingly, no large differences can be found in the overall perception of the empathic and non-empathic chatbot (friendliness, appropriate answers, adequate interaction). However, only one participant in condition NonE mentioned that they felt helped, compared to 4 in condition E. After participants were debriefed and told about the different conditions, 4 participants in NonE indicated that they would have preferred to talk to an empathic agent. Obviously, no participants from condition E made this comment, but they did not express any contrary preference for a non-empathic agent, and did positively comment on how they felt like the chatbot was 'compassionate and understanding', with one participant even commenting 'This was the sweetest chatbot I have ever talked to.'. Another interesting finding was that some participants (in either condition) immediately guessed in what condition they were, and each participant guessed correctly.

Response	Empathic (n=13)	Non-empathic (n=12)
Friendly chatbot	7	6
Good, elaborate answers	4	4
Smooth, clear interaction	4	4
Felt understood	4	1
Felt human-like	3	0
Nice to whine	2	0
Felt I was helped	2	0
Prefer empathic behaviour*	N/A	5

TABLE 7.1: Responses on chatbot perception (positive)

*Only relevant for NonE condition

In addition, a notable observation was that one participant in the E condition told the researcher 'At some point I even started feeling bad for the chatbot', after they acted (in their words) rudely to get a compensation. The opposite was reported by 2 other participants in the same condition, who enjoyed the possibility 'to whine' to a chatbot without hurting a human worker's feelings.

The most frequent remarks on negative perceptions are shown in Table 7.2. Most participants unpromptedly commented that the scenario was frustrating, a little over half of the participants in condition E, and 75% of participants in condition NonE. Most participants in the NonE condition appeared properly frustrated at the end, some of them audibly swearing or commenting 'This chatbot is stupid', and 'It's like talking to a wall'. Whereas some participants in condition E were also frustrated, the levels and frequency appeared to be much higher in condition NonE.

Another critique was that the chatbot was not able to be in-depth enough about certain topics and requests, in both conditions. When this was mentioned, most participants expressed their desire for more follow-up questions and in-depth conversation. Often, participants would make the comparison with a human agent, which could give more in-depth responses. Out of 25 participants, 8 indicated that they would rather talk to a human, and the majority of them came from the NonE condition. Another interesting finding is the remark of two participants in the E condition, who mentioned that they disliked the empathy and rather would had just have compensation. Lastly, almost half the participants in the NonE condition mentioned that they felt like the chatbot could not help them, whereas none of the participants in E mentioned this.

Response	Empathic (n=13)	Non-empathic (n=12)
Frustrating scenario	6	9
Lack of in-depth conversation	5	4
Rather talk to a human	3	5
Would have stopped interacting sooner (if not experiment)	2	4
Misunderstandings are frustrating	3	1
Repetition is frustrating	1	2
Don't want sympathy, I want compensation	2	0
Usually never talk to chatbots	1	2
Felt like the chatbot would never help me	0	5

TABLE 7.2: Responses on chatbot perception (negative)

Overall, three trends can be found in the answers to this question. First, in terms of functionality, both chatbots received similar responses. Second, positive comments were most frequently made by participants in condition E, whereas negative comments were most frequently made by participants in condition NonE, which supports H2. Lastly, participants seemed to become much more frustrated in the NonE condition than when they interacted with the empathic chatbot (E condition).

7.2.2 Realism of scenario and chatbot behaviour

Responses on this question are split into two categories; realism of the scenario and realism of the chatbot behaviour. The scenario was deemed realistic by all participants except 2 from the NonE condition, who added that realistically they would have gotten a refund. Contrarily, 6 and 8 participants (E and NonE respectively) commented that the lack of compensation was realistic. The interviews also revealed that participants had two different interpretations of 'realism' for chatbot behaviour: 1. behaviour like current chatbots, 2. behaviour like humans.

The comments on chatbot behaviour are shown in Table 7.3. Again, most participants deemed the chatbot's responses and behaviour realistic, which was meant both positively and negatively. To illustrate, some participants compared the chatbot to existing company chatbots to show that they were impressed with its realistic behaviour, whereas other participants compared the chatbot to Bol.com's Billie chatbot¹ or PostNL's Daan chatbot² while expressing their discontent with those: 'Similar to the Billie bot, this one was also useless'. In most cases, participants who referred to the chatbot as bot-like emphasized that this was meant as critique.

The level of advancement of the chatbot also stirred different reactions. Some participants said that the chatbot was not realistic as it was way more advanced than current chatbots, while others explained that this aspect made it more realistic for them since it was more human-like. Another finding was that participants in the E condition particularly commented on the level of empathy during this question, for example 'I had no idea chatbots were already capable of being so empathic' or 'I was surprised that the chatbot cared about my feelings'. Whereas participants in

¹www.bol.com

²www.postnl.nl

the NonE commented on this in terms of understanding and appropriateness of the answers.

Response	Empathic (n=13)	Non-empathic (n=12)
Realistic responses	8	5
Acts like a bot	1	5
Similar to company bots	1	3
More advanced than company bots	4	3

TABLE 7.3: Responses on realism chatbot behaviour

7.2.3 Comparison to human customer service

Both the participants' comments regarding their own behavior and their comparisons of the chatbot's behavior to that of a human agent will be addressed next. First, the user's comments comparing the interaction to interacting with a human agent will be discussed. The results are shown in Table 7.4.

Response	Empathic (n=13)	Non-empathic (n=12)
Better understanding of situation	3	4
More in-depth answers	1	5
No repetition	2	2
Human has (real) feelings	1	2
Human can be grumpy	1	1
Less smooth and straightforward interaction	2	0

TABLE 7.4: Responses on human agent behaviour compared to chatbot

A big difference between the conditions is that participants in the NonE condition more often highlighted the chatbot's limited ability to provide in-depth answers compared to a human agent, while this was only noted by 1 participant in the E condition. Additionally, 2 participants in the E condition praised the chatbot for the fast and straightforward interaction, which they believed to be less likely when talking to a human. However, no participant in NonE mentioned this. An interesting comment made by participants in both conditions was that human agents can be grumpy and rude, whereas they thought the chatbot was always friendly: 'the interaction would be the same if the human was an empathic person, but sometimes they are grumpy and rude'.

An overview of the most frequent responses regarding the user's behavior in comparison to interacting with a human agent is shown in Table 7.5. Participants in the E condition mostly answered that they would have acted the same to a human agent, whereas participants in the NonE condition emphasized the differences more frequently. The results provide valuable insights into the behaviour of humans, with people stating that they were ruder to the chatbot since 'I don't have to care about their feelings (i.e. they have none)', and participants explaining that they used shorter and direct sentences to help the chatbot's understanding. Notably, most of the participants in the NonE condition mentioned that they would be more emotional to a human agent: 'in this case I turned off my own emotions because he doesn't have any either', whereas only a few mentioned this in the E condition.

Response	Empathic (n=13)	Non-empathic (n=12)
Yes, similar	9	2
More direct to bot	2	3
Am ruder to bot	2	1
Would persist longer to human	1	3
Would be more emotional to a human	3	9

TABLE 7.5: Responses on user behaviour compared to interaction with human agent

Concluding remarks

The interview questions provided rich insights into the interaction as perceived by the participants. Additionally, they showed that the scenario and responses were realistic, and that the participants had different opinions on the realism of the behaviour of the chatbot compared both other chatbots and humans. In general, participants interacting with the empathic bot were more positive in their comments, whereas participants interacting with the non-empathic chatbot were perceived as more frustrated and commented on flaws more frequently. This was true for both the first question about the interaction and the comparison to a human agent. Additionally, these answers showed that the chatbot's functionality was deemed adequate by both conditions, and highlighted the variability between participants.

7.3 Relations qualitative and quantitative results

To gain a comprehensive understanding of the effect on perceived interactional justice and user perceptions of empathic chatbots, a comparison between the quantitative and qualitative results will be briefly discussed.

Overall, both results showed similar trends. The quantitative results showed that the E condition scored higher in terms of empathy, which was reflected in the comments of participants on how they found the chatbot 'caring' and 'human-like'. Also, the frustration of the participants during and after the conversation was perceived as higher for those in the NonE condition than in the E condition, reflecting the quantitative results for perceived interactional justice. Additional comments also showed that user's often preferred empathic behaviour during the interaction. The quantitative results did not show significant difference for usability, which is also indicated by user among all conditions deeming the chatbot as 'friendly' and commenting on the appropriateness of the answers.

The only notable difference between the quantitative and qualitative results is in terms of the additional quantitative data. Whereas the amount of words used by participants did not differ statistically between conditions, the NonE condition showed a slightly higher count than the E condition. However, this is not reflected in the qualitative analysis, as a greater number of participants expressed being direct and using fewer words when interacting with the chatbot in the NonE condition compared to the E condition. This unexpected finding will be further explored in chapter 8.

7.4 Conclusion

This chapter discussed the quantitative and qualitative results of the experiments. The quantitative analysis focused on perceived interactional justice, perceived empathy, and perceived usability. All data was pre-processed and transformed into percentages, and normality and equal variances were assumed based on statistical tests. To compare the results between the empathic and non-empathic chatbot conditions, an independent samples T-Test was conducted. The quantitative results supported hypotheses 1, 3, and 4. The support for H3 and H4 shows that the prototypes were successfully manipulated in terms of empathy while keeping similar functionality. As expected, participants in the empathic chatbot condition reported significantly higher levels of interactional justice compared to the non-empathic chatbot condition, with a large effect size.

Additionally, the study examined three quantitative measures to gain more insight into the interaction dynamics. This included the amount of user turns, the number of words used by the user, and the total interaction time. These measures did not show statistically significant differences between the empathic and non-empathic chatbot conditions. However, visual inspection of the data indicated small observable differences between conditions for these measures.

The qualitative analysis focused on open-ended questions related to the participants' initial impression of the chatbot, the realism of the scenario and chatbot behavior, and a comparison to human customer service. The findings of this analysis supported hypothesis 2: participants interacting with the empathic chatbot generally had a more positive perception of the chatbot, while participants engaging with the non-empathic chatbot more often expressed frustration and dissatisfaction. Interestingly, participants interacting with the empathic chatbot would more frequently comment on the helpfulness of the chatbot than those in the non-empathic condition. In terms of the realism of the scenario and chatbot behavior, most participants found the scenario realistic. When comparing the interaction to human customer service, participants engaging with the empathic agent mentioned that the chatbot appeared to have similar understanding of the situation and a smoother and more straightforward interaction compared to a human agent. On the other hand, participants interacting with the non-empathic chatbot more frequently highlighted the limitations of the chatbot, such as its inability to provide in-depth answers.

Lastly, despite some small differences in terms of *use of words*, the quantitative and qualitative results of the study on perceived interactional justice and user perceptions of empathic chatbots showed consistent findings.

Chapter 8

Discussion

The outcomes of this research provide insight into the effect of empathic behaviour in chatbots on customers' perceived interactional justice during complaint handling. The findings revealed several relationships in the data. This chapter will discuss the analysis and implications of the results. As well as a discussion on the limitations of the study. Lastly, the recommendations for future work will be discussed.

8.1 Analysis of results

The analysis of the perceived empathy results indicated a significant difference between the two conditions. As expected, participants interacting with the empathic chatbot reported higher perceived empathy scores than those interacting with the non-empathic chatbot. The results also indicate that the confounding variable 'usability' was successfully controlled: there were no significant differences between conditions despite the presence or absence of empathy in the prototypes. The results from the interviews also indicated that the functionality of the chatbot was adequate for the interaction.

8.1.1 Effect empathy on user experience

The results indicated that incorporating empathic behaviour in a customer service chatbot can positively affect the user's feeling of interactional justice, supporting Hypothesis 1. This aligns with the literature on the effect of (human) empathy on perceived interactional justice [17, 18, 43, 61, 63]. Given that interactional justice influences overall justice, which in turn leads to increases user satisfaction and loyalty [9, 108], the use of empathic chatbots can have substantial benefits for companies. The positive effect of empathy during the interaction was further supported by the interviews. Those interacting with the empathic chatbot generally made more positive comments, emphasizing the friendliness, human-likeness, and expressed feeling understood and appreciative of the chatbot's sympathy. On the other hand, participants in the non-empathic condition more often expressed negative comments and indicated a preference for an empathic chatbot after being debriefed about the conditions. The indication that the empathic chatbot resulted in a more positive interaction supports Hypothesis 2.

The greatest difference between conditions was in regards to user comments on their perception of the likeliness of the chatbot helping them. Almost half the participants in the non-empathic condition expressed that they thought the chatbot would never help them, while none of the participants in the empathic condition expressed this. This phenomenon can be explained by research from Bove [65], who found that

'motivating helping behaviour' is a benefit of empathy. Whereas this research was focused on human behaviour, it is possible that the participants unconsciously held the belief that the (empathic) chatbot would experience this motivation as well due to its empathic clues. This idea is supported by Simon [43] who found that customers may perceive empathic behaviour in a (human) employee as an act of benevolence, indicating a willingness to help them. Furthermore, both groups mentioned the lack of in-depth conversation as a limitation. This was also found by Tsai et al. [19], whose study including chatbots showed that customers preferred the human ability to get into more details. It is interesting to further explore this comment, to see whether this is actually affected by the interaction, or more dependent on human's current perceptions of chatbots. Overall, the positive effect of empathy on the customer experience was supported by both quantitative (interactional justice) and qualitative (user perception) results.

8.1.2 Effect empathy on user behaviour

The qualitative results of this study also revealed unexpected findings related to customer behaviour to chatbots compared to humans. A significant number of participants in the empathic condition explained that they behaved similarly to how they would with a human agent, which suggests that the presence of empathy in chatbot interactions can elicit more natural and genuine responses from users. This may indicate that the presence of empathy in chatbot interactions might contribute to participants perceiving the chatbot as more human-like which could in turn have influenced their behavior to be similar with how they would typically interact with a human. The idea that empathy increases the feeling of human-likeness in chatbots aligns with the perspective expressed by Bošnjaković and Radionov [59], who emphasize that empathy is a vital human ability. This could mean that empathy and the perception of human-likeness influence each other: the presence of empathy increases the feeling of human-likeness, and human-likeness may increase the feeling of empathy. However, further research is necessary to explore this hypothesis.

Furthermore, the participants in the non-empathic condition more often commented that they avoided emotional language as the chatbot expressed none either, whereas participants in the empathic condition mentioned being more rude (including angry language) to the chatbot. So while participants in both conditions indicated a change in behavior to the chatbot compared to humans, the group interacting with the empathic chatbot mentioned an increase in their expression of emotion and self-reported rudeness, whereas the group interacting with the non-empathic chatbot expressed a decrease in emotional expressions. These trends could suggest that the empathic clues from the chatbot resulted in an increase in the emotion expressed by the participants. This could potentially account for the difference in frustration during and after the interaction as well. While no conclusive answer can be drawn from these results, it is possible that the empathic behaviour of the chatbot encouraged emotional expressions by participants and then regulated these emotions. This could be possible since emotion validation techniques were incorporated in the chatbot, and Lambie et al. [72] already showed that this positively influences emotion regulation in humans. If that were the case, the reason that participants in the non-empathic condition were significantly more frustrated than those in the empathic condition could be attributed to the lack of emotion regulation in their condition. However, the increased frustration could also be attributed to the fact that participants in the non-empathic condition more frequently felt like the chatbot could

never help them, or due to their overall more negative perception of the chatbot.

A few participants from both conditions expressed that they enjoyed talking to a chatbot since they could freely whine and be more rude without having to care about a human employee's feelings. Whereas this contradicts earlier statements of (a larger number of) participants explaining that they behaved similar as they would to a human agent, it should be noted that there is no overlap between participants who made either comment, indicating a personal preference. This corresponds to studies which found that humans feel less judged when talking to a chatbot [19, 33]. Also, the notion of people being more rude towards bots is not new, as Bartneck and Keijsers' [109] research into robot bullying has shown. However, whether this advantage over talking to a human employee is ethically desirable is another point of discussion.

8.1.3 Effect empathy on user engagement

An additional consideration is that from the additional results on interaction dynamics it did not appear as if empathy had a significant effect on user engagement but some trends could be observed. Both in terms of turns and interaction length, the participants in the empathic condition had a larger score, which corresponds to the interviews where participants in the non-empathic condition mentioned they usually would have continued arguing with a human, but gave up relatively quickly with the chatbot. However, participants in the non-empathic condition actually used more words than participants in the empathic condition, while the opposite was expected based upon their comments of being more direct in the interview. There are a few possible explanations for this. First, the small sample size makes that the results are very sensitive to inter-person variability, meaning that the participants in the non-empathic condition could just be more talkative (e.g. more words per turn) than the participants in the empathic condition. Second, participants in the non-empathic condition might have felt the need to provide more detailed information or explanations to try to compensate for the lack of empathic cues or even elicit empathy from the chatbot. This can be characterized as over-accommodation which relates to the theory of communication accommodation [110], a detailed discussion of this theory is outside the scope of this research, but it may be interesting to investigate for future research. In conclusion, while the results on user engagement were not statistically significant, they do show interesting trends which can be further explored.

8.1.4 Implications of results

The results have a variety of implications for current implementations of chatbots in customer service. The significant positive effect of empathy on interactional justice and overall experience shows that this is a promising method to increase the customer's feeling of justice during a service recovery. As discussed before, this has great implications for overall user satisfaction and customer loyalty [9, 108]. By integrating empathic models into their existing chatbot systems, companies can improve the feeling of justice during service recovery, leading to increased customer satisfaction levels and loyalty.

The use of an empathic chatbot also has benefits for the users. Participants who interacted with the empathic agent expressed a greater sense of being understood and helped during the service interaction. The use of empathy increases the feeling of human-likeness and natural conversation flow, which enhances the overall service

experience for users. Additionally, the incorporation of emotion regulation techniques by an empathic chatbot can positively impact users' emotional well-being by providing a sense of emotional relief and support. This emotional support may lead to a more positive emotional state for the user, possibly enhancing their emotional well-being and satisfaction with the chatbot interaction.

When implementing empathic chatbots in customer service it is crucial to manage expectations and increase customer's acceptance of chatbots. During this study participants in the non-empathic condition mentioned that they did not expect the chatbot to be able to help them, since a chatbot does not have the authority to give compensation. Since their task was to get compensation, this knowledge may negatively influence their willingness to contact a chatbot outside of the experimental setting. Additionally, some participants commented negatively on existing chatbots, deeming them useless. For similar reasons, participants noted that they would usually rather talk to a human. In order to make full use of the advantages of chatbots in customer service, it is crucial to take this expectation and lack of acceptance into account. Since most companies have a policy on when to provide small refunds/-coupons, it could be a possibility to add this functionality to the chatbot. This way, the chatbot would be able to better assist the customers, resulting in (1) more satisfied customers, (2) less workforce bothered with simple requests, and (3) increased likelihood of customer's contacting a chatbot.

It should be noted that the use of empathy into chatbot interactions can contribute to building trust between users and technology. Research has shown that empathic agents have a positive effect on user's perception of trust [111, 112]. Given the importance of trust in human-robot interaction [113] the increased trust could in turn lead to a stronger relationship between users and technology [111], and consequently customer loyalty [114]. On the other hand, the presence of empathy may lead to overestimation of the chatbots abilities, resulting in misuse [113]. It is possible that users then rely solely on chatbots for problem resolution which may result in unsatisfactory solutions, leading to user frustration and a negative service experience. Therefore, it is crucial that the user's trust is well calibrated to the chatbot's abilities. Lastly, while chatbots can simulate empathy, they lack genuine emotional understanding, which may result in users feeling misunderstood or emotionally disconnected. Future researchers should keep this in mind when designing empathic agents for customer service.

8.2 Limitations

Whereas the study has successfully shown a positive relation between empathy in chatbots and perceived interactional justice, there are some limitations to the study. One limitation of the study is with regards to the research setup. Due to the online nature of the experiment, the participants knew that their answers were observed. This may have influenced them to portray more socially accepted behaviour, due to the social desirability bias [115]. However, as an initial study into this specific topic, the constant observation helped to collect more data about the user's behaviour (e.g. swearing, sighing, looking angry) and the online nature helped reaching more participants. However, when repeating this study on a larger scale and with more resources, it is suggested to do the experiment offline, so the user's answers appear hidden from the experimenter. For an even larger scale research, it would be useful

to conduct a fully automated experiment, to also avoid effects of social desirability bias.

Another limitation was the functionality of the chatbot. Due to time constraints in combination with the complexity of natural language understanding, the chatbot was prone to making intent classification mistakes. However, by adding the opportunity for the experimenter to interfere, this limitation was largely accounted for. Unfortunately as a result, the current chatbot prototype is not scalable and needs further improvement to run in a fully automated experiment. Furthermore, the chatbot had a limited knowledge space, and was thus unable to answer every request. As expected due to inter-person variability, some users asked more information outside of the chatbot's knowledge than others, which might have influenced their perception of the chatbot. The measure of usability however showed that the differences in interaction did not influence the perceived usability significantly. Additionally, it is possible that some participants received more 'empathic' answers than others due to the length of the conversation or type of requests they made to the chatbot. For example, asking for a manager would result in a neutral response, whereas asking for a refund would yield an empathic response. This may have influenced their perception of empathy in the chatbot.

As explained before, the responses by the chatbot were pre-designed, this means that the chatbot is currently not a scalable prototype. Also, whereas (all variations of) the responses were carefully based upon existing research and example phrases, no additional measure of the validity of the rephrasing of the responses according to the empathic intent mapping was done. An additional validation check may be useful to make this part of the empathic model more robust. Furthermore, the small sample size made the results sensitive to inter-personal differences. Larger scale studies should be conducted to further prove the findings from this study.

Lastly, during the course of this research the OpenAI's ChatGPT¹ was released, resulting in a surge of interest in chatbots and their possible applications. This also meant that more people suddenly had experience in using a completely functional chatbot. Whereas this may have influenced some user's perceptions, most of the recruited participants were not acquainted with ChatGPT.

8.3 Future Work

A variety of possibilities for future work have already been discussed in the previous sections. This section will briefly go over opportunities for future research which have not yet been discussed in detail. First of all, the effect of chatbot empathy on user engagement could be further explored. Whereas this research did not yield significant results, small trends were observable which could be further explored. This could provide more insight in whether participants would really interact longer with an empathic chatbot, or whether this was just coincidence. To do so, it is crucial to repeat the study on a larger scale or to use within-subject design, to account for inter-personal differences. Also, more research is necessary to determine the connection between emotion regulation and chatbot empathy. While some hypotheses were made based on the results of this research, there is no evidence of a direct relationship. If such a connection is proven, it could create more opportunities for chatbots in customer service or emotional support roles.

¹<https://openai.com/>

Another interesting area for future research is into the effects and empathy mechanisms of ChatGPT. This research was based upon a custom chatbot, with limited functionality but with a heavy focus on empathy mechanisms. On the other hand, the mechanisms of ChatGPT are unknown, but it comes with extensive functionality. A similar research could be conducted using ChatGPT, by first evaluating its possibilities for empathic behaviour, and then assessing the influence of this behaviour in combination with its broad functionality on interactional justice and customer satisfaction. Small tests have already been done by the experimenter which indicate that ChatGPT is able to follow the complaint handling framework and respond in a similar empathic manner as the prototype used for this study. Additionally, it is interesting to investigate what empathy mechanisms are used by this chatbot during complaint handling. Given the fact that it is based upon a "mixture of licensed data, data created by human trainers, and publicly available data" (source: ChatGPT²), the empathic mechanisms used in ChatGPT could represent human empathic behaviour and also a way to incorporate this in a chatbot.

²<https://openai.com/blog/chatgpt>

Chapter 9

Conclusion

During this research the effect of the presence of empathy in a chatbot on customers' perceived interactional justice during service recovery was investigated. The research question was:

What is the effect of empathy in chatbots on perceived interactional justice during service recovery?

A complaint handling chatbot was designed and implemented to find an answer to the research question. A between-participant study using a mixed-methods approach was conducted with 25 participants, with 12 participants interacting with a non-empathic chatbot and 13 with the empathic version. The experimental results suggest that interacting with an empathic chatbot resulted in significantly higher scores of interactional justice compared to a non-empathic chatbot. Furthermore, the overall perception of the chatbot was more positive when participants interacted with the empathic chatbot. Additional insights from the interviews show that participants interacting with the empathic chatbot more often indicated that they felt understood and that chatbot wanted to help them than participants interacting with the non-empathic chatbot. Also, more than half of the participants interacting with the non-empathic chatbot indicated that they would use more emotional language with a human agent, assuming the non-empathic chatbot wouldn't respond to it. This was not the case for participants in the empathic condition, implying that the empathic cues of the chatbot triggered more emotional responses.

This research has shown that empathic behaviour has a positive effect on the perceived interactional justice and perception of the chatbot during complaint handling scenarios. Whereas various studies have investigated the effect of empathy in service recovery and the effect of empathy in chatbots, only limited research has combined the two topics, making this research one of the first studies to show that empathy in chatbots has a positive effect on customers during complaint handling. Whereas further research should be conducted due to limitations of this study, the implications of this research are already significant; they highlight the need for companies to place a more emphasis on incorporating empathy into their customer service chatbots. By adding empathic strategies such as empathic response mapping, emotion validation, personalizing and empathic questions, companies could greatly enhance two key components: customer satisfaction and in turn customer loyalty. Another contribution is that the proposed framework based upon extensive literature research can in itself serve as a basis for future research into empathy and complaint handling for chatbots. Interesting paths for future research include investigating the relation between empathy and user engagement and whether there indeed exists a direct relation between chatbot empathy and emotion regulation.

References

- [1] A. Naikoo, S. Thakur, D. Guroo, and A. Altaf, "Development of society under the modern technology - a review," *scholedge international journal of business policy and governance*, vol. 5, pp. 1–8, Apr. 2018. DOI: 10 . 19085 / journal . sijbpg050101.
- [2] S. Campbell and N. Kwak, "Mobile communication and civil society: Linking patterns and places of use to engagement with others in public," *Human Communication Research*, vol. 37, pp. 207–222, Mar. 2011. DOI: 10 . 1111 / j . 1468 - 2958 . 2010 . 01399 . x.
- [3] *Chatbot english definition and meaning*. [Online]. Available: <https://www.lexico.com/en/definition/chatbot>.
- [4] P. Costa, "Conversing with personal digital assistants: On gender and artificial intelligence," *Journal of Science and Technology of the Arts*, vol. 10, p. 2, Dec. 2018. DOI: 10.7559/citarj.v10i3.563.
- [5] E. Adamopoulou and L. Moussiades, "An overview of chatbot technology," in *IFIP International Conference on Artificial Intelligence Applications and Innovations*, May 2020, pp. 373–383, ISBN: 978-3-030-49185-7. DOI: 10 . 1007 / 978 - 3 - 030 - 49186 - 4 _ 31.
- [6] P. Brandtzaeg and A. Følstad, "Why people use chatbots," in *The 4th International Conference on Internet Science*, Nov. 2017, ISBN: 978-3-319-70283-4. DOI: 10 . 1007 / 978 - 3 - 319 - 70284 - 1 _ 30.
- [7] A. Augello, M. Gentile, L. Weideveld, and F. Dignum, "A model of a social chatbot," in *Smart Innovation, Systems and Technologies*, vol. 55, Jun. 2016, pp. 637–647, ISBN: 978-3-319-39344-5. DOI: 10 . 1007 / 978 - 3 - 319 - 39345 - 2 _ 57.
- [8] A. P. Chaves and M. A. Gerosa, "How should my chatbot interact? a survey on social characteristics in human–chatbot interaction design," *International Journal of Human–Computer Interaction*, vol. 37, no. 8, pp. 729–758, 2021. DOI: 10 . 1080 / 10447318 . 2020 . 1841438. eprint: <https://doi.org/10.1080/10447318.2020.1841438>. [Online]. Available: <https://doi.org/10.1080/10447318.2020.1841438>.
- [9] L. Jenneboer, C. Herrando, and E. Constantinides, "The impact of chatbots on customer loyalty: A systematic literature review," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 17, pp. 212–229, Jan. 2022. DOI: 10 . 3390 / jtaer17010011.
- [10] D. Bendall-Lyon and T. Powers, "The role of complaint management in the service recovery process," *The Joint Commission journal on quality improvement*, vol. 27, pp. 278–86, Jun. 2001. DOI: 10 . 1016 / S1070 - 3241 (01) 27024 - 2.
- [11] V. Singh, A. Jain, and S. Choraria, "Exploring the role of complaint handling among complaining consumers," *Vision: The Journal of Business Perspective*, vol. 20, pp. 331–344, Dec. 2016. DOI: 10 . 1177 / 0972262916668739.
- [12] A. Trappey, C. Lee, W.-P. Chen, and C. Trappey, "A framework of customer complaint handling system," in *7th International Conference on Service Systems*

- and Service Management*, Jul. 2010, pp. 1–6. DOI: 10.1109/ICSSSM.2010.5530119.
- [13] B. L. Chua, M. Othman, H. C. Boo, M. S. Abkarim, and S. Ramachandran, "Servicescape failure and recovery strategy in the food service industry: The effect on customer repatronization," *Journal of Quality Assurance in Hospitality & Tourism*, vol. 11, no. 3, pp. 179–198, 2010. DOI: 10.1080/1528008X.2010.483419. eprint: <https://doi.org/10.1080/1528008X.2010.483419>. [Online]. Available: <https://doi.org/10.1080/1528008X.2010.483419>.
- [14] A. Smith and R. Bolton, "The effect of customers' emotional responses to service failures on their recovery effort evaluations and satisfaction judgments," *Journal of The Academy of Marketing Science*, vol. 30, pp. 5–23, Dec. 2002. DOI: 10.1177/03079450094298.
- [15] C. Boshoff, "An experimental study of service recovery options," *International Journal of Service Industry Management*, vol. 8, no. 2, pp. 110–130, 1997, ISSN: 0956-4233. DOI: <https://doi-org.ezproxy2.utwente.nl/10.1108/09564239710166245>.
- [16] S. Tax and S. Brown, *Recovering and learning from service failure*, 1998. [Online]. Available: <https://sloanreview.mit.edu/article/recovering-and-learning-from-service-failure/>.
- [17] R. Van Herck, S. Decock, B. Clerck, and L. Hudders, "The impact of employee empathy on brand trust in organizational complaint response emails: A closer look at linguistic realization," *International Journal of Business Communication*, Jul. 2021. DOI: 10.1177/23294884211032316.
- [18] C. Miao, M. Barone, S. Qian, and R. Humphrey, "Emotional intelligence and service quality: A meta-analysis with initial evidence on cross-cultural factors and future research directions," *Marketing Letters*, vol. 30, Dec. 2019. DOI: 10.1007/s11002-019-09495-7.
- [19] W.-H. S. Tsai, D. Lun, N. Carcioppolo, and C.-H. Chuan, "Human versus chatbot: Understanding the role of emotion in health marketing communication for vaccines," *Psychology & Marketing*, vol. 38, no. 12, pp. 2377–2392, 2021. DOI: <https://doi.org/10.1002/mar.21556>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mar.21556>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/mar.21556>.
- [20] S. S. I. Lei, H. Shen, and S. Ye, "A comparison between chatbot and human service: Customer perception and reuse intention," *International Journal of Contemporary Hospitality Management*, pp. 3977–3995, Aug. 2021. DOI: 10.1108/IJCHM-12-2020-1399.
- [21] A. Miklosik, N. Evans, and A. Qureshi, "The use of chatbots in digital business transformation: A systematic literature review," *IEEE Access*, vol. 9, pp. 106 530–106 539, Jul. 2021. DOI: 10.1109/ACCESS.2021.3100885.
- [22] C. Versteegen, *The pros and cons of chatbots*, Chatdesk, Jun. 2022. [Online]. Available: <https://www.chatdesk.com/blog/pros-and-cons-of-chatbots>.
- [23] A. M. Turing, "I.—computing machinery and intelligence," *Mind*, vol. LIX, no. 236, pp. 433–460, Oct. 1950, ISSN: 0026-4423. DOI: 10.1093/mind/LIX.236.433. eprint: <https://academic.oup.com/mind/article-pdf/LIX/236/433/30123314/lix-236-433.pdf>. [Online]. Available: <https://doi.org/10.1093/mind/LIX.236.433>.
- [24] J. Weizenbaum, "Eliza — a computer program for the study of natural language communication between man and machine," *Commun. ACM*, vol. 26, no. 1, pp. 23–28, Jan. 1983, ISSN: 0001-0782. DOI: 10.1145/357980.357991. [Online]. Available: <https://doi.org/10.1145/357980.357991>.

- [25] H.-Y. Shum, X. He, and D. Li, "From Eliza to XiaoIce: Challenges and opportunities with social chatbots," *Frontiers of Information Technology and Electronic Engineering*, vol. 19, Jan. 2018. DOI: 10.1631/FITEE.1700826.
- [26] R. S. Wallace, "The anatomy of a.l.i.c.e.," in *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer*, R. Epstein, G. Roberts, and G. Beber, Eds. Dordrecht: Springer Netherlands, 2009, pp. 181–210, ISBN: 978-1-4020-6710-5. DOI: 10.1007/978-1-4020-6710-5_13. [Online]. Available: https://doi.org/10.1007/978-1-4020-6710-5_13.
- [27] M. Rijhwani, *AIML: What is AIML: Artificial intelligence markup language*, May 2021. [Online]. Available: <https://www.analyticsvidhya.com/blog/2021/05/aiml-a-language-for-chatbots/>.
- [28] Y. Leviathan, *Google Duplex: An AI system for accomplishing real-world tasks over the phone*, 2022. [Online]. Available: <https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html>.
- [29] L. Klopfenstein, S. Delpriori, S. Malatini, and A. Bogliolo, "The rise of bots: A survey of conversational interfaces, patterns, and paradigms," in *Proceedings of the 2017 Conference on Designing Interactive Systems*, Jun. 2017, pp. 555–565. DOI: 10.1145/3064663.3064672.
- [30] S. Devaram, "Empathic chatbot: Emotional intelligence for empathic chatbot: Emotional intelligence for mental health well-being," *CoRR*, vol. abs/2012.09130, Dec. 2020. DOI: 10.13140/RG.2.2.16077.46564. arXiv: 2012.09130. [Online]. Available: <https://arxiv.org/abs/2012.09130>.
- [31] A. Xu, Z. Liu, Y. Guo, V. Sinha, and R. Akkiraju, "A new chatbot for customer service on social media," in *Proceedings of the 2017 CHI conference on human factors in computing*, May 2017. DOI: 10.1145/3025453.3025496.
- [32] J. Lee, *Brands expected to respond within an hour on twitter [study] - search engine watch*, 2013. [Online]. Available: <https://www.searchenginewatch.com/2013/11/01/brands-expected-to-respond-within-an-hour-on-twitter-study/>.
- [33] A. Vaidyam, H. Wisniewski, J. Halamka, M. Keshavan, and J. Torous, "Chatbots and conversational agents in mental health: A review of the psychiatric landscape," *The Canadian Journal of Psychiatry*, vol. 64, pp. 456–464, Mar. 2019. DOI: 10.1177/0706743719828977.
- [34] S. Narasiman, T. Srinivassababu, S. Raja, and R. Babu, "Indquery - an online portal for registering e-complaints integrated with smart chatbot," in Jan. 2020, pp. 1286–1294, ISBN: 978-3-030-32149-9. DOI: 10.1007/978-3-030-32150-5_130.
- [35] P. B. Brandtzaeg and A. Følstad, "Chatbots: Changing user needs and motivations," *Interactions*, vol. 25, no. 5, pp. 38–43, Aug. 2018, ISSN: 1072-5520. DOI: 10.1145/3236669. [Online]. Available: <https://doi.org/10.1145/3236669>.
- [36] Q. Ding, V. Carmona, M. Liu, F. Peng, Y. Zhang, and C. Hu, "Towards a six-level framework of emotional intelligence for customer service chatbots," in *AffCon@AAAI*, 2021.
- [37] E. Go and S. S. Sundar, "Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions," *Computers in Human Behavior*, vol. 97, Aug. 2019. DOI: 10.1016/j.chb.2019.01.020.
- [38] Y. Duan, M. Yoon, Z. Liang, and J. Hoorn, "Self-disclosure to a robot: Only for those who suffer the most," *Robotics*, vol. 10, p. 98, Jul. 2021. DOI: 10.3390/robotics10030098.
- [39] J. Othlinghaus-Wulhorst, A. Mainz, and H. U. Hoppe, "Training customer complaint management in a virtual role-playing game: A user study," *Lecture*

- Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11722 LNCS, pp. 436–449, 2019. DOI: 10.1007/978-3-030-29736-7_33.
- [40] A. Rossmann, A. Zimmermann, and D. Hertweck, “The impact of chatbots on customer service performance,” in *Advances in the Human Side of Service Engineering*. Jul. 2020, pp. 237–243, ISBN: 978-3-030-51056-5. DOI: 10.1007/978-3-030-51057-2_33.
- [41] C. Fornell, “A national customer satisfaction barometer: The Swedish experience,” *Journal of Marketing*, vol. 56, no. 1, pp. 6–21, 1992. DOI: 10.1177/002224299205600103. eprint: <https://doi.org/10.1177/002224299205600103>. [Online]. Available: <https://doi.org/10.1177/002224299205600103>.
- [42] A. Filip, “Complaint management: A customer satisfaction learning process,” *Procedia - Social and Behavioral Sciences*, vol. 93, pp. 271–275, 2013, 3rd World Conference on Learning, Teaching and Educational Leadership, ISSN: 1877-0428. DOI: <https://doi.org/10.1016/j.sbspro.2013.09.188>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877042813032916>.
- [43] F. Simon, “The influence of empathy in complaint handling: Evidence of gratitude and transactional routes to loyalty,” *Journal of Retailing and Consumer Services*, vol. 20, no. 6, pp. 599–608, 2013, ISSN: 0969-6989. DOI: <https://doi.org/10.1016/j.jretconser.2013.05.003>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S096969891300060X>.
- [44] C. Fornell and R. A. Westbrook, “The vicious circle of consumer complaints,” *Journal of Marketing*, vol. 48, no. 3, pp. 68–78, 1984, ISSN: 00222429. [Online]. Available: <http://www.jstor.org/stable/1251330> (visited on 06/27/2022).
- [45] O. Iyiola and O. S. Ibidunni, “The relationship between complaints, emotion, anger, and subsequent behavior of customers,” *IOSR Journal Of Humanities And Social Science*, vol. 17, pp. 34–41, Jan. 2013. DOI: 10.9790/0837-1763441.
- [46] J. Barsky and L. Nash, “Evoking emotion: Affective keys to hotel loyalty,” *The Cornell Hotel and Restaurant Administration Quarterly*, vol. 43, no. 1, pp. 39–46, 2002, ISSN: 0010-8804. DOI: [https://doi.org/10.1016/S0010-8804\(02\)80007-6](https://doi.org/10.1016/S0010-8804(02)80007-6). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0010880402800076>.
- [47] R. L. Oliver, R. T. Rust, and S. Varki, “Customer delight: Foundations, findings, and managerial insight,” *Journal of Retailing*, vol. 73, no. 3, pp. 311–336, 1997, ISSN: 0022-4359. DOI: [https://doi.org/10.1016/S0022-4359\(97\)90021-X](https://doi.org/10.1016/S0022-4359(97)90021-X). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S002243599790021X>.
- [48] C. Boshoff and J. Allen, “The influence of selected antecedents on frontline staff’s perceptions of service recovery performance,” *International Journal of Service Industry Management*, vol. 11, pp. 63–90, Mar. 2000. DOI: 10.1108/09564230010310295.
- [49] S. Michel and D. E. Bowen, “Service recovery management: Closing the gap between best practices and actual practices,” in *Academy of Management Perspectives*, 2006.
- [50] A. del Rio, R. Vázquez, and A. Díaz-Martín, “Satisfaction with service recovery: Perceived justice and emotional responses,” *Journal of Business Research*, vol. 62, pp. 775–781, Aug. 2009. DOI: 10.1016/j.jbusres.2008.09.015.
- [51] D. T. Nguyen and J. R. McColl-Kennedy, “Diffusing customer anger in service recovery: A conceptual framework,” *Australasian Marketing Journal (AMJ)*,

- vol. 11, no. 2, pp. 46–55, 2003, *Emerging Issues in Services Marketing, emotions, e-marketing and encounters*, ISSN: 1441-3582. DOI: [https://doi.org/10.1016/S1441-3582\(03\)70128-1](https://doi.org/10.1016/S1441-3582(03)70128-1). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1441358203701281>.
- [52] C. Homburg, A. Fürst, and N. Koschate, "On the importance of complaint handling design: A multi-level analysis of the impact in specific complaint situations," *Journal of the Academy of Marketing Science*, vol. 38, pp. 265–287, Jun. 2009. DOI: 10.1007/s11747-009-0172-y.
- [53] S. W. Kelley, K. Hoffman, and M. A. Davis, "A typology of retail failures and recoveries," *Journal of Retailing*, vol. 69, no. 4, pp. 429–452, 1993, ISSN: 0022-4359. DOI: [https://doi.org/10.1016/0022-4359\(93\)90016-C](https://doi.org/10.1016/0022-4359(93)90016-C). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S002243599390016C>.
- [54] L. Voinea, "Mutations in new consumer buying behaviour," *Quality - Access to Success*, vol. 12, pp. 47–51, Aug. 2011.
- [55] M. Constantinescu, "Relationship between quality of life and ethics in marketing," *Romanian Marketing Journal*, vol. 3, pp. 37–44, Sep. 2011.
- [56] K. Seawright, K. Detienne, M. Preston, B. Baker, H. P, and C. Larson, "An empirical examination of service recovery design," *Marketing Intelligence Planning*, vol. 26, May 2008. DOI: 10.1108/02634500810871320.
- [57] N. Gidener, "The service failure and recovery strategies in logistics service sector," *Celal Bayar Üniversitesi Sosyal Bilimler Dergisi*, vol. 14, Jan. 2016. DOI: 10.18026/cbusos.03006.
- [58] M. A. Clark, M. M. Robertson, and S. Young, "'I feel your pain': A critical review of organizational research on empathy," *Journal of Organizational Behavior*, vol. 40, no. 2, pp. 166–192, 2019. DOI: <https://doi.org/10.1002/job.2348>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/job.2348>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/job.2348>.
- [59] J. Bošnjaković and T. Radionov, "Empathy: Concepts, theories and neuroscientific basis," *Alcoholism and psychiatry research : Journal on psychiatric research and addictions*, vol. 54, pp. 123–150, Dec. 2018. DOI: 10.20471/dec.2018.54.02.04.
- [60] S. Preston and F. Waal, "Empathy: Its ultimate and proximate bases," *The Behavioral and brain sciences*, vol. 25, pp. 1–20, Mar. 2002. DOI: 10.1017/S0140525X02000018.
- [61] C. Clark, U. Murfett, P. Rogers, and S. Ang, "Is empathy effective for customer service? Evidence from call center interactions," *Journal of Business and Technical Communication*, vol. 27, pp. 123–153, Apr. 2013. DOI: 10.1177/1050651912468887.
- [62] B. Cuff, S. Brown, L. Taylor, and D. Howat, "Empathy: A review of the concept," *Emotion Review*, vol. 8, pp. 144–153, Apr. 2016. DOI: 10.1177/1754073914558466.
- [63] J. Wieseke, A. Geigenmüller, and F. Kraus, "On the role of empathy in customer-employee interactions," *Journal of Service Research*, vol. 15, pp. 316–331, Aug. 2012. DOI: 10.1177/1094670512439743.
- [64] J. Dovidio, "Adulthood: Prosocial behavior and empathy," in *International Encyclopedia of the Social Behavioral Sciences*, N. J. Smelser and P. B. Baltes, Eds., Oxford: Pergamon, 2001, pp. 159–162, ISBN: 978-0-08-043076-8. DOI: <https://doi.org/10.1016/B0-08-043076-7/01740-X>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B008043076701740X>.

- [65] L. Bove, "Empathy for service: Benefits, unintended consequences, and future research agenda," *Journal of Services Marketing*, vol. 33, Apr. 2019. DOI: 10.1108/JSM-10-2018-0289.
- [66] R. Fisher and D. Shapiro, *Beyond reason: Using emotions as you negotiate*. Penguin, 2005.
- [67] J. Parlamis, "Venting as emotion regulation," *International Journal of Conflict Management*, vol. 23, pp. 77–96, 2012.
- [68] A. Patel and L. Reinsch, "Companies can apologize: Corporate apologies and legal liability," *Business Communication Quarterly*, vol. 66, no. 1, pp. 9–25, 2003. DOI: 10.1177/108056990306600103. eprint: <https://doi.org/10.1177/108056990306600103>. [Online]. Available: <https://doi.org/10.1177/108056990306600103>.
- [69] M. Fjelstad, *Regret vs. remorse*, Psychology Today, 2022. [Online]. Available: <https://www.psychologytoday.com/us/blog/stop-caretaking-the-borderline-or-narcissist/201507/regret-vs-remorse>.
- [70] K. Salters-Pedneault, *What is emotional validation?* Very Well Mind, Mar. 2022. [Online]. Available: <https://www.verywellmind.com/what-is-emotional-validation-425336>.
- [71] J. Lambie and A. Lindberg, "The role of maternal emotional validation and invalidation on children's emotional awareness," *Merrill-Palmer Quarterly*, vol. 62, p. 129, Apr. 2016. DOI: 10.13110/merrpalquar1982.62.2.0129.
- [72] J. Lambie, H. Lambie, and S. Sadek, "'My child will actually say 'I am upset' before all they would do was scream": Teaching parents emotion validation in a social care," *Child: Care, Health and Development*, vol. 46, May 2020. DOI: 10.1111/cch.12770.
- [73] R. E. Plank, A. P. Minton, and D. A. Reid, "A short measure of perceived empathy," *Psychological Reports*, vol. 79, no. 3_suppl, pp. 1219–1226, 1996, PMID: 9009768. DOI: 10.2466/pr0.1996.79.3f.1219. eprint: <https://doi.org/10.2466/pr0.1996.79.3f.1219>. [Online]. Available: <https://doi.org/10.2466/pr0.1996.79.3f.1219>.
- [74] D. Lishner, E. Stocks, and S. Steinert, "Empathy," in Jan. 2017, pp. 1–8. DOI: 10.1007/978-3-319-28099-8_513-1.
- [75] D. Delpechitre, B. Rutherford, and L. Comer, "The importance of customer's perception of salesperson's empathy in selling," *Journal of Business Industrial Marketing*, vol. 34, Oct. 2018. DOI: 10.1108/JBIM-03-2017-0073.
- [76] L. Charrier, A. Rieger, A. Galdeano, A. Cordier, M. Lefort, and S. Hassas, "The rope scale: A measure of how empathic a robot is perceived," Mar. 2019. DOI: 10.1109/HRI.2019.8673082.
- [77] M. Waqas, H. Ali, and M. A. Khan, "An investigation of effects of justice recovery dimensions on students' satisfaction with service recovery in higher education environment," *International Review on Public and Nonprofit Marketing*, vol. 11, no. 3, pp. 263–284, 2014, Cited by: 10. DOI: 10.1007/s12208-014-0120-5. [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85027916774&doi=10.1007%2fs12208-014-0120-5&partnerID=40&md5=bd17d1495b191409f6187876548e5604>.
- [78] C. Boshoff, "Recovsat: An instrument to measure satisfaction with transaction-specific service recovery," *Journal of Service Research*, vol. 1, no. 3, pp. 236–249, 1999. DOI: 10.1177/109467059913005. eprint: <https://doi.org/10.1177/109467059913005>. [Online]. Available: <https://doi.org/10.1177/109467059913005>.

- [79] Y. Ma, K. L. Nguyen, F. Z. Xing, and E. Cambria, "A survey on empathetic dialogue systems," *Information Fusion*, vol. 64, pp. 50–70, 2020, ISSN: 1566-2535. DOI: <https://doi.org/10.1016/j.inffus.2020.06.011>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1566253520303092>.
- [80] L. Wang *et al.*, *Empathetic dialogue generation via sensitive emotion recognition and sensible knowledge selection*, 2022. DOI: 10.48550/ARXIV.2210.11715. [Online]. Available: <https://arxiv.org/abs/2210.11715>.
- [81] C. Lisetti, R. Amini, U. Yasavur, and N. Rishe, "I can help you change! an empathic virtual agent delivers behavior change health interventions," *ACM Transactions on Management Information Systems (TMIS)*, vol. 4, Dec. 2013. DOI: 10.1145/2544103.
- [82] M. Ochs, D. Sadek, and C. Pelachaud, "A formal model of emotions for an empathic rational dialog agent," *Autonomous Agents and Multi-Agent Systems*, vol. 24, pp. 410–440, May 2012. DOI: 10.1007/s10458-010-9156-z.
- [83] H. Boukricha, I. Wachsmuth, M. Carminati, and P. Knoeferle, "A computational model of empathy: Empirical evaluation," Sep. 2013. DOI: 10.1109/ACII.2013.7.
- [84] O. Yalcin, "Empathy framework for embodied conversational agents," *Cognitive Systems Research*, vol. 59, Sep. 2019. DOI: 10.1016/j.cogsys.2019.09.016.
- [85] Ö. N. Yalçın and S. DiPaola, "M-path: A conversational system for the empathic virtual agent," in *Biologically Inspired Cognitive Architectures 2019*, A. V. Samsonovich, Ed., Cham: Springer International Publishing, 2020, pp. 597–607, ISBN: 978-3-030-25719-4.
- [86] Ö. Yalçın and S. R. DiPaola, "Modeling empathy: Building a link between affective and cognitive processes," *Artificial Intelligence Review*, vol. 53, pp. 2983–3006, 2019.
- [87] S. Rodrigues, S. Mascarenhas, J. Dias, and A. Paiva, "A process model of empathy for virtual agents," *Interacting with Computers*, vol. 27, Feb. 2014. DOI: 10.1093/iwc/iwu001.
- [88] A. Welivita and P. Pu, "A taxonomy of empathetic response intents in human social conversations," in *Proceedings of the 28th International Conference on Computational Linguistics*, Barcelona, Spain (Online): International Committee on Computational Linguistics, Dec. 2020, pp. 4886–4899. DOI: 10.18653/v1/2020.coling-main.429. [Online]. Available: <https://aclanthology.org/2020.coling-main.429>.
- [89] J. Gao, M. Galley, and L. Li, "Neural approaches to conversational AI," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts*, Melbourne, Australia: Association for Computational Linguistics, Jul. 2018, pp. 2–7. DOI: 10.18653/v1/P18-5002. [Online]. Available: <https://aclanthology.org/P18-5002>.
- [90] Z. Lin *et al.*, "Caire: An end-to-end empathetic chatbot," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 13 622–13 623, Apr. 2020. DOI: 10.1609/aaai.v34i09.7098.
- [91] A. Radford, K. Narasimhan, T. Salimans, I. Sutskever, *et al.*, "Improving language understanding by generative pre-training," 2018.
- [92] S. Zhang, E. Dinan, J. Urbanek, A. Szlam, D. Kiela, and J. Weston, "Personalizing dialogue agents: I have a dog, do you have pets too?" In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume*

- 1: *Long Papers*), Melbourne, Australia: Association for Computational Linguistics, Jul. 2018, pp. 2204–2213. DOI: 10.18653/v1/P18-1205. [Online]. Available: <https://aclanthology.org/P18-1205>.
- [93] H. Rashkin, E. M. Smith, M. Li, and Y.-L. Boureau, “Towards empathetic open-domain conversation models: A new benchmark and dataset,” in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy: Association for Computational Linguistics, Jul. 2019, pp. 5370–5381. DOI: 10.18653/v1/P19-1534. [Online]. Available: <https://aclanthology.org/P19-1534>.
- [94] T. Hu *et al.*, “Touch your heart: A tone-aware chatbot for customer care on social media,” Apr. 2018, pp. 1–12. DOI: 10.1145/3173574.3173989.
- [95] H. Zhou, M. Huang, T. Zhang, X. Zhu, and B. Liu, “Emotional chatting machine: Emotional conversation generation with internal and external memory,” *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, Apr. 2017. DOI: 10.1609/aaai.v32i1.11325.
- [96] R. Hoegen, D. Aneja, D. McDuff, and M. Czerwinski, “An end-to-end conversational style matching agent,” in *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, ser. IVA ’19, Paris, France: Association for Computing Machinery, 2019, pp. 111–118, ISBN: 9781450366724. DOI: 10.1145/3308532.3329473. [Online]. Available: <https://doi.org/10.1145/3308532.3329473>.
- [97] B. Zhang, G. Essl, and E. Mower Provost, “Automatic recognition of self-reported and perceived emotion: Does joint modeling help?,” Oct. 2016, pp. 217–224. DOI: 10.1145/2993148.2993173.
- [98] E. Svikhnushina, I. Voinea, A. Welivita, and P. Pu, “A taxonomy of empathetic questions in social dialogs,” Jan. 2022, pp. 2952–2973. DOI: 10.18653/v1/2022.acl-long.211.
- [99] R. Razali and J. Jaafar, “Complaint handling theoretical framework,” in *2012 International Conference on Computer Information Science (ICCIS)*, vol. 1, 2012, pp. 382–385. DOI: 10.1109/ICCISci.2012.6297275.
- [100] B. Stauss and W. Seidel, *Effective Complaint Management (Management for Professionals 978-3-319-98705-7)*, 2nd. Springer, Dec. 2019, ISBN: ARRAY(0x45d0c260). DOI: 10.1007/978-3-319-98705-7. [Online]. Available: <https://ideas.repec.org/b/spr/mgmtpr/978-3-319-98705-7.html>.
- [101] M. Zairi, “Managing customer dissatisfaction through effective complaint management systems,” *The TQM Magazine*, vol. 12, pp. 331–337, Oct. 2000. DOI: 10.1108/09544780010341932.
- [102] H.-H. Lin, Y.-S. Wang, and L.-K. Chang, “Consumer responses to online retailer’s service recovery after a service failure: A perspective of justice theory,” *Managing Service Quality*, vol. 21, pp. 511–534, Sep. 2011. DOI: 10.1108/09604521111159807.
- [103] S. Borsci *et al.*, “The chatbot usability scale: The design and pilot of a usability scale for interaction with ai-based conversational agents,” *Personal and Ubiquitous Computing*, vol. 26, pp. 1–25, Feb. 2022. DOI: 10.1007/s00779-021-01582-9.
- [104] A. F. Bulagang, N. G. Weng, J. Mountstephens, and J. Teo, “A review of recent approaches for emotion classification using electrocardiography and electrodermography signals,” *Informatics in Medicine Unlocked*, vol. 20, p. 100363, 2020, ISSN: 2352-9148. DOI: <https://doi.org/10.1016/j.imu.2020.100363>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352914820301040>.

- [105] H. Crijns, V. Cauberghe, L. Hudders, and A.-S. Claeys, "How to deal with online consumer comments during a crisis? the impact of personalized organizational responses on organizational reputation," *Computers in Human Behavior*, vol. 75, pp. 619–631, 2017, ISSN: 0747-5632. DOI: <https://doi.org/10.1016/j.chb.2017.05.046>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0747563217303679>.
- [106] R. PLUTCHIK, "A general psychoevolutionary theory of emotion," in *Theories of Emotion*, R. Plutchik and H. Kellerman, Eds. Academic Press, 1980, ISBN: 978-0-12-558701-3. DOI: <https://doi.org/10.1016/B978-0-12-558701-3.50007-7>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780125587013500077>.
- [107] G. Feigenblat, C. Gunasekara, B. Sznajder, S. Joshi, D. Konopnicki, and R. Aharonov, "TWEETSUMM - a dialog summarization dataset for customer service," in *Findings of the Association for Computational Linguistics: EMNLP 2021*, Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 245–260. [Online]. Available: <https://aclanthology.org/2021.findings-emnlp.24>.
- [108] N. Siu, J. Zhang, and C.-Y. Yau, "The roles of justice and customer satisfaction in customer retention: A lesson from service recovery," *Journal of Business Ethics*, vol. 114, Jun. 2013. DOI: 10.1007/s10551-013-1713-3.
- [109] C. Bartneck and M. Keijsers, *Paladyn, Journal of Behavioral Robotics*, vol. 11, no. 1, pp. 271–283, 2020. DOI: [doi:10.1515/pjbr-2020-0017](https://doi.org/10.1515/pjbr-2020-0017). [Online]. Available: <https://doi.org/10.1515/pjbr-2020-0017>.
- [110] A. Elhami, "Communication accommodation theory: A brief review of the literature," *Journal of Advances in Education and Philosophy*, vol. 4, pp. 192–200, May 2020. DOI: 10.36348/jaep.2020.v04i05.002.
- [111] S. Park and M. Whang, "Empathy in human-robot interaction: Designing for social robots," *International Journal of Environmental Research and Public Health*, vol. 19, p. 1889, Feb. 2022. DOI: 10.3390/ijerph19031889.
- [112] T. Bickmore and R. Picard, "Establishing and maintaining long-term human-computer relationships," *ACM Trans. Comput.-Hum. Interact.*, vol. 12, pp. 293–327, Jun. 2005. DOI: 10.1145/1067860.1067867.
- [113] B. Pompe, E. Velnor, and K. Truong, "The robot that showed remorse: Repairing trust with a genuine apology," Aug. 2022, pp. 260–265. DOI: 10.1109/RO-MAN53752.2022.9900860.
- [114] A. Følstad, C. Nordheim, and C. Bjørkli, "What makes users trust a chatbot for customer service? an exploratory interview study," Oct. 2018, pp. 194–208, ISBN: 978-3-030-01436-0. DOI: 10.1007/978-3-030-01437-7_16.
- [115] M. King and G. Bruner, "Social desirability bias: A neglected aspect of validity testing," *Psychology and Marketing*, vol. 17, pp. 79–103, Feb. 2000. DOI: 10.1002/(SICI)1520-6793(200002)17:2<79::AID-MAR2>3.0.CO;2-0.

Appendix A

Implementation details

A.1 Models

A.1.1 Emotion recognizer

For emotion recognition NRCLex¹ was used. To fit the purpose of the study, the lexicon was adapted slightly. For example, words as 'birthday' would usually have a positive connotation, but in this case rather express a negative emotion.

```

1 def emotion_recognizer(msg):
2     # Take message as input and return highest ranking emotion
3
4     emotion = NRCLex(msg)
5     aff_freq = emotion.affect_frequencies
6
7     # Remove the unnecessary dictionary items
8     del aff_freq['negative'], aff_freq['positive']
9
10    # Necessary debug
11    if 'anticipation' in aff_freq:
12        del aff_freq['anticipation']
13
14    max_value = max(aff_freq.values())
15    top_emo = {k: v for k, v in aff_freq.items() if v == max_value}
16
17    # Replace 'anticip' with 'anticipation'
18    if 'anticip' in top_emo.keys():
19        top_emo['anticipation'] = top_emo.pop('anticip')
20
21    return top_emo

```

LISTING A.1: Python code for emotion recognizer function

A.1.2 Intent classifier

An intent classifier was custom made for the purpose of this study. The library 'sklearn'² was used to train a model on custom data. First, 'CountVectorizer' converted the text into a numerical representation. Then, 'MultinomialNB' was used as a classifier based upon the multinomial naive Bayes algorithm. Additionally, Grid-SearchCV was used for hyperparameter tuning to find the best set of hyperparameters for the classifier.

¹<https://pypi.org/project/NRCLex/>

²<https://scikit-learn.org/stable/>

The data used for training was custom made. All necessary intents were identified, and example sentences for each intent were added. ChatGPT was used to expand the list of example sentences, by giving commands as 'Give me 15 sentences similar to *'neutral sentence example'*.' or 'Give me 15 sentences expressing the intent *'bot intent'*'. Since ChatGPT is trained upon an enormous dataset based upon human data, this could be used to provide more data samples. As a result, each intent had 15-30 (depending on the complexity of the intent) sample sentences for training.

```
1 def train_model(data_file):
2     # Function to train an intents classifier model using the data in the
3     # JSON file
4     with open(data_file) as f:
5         data = json.load(f)
6
7     training_data = []
8     labels = []
9
10    for intent in data['intents']:
11        for phrase in intent['training_phrases']:
12            training_data.append(phrase)
13            labels.append(intent['intent_name'])
14
15    pipeline = Pipeline([
16        ('vect', CountVectorizer(ngram_range=(1,2), min_df=2)),
17        ('clf', MultinomialNB(alpha=0.1))
18    ])
19
20    model = pipeline.fit(training_data, labels)
21
22    joblib.dump(model, CLF_DIR)
23
24    return model
```

LISTING A.2: Python code for intent classification function

A.1.3 Entity extraction

For entity extraction the library 'spacy'³ was used. The code was based upon the tutorial from Medium⁴. Again, the samples for training data were first made manually, and expanded using ChatGPT.

A.2 User interface

For the user interface, the template from Studygyaan⁵ was used and adapted (e.g. colors and background). To this template the chatbot and user avatar were added, as well as 'typing-indicators' to make it seem as if the chatbot was typing a response.

³<https://spacy.io/>

⁴<https://medium.com/mlearning-ai/named-entity-recognition-with-spacy-fd834ff84b86>

⁵https://studygyaan.com/python/create-web-based-chatbot-in-python-django-flask?utm_content=cmp-true

Appendix B

Complaint handling framework

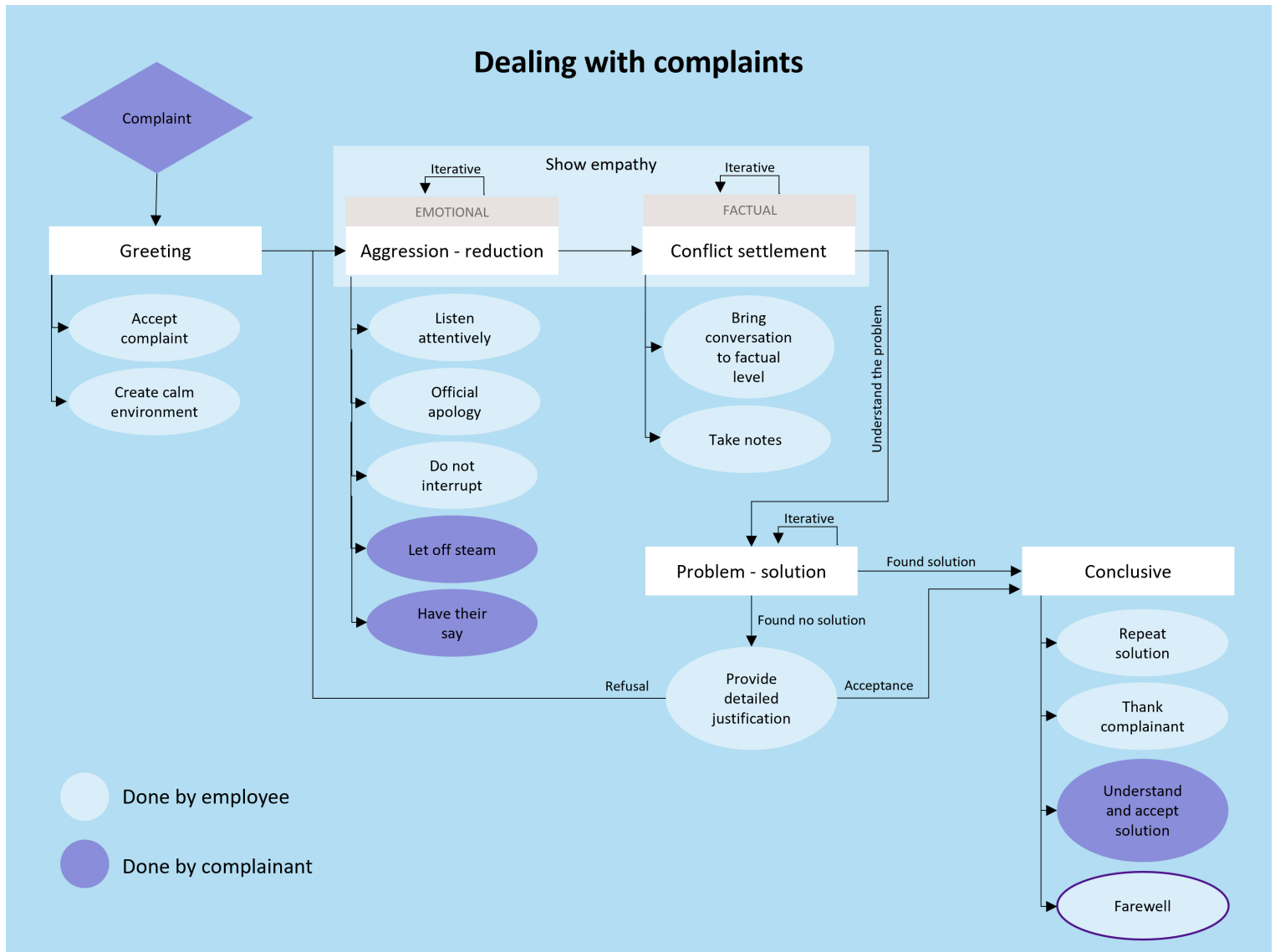


FIGURE B.1: Model of dealing with complaints based upon Stauss and Seidel [100]

Appendix C

Complete decision tree

The following figures show the decision tree. For size it was split up into four parts, showing the greeting phase, aggression reduction and conflict settlement phase, problem solution phase, and conclusive phase. The dashed outline indicates where each phase starts and where from where it continues in a new figure.

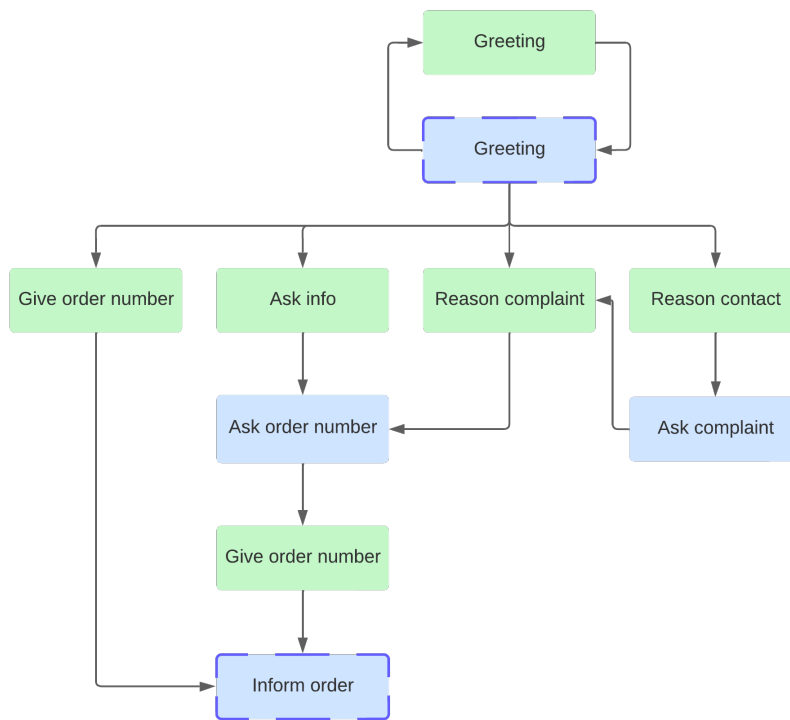


FIGURE C.1: Greeting phase

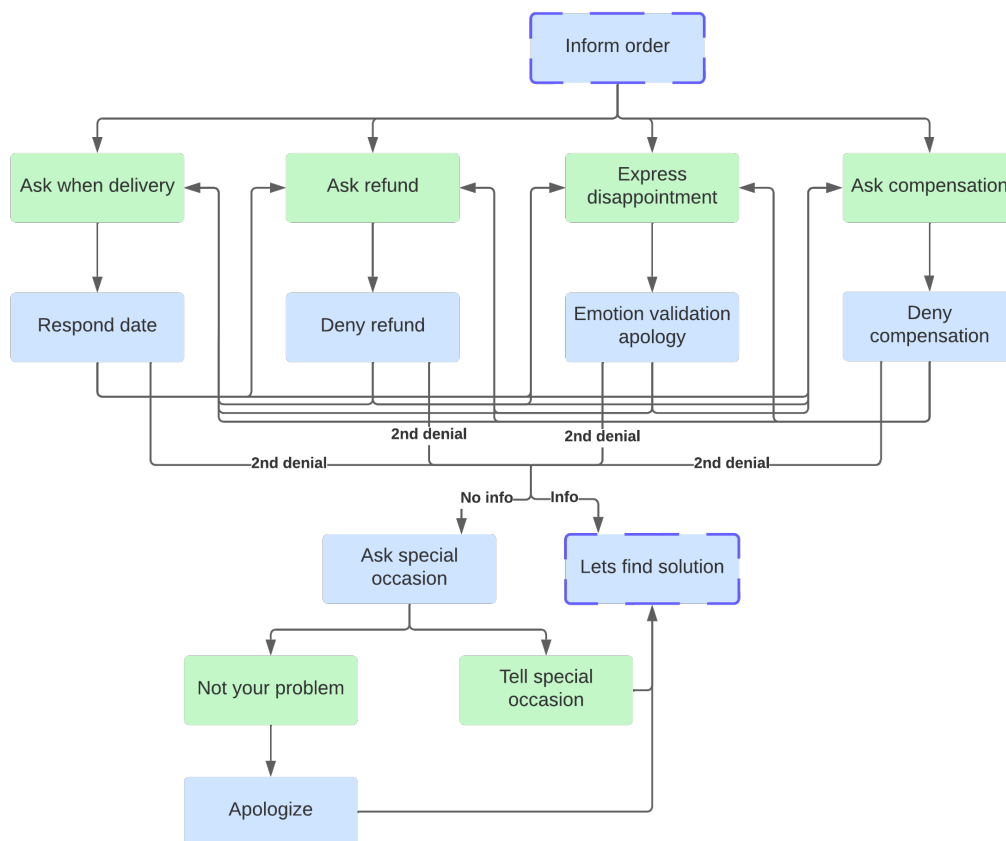


FIGURE C.2: Aggression reduction + conflict settlement phase

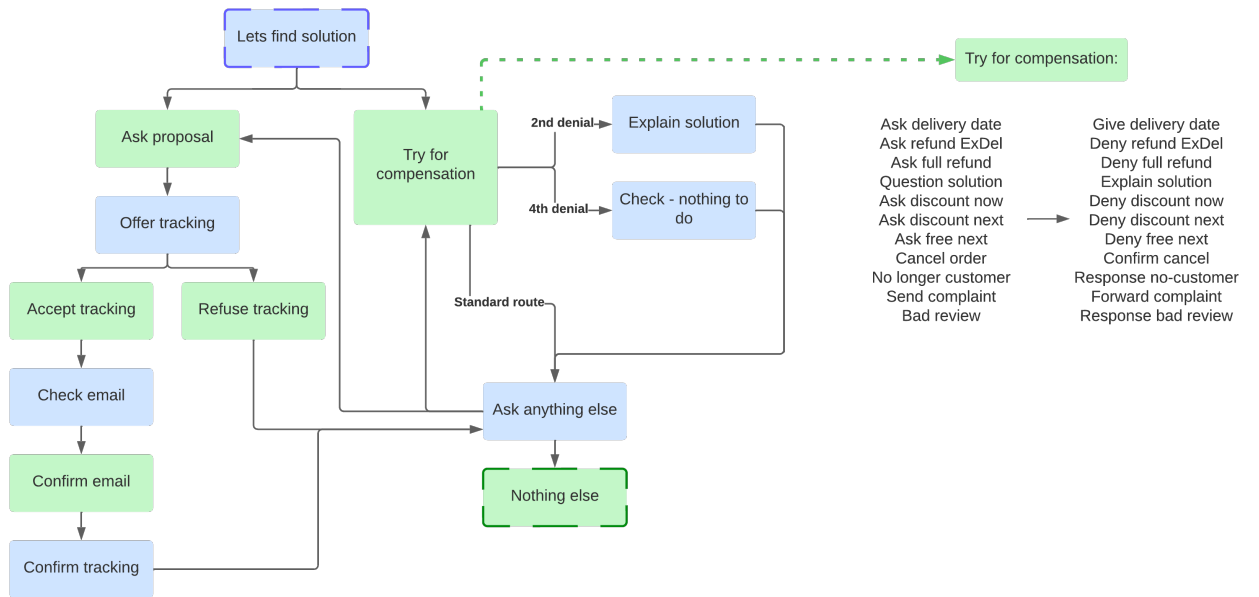


FIGURE C.3: Problem solution phase

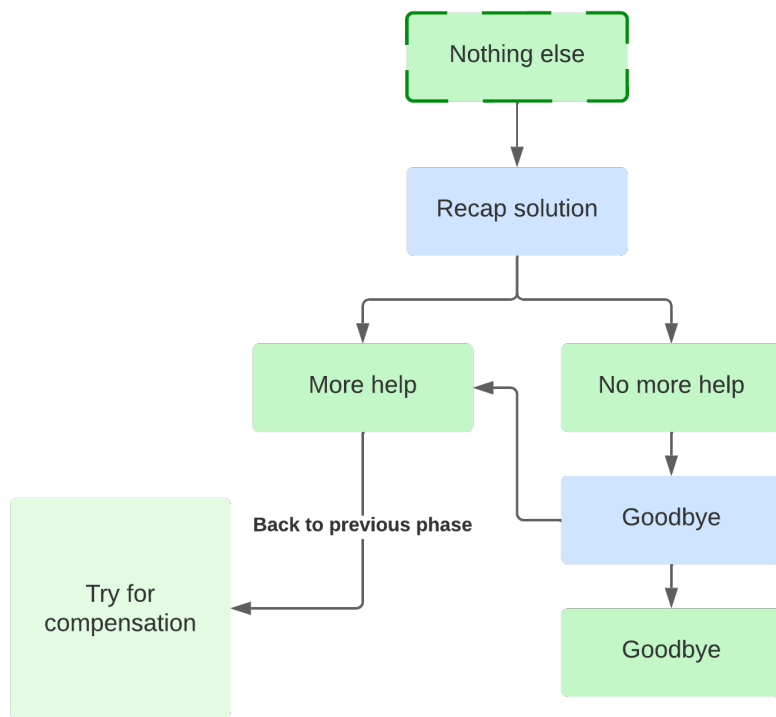


FIGURE C.4: Conclusive phase

Appendix D

Experiment Instructions

D.1 Informatie voor deelnemers (NL)

Bedankt voor je deelname aan dit onderzoek. Je zal een gesprekje aangaan met een customer service chatbot. Stel je voor dat je je bevindt in de volgende situatie:

Je bent een klant die praat met customer support. Je bent gefrustreerd want je hebt een pakketje besteld met Express Delivery maar het is alsnog te laat, het was de bedoeling dat het een cadeautje was voor je favoriete nichtje. Je hebt zelfs extra betaald voor de Express Delivery zodat het op tijd zou zijn. Helaas was het dat niet, de verjaardag was gisteren en het pakketje is er nog steeds niet. Je wil iets van compensatie krijgen, en wil alsnog dat het pakketje bezorgd wordt zodat je het aan je nichtje kunt geven.

Extra informatie voor het gesprek met de chatbot:

- De chatbot is puur gebouwd als prototype voor dit onderzoek, en maakt soms fouten. Wanneer de chatbot je verkeerd begrijpt kun je het aan hem laten weten, en dan zal hij proberen je alsnog te begrijpen. Soms geeft de chatbot ook opties om uit te kiezen. Bijvoorbeeld: "I did not quite get that, did you mean to say hi [HI] or ask information [INFO]?". Dan kun je de tekst in de vierkante haken sturen in CapsLock (in dit geval HI of INFO) om hem te laten weten wat je bedoelt.
- Je naam is John or Jane Smith
- Je package number is ABC123

Extra hulp nodig?

De onderzoeker is het hele onderzoek beschikbaar. Voel je vrij om vragen te stellen tijdens het onderzoek.

D.2 Information for participants (ENG)

Thank you for participating in this study. You will have a conversation with a customer service chatbot. Imagine you are in the following scenario:

You are a customer chatting to customer support. You are frustrated because you ordered a package with express delivery but it is too late, it was supposed to be a birthday present for your favourite cousin. You even paid express delivery so it would be in time. Unfortunately it was not, the birthday was yesterday and the package has still not arrived. You want to receive some sort of compensation, but still want to get the package delivered to gift to your cousin.

Additional information for interacting with the chatbot:

- The chatbot is built for research purposes only, and may make mistakes. When the chatbot misunderstands you, you can inform them and they will try to improve. Sometimes the chatbot will suggest options for you to choose from. For example: "I did not quite get that, did you mean to say hi [HI] or ask information [INFO]?". Then you can type the text in the square brackets (in this case HI or INFO) in CapsLock to let them know what you meant.
- Your name is John or Jane Smith
- Your package number is ABC123

Need extra help?

The researcher will be available at all times. Please feel free to ask questions when they arise during the experiment.

Appendix E

Extra (Fallback) Response Intents

bot_intent	variation_1	variation_2	variation_3
ask manager	I'm sorry to inform you that the manager is unavailable today. I apologize for any inconvenience. Is there anything else I can assist you with?	Regrettably, the manager is not available today. I apologize for any inconvenience this may cause. Can I help you with anything else?	I regret to inform you that the manager is unavailable today. I apologize for the inconvenience. Do you require any further assistance?
misunderstood	I apologize, but I'm unable to assist with that. Could you please try asking for something else?	I'm sorry, my capabilities don't cover that. Could you please ask for something else?	Unfortunately, I'm not equipped to help with that. Can you please try asking for something different?
info something	I apologize, but I'm unable to assist with that. Could you please try asking for something else?	I'm sorry, my capabilities don't cover that. Could you please ask for something else?	Unfortunately, I'm not equipped to help with that. Can you please try asking for something different?
blank	Okay, what further assistance can I provide for you?	Okay, in what other ways can I be of assistance to you?	Okay, is there anything else I can do to help you?

error	I'm sorry, I'm not quite following you. Can you explain what I can do to assist you at this moment?	I apologize, but I'm having trouble understanding your needs. Could you clarify how I can be of assistance right now?	It seems like we're not on the same wavelength. Can you please elaborate on what I can do to help you currently?
insult	I'm giving my best effort, I apologize if it's not enough.	I'm sorry, I am trying my best.	I'm trying my hardest, but I understand if it's not satisfactory.
dont offer	I sympathize with your perspective, and we're actively exploring solutions to address the issue. But unfortunately, the policy cannot be altered right now.	I acknowledge your point of view, and we're currently investigating options to resolve the matter. Nonetheless, the policy remains unchangeable at the moment.	Yes, I understand where you're coming from, and I can assure you that we are looking into a way to fix that as well. However, at this moment I cannot change the policy.
cg delivery company	You can contact the delivery company at 0614582413, if you tell them your order number they should further help you.	To further assist you, please contact the delivery company at 0614582413 and provide them with your order number.	If you need help with your order, feel free to call the delivery company at 0614582413 and provide your order number when prompted.
cg cs	You can find all the contact information of our different customer support on www.customersupportFlow.com .	Our www.customersupportFlow.com website, has all the contact information for our various customer support options.	If you need to contact customer support, you can find all of our different options' contact information on www.customersupportFlow.com .

Appendix F

Response Dataset

bot_intent	neutral	questioning	acknowledging	wishing	sympathizing	agreeing	suggesting
greeting	Hey \$PERSON!	Hey \$PERSON!	Hey \$PERSON!	Hey \$PERSON!	Hey \$PERSON!	Hey \$PERSON!	Hey \$PERSON!
ask order number	I'm sorry to hear that your \$PACKAGE hasn't arrived yet. Could you give me your order number?	I'm sorry to hear that your \$PACKAGE hasn't arrived yet. Do you have an order number for me so I can look up your information?	I'm sorry to hear that your \$PACKAGE has not yet arrived, I can understand that is inconvenient. Can you provide me with your order number?	I'm sorry to hear that you \$PACKAGE has not been delivered yet, I hope it arrives soon! Could you please share your order number with me so I can track it for you?	I'm sorry to hear that your \$PACKAGE hasn't arrived yet. That must be frustrating. Can you please provide me with your order number so that I can investigate the issue?	Sorry to hear that, it's always frustrating when a \$PACKAGE doesn't arrive on time. Do you have your order number for me?	Sorry to hear that, if you have your order number available, I can try to track down your \$PACKAGE. Would you mind sharing it with me?
ask complaint	I'm sorry to hear that you experienced issues with our service, what happened?	I completely understand how frustrating it can be to experience issues with our service. Would you mind sharing what happened?	I bet it must be frustrating to experience issues with our service. What happened?	I completely understand how frustrating it can be to experience issues with our service. Hopefully we can find a solution. What happened?	I'm sorry to hear that you experienced issues with our service. What happened?	I completely understand how frustrating it can be to experience issues with our service. What happened?	I completely understand how frustrating it can be to experience issues with our service. Perhaps we can work it out together. What happened?

emotion validation apology	I apologize for the inconvenience that this has caused you.	I understand that you can feel \$EMOTION when a \$PACKAGE is delayed. I apologize for the inconvenience that this has caused you. How can I help you further?	I can understand that you can feel \$EMOTION when a \$PACKAGE is delayed. I apologize for the inconvenience that this has caused you. I hope that we can get your \$PACKAGE to you as soon as possible.	I understand that you feel \$EMOTION when a \$PACKAGE is delayed. I hope we can quickly resolve the issue, and I'm sorry you're experiencing this.	I'm sorry that you feel \$EMOTION due to the delayed \$PACKAGE. I apologize for the inconvenience that this has caused you. How can I help you further?	I completely understand that you can feel \$EMOTION when a \$PACKAGE is delayed. I apologize for the inconvenience that this has caused you. I hope that we can get your \$PACKAGE to you as soon as possible.	I understand that you can feel \$EMOTION when a \$PACKAGE is delayed. I apologize for the inconvenience that this has caused you. I suggest we work together to resolve this issue.
deny refund	I'm sorry \$PERSON, but unfortunately I am not allowed to give any refunds for delivery as this is out of our control. However, I understand that it may be disappointing.	Unfortunately, we do not give refunds for deliveries as it is beyond our control. Since the packages are delivered by an external company, we have to adhere to this policy. However, I understand that it may be disappointing. Is there something else I can do for you?	I get your disappointment \$PERSON, but unfortunately I cannot provide a refund for deliveries. Since an external company delivers the packages, we have no control over this situation. However, I can understand that this may be disappointing.	I'm sorry to hear about your experience \$PERSON, but we are not allowed to offer refunds for deliveries. Since the packages are delivered by an external company, we have no choice but to follow this policy. I wish I could help more.	I sympathize with your situation \$PERSON, but unfortunately, we do not give refunds for deliveries. Since the packages are delivered by an external company, we have to abide by this rule. However, I understand that this may be disappointing	I agree that it's frustrating \$PERSON, but unfortunately, we cannot offer refunds for deliveries. Since the packages are delivered by an external company, we are unable to make any exceptions to this policy. However, I understand that this may be disappointing.	I'm sorry \$PERSON, but unfortunately I am not allowed to give any refunds for delivery as this is out of our control. However, I understand that it may be disappointing. Perhaps there is something else I can do for you?

deny refund ask special	I apologize \$PERSON, but I am also not allowed to give any refunds on deliveries. Since an external company delivers the packages. I understand the disappointment, was the package for a special occasion?	I apologize \$PERSON, but I am also not allowed to give any refunds on deliveries. Since an external company delivers the packages. I understand the disappointment, was the package for a special occasion?	I get that this is disappointing again \$PERSON, but I am also not allowed to give any refunds on deliveries. Since an external company delivers the packages. I can understand the disappointment, was the package for a special occasion?	I wish I could do more \$PERSON, but I am also not allowed to give any refunds on deliveries. Since an external company delivers the packages. I understand the disappointment, was the package for a special occasion?	I apologize \$PERSON, but I am also not allowed to give any refunds on deliveries. Since an external company delivers the packages. I understand how disappointing this must be, was the package for a special occasion?	I completely understand that its disappointing \$PERSON, but I am also not allowed to give any refunds on deliveries. Since an external company delivers the packages. I understand that this may not be what you wanted to hear, was the package for a special occasion?	I apologize \$PERSON, but I am also not allowed to give any refunds on deliveries. Since an external company delivers the packages. I understand how disappointing this must be, was the package for a special occasion?
deny refund lets find solution	I'm sorry \$PERSON, but we also can't provide a refund, let's see what we can do to find another solution to this problem.	I'm sorry \$PERSON, but we also can't provide a refund. Shall we try to find another solution to this issue together?	I get that this is not what you want to hear \$PERSON, but we also can't provide a refund. Let's try to find another solution to this issue.	I wish I could give you a refund \$PERSON, but we can't provide a refund due to our policy. Let's try to find another solution to this issue.	I'm sorry \$PERSON, this is disappointing, but we also can't provide a refund. Let's try to find another solution to this issue.	I absolutely understand that this may be disappointing \$PERSON, but we also can't provide a refund. Let's try to find another solution to this issue.	I'm sorry \$PERSON, but we also can't provide a refund. Maybe we can try to find another solution to this issue?

deny compensation	I'm sorry, but unfortunately I am not allowed to give compensation for delayed packages as this is out of our control. However, I understand that it may be disappointing.	I understand your frustration, but unfortunately, we do not offer compensation for delayed packages as it is beyond our control. Since the packages are delivered by an external company, we have to adhere to this policy. However, I understand that it may be disappointing. Can I help you with anything else?	I get your disappointment, but unfortunately, I cannot provide compensation for delayed packages as it is out of our control. Since an external company delivers the packages, we have no control over this situation. However, I can understand that it may be disappointing.	I'm sorry to hear about your experience, but we are not allowed to offer compensation for delayed packages as this is beyond our control. Since the packages are delivered by an external company, we have no choice but to follow this policy. I hope for your understanding. However, I understand that it may be disappointing.	I sympathize with your situation, but unfortunately, we do not provide compensation for delayed packages as it is beyond our control. Since the packages are delivered by an external company, we have to abide by this rule. However, I understand that it may be disappointing.	I agree that it's frustrating, but unfortunately, we cannot offer compensation for delayed packages as it is out of our control. Since the packages are delivered by an external company, we are unable to make any exceptions to this policy. However, I understand that it may be disappointing.	I wish I could help you more, but unfortunately, we are not allowed to offer compensation for delayed packages as this is outside our control. Since an external company delivers the packages, we have to follow this policy. However, I understand that it may be disappointing. Perhaps there is something else I can do for you?
deny compensation ask special	I apologize, but I am also not allowed to give any compensation for delayed packages. Since an external company delivers the packages. I understand the disappointment, was the package for a special occasion?	I apologize, but I am also not allowed to give any compensation for delayed packages. Since an external company delivers the packages. I understand the disappointment, was the package for a special occasion?	I'm sorry for the delayed delivery and understand the inconvenience. However, as per policy, I cannot offer compensation for external company's delayed packages. Was the package meant for a special occasion?	I'm sorry to tell you that I cannot give compensation for delayed deliveries as it's outside of our control. However, I wish it had arrived on time for you. Was the package for a special occasion?	I'm sorry to disappoint you, but I am also not allowed to give any compensation for delayed packages. Since an external company delivers the packages. I understand the disappointment, was the package for a special occasion?	I apologize, but I am also not allowed to give any compensation for delayed packages. Since an external company delivers the packages. I get the disappointment, was the package for a special occasion?	I apologize, but I am also not allowed to give any compensation for delayed packages. Since an external company delivers the packages. I understand the disappointment, was the package for a special occasion?

<p>explain solution ask anything else</p>	<p>Unfortunately I am really not allowed to give any kind of compensation for delayed packages. I understand that your situation is delicate as it was supposed to be a present, but unfortunately I am still not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. I hope for your understanding. Is there anything else I can help you with today?</p>	<p>Unfortunately I am really not allowed to give any kind of compensation for delayed packages. I understand that your situation is delicate as it was supposed to be a present, but unfortunately I am still not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. I hope for your understanding. Is there anything else I can help you with today?</p>	<p>While I completely understand your request, I am really not allowed to give any kind of compensation for delayed packages. I can understand that your situation is delicate as it was supposed to be a present, but unfortunately I am still not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. I hope for your understanding.</p>	<p>I wish I could give a different answer, but unfortunately I am really not allowed to give any kind of compensation for delayed packages. I understand that your situation is delicate as it was supposed to be a present, but unfortunately I am still not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. I hope for your understanding.</p>	<p>I'm sorry, but unfortunately I am really not allowed to give any kind of compensation for delayed packages. I understand that your situation is delicate as it was supposed to be a present, but unfortunately I am still not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. I hope for your understanding.</p>	<p>I completely understand your request but unfortunately I am really not allowed to give any kind of compensation for delayed packages. I understand that your situation is delicate as it was supposed to be a present, but unfortunately I am still not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. I hope for your understanding.</p>	<p>Unfortunately I am really not allowed to give any kind of compensation for delayed packages. I understand that your situation is delicate as it was supposed to be a present, but unfortunately I am still not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. Perhaps this makes it more understandable?</p>
---	--	--	--	--	---	---	--

<p>check nothing todo ask any- thing else</p>	<p>I'm afraid I cannot do anything in terms of compensation for you, but I just checked the policies to see if there are any exceptions. Unfortunately I was correct and I am not able to give any compensations or refunds. Is there something else I can help you with today?</p>	<p>I'm afraid I cannot do anything in terms of compensation for you, but I just checked the policies to see if there are any exceptions. Unfortunately I was correct and I am not able to give any compensations or refunds. Is there something else I can help you with today?</p>	<p>I'm afraid I cannot do anything in terms of compensation for you, but I just checked the policies to see if there are any exceptions. Unfortunately I was correct and I am not able to give any compensations or refunds.</p>	<p>I'm afraid I cannot do anything in terms of compensation for you, but I just checked the policies to see if there are any exceptions. Unfortunately I was correct and I am not able to give any compensations or refunds.</p>	<p>I'm afraid I cannot do anything in terms of compensation for you, but I just checked the policies to see if there are any exceptions. Unfortunately I was correct and I am not able to give any compensations or refunds.</p>	<p>I'm afraid I cannot do anything in terms of compensation for you, but I just checked the policies to see if there are any exceptions. Unfortunately I was correct and I am not able to give any compensations or refunds.</p>	<p>I'm afraid I cannot do anything in terms of compensation for you, but I just checked the policies to see if there are any exceptions. Unfortunately I was correct and I am not able to give any compensations or refunds.</p>
<p>give delivery date</p>	<p>I see here that your order is scheduled to be delivered in 3-4 business days. Is there anything else I can help you with right now?</p>	<p>I see here that your order is scheduled to be delivered in 3-4 business days. Is there anything else I can help you with right now?</p>	<p>I see here that your order is scheduled to be delivered in 3-4 business days. Is there anything else I can help you with right now?</p>	<p>I see here that your order is scheduled to be delivered in 3-4 business days. Is there anything else I can help you with right now?</p>	<p>I see here that your order is scheduled to be delivered in 3-4 business days. Is there anything else I can help you with right now?</p>	<p>I see here that your order is scheduled to be delivered in 3-4 business days. Is there anything else I can help you with right now?</p>	<p>I see here that your order is scheduled to be delivered in 3-4 business days. Is there anything else I can help you with right now?</p>

deny some compensation	Unfortunately no... Since the deliveries are outsourced to another company, our policy is not to give any compensation for late deliveries as they are out of our control. Can I do something else to help?	Unfortunately no... Since the deliveries are outsourced to another company, our policy is not to give any compensation for late deliveries as they are out of our control. Is there anything else you'd like me to check?	Unfortunately no... I'd be disappointed too, but since the deliveries are outsourced to another company, our policy is not to give any compensation for late deliveries as they are out of our control.	I wish I could give a different answer but unfortunately no... Since the deliveries are outsourced to another company, our policy is not to give any compensation for late deliveries as they are out of our control.	I feel sorry for you but unfortunately no... Since the deliveries are outsourced to another company, our policy is not to give any compensation for late deliveries as they are out of our control.	Unfortunately no... I completely understand that this may not be the solution you hoped for, but since the deliveries are outsourced to another company, our policy is not to give any compensation for late deliveries as they are out of our control.	Unfortunately no... Since the deliveries are outsourced to another company, our policy is not to give any compensation for late deliveries as they are out of our control. Perhaps we can find another way to resolve the issue?
deny refund exdel	Unfortunately I cannot do that, the deliveries are done by an external company, so we have the policy not to give a refund for the delivery costs. I understand this may be frustrating. Is there something else I can do for you?	Unfortunately I cannot do that, the deliveries are done by an external company, so we have the policy not to give a refund for the delivery costs. I understand this may be frustrating. Is there something else I can do for you?	I can understand that you were hoping for a different outcome, but unfortunately I cannot do that. The deliveries are done by an external company, so we have the policy not to give a refund for the delivery costs. I can understand this may be frustrating.	Unfortunately I cannot do that, I wish there was something more for me to assist you but since the deliveries are done by an external company, we have the policy not to give a refund for the delivery costs. I understand this may be frustrating.	I sympathize with your situation, but unfortunately, I cannot do that. The deliveries are done by an external company, so we have the policy not to give a refund for the delivery costs. I understand this may be frustrating.	Unfortunately I cannot do that, the deliveries are done by an external company, so we have the policy not to give a refund for the delivery costs. I understand this may be frustrating.	Unfortunately I cannot do that, the deliveries are done by an external company, so we have the policy not to give a refund for the delivery costs. Why don't we consider a different approach that might work better?

deny full refund	The \$PACKAGE is already in transit so I cannot give you a full refund for the order. Of course you may choose to cancel, and I could help with that. Is there anything else I can do for you?	The \$PACKAGE is already in transit so I cannot give you a full refund for the order. Of course you may choose to cancel, and I could help with that. Is there anything else I can do for you?	I can understand that you were expecting a full refund for the order but the \$PACKAGE is already in transit so I cannot give you a full refund for the order. Of course you may choose to cancel, and I could help with that. Is there anything else I can do for you?	I wish there was more we could do to assist you in this situation. However, the \$PACKAGE is already in transit so I cannot give you a full refund for the order. Of course you may choose to cancel, and I could help with that. Is there anything else I can do for you?	I understand that this may be frustrating for you, and I sympathize with your situation but the \$PACKAGE is already in transit so I cannot give you a full refund for the order. Of course you may choose to cancel, and I could help with that. Is there anything else I can do for you?	The \$PACKAGE is already in transit so I cannot give you a full refund for the order. Of course you may choose to cancel, and I could help with that. Is there anything else I can do for you?	The \$PACKAGE is already in transit so I cannot give you a full refund for the order. Of course you may choose to cancel, and I could help with that. Is there anything else I can do for you?
------------------	--	--	---	--	--	--	--

<p>explain solution</p>	<p>Unfortunately there is nothing I can change. I understand that your situation is delicate as it was supposed to be a present, but unfortunately I am not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. I hope for your understanding.</p>	<p>Unfortunately there is nothing I can change. I understand that your situation is delicate as it was supposed to be a present, but unfortunately I am not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. Does this explain it?</p>	<p>Unfortunately there is nothing I can change. I can understand that your situation is delicate as it was supposed to be a present, but unfortunately I am not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. I hope for your understanding.</p>	<p>Unfortunately there is nothing I can change. I understand that this is an important situation for you, and I regret not being able to make exceptions to our policy to help you. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. I hope for your understanding.</p>	<p>Unfortunately there is nothing I can change. I'm sorry to say that we cannot make an exception in this situation, and I understand how important it is to you. But if I do everybody will ask for compensation when the deliveries are delayed, while it is in our policy not to give compensation for things out of our control, like the deliveries. I hope for your understanding.</p>	<p>Unfortunately there is nothing I can change. I agree that it must be a difficult situation to be in, and we apologize for any inconvenience caused as it was supposed to be a present, but unfortunately I am not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. I hope for your understanding.</p>	<p>Unfortunately there is nothing I can change. I understand that your situation is delicate as it was supposed to be a present, but unfortunately I am not allowed to make exceptions. If I do everybody will ask for compensation when the deliveries are delayed, but it is in our policy not to give compensation for things out of our control, like the deliveries. Maybe we can find another solution that meets your needs while still adhering to our policy. I hope for your understanding.</p>
-------------------------	--	---	--	--	--	---	---

deny now	discount	Sorry but I am not allowed to do that. Due to our policy I cannot give you a discount on this delivery. I understand the frustration, but we cannot compensate for things out of our control. Can I do something else for you?	Sorry but I am not allowed to do that. Due to our policy I cannot give you a discount on this delivery. I understand the frustration, but we cannot compensate for things out of our control. Can I do something else for you?	I get that you were hoping for a discount on this delivery and I apologize that I cannot offer one due to our policy. We cannot compensate for things out of our control. Can I do something else for you?	I wish we could offer you a discount on this delivery, but our policy does not allow for it. I understand the frustration, but we cannot compensate for things out of our control. Can I do something else for you?	I'm sorry to say that we are unable to provide a discount on this delivery and I understand how frustrating this may be for you, but we cannot compensate for things out of our control. Can I do something else for you?	I understand your frustration, but unfortunately, I cannot offer a discount on this delivery due to our policy. We cannot compensate for things out of our control. Can I do something else for you?	Sorry but I am not allowed to do that. Due to our policy I cannot give you a discount on this delivery. I understand the frustration, but we cannot compensate for things out of our control. Let's look at other ways that we can help you today?
deny next	discount	I'm sorry but I cannot give you a coupon for a next order. Late deliveries unfortunately happen, but since they are done by an external company we have no control over those. Of course you could always contact the delivery company. Can I help you with anything else today?	I'm sorry but I cannot give you a coupon for a next order. Late deliveries unfortunately happen, but since they are done by an external company we have no control over those. Of course you could always contact the delivery company. Can I help you with anything else today?	I know this is not what you want to hear but I cannot give you a coupon for a next order. Late deliveries unfortunately happen, but since they are done by an external company we have no control over those. Of course you could always contact the delivery company.	I'm sorry but I cannot give you a coupon for a next order. Late deliveries unfortunately happen, but since they are done by an external company we have no control over those. Of course you could always contact the delivery company and hopefully they can further help you.	I understand how important it is to feel appreciated as a customer, so I am sorry that we cannot offer you a coupon for your next order. Late deliveries unfortunately happen, but since they are done by an external company we have no control over those. Of course you could always contact the delivery company.	I get that a coupon would have been helpful, but unfortunately, I cannot offer one at this time. Late deliveries unfortunately happen, but since they are done by an external company we have no control over those. Of course you could always contact the delivery company.	I'm sorry but I cannot give you a coupon for a next order. Late deliveries unfortunately happen, but since they are done by an external company we have no control over those. Of course you could always contact the delivery company. While I am not able to provide you with a coupon for your next order, perhaps there are other ways I can help you?

deny free next	<p>Unfortunately no. We don't do free deliveries, so unfortunately I cannot do that for you. Late deliveries are generally not compensated by our company as they are out of our control. Please let me know if I can help with anything else?</p>	<p>Unfortunately no. We don't do free deliveries, so unfortunately I cannot do that for you. Late deliveries are generally not compensated by our company as they are out of our control. Please let me know if I can help with anything else?</p>	<p>I completely understand that you were hoping for a free delivery, and I'm very sorry that I have to inform you that it is not something I can offer. Late deliveries are generally not compensated by our company as they are out of our control.</p>	<p>Unfortunately no. We don't do free deliveries, so unfortunately I cannot do that for you. Late deliveries are generally not compensated by our company as they are out of our control. I hope for your understanding.</p>	<p>Unfortunately no. We don't do free deliveries, so unfortunately I cannot do that for you. Late deliveries are generally not compensated by our company. I'm sorry to dissappoint you.</p>	<p>Unfortunately no. I'm sorry to say it but we cannot offer you a free delivery, and I do understand that this may be disappointing for you. Late deliveries are generally not compensated by our company since they are out of our control. Please let me know if I can help with anything else?</p>	<p>Unfortunately no. We don't do free deliveries, so unfortunately I cannot do that for you. Late deliveries are generally not compensated by our company as they are out of our control. Please let me know if I can perhaps help with anything else?</p>
----------------	--	--	--	--	--	--	--

confirm cancel	I'm sorry to hear that you want to cancel \$PERSON. I will cancel your order \$ORDER NUMBER immediately, and the \$PACKAGE will be sent back to us directly. The money of your purchase will be back in your account in 3-5 business days. Unfortunately, we do not refund the delivery costs for returns, as stated in our terms and conditions. Please let me know if I can help you with anything else?	I'm sorry to hear that you want to cancel \$PERSON. I will cancel your order \$ORDER NUMBER immediately, and the \$PACKAGE will be sent back to us directly. The money of your purchase will be back in your account in 3-5 business days. Unfortunately, we do not refund the delivery costs for returns, as stated in our terms and conditions. Please let me know if I can help you with anything else?	I'm sorry to hear that you want to cancel \$PERSON. I will cancel your order \$ORDER NUMBER immediately, and the \$PACKAGE will be sent back to us directly. The money of your purchase will be back in your account in 3-5 business days. Unfortunately, we do not refund the delivery costs for returns, as stated in our terms and conditions. Please let me know if I can help you with anything else?	I'm sorry to hear that you want to cancel \$PERSON. I will cancel your order \$ORDER NUMBER immediately, and the \$PACKAGE will be sent back to us directly. The money of your purchase will be back in your account in 3-5 business days. Unfortunately, we do not refund the delivery costs for returns, as stated in our terms and conditions. Please let me know if I can help you with anything else?	I'm sorry to hear that you want to cancel \$PERSON. I will cancel your order \$ORDER NUMBER immediately, and the \$PACKAGE will be sent back to us directly. The money of your purchase will be back in your account in 3-5 business days. Unfortunately, we do not refund the delivery costs for returns, as stated in our terms and conditions. Please let me know if I can help you with anything else?	I'm sorry to hear that you want to cancel \$PERSON. I will cancel your order \$ORDER NUMBER immediately, and the \$PACKAGE will be sent back to us directly. The money of your purchase will be back in your account in 3-5 business days. Unfortunately, we do not refund the delivery costs for returns, as stated in our terms and conditions. Please let me know if I can help you with anything else?	I'm sorry to hear that you want to cancel \$PERSON. I will cancel your order \$ORDER NUMBER immediately, and the \$PACKAGE will be sent back to us directly. The money of your purchase will be back in your account in 3-5 business days. Unfortunately, we do not refund the delivery costs for returns, as stated in our terms and conditions. Please let me know if I can help you with anything else?
----------------	--	--	--	--	--	--	--

response no customer	I'm sorry to hear that we disappointed you \$PERSON. Of course I hope we will see you in the future and show you the service you expect from us. I hope you have a good day, and please let me know if I can help you with anything else?	I'm sorry to hear that we disappointed you \$PERSON. Of course I hope we will see you in the future and show you the service you expect from us. I hope you have a good day, and please let me know if I can help you with anything else?	I can understand that you were disappointed with our service \$PERSON, and I want to assure you that we take customer satisfaction seriously. Of course I hope we will see you in the future and show you the service you expect from us. I hope you have a good day, and please let me know if I can help you with anything else.	I'm sorry to hear that we disappointed you \$PERSON. However, I hope we get the opportunity to provide you with the exceptional service that you expect from us. I hope you have a good day, and please let me know if I can help you with anything else.	I'm sorry to hear that we disappointed you \$PERSON. Of course I hope we will see you in the future and show you the service you expect from us. I hope you have a good day, and please let me know if I can help you with anything else.	I'm sorry to hear that I disappointed you \$PERSON, I understand how frustrating this process can be. Of course I hope we will see you in the future and show you the service you expect from us. I hope you have a good day, and please let me know if I can help you with anything else.	I'm sorry to hear that I disappointed you \$PERSON. Perhaps I will see you in the future and show you the service you expect from us. I hope you have a good day, and please let me know if I can help you with anything else.
forward complaint	We are already in contact with them, but I'll be sure to forward your complaint as well. Is there anything else I can do for you?	We are already in contact with them, but I'll be sure to forward your complaint as well. Is there anything else I can do for you?	We are already in contact with them, but I'll be sure to forward your complaint as well. Is there anything else I can do for you?	We are already in contact with them, but I'll be sure to forward your complaint as well. Is there anything else I can do for you?	We are already in contact with them, but I'll be sure to forward your complaint as well. Is there anything else I can do for you?	We are already in contact with them, but I'll be sure to forward your complaint as well. Is there anything else I can do for you?	We are already in contact with them, but I'll be sure to forward your complaint as well. Is there anything else I can do for you?

response bad review	I'm sorry to hear that \$PERSON. I hoped we could resolve the issue together and I'm sorry that our service was not what you expected. If there is anything else I can do for you please let me know.	I'm sorry to hear that \$PERSON. I hoped we could resolve the issue together and I'm sorry that our service was not what you expected. Can I help with something else today?	I'm sorry that we did not meet your expectations \$PERSON, I would be \$EMOTION too and I would love the opportunity to make things right. If there is anything else I can do for you please let me know.	I'm sorry that we fell short of your expectations \$PERSON, and I sincerely hope that we can earn back your trust in the future. If there is anything else I can do for you please let me know.	I'm sorry to hear that our service did not meet your expectations \$PERSON, I understand how disappointing that can be. If there is anything else I can do for you please let me know.	I am sorry to hear that we fell short of your expectations \$PERSON, and I appreciate your feedback. If there is anything else I can do for you please let me know.	I'm sorry to hear that \$PERSON. I hoped we could resolve the issue together and I'm sorry that our service was not what you expected. If there is anything else I can do for you please let me know.
future plans	Since this is a responsibility of the delivery company, we cannot physically expedite the delivery in the future. However, we can improve our communication, so you'll be informed sooner if there are any delays.	Since this is a responsibility of the delivery company, we cannot physically expedite the delivery in the future. However, we can improve our communication, so you'll be informed sooner if there are any delays.	Since this is a responsibility of the delivery company, we cannot physically expedite the delivery in the future. However, we can improve our communication, so you'll be informed sooner if there are any delays.	Since this is a responsibility of the delivery company, we cannot physically expedite the delivery in the future. However, we can improve our communication, so you'll be informed sooner if there are any delays.	Since this is a responsibility of the delivery company, we cannot physically expedite the delivery in the future. However, we can improve our communication, so you'll be informed sooner if there are any delays.	Since this is a responsibility of the delivery company, we cannot physically expedite the delivery in the future. However, we can improve our communication, so you'll be informed sooner if there are any delays.	Since this is a responsibility of the delivery company, we cannot physically expedite the delivery in the future. However, we can improve our communication, so you'll be informed sooner if there are any delays.

recap solution	Okay, then to recap; I will keep you updated on the \$PACKAGE. And we will work on improving our communication about delayed packages in the future. Is that okay, or do you want to talk about something else?	Okay, then I will keep you updated on the \$PACKAGE and let you know when it is in transit. Can I help you with anything else today?	Okay, then I will keep you updated on the \$PACKAGE and let you know when it is in transit. Can I help you with anything else today?	Okay, then I will keep you updated on the \$PACKAGE and let you know when it is in transit. Can I help you with anything else today?	Okay, then I will keep you updated on the \$PACKAGE and let you know when it is in transit. Can I help you with anything else today?	Okay, then I will keep you updated on the \$PACKAGE and let you know when it is in transit. Can I help you with anything else today?	Okay, then I will keep you updated on the \$PACKAGE and let you know when it is in transit. Can I help you with anything else today?
goodbye	Okay, please feel free to contact me when you come up with any other questions. Again, my apologies for the delay of the \$PACKAGE. I hope you have a great day \$PERSON!	Okay, please feel free to contact me when you come up with any other questions. Again, my apologies for the delay of the \$PACKAGE. I hope you have a great day \$PERSON!	Okay, please feel free to contact me when you come up with any other questions. Again, my apologies for the delay of the \$PACKAGE. I hope you have a great day \$PERSON!	Okay, please feel free to contact me when you come up with any other questions. Again, my apologies for the delay of the \$PACKAGE. I hope you have a great day \$PERSON!	Okay, please feel free to contact me when you come up with any other questions. Again, my apologies for the delay of the \$PACKAGE. I hope you have a great day \$PERSON!	Okay, please feel free to contact me when you come up with any other questions. Again, my apologies for the delay of the \$PACKAGE. I hope you have a great day \$PERSON!	Okay, please feel free to contact me when you come up with any other questions. Again, my apologies for the delay of the \$PACKAGE. I hope you have a great day \$PERSON!