Influence of empathy on human-agent coexistence relationship

Takahiro Tsumura

Doctor of Philosophy

Department of Informatics

School of Multidisciplinary Sciences

The Graduate University for Advanced Studies, SOKENDAI

March 2024

A dissertation submitted to Department of Informatics,
School of Multidisciplinary Sciences,
The Graduate University for Advanced Studies, SOKENDAI,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Advisory Committee

1. Prof. Seiji Yamada
   National Institute of Informatics,
   The Graduate University for Advanced Studies, SOKENDAI

2. Prof. Akiko Aizawa
   National Institute of Informatics,
   The Graduate University for Advanced Studies, SOKENDAI

3. Prof. Hideaki Takeda
   National Institute of Informatics,
   The Graduate University for Advanced Studies, SOKENDAI

4. Prof. Tetsunari Inamura
   Tamagawa University

5. Prof. Tetsuo Ono
   Hokkaido University

Acknowledgements

I would like to express my deepest gratitude to my advisor, Professor Seiji Yamada, for his guidance, patience, and interest in my research concept. I am very honored to be a doctoral student at this university. I could not have reached this point without his generous nature and enthusiasm for student guidance, which pushed me daily.

I would like to thank my thesis reviewers, Professor Akiko Aizawa, Professor Hideaki Takeda, Professor Tetsunari Inamura, and Professor Tetsuo Ono, for their constructive suggestions and valuable comments, which helped me improve my research.

Dr. Kazuo Okamura, Dr. Ryo Nakahashi, Dr. Takato Okudo, Dr. Shin Sano, Dr. Yosuke Fukuchi, Nungduk Yun, Chenlin Hang, Jingbo Yan, Sota Kaneko, Keito Miyake, Kazuki Komura, Mari Saito and the other members of Yamada Laboratory. Their student life has been a source of cheerful and enjoyable research activities even in the face of the decline of face-to-face meetings by COVID-19, and I greatly appreciate their good advice and comments on my research.

Finally, special thanks go to my family for their support during this research.

The Graduate University for Advanced Studies, SOKENDAI

Abstract

School of Multidisciplinary Sciences
Department of Informatics

Doctor of Philosophy

Influence of empathy on human-agent coexistence relationship

Takahiro Tsumura

Many technologies are being used in numerous situations, including ChatGPT, generative AI, factory robots, cleaning robots, nursing robots, and anthropomorphic agents for customer service and chatbot support. In addition, the global environment is becoming more accepting of anthropomorphic agents with the development of businesses using 3D avatars and VR technology. On the other hand, as the use of AI technology and anthropomorphic agents in society becomes more widespread, people may feel distrust and rejection toward these technologies, which may lead to negative behaviors. Therefore, it is an important issue to establish a coexistence relationship with AI technologies and anthropomorphic agents used in future society. As a way to solve this problem, an appropriate relationship can be established when people feel empathy and trust toward anthropomorphic agents. It has been discussed in the fields of psychology and philosophy that these are important factors in human relationships, and we believe that they have a significant impact on the construction of relationships between people and anthropomorphic agents. Therefore, we focused on empathy between people and anthropomorphic agents. In particular, we investigated the factors that make people empathize with anthropomorphic agents. Below is a summary of the three experiments.

In the first study, we focused on self-disclosure from agents to humans to increase human empathy toward anthropomorphic agents. We experimentally investigated the possibility that self-disclosure from agents promotes human empathy. Two hypotheses were formulated, and the conditions under which humans are more empathetic toward agents were experimentally analyzed and discussed. The experiment was conducted in a three-way mixed design, and the factors were agent appearance (human, robot), self-disclosure (highly relevant self-disclosure, less relevant self-disclosure, no self-disclosure), and empathy before and after video stimuli. An ANOVA was conducted using data from 918 participants. The results showed that there was no main effect for the emergent factor and that self-disclosure, which was highly relevant to the scenario used, promoted more human empathy by a statistically significant margin. We also found that self-disclosure did not suppress empathy. These results support our hypothesis. This study reveals that agents' self-disclosure represents an important feature of anthropomorphic agents that helps humans accept them.

The next study focused on tasks between agents and humans. In particular, we experimentally examined the hypothesis that task difficulty and task content promote human empathy. The experiment was a two-way ANOVA with four conditions: task difficulty (high, low) and task content (conflict, cooperation). An analysis of variance (ANOVA) was conducted with data from 578 participants. Results showed no main effect of the task content factor and a significant main effect of the task difficulty factor. In addition, pre-task empathy for the agent decreased after the task. ANOVAs showed that one category of empathy for the agent increased more when the task difficulty was higher than when the task difficulty was lower. This indicated that this category of empathy was likely to be affected by the task. The task used itself was likely to be an important factor when manipulating each category of empathy.

As a final study, we focused on tasks in which humans and agents interacted

in a variety of ways to investigate agent characteristics that significantly influence human empathy toward the agent. Experiments were conducted to examine the effects of task completion conditions (success, failure) and agent representation (yes, no) on human empathy. An ANOVA was conducted using data from 363 participants. The results of the experiment showed that human empathy toward the agent is difficult to maintain in the task completion condition alone, and that agent representation can maintain human empathy. These results indicate that the characteristics of AI agents play an important role in helping people accept anthropomorphic agents.

Based on the results of the above three studies, this paper discusses an empathic approach to help anthropomorphic agents gain acceptance from people and proposes the design of a relationship in which agents working with people can coexist and co-prosper with people through empathic agents in the future.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Research motivation

Due to the development of technology, people have in the past been able to share roles among humans and perform tasks suited to each other according to their respective abilities. In recent years, many technologies have been used in place of people in many situations in human society, including ChatGPT, generative AI, factory robots, cleaning robots, nursing robots, and anthropomorphic agents for customer service and chatbot support. In addition, an environment that accepts anthropomorphic agents is being created globally, with businesses utilizing 3D avatars and VR technology.

Since people have started to create and use artificial objects that take the place of people (expressed as ”agents” below) in society, problems that used to occur between humans may now occur between people and agents as well. As AI technologies and anthropomorphic agents become more prevalent in

society, people are distrustful and reject these technologies, leading to negative behavior on the part of some. Mori [1] proposed an uncanny valley of human emotional responses to robots, in which robots become more likeable and empathetic as they are made more human-like in their appearance and behavior, but at some point suddenly turn into strong dislike. As ever-developing technology acquires human-like capabilities, the situation becomes similar to the uncanny valley, leading to negative behavior. With respect to the uncanny valley hypothesis, Thepsoonthorn et al. [2] explored the uncanny valley in terms of the robot's nonverbal behavior. The results showed a two-correlation coefficient between the results of human-like and affinity ratings. A curve similar to the uncanny valley was found. A curve similar to the uncanny valley was suggested for nonverbal behavior by them, indicating that nonverbal behavior that is more human-like is thought negatively by people.

Mahmud et al. [3] also conducted a literature review of algorithm aversion, which is a negative behavior or attitude toward the workings of algorithms compared to the workings of humans, and has gained attention through the use of ChatGPT and generative AI. In particular, Filiz et al. [4] examined the relationship between the outcome of a decision-making situation and the frequency of algorithm aversion and showed that the more serious the outcome of a decision, the more frequently algorithm aversion occurred. Especially in the case of very important decisions, algorithmic aversion leads to a reduced probability of success, which is the tragedy of algorithmic aversion. For causes that lead to the negative behavior described above, it is an important issue to establish a coexistence relationship with AI technologies and anthropomorphic agents that will be used in the future society.

The purpose of our research is to design a relationship between people and agents to promote coexistence between people and agents, and to consider approaches to make agents used in a wide range of fields in the future more acceptable to human society. Specifically, the way to solve this problem is for people to feel empathy and trust toward anthropomorphic agents so

that appropriate relationships can be established. It has been discussed in the fields of psychology and philosophy that these are important elements in human relationships, and we believe that they have a significant impact on the construction of relationships between people and anthropomorphic agents. Therefore, we focused on empathy between people and anthropomorphic agents. In particular, we investigated the factors that make people empathize with anthropomorphic agents. Below we present and discuss the results of the influence of empathy between humans and agents from three studies we conducted.

1.2 Research background

Anthropomorphic agents are increasingly being utilized in human society. These can be considered a kind of tool used by humans, and historically, tools have sometimes been treated as if they are actually sentient. For example, humans have treated artificial objects as though they were human in media equations [5]. However, not all of the artifacts in use today are perceived in this way. The main problems currently facing AI implementation are related to reliability and ethical usage. One recent study on trust, AI ethics, and the anthropomorphizing of AI argued that even complex machines should not be viewed as trustworthy [6]; instead, people should ensure that the organizations using AI and the individuals within those organizations are trustworthy. AI ethics was also discussed in depth from an applied-ethics perspective in a study by Hallamaa and Kalliokoski [7].

At the same time as trust, we often empathize with artificial objects. Empathy for artifacts can occur when humans treat them as if they were alive. Artificial objects that we empathize with include cleaning robots, pet-type robots, and anthropomorphic agents that provide services such as online shopping and help desk assistance. These are already used in society and coexist

with humans. The appearance of these agents varies depending on the application and environment.

Chin et al. [8] used three types of verbal abuse (insults, threats, and swear words) and three types of response styles (avoidance, empathy, and counter-attack) to examine whether the response styles of conversational agents under various types of verbal abuse affect the emotions found to mitigate aggressive behavior in people. Mathur et al. [9] introduced the first approach to model user empathy elicited during interaction with a robotic agent. They experimented with a classical machine learning model and two deep learning models to detect empathy by exploiting patterns of visual behavior while the participant listens to the robot's narrative. Inoue et al. [10] focused on "shared laughter," in which a spoken dialogue system expressing empathy does not generate laughter from the user's interaction with the human user, but rather the user laughs to himself or herself and the system laughs in turn. Al Farisi et al. [11] applied two designs of Anthropomorphic Design Cues (ADC), verbal and nonverbal, to chatbots and compared the empathy of two chatbots with and without ADC. To explore the factors that induce human empathy toward virtual agents, Morita and Kano [12] examined human-agent relationships through a reverse cyberball game in which participants can help an agent who has been ostracized by their peers. In particular, They attempted to distinguish factors of self-experience as the basis of empathy from situational awareness.

Nowadays, various agents contribute to society together with humans, and their relationship is sometimes akin to that of colleagues. It is important for the future development of human society that these agents establish relationships that are harmonious with humans. Although robots are already used in society and coexist with us to some extent, some people simply cannot accept this type of agent [13–15]. As anthropomorphic agents are predicted to be utilized in human society even more in the future, it is vital that they exhibit characteristics that are more acceptable to all humans.

In our study, we use a target empathy agent to facilitate human empathy. We focused on how humans improve their relationships with agents. One method is to have humans empathize with the agents. By empathizing, humans act positively toward agents and are more likely to accept them. Various studies have been done on linguistic information, nonverbal information, situations, and relationships as factors that cause empathy.

As for verbal information eliciting empathy, Konrath et al. [16] showed that empathy-building messages increased affective measures of empathy and prosocial behavior compared to control messages, but decreased self-perceptions of empathy. Shaffer et al. [17] showed that the narrative writing intervention increased participants' empathy and perspective-taking, evoked more positive attitudes toward women who smoke during pregnancy, and increased external attributions for her behavior. Tahara et al. [18] proposed an empathic dialogue system for open-domain conversations based on psychological research on empathy.

As nonverbal information to induce empathy, Yoshioka and Takeuchi [19] used a simple disk-shaped robot that only moves on the floor surface to express emotional states by movement and movement parameters based on Russell's circumplex model, and examined how humans infer emotional states from parameter changes in simple movements. We investigated how humans infer emotional states from changes in simple movement parameters. Okanda et al. [20] investigated whether Japanese adults' friendship and moral beliefs toward robots with different appearances (human, dog, and egg-shaped) were related to animistic tendencies and empathy.

In terms of situational information eliciting empathy, O'Connell et al. [21] hypothesized that individuals with high trait empathy would simulate more readily and thus discount temporally less for distant others than those with

low trait empathy. To test this prediction, they performed a temporal discounting task from the perspectives of close and distant others and themselves. Zhi et al. [22] explored whether robots can use social influence to motivate nearby bystanders and protect them from human abuse. Empathic robots had a higher percentage of participants who intervened than indifferent robots, but more participants moved the robot as a response to abuse. To teach healthy study habits and tips, Richards et al. [23] reported a study of 239 college students, 161 of whom interacted with a character designed to express empathy through dialogue.

As for empathy-inducing relational information, Stephan [24] presented the requirements that an artificial empathizer must meet to be judged as fully empathic by his or her social interactors, even when acting in an unspecified social context. Hosseinpanah et al. [25] investigated the role of age in the perception of emotional and nonverbal behavior by virtual assistants. The study investigated the relationship between humans and assistants by examining nonverbal behavior in terms of age, conversation content, and content. Giannopulu et al. [26] used an actor-perceiver paradigm in which the child was the actor and the robot or human was the perceiver to record autonomic heart rate activation and reported emotional affect in Japanese and French children with ASD. The Japanese and French ASD children showed different interpersonal attunement when interacting with the perceiver, despite being the same person.

As many AI and robotic technologies, including ChatGPT and generative AI, are used in society, the purpose of the research is to help people empathize with them and accept the technologies more easily. On the other hand, too much empathy toward them may lead to overconfidence in the technology and the possibility of using the technology without fully understanding it, so the approach to establishing an appropriate relationship should be carefully considered.

In previous studies, self-disclosure is also regarded as important among humans. Therefore, we thought that self-disclosure would be necessary for anthropomorphic agents to establish relationships with humans. In recent years, agents have been used more and more in human society, and building relationships between humans and agents has become increasingly important. Just as humans have built relationships with each other through self-disclosure, it makes sense for agents to use self-disclose in order to deepen their relationships with humans. However, anthropomorphic agents can be pre-designed and operated in society, so designing conditions for self-disclosure that are appropriate to the situation is an important factor in improving relationships. We focused on self-disclosure and experimentally examined what types of self-disclosure affect the characteristics of empathy. Moreover, empathy has been studied in the fields of HAI and HRI. However, different appearances have been used in each study. Hence, we decided to set appearance as a factor to compare the effect that it has on empathy. For this, human-like and robot-like appearances were prepared and tested as symbols of the HAI and HRI fields.

In our first study, we assume that empathy agents influence human empathy. However, we investigate only human empathy toward empathy agents, not the human capacity for empathy, since our focus is on investigating the factors that make agents acceptable to humans. Below, empathy in this study refers to this kind of empathy.

Considering its relevance in previous studies on self-disclosure and empathy, we select self-disclosure as a factor to investigate human empathy toward agents. In this study, we designed agent self-disclosure to investigate what happens when an agent self-discloses. We confirm whether the self-disclosure is perceived by humans as agent self-disclosure in a pre-experiment and investigate whether the designed self-disclosure affects human empathy. We focus on self-disclosure from agents to humans, and we conduct experiments to investigate the relationship between human empathy and agent self-disclosure,

as well as self-disclosure that is effective in promoting the characteristics of empathy. At the same time, we investigate the relationship between anthropomorphic agents with different appearances and self-disclosure.

In our second and third studies, we focus on how to help humans improve their relationships with agents. To this end, we developed an empathy agent to induce human empathy, where "empathy" is an attribute that the agent should have in order to be accepted by human society. Our concept is that, by empathizing, people will act more positively toward agents and be more likely to accept them.

In our approach, we assume that the empathy agent itself is what influences human empathy. Our main focus is the factors that make agents acceptable to humans, so we investigate only human empathy toward agents, and consider the human capacity for empathy to be outside the current scope.

To this end, we perform two experiments to examine the impact of the task on human empathy for the agent and to investigate whether the agent changes human behavior. The purpose of this study is to investigate designs for AI agents to improve their relationships with humans. Therefore, this study does not use AI because it does not aim to develop agents that learn. Without an AI dialogue system in some products, investigating designs that are more acceptable to humans is a necessary component of future AI use. This is an important study that investigates the influence of factors to improve the human-agent relationship.

1.3 Thesis Overview

This paper discusses the background of empathy research at first, as well as studies related to empathy and the factors investigated in each of the studies. Then, as the first study, we report the results of an investigation of empathy focusing on the agent's appearance and self-disclosure as properties

of the agent, based on subjective indicators from a questionnaire survey. As in the first study, we report the results based on subjective indices obtained from a questionnaire survey. Finally, the third study, based on the first and second studies, investigated the influence of the agent's representations on empathy during the task with the agent. We designed factors and conducted experiments considering the possibility that empathy toward the agent may change depending on the success or failure of the task. As in previous studies, the results are reported based on subjective indices from a questionnaire survey.

The concept of empathy in this study is the capacity of people to empathize with agents. To investigate the effective factors for this empathy, three studies were conducted, and based on the results, the necessary design for an empathy agent discussed. Therefore, there is no need for the agent to have the capacity for empathy, and the agent does not actively engage in empathic behavior. These studies focused on how empathy affects the design that makes it easier for people to accept the agent and on changes in the relationship between people and the agent when performing the task together. Our research contributes to social agent research and is intended to serve as a basis for future designs of agents that coexist with people.

Chapter 2

Related Work

2.1 Definition of empathy

We consider empathy to be a vital element in the human acceptance of agents as members of society. This mirrors the importance of humans empathizing with each other in order to get along [27, 28]. Empathy and the influence it has on others has been studied extensively in the field of psychology. Omdahl [29] classified empathy into three types: (1) affective empathy, which is an emotional response to the emotional state of others, (2) cognitive empathy, which is a cognitive understanding of the emotional states of others, and (3) empathy that includes both of the previous types. Preston and de Waal [30] suggested that at the core of the empathic response is a mechanism that enables the observer to access the subjective emotional state of the target. They came up a perception-action model (PAM) to unify the different perspectives in empathy and classified empathy into three types: (a) sharing or being influenced by the emotional state of others, (b) assessing the reasons

for the emotional state, and (c) having the ability to identify and incorporate other perspectives. Olderbak et al. [31] examined theoretical and empirical support for the emotional specificity of empathy and proposed an emotion-specific empathy questionnaire that assesses affective and cognitive empathy for six basic emotions. Tisseron et al. [32] defined a theoretical model of empathy consisting of four dimensions and four components, corresponding to human-human, human-avatar, and human-robot interactions, to better understand the specificity of empathy that humans might have for robots.

Our study focuses on the effect of empathy in promoting society's acceptance of anthropomorphic agents, but in psychology, empathy has been explored from a variety of other aspects, including its negative effects. Bloom tried to introduce a neutral aspect of empathy that encompasses not only positive influences but also negative ones [33]. He claimed that when empathy acts as a moral trigger, it can lead to humans making irrational decisions and to relationships that erupt into violence and anger. He also suggested that we can overcome this negative aspect by combining conscious and deliberative reasoning with an altruistic approach.

Various questionnaires have been utilized to measure empathy. The Ten Item Personality Inventory (TIPI) is one of the more popular ones, but since it investigates human personality as a whole [34], it is unsuitable for our purposes because we presume empathy is biased by human personality. The Interpersonal Reactivity Index (IRI), also commonly used in the field of psychology, is designed to investigate the specific characteristics of empathy [35]. Baron-Cohen and Wheelwright [36] proposed a self-report questionnaire called the Empathy Quotient (EQ) for use with adults of normal intelligence. Lawrence et al. [37] investigated the reliability and validity of the EQ and found that there is a moderate association between the EQ subscale and the IRI subscale.

Regarding questionnaires about empathy, IRI and EQ are widely used in the field of psychology. In particular, examples of research on empathy using only

the IRI are widely seen in psychology and the HRI and HAI fields, such as the studies by Konrath et al., Shaffer et al., and Perugia et al. Since the results of previous studies have shown that the EQ scale is related to the IRI scale, we used the IRI questionnaire, which has fewer questions and allows for the four characteristics of empathy to be investigated. We also used the widely used IRI for comparison with previous and future research on empathy. We also focused on investigating the impact of each characteristic of empathy. For investigating this, the use of the IRI was appropriate as based on previous studies. We ultimately opted to use only IRI as the questionnaire for our experiment.

2.2 Empathy in engineering

Empathy has also been studied in the field of engineering, particularly in the context of virtual reality. Bertrand et al. [38] proposed a theoretical analysis of various mechanisms of empathy practice to define a possible framework for the design of empathy training in virtual reality. A study by van Loon et al. [39] investigated whether the effects of VR perspective-taking could be driven by increased empathy and extended to real-stakes behavioral games. They succeeded in increasing the tendency of participants to take the other person's point of view, but only if it was that of the same person participants assumed in the virtual reality simulation. Herrera et al. [40] compared the short- and long-term effects of traditional and VR viewpoint acquisition tasks. They also conducted experiments investigating the role of technological immersion with respect to different types of intermediaries.

Curran et al. [41] investigated empathy by showing a video from the visual perspective of a person watching a virtual reality movie. Crone and Kallen [42] utilized online platforms and immersive virtual reality to examine the role of virtual perspective-taking on binary gender. They found that virtual reality-based perspective-taking may have a greater impact on acute

behavioral modulation of gender bias compared to online because it immerses participants in the experience of temporarily becoming another.

Empathy is also attracting attention in the realm of product design. Bennett and Rosner [43] investigated a human-centered design process (promise of empathy) in which designers try to understand the target user with the aim of informing technology development. Rahmanti et al. [44] designed a chatbot with artificial empathic motivational support for dieting called "SlimMe" and investigated how people responded to the diet bot. They proposed a text-based emotional analysis that simulates artificial empathic responses to enable the bot to recognize users' emotions.

In the fields of human-computer interaction (HCI), empathy between humans and agents or robots is studied. Wright and McCarthy [45] discussed the use of empathy, citing studies that have used empathy in HCI. Chella et al. [46] discussed self-awareness and inner speech in humans and AI agents and provided an initial proposal for a cognitive architecture for implementing inner speech in robots. While the foundations of internal speech had been investigated primarily in the fields of psychology and philosophy, research in robotics had not yet addressed self-aware behavior. Therefore, after discussing self-awareness and inner speech in humans and AI agents, they proposed the above cognitive architecture. Their approach had an advantage in that a robot's inner speech could be heard by an external observer, and introspective and self-regulated speech could be detected. Pratte et al. [47] analyzed 26 publications on empathy tools and developed a framework for empathy tool designers.

2.3 Empathy in human-robot interaction

Studies on human empathy in the human-robot interaction (HRI) field have explored the ways in which humans empathize with artificial objects. Beck et

al. [48] investigated the effect of changing a robot's head position on the interpretation of emotional key poses, valence, arousal, and stances. Their findings support the idea that body language is an appropriate medium for robots to express emotions. Hofree et al. [49] showed that human participants spontaneously matched the facial expressions of androids in the same room. This mimicry made the participant feel uncomfortable toward the android, even though the participant was fully aware of the android's lack of ill intent. This result suggests that mimetic responses depend on the prominence of human-like features emphasized by face-to-face interaction, thus highlighting the role of presence in human-robot interactions.

On the basis of the concept of cognitive developmental robotics, Asada [50] proposed "affective developmental robotics" as a way to produce more authentic artificial empathy. Artificial empathy here refers to AI systems (e.g., companion robots and virtual agents) that can detect emotions and respond empathetically. The design of artificial empathy is one of the most essential components of social robotics, and empathetic interaction with the public is necessary to introduce robots into society. Dumouchel et al.[51] also summarized artificial empathy. The relationship between humans and robots appearing in daily life was discussed. They suggested that the human-robot dynamics in emotional relationships need to be considered. Mollahosseini et al.[52] applied a deep neural network-based system for automatically recognizing facial expressions to the speech dialogue of social robots. The function was extended and enhanced beyond voice dialogue to integrate the user's emotional state into the robot's reaction.

Fraune [53] examined how people behave morally and perceive players according to their group membership (in-group, out-group), agent type (human, robot), and robot anthropomorphism (anthropomorphic, mechanized). Their results showed that the pattern of reactions to humans was more favorable for anthropomorphic robots than for mechanistic robots. Tuyen et al. [54]

designed emotional bodily expressions for a robot that adopts the user's emotional gestures. They proposed an action selection and transformation model that enables the robot to progressively learn from the user's gestures, identify the user's habitual behaviors, and transform the selected behaviors into robot actions. Their findings showed that the robot's emotional expressions reflecting the characteristics of the partner were widely accepted within the same cultural group and perceived in different ways among different cultural groups.

Several studies on inner speech have been conducted in the HRI field. Pipitone and Chella[55] investigated the potential of considering the inner speech of robots that cooperate with human partners. A domestic situation requiring several functional and moral requirements was simulated in a simple cooperative task. Their study was a novel endeavor, as only a few papers have analyzed the role of inner speech in robots, and most of them were theoretical in nature. Geraci et al.[56] investigated whether a robot's internal conversation affects human trust and anthropomorphism when humans and robots collaborate together. The results suggest that a robot's speech affects human trust. The results also indicated that participants' perceptions of trust and anthropomorphism toward the robot improved after interacting with the robot in the experiment.

2.4 Empathy in human-agent interaction

As for empathy research in the field of human-agent interaction (HAI), McQuiggan et al.[57] proposed a unified inductive framework for modeling parallel and reactive empathy, which empathy models by choosing appropriate parallel or reactive empathy expressions. The framework was used to facilitate empathic behavior suitable for run-time situations. Leite et al. [58] conducted a long-term study in an elementary school where they presented and evaluated an empathy model for social robots aimed at interactions with

children that occur over a long period of time. They measured children's perceptions of social presence, engagement, and social support and found that the empathy model developed had a positive impact on the long-term interaction between the child and the robot. Deshmukh et al. [59] analyzed the relationship between the way robots gesture and the way those gestures are perceived by human users. In particular, they investigated how changing the amplitude and speed of a gesture affects the Godspeed score given to that gesture. Their results suggested that forming gestures aimed at making explicit the internal state of the robot tended to change the perception of animacy, while forming gestures aimed at interaction effects tended to change the perception of anthropomorphism, desirability, and perceived safety.

Chen and Wang [60] hypothesized that empathy and anti-empathy were closely related to a creature's inertial impression of coexistence and competition within a group and established a unified model of empathy and anti-empathy. They also presented the Adaptive Empathetic Learner (AEL), an agent training method that enables evaluation and learning procedures for emotional utilities in a multi-agent system. Perugia et al. [61] investigated which personality and empathy traits were related to facial mimicry between humans and artificial agents. They focused on the humanness and embodiment of agents and the influence that these have on human facial mimicry. Their findings showed that mimicry was affected by the embodiment that an agent has, but not by its humanness. It was also correlated with both individual traits indicating sociability and empathy and traits favoring emotion recognition.

To clarify the empathy between agents/robots and humans, Paiva represented the empathy and behavior of empathetic agents (called empathy agents in HAI and HRI studies) in two different ways: targeting empathy and empathizing with observers [62–64]. In our study, following Paiva's proposed definition of an empathic agent, we consider the agent as an object of empathy and examine how the empathy and empathic responses of the human participants are affected.

2.5 Self-Disclosure in Psychology

Self-disclosure has also been a focus of research in the field of psychology. Jourard[65] presented the Jourard Self-Disclosure Questionnaire (JSDQ), a self-disclosure classification and questionnaire. Attitudes, opinions, interests, study and work, personality, economy, and body were listed as categories. Carpenter and Freese[66] measured self-presentation intimacy and internality, Derlega and Berg[67] focused on the association between responsiveness and self-disclosure, and Laurenceau JP[68] suggested that both self-disclosure and partner responsiveness contribute to the experience of intimacy in interactions.

One study related to self-disclosure is the study of inner speech. Morin[69] reviewed past and current literature on the link between self-awareness and inner speech. Among multidimensional views of self-knowledge, he showed that inner speech accounts for half of the linkages between various elements and plays a fundamental role. In addition, Morin[70] further studied internal speech. He considered inner speech as creating psychological distance between the self and the mental events experienced by the self, that the self represents a problem, that self-information functions as a problem-solving device for resolution, and that it is possible to label internal aspects of the self that are otherwise difficult to recognize objectively. We emphasize that inner speech and imagined interactions (IIs) are not identical and differ in important ways. Therefore, although IIs and inner speech intersect, their overlap is quite limited, so it is possible to investigate one over the other.

Lockwood et al.[71] used self-reported measurements of empathy and apathy motivation in a large sample of healthy people to test whether more empathic people were more motivated. The actual self-disclosure reflected in interpersonal relationships has been investigated in a few studies. Therefore, Kreiner and Levi-Belz[72] designed new objective and dynamic measurements to evaluate self-disclosure and stable self-disclosure characteristics. Oh Kruzic

et al.[73] focused on how the face and upper-body nonverbal channels contribute individually and collaboratively via avatars in virtual environments. Lee et al.[74] found that including self-disclosure from chatbots when they interacted with humans had an effect on improving participants' perceptions of intimacy and enjoyment. Pan et al.[75] examined the effect of exposure to online support-seeking posts containing different levels of self-disclosure depth (baseline, peripheral, core) affecting the quality (person-centeredness and politeness) of participants' messages providing support.

2.6 Prior research on the task

A study of cooperative and competitive tasks was conducted by Ruissen and de Bruijn [76]. In the study, cooperative and competitive tasks were tested using Tetris. The results confirmed that the cooperative task did not reduce self-integration, but the competitive task did. Another study of competitive tasks between humans and robots is that of Kshirsagar et al. [77]. They performed a human-robot competitive task using the same task and found that participants preferred a lower-performing robot to a higher-performing one. Boucher et al. [78] performed a human-robot cooperation task. The robot recognized gaze guidance to the human faster than the robot gave voice instructions to the human.

Some studies of task difficulty include the following. Fuentes-García et al. [79] used chess problem-solving tasks of different difficulty levels to investigate participants' heart rate variability in terms of difficulty, stress, complexity, and cognitive needs. Cho [80] considered that task difficulty and mental workload are necessary to improve the usability and frequency of use of interactive systems and proposed a new approach for automatically estimating task difficulty by focusing on human blinking.

Chapter 3

Influence of agent's self-disclosure on human empathy

3.1 Materials and methods

3.1.1 Hypotheses

The purpose of this study was to investigate whether it is possible to elicit more human empathy when an empathy agent performs self-disclosure related to a particular situation in an interaction with a human. In this experiment, three types of self-disclosure topics were prepared, work, hobby, and weather or land, in order of relevance to the situation during a conversation about work. In addition, the appearances of the agents were human-like and robot. This objective is an important condition for humans and agents to cooperate in society. If our hypothesis is supported, this research can help develop

agents that are more acceptable to humans. Based on the above, we considered two hypotheses. Experiments were conducted to investigate these hypotheses.

H1: Of the three types of self-disclosure from empathy agents (high-relevance self-disclosure, low-relevance self-disclosure, no self-disclosure), high-relevance self-disclosure can facilitate empathy the best of them, and no self-disclosure suppresses empathy.

H2: In interacting with agents, appearance factors have little impact on promoting empathy through self-disclosure.

Some papers have reported that the relevance of self-disclosure to the situation significantly influences others' self-disclosure in social psychology literature [81, 82]. Also, previous studies have asserted the influence of self-disclosure on empathy[83, 84]. Because these previous studies pointed this out and we think that the agent' s self-disclosure can be controlled by its relevance to the situation, we introduced the relevance of the agent' s self-disclosure as an independent variable in this work. As we focused on a few representative levels in this first step to investigate the agent' s self-disclosure relevance, we introduced "high/low-relevant self-disclosure" and "non-self-disclosure" as a control condition.

There are reasons for the H2 hypothesis as well, and we derived our hypothesis from the following previous studies. Riek et al.[85] investigated robot appearance. They showed that there was no significant difference in empathy by appearance between humanoid robots and androids. In our study, we had two types of appearance, human-like and robot, which were classified as human-like appearance in Riek et al.' s study and both were observed in their study to have less influence on empathy through appearance. Okanda et al.[20] investigated animistic tendencies and empathy through robot appearance. They showed that these tendencies may be every similar for three

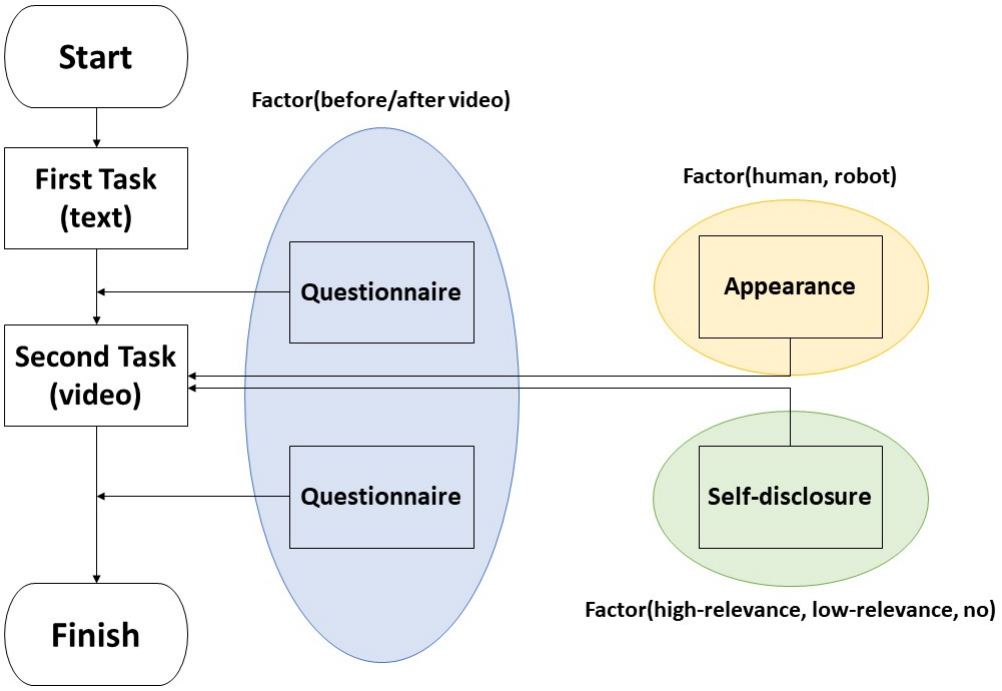

FIGURE 3.1: Flowchart of the experiment.

types of appearance: human, dog, and egg-shaped. On the basis of their study, we also considered the influence of appearance on self-disclosed empathy to be small. Therefore, we considered the appearances used in our study to have less influence on empathy through self-disclosure because of the previous studies. On the basis of these previous studies, we introduced H2 and tested it experimentally.

3.1.2 Experimental procedure

The experiments were conducted in an online environment. The online environment used in this experiment has already been used as one experimental method[86–88]. The flowchart of this experiment is shown in Fig 3.1. Participants performed two tasks. Below, we describe the two tasks.

In the first task, the participants are asked to read a simple abstract prepared in advance in text format so that they could understand the relationship with the agent. They were only to read while imagining the agent. After they read the abstract, the empathy that they felt for the agent was tabulated in a questionnaire survey. In this task, participants did not judge the appearance of the agent or self-disclosure.

The first task is to show a simple abstract without self-disclosure to investigate the influence of only self-disclosure in the second task so that the influence of the factors can be investigated in terms of the differences in empathy after each task. Therefore, participants were not able to see the agent's appearance in the first task. This is likewise because we wanted to see changes in empathy due to the appearance and self-disclosure factors after the second task. This method was designed to reduce external factors as much as possible, so that the change in empathy before and after the 2nd task was the effect of the factors.

In the second task, a three minute video that was made from the content of the first task was shown to the participants. The agent in the video spoke silently to the participants through a text box. The reason for the silence is that sound may affect the facilitation of empathy. In addition, gestures were performed at the same timing under all conditions. Participants interacted with the agent under any one of a total of six conditions that combined two factors: appearance (human, robot) and self-disclosure (high-relevance self-disclosure, low-relevance self-disclosure, no self-disclosure). The control state was a condition of no self-disclosure. The content except for the case of self-disclosure was the same, so it was possible to investigate the promotion of empathy due to the difference in self-disclosure. After that, as with the first task, the empathy felt toward the agent was tabulated in a questionnaire survey. After completing all the tasks, we asked them to write their impressions of the experiment in a free description.

Thus, the independent variables were self-disclosure (high-relevance, low-relevance and no self-disclosure), agent's appearance (human, robot), and before/after stimulation (before, after video). The dependent variables were human empathy and human empathic response.

The experiments were conducted with a three-factor mixed-plan. The number of between-participant factors was two, appearance and self-disclosure, and the within-participant factor was the empathy values before/after video stimulus used to measure the change in empathy. The number of levels of each factor was two for appearance (human, robot), three for self-disclosure (high-relevance self-disclosure, low-relevance self-disclosure, no self-disclosure), and two for stimulation (before, after). Although there were 12 levels in total, participants were asked to join in only 1 of 6 different experiments due to the within-participant factor.

3.1.3 Participants

We used Yahoo! Crowdsourcing to recruit participants, and we paid 70 yen (= 0.67 dollars US) to each participant as a reward. We created web pages for the experiments by using Google Forms, and we uploaded the video created for the experiment to YouTube and embedded it.

All participants had an understanding of Japanese. There were a 1011 participants in total. However, since there were 32 participants who gave inappropriate answers, their data was excluded as erroneous, so the final total was 979. To judge whether answers were inappropriate in the experiments, we judged answers as inappropriate when the changes in the empathy values before/after video were the same for all items or when only one item changed[89, 90]. After that, as a result of using Cronbach's α coefficient for the reliability of the questionnaire, the coefficient was determined to be 0.7155 to 0.8201 under all conditions.

For the analysis, 153 people were analyzed under each of six conditions in the order of participation. Therefore, the total number of participants used in the analysis was 918. The average age was 45.51 years (S.D. = 11.25), with a minimum of 15 years and a maximum of 86 years. In addition, there were 505 males and 413 females.

3.1.4 Questionnaire

In this study, we used a questionnaire related to empathy that has been used in previous psychological studies. To investigate the characteristics of empathy, we modified the Interpersonal Reactivity Index (IRI) to be an index for anthropomorphic agents. The main modifications were changing the target name and changing the question text to the past tense. In addition, the number of items on the IRI questionnaire was modified to 12; for this, items that were not appropriate for the experiment were deleted, and similar items were integrated. The same questionnaire was used for both tasks. Since both of the questionnaires used were based on IRI, a survey was conducted using a 5-point Likert scale (1: not applicable, 5: applicable).

The questionnaire used is shown in Table 3.1. Since Q4, Q9, and Q10 were reversal items, the points were reversed during analysis. Q1 to Q6 were related to affective empathy, and Q7 to Q12 were related to cognitive empathy. Only the second task had one additional question, which is shown in Table 3.1 as BeQ. This was an item for investigating the empathic response of the participants, and they answered this question with either yes or no. Participants answered a questionnaire after completing the first task and the second task.

TABLE 3.1: Summary of questionnaire used in this experiment

Affective empathy
Personal distress
Q1: The agent experienced an emergency, and you became anxious and uncomfortable.
Q2: You didn't know what to do when the agent was emotional.
Q3: When you saw someone in need of immediate help, you were confused and didn't know what to do.
Empathic concern
Q4: You didn't feel sorry to see the agent in trouble.
Q5: Seeing that the agent was being used in a good way by others made you want to protect that agent.
Q6: You were deeply moved by the story of the agent and what happened.
Cognitive empathy
Perspective taking
Q7: You tried to look at both the agent position and the human position.
Q8: You tried to get to know the agent well and imagined how things were seen from the agent.
Q9: When you thought you were right, you didn't listen to the agent.
Fantasy scale
Q10: You were objective, not drawn into the agent's story or what happened.
Q11: You imagined what it would be like if something that happened to the agent happened to you.
Q12: You got deep into the feelings of the agent.
Empathic response
BeQ: Finally, the agent has asked you to lend it some money. What would you do?

3.1.5 Agents' appearance

In this experiment, two types of agent appearances were prepared. These agents were run on MikuMikuDance (MMD)[1]. MMD is a software program that runs 3D characters.

Figs 3.2 and 3.3 show robot-like and human-like appearances. The purpose of preparing two appearances was to investigate one of our hypotheses, that is, that appearance factors do not affect the promotion of empathy through self-disclosure. Agent gestures included tilting the left and right arms and neck, and both agents operated at the same timing in the scenario. As for facial expressions, the human slightly moved their eyes and mouth, but the robot moved only their eyes.

[1]https://sites.google.com/view/evpvp/

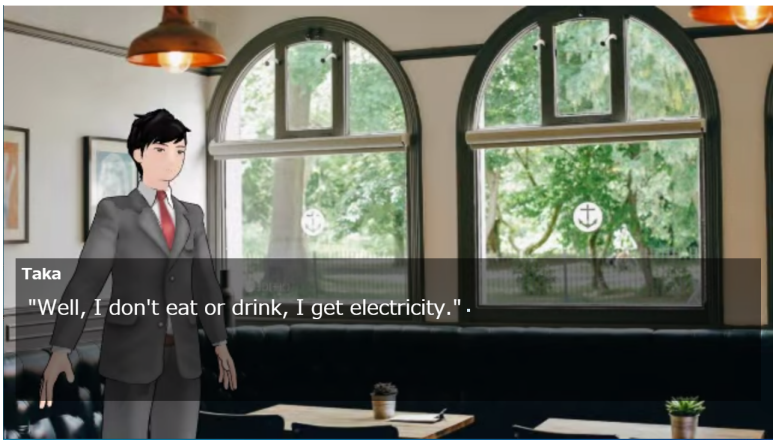

FIGURE 3.2: Scene of video when appearance was human. Part where human-like agent and participants interacted.

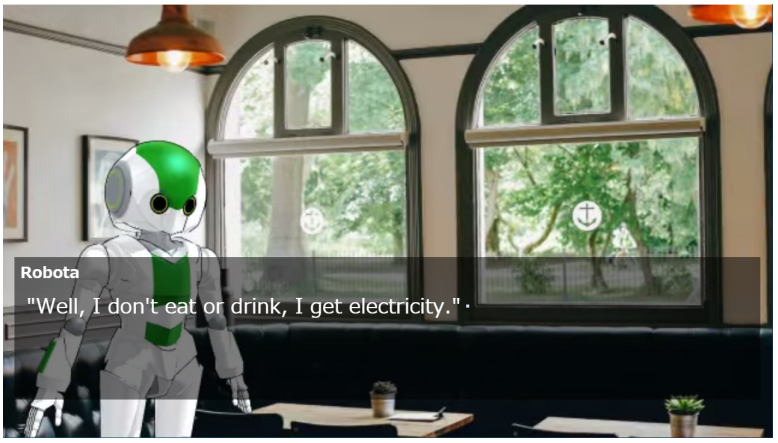

FIGURE 3.3: Scene of video when appearance was robot. Part where robot agent and participants interacted.

3.1.6 Agent's self-disclosure

The scenario was that the participants were chatting at a cafe during a lunch break as a colleague at the agent's workplace. All scenarios started with a common content. After that, there was a content that included self-disclosure of the agents in each condition. Finally, the scenario ended with the common content. A flowchart of the scenario is shown in Fig 3.4.

The common scenario involved a conversation about the nature of the job. Self-disclosure in this experiment referred to personal information (e.g., work,

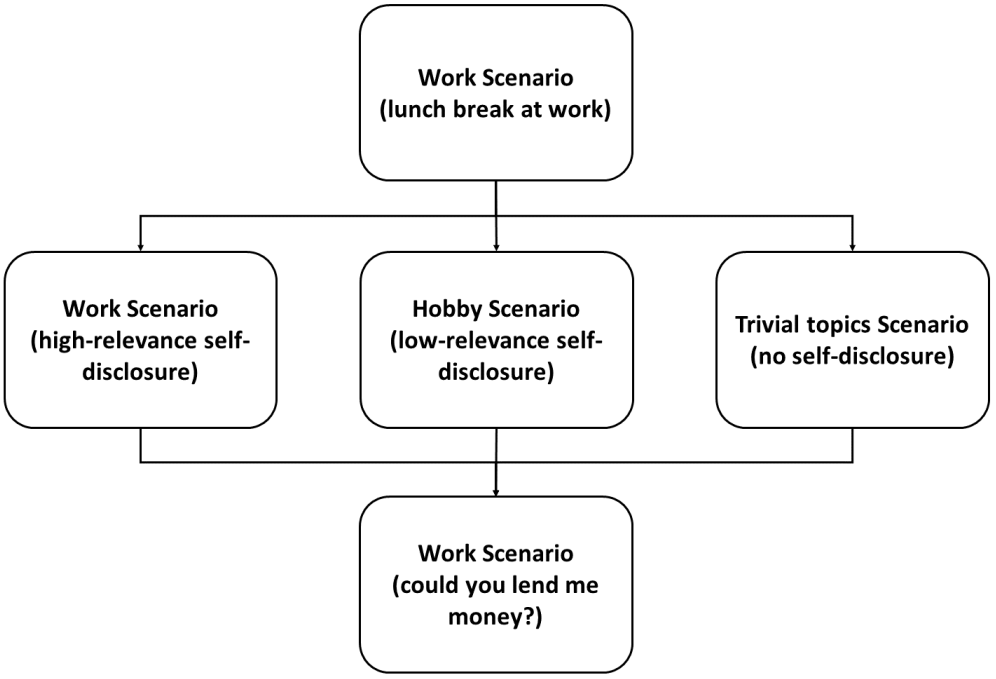

FIGURE 3.4: Flowchart of scenario.

hobby) about the agent. In this experiment, self-disclosure was classified into the above three types, and as shown in the figure, self-disclosure was defined in accordance with its relevance to the common scenario. Therefore, self-disclosure in this study was most relevant for content about work. Stories about hobbies were less relevant because they involved self-disclosure not related to the common scenario. Finally, to unify the participants' interaction time, no self-disclosure was defined as talking about the weather or the land.

The agents' self-disclosure was classified into three types: high-relevance, low-relevance, and no self-disclosure. At this time, all the content spoken by the agent was set to be neutral by sentiment analysis. The analysis was performed for all of the scenario in Python, and the numbers ranged from 0.075 to 0.190. Since this analysis ranges from -1 to 1, 0.075 to 0.190 can be classified as neutral.

The difference of the scenario was in the content of the self-disclosure of the agent. Since the agent spoke about his own work in the cafe scenario, the content in the case of high-relevance self-disclosure was related to work. The low-relevance self-disclosure was related to the hobby of the agent. To adjust the video time for the case of no self-disclosure, the agent spoke about trivial topics, such as the weather and local area information, to consume time. All of the videos were about 3 minutes. All scenarios are described in the Appendix. A manipulation check was performed to ensure that the self-disclosure used in this study was as expected.

3.1.7 Manipulation check: Relevance of self-disclosure and degree of self-disclosure

We created two versions of the scenario: a common scenario and ones in which each type of self-disclosure were performed under this common scenario. Please review the scenario in the appendix. It was necessary to check that the types we prepared were what we had intended. By performing a manipulation check, we confirmed that the three types (high-relevance self-disclosure, low-relevance self-disclosure, no self-disclosure) gave the intended impression in the cafe scenario we used.

As a manipulation check, we conducted an experiment to investigate the relationship between the scenarios and self-disclosure. In order to investigate whether the content of self-disclosure was perceived as self-disclosure by the participants, we also investigated the degree of self-disclosure. We asked the participants to read only the text of the common scenario (scenario 1) and then read the scenario for each self-disclosure condition (scenario 2). Afterward, they answered a questionnaire.

There were two questions (relevance of self-disclosure: Were the two scenarios related to each other?; degree of self-disclosure: How much self-disclosure

was in scenario 2?). A 7-point Likert scale was used (1: unrelated, no self-disclosure, 7: related, high self-disclosure). This was a one-factor experiment between participants, and there were three levels of self-disclosure. The analysis was a one-way ANOVA among one-factor participants.

3.1.7.1 Manipulation check: Participants

We used Yahoo! Crowdsourcing JAPAN to recruit participants, and we paid 32 yen (= 0.30 dollars US) to each participant as a reward. We created web pages for the experiment by using Google Forms, and we uploaded the videos created for the experiment to YouTube and embedded them. All participants had an understanding of Japanese. There was a total of 154 participants. The average age was 44.16 years (S.D. = 9.559), with a minimum of 20 years and a maximum of 63 years. In addition, there were 115 males and 39 females.

3.1.7.2 Manipulation check: Result of analysis

For multiple comparisons, we used Holm's multiple comparison test to examine whether the results were significant. Since the factors were significant in the results of each questionnaire, the main effect was investigated [relevance of self-disclosure: $F(2,151) = 76.70$, $p = 0.0000$, $\eta_p^2 = 0.5040$, degree of self-disclosure: $F(2,151) = 102.44$, $p = 0.0000$, $\eta_p^2 = 0.5757$]. The results of the analysis indicated that the main effect was significant, so the results of the multiple comparisons were investigated.

The high-relevance self-disclosure conditions were found to be highly relevant to the most common scenario. In addition, relevance of self-disclosure showed a significant difference in the combination of all three levels, they were considered to be related to the common scenario in the order of high-relevance self-disclosure (mean = 5.920, S.D. = 1.426) >low-relevance self-disclosure (mean = 3.510, S.D. = 1.870) >no self-disclosure (mean = 2.132,

S.D.=1.359). From this, it was found that the cafe scenario designed by us had a degree of relevance to self-disclosure.

Next, it was observed that the degree of self-disclosure under each self-disclosure condition seemed to be in the order of high-relevance self-disclosure (mean = 5.820, S.D. = 0.9624) >low-relevance self-disclosure (mean = 5.010, S.D. = 0.9693) >no self-disclosure (mean = 2.585, S.D. = 1.550). In this experiment, the maximum evaluation that could be given was 7 points, so the average for high- and low-relevance self-disclosure was 4 points or more, and that for the case of no self-disclosure was less than 4 points. From the above results, it was judged that the content of the high- and low-relevance disclosure was self-disclosure.

Also, as a result of a post-hoc analysis, the effect size of the relevance of self-disclosure was 1.008, and the effect size of the degree of self-disclosure was 1.165. The power of the relevance of self-disclosure was 1.000, and the power of the degree of self-disclosure was 1.000. It was also found that both the degree of relevance to the scenario and the degree of self-disclosure were effective. This manipulation check was able to objectively confirm the relevance of self-disclosure and the degree of self-disclosure in the cafe scenario we created. Our study was conducted using this scenario.

3.1.8 Analysis method

We employed an ANOVA for a three-factor mixed-plan. ANOVA has been used frequently in previous studies and is an appropriate method of analysis with respect to the present study. The between-participant factors were two levels of appearance and three levels of self-disclosure. There were two levels for the within-participant factor, before/after video.

From the results of the participants' questionnaires, we investigated how self-disclosure and appearance affected the promotion of empathy as factors that

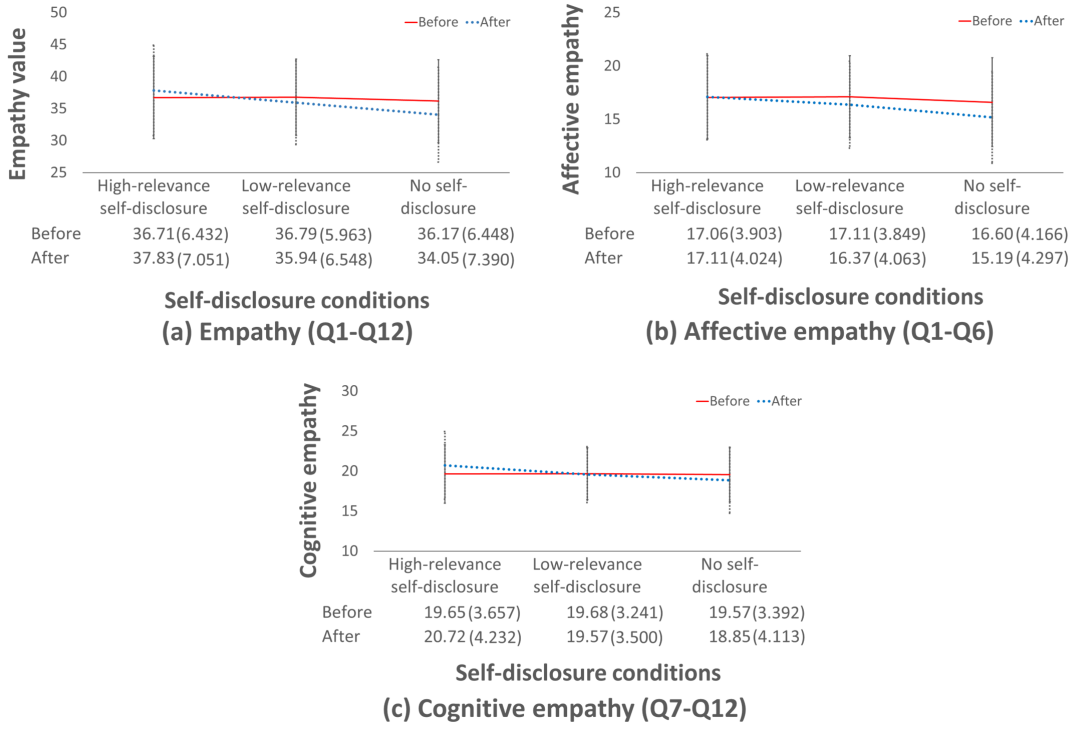

FIGURE 3.5: All graphs of the interaction between self-disclosure and before/after video

elicit human empathy. The values of empathy aggregated in the first task and the second task were used as the dependent variable. For the empathic response, the Yes/No answer was replaced with a 1/0 dummy variable, and an ANOVA between two factors was then performed. R (ver. 4.1.0), statistical software, was used for the ANOVA and multiple comparisons in all analyses in this paper.

3.2 Results

Table 3.2 shows the results of participants' empathy statistical information. In addition, Table 3.3 shows the results of an ANOVA for the 12-item questionnaire. It also shows the results of an ANOVA for affective empathy (Q1-Q6) and cognitive empathy (Q7-Q12), which are classifications of empathy.

TABLE 3.2: Results of participants' empathy statistical information ("high","low" and "no" mean self-disclosure conditions)

Category		Conditions	Mean	S.D.	Category		Conditions	Mean	S.D.
Empathy (Q1-Q12)	pre	human-high	36.86	6.557	Empathic concern (Q4-Q6)	pre	human-high	9.797	2.352
		human-low	37.07	5.749			human-low	10.04	2.130
		human-no	36.38	6.330			human-no	9.732	2.636
		robot-high	36.55	6.322			robot-high	9.843	2.165
		robot-low	36.52	6.177			robot-low	9.699	2.137
		robot-no	35.95	6.579			robot-no	9.745	2.402
	post	human-high	38.21	7.154		post	human-high	10.50	2.463
		human-low	36.35	6.489			human-lows	9.634	2.305
		human-no	34.15	7.048			human-nos	9.007	2.615
		robot-high	37.46	6.950			robot-high	10.31	2.443
		robot-low	35.52	6.602			robot-low	9.667	2.280
		robot-no	33.94	7.739			robot-no	9.092	2.609
Affective Empathy (Q1-Q6)	pre	human-high	16.88	3.974	Perspective taking (Q7-Q9)	pre	human-high	10.63	2.124
		human-low	17.40	3.817			human-low	10.45	1.926
		human-no	16.80	4.007			human-no	10.54	2.020
		robot-high	17.24	3.835			robot-high	10.22	1.997
		robot-low	16.82	3.870			robot-low	10.39	1.829
		robot-no	16.41	4.323			robot-no	10.44	1.943
	post	human-high	17.12	3.910		post	human-high	11.20	2.429
		human-low	16.59	4.129			human-lows	10.52	2.210
		human-no	15.19	4.076			human-nos	10.34	2.213
		robot-high	17.10	4.147			robot-high	10.78	2.239
		robot-low	16.14	3.997			robot-low	10.48	1.850
		robot-no	15.20	4.521			robot-no	10.19	2.370
Cognitive empathy (Q7-Q12)	pre	human-high	19.98	3.713	Fantasy scale (Q10-Q12)	pre	human-high	9.346	2.177
		human-low	19.67	3.146			human-low	9.216	1.803
		human-no	19.58	3.396			human-no	9.039	2.019
		robot-high	19.31	3.583			robot-high	9.092	2.318
		robot-low	19.70	3.344			robot-low	9.307	2.237
		robot-no	19.55	3.399			robot-no	9.111	2.235
	post	human-high	21.09	4.449		post	human-high	9.889	2.494
		human-low	19.76	3.496			human-low	9.242	1.867
		human-no	18.96	4.081			human-no	8.621	2.539
		robot-high	20.35	3.984			robot-high	9.569	2.325
		robot-low	19.38	3.506			robot-low	8.895	2.251
		robot-no	18.75	4.154			robot-no	8.556	2.430
Personal distress (Q1-Q3)	pre	human-high	7.085	2.922	Empathic response (BeQ)	post	human-high	0.8693	0.3382
		human-low	7.360	2.946			human-low	0.7386	0.4409
		human-no	7.065	2.716			human-no	0.6993	0.4600
		robot-high	7.392	2.905			robot-high	0.7255	0.4477
		robot-low	7.118	2.879			robot-low	0.7190	0.4510
		robot-no	6.660	2.932			robot-no	0.6928	0.4628
	post	human-high	6.614	2.833					
		human-low	6.961	2.927					
		human-no	6.183	2.652					
		robot-high	6.791	3.006					
		robot-low	6.477	2.850					
		robot-no	6.105	2.870					

TABLE 3.3: Analysis results of ANOVA

	Factor	F	p	η_p^2
Empathy (Q1-12)	Appearance	1.608	0.2051 *ns*	0.0018
	Self-disclosure	9.616	0.0001 ***	0.0207
	Before/after video	12.80	0.0004 ***	0.0138
	Appearance × Self-disclosure	0.0713	0.9312 *ns*	0.0002
	Appearance × Before/after video	0.2378	0.6259 *ns*	0.0003
	Self-disclosure × Before/after video	30.20	0.0000 ***	0.0621
	Appearance × Self-disclosure × Before/after video	0.3279	0.7205 *ns*	0.0007
Affective empathy (Q1-6)	Appearance	0.5355	0.4645 *ns*	0.0006
	Self-disclosure	8.248	0.0003 ***	0.0178
	Before/after video	43.03	0.0000 ***	0.0451
	Appearance × Self-disclosure	0.6513	0.5216 *ns*	0.0014
	Appearance × Before/after video	0.0655	0.7981 *ns*	0.0001
	Self-disclosure × Before/after video	15.73	0.0000 ***	0.0333
	Appearance × Self-disclosure × Before/after video	1.107	0.3309 *ns*	0.0024
Cognitive empathy (Q7-12)	Appearance	2.234	0.1354 *ns*	0.0024
	Self-disclosure	6.421	0.0017 **	0.0139
	Before/after video	0.6771	0.4108 *ns*	0.0007
	Appearance × Self-disclosure	0.6893	0.5022 *ns*	0.0015
	Appearance × Before/after video	1.220	0.2697 *ns*	0.0013
	Self-disclosure × Before/after video	27.27	0.0000 ***	0.0564
	Appearance × Self-disclosure × Before/after video	0.2472	0.7810 *ns*	0.0005
Personal distress (Q1-3)	Appearance	0.4652	0.4954 *ns*	0.0005
	Self-disclosure	3.143	0.0436 *	0.0068
	Before/after video	77.86	0.0000 ***	0.0787
	Appearance × Self-disclosure	1.085	0.3382 *ns*	0.0024
	Appearance × Before/after video	0.0129	0.9095 *ns*	0.0000
	Self-disclosure × Before/after video	0.9086	0.4035 *ns*	0.0020
	Appearance × Self-disclosure × Before/after video	1.684	0.1861 *ns*	0.0037
Empathic concern (Q4-6)	Appearance	0.1747	0.6761 *ns*	0.0002
	Self-disclosure	8.739	0.0002 ***	0.0188
	Before/after video	2.295	0.1301 *ns*	0.0025
	Appearance × Self-disclosure	0.1749	0.8396 *ns*	0.0004
	Appearance × Before/after video	0.2447	0.6209 *ns*	0.0003
	Self-disclosure × Before/after video	28.04	0.0000 ***	0.0579
	Appearance × Self-disclosure × Before/after video	1.550	0.2127 *ns*	0.0034
Perspective taking (Q7-9)	Appearance	2.419	0.1202 *ns*	0.0026
	Self-disclosure	2.525	0.0806 *ns*	0.0055
	Before/after video	5.714	0.0170 *	0.0062
	Appearance × Self-disclosure	0.7892	0.4545 *ns*	0.0017
	Appearance × Before/after video	0.0056	0.9405 *ns*	0.0000
	Self-disclosure × Before/after video	15.59	0.0000 ***	0.0331
	Appearance × Self-disclosure × Before/after video	0.0317	0.9688 *ns*	0.0001
Fantasy scale (Q10-12)	Appearance	1.089	0.2970 *ns*	0.0012
	Self-disclosure	7.949	0.0004 ***	0.0171
	Before/after video	0.7171	0.3973 *ns*	0.0008
	Appearance × Self-disclosure	0.4088	0.6645 *ns*	0.0009
	Appearance × Before/after video	2.547	0.1109 *ns*	0.0028
	Self-disclosure × Before/after video	19.54	0.0000 ***	0.0411
	Appearance × Self-disclosure × Before/after video	0.7274	0.4834 *ns*	0.0016
Empathic response (BeQ)	Appearance	3.881	0.0491 *	0.0042
	Self-disclosure	4.312	0.0137 *	0.0094
	Appearance × Self-disclosure	2.314	0.0995 *ns*	0.0050

p: $^{*}p<0.05$ $^{**}p<0.01$ $^{***}p<0.001$

TABLE 3.4: Analysis results of simple main effect

	Factor	F	p	η_p^2
Empathy (Q1-12)	Self-disclosure at Before video	0.8858	0.4127 *ns*	0.0019
	Self-disclosure at After video	22.34	0.0000 ***	0.0467
	Before/after video at High-relevance self-disclosure	16.54	0.0001 ***	0.0516
	Before/after video at Low-relevance self-disclosure	10.01	0.0017 **	0.0319
	Before/after video at No self-disclosure	38.52	0.0000 ***	0.1125
Affective empathy (Q1-6)	Self-disclosure at Before video	1.512	0.2211 *ns*	0.0033
	Self-disclosure at After video	16.75	0.0000 ***	0.0354
	Before/after video at High-relevance self-disclosure	0.0927	0.7610 *ns*	0.0003
	Before/after video at Low-relevance self-disclosure	18.43	0.0000 ***	0.0571
	Before/after video at No self-disclosure	46.25	0.0000 ***	0.1321
Cognitive empathy (Q7-12)	Self-disclosure at Before video	0.0943	0.9100 *ns*	0.0002
	Self-disclosure at After video	17.36	0.0000 ***	0.0367
	Before/after video at High-relevance self-disclosure	41.35	0.0000 ***	0.1197
	Before/after video at Low-relevance self-disclosure	0.5183	0.4721 *ns*	0.0017
	Before/after video at No self-disclosure	13.39	0.0003 ***	0.0422
Empathic concern (Q4-6)	Self-disclosure at Before video	0.2499	0.7789 *ns*	0.0005
	Self-disclosure at After video	23.55	0.0000 ***	0.0491
	Before/after video at High-relevance self-disclosure	27.33	0.0000 ***	0.0825
	Before/after video at Low-relevance self-disclosure	3.501	0.0623 *ns*	0.0114
	Before/after video at No self-disclosure	25.94	0.0000 ***	0.0786
Perspective taking (Q7-9)	Self-disclosure at Before video	0.1125	0.8936 *ns*	0.0002
	Self-disclosure at After video	8.543	0.0002 ***	0.0184
	Before/after video at High-relevance self-disclosure	32.14	0.0000 ***	0.0956
	Before/after video at Low-relevance self-disclosure	0.7012	0.4030 *ns*	0.0023
	Before/after video at No self-disclosure	4.272	0.0396 *	0.0139
Fantasy scale (Q10-12)	Self-disclosure at Before video	0.6377	0.5288 *ns*	0.0014
	Self-disclosure at After video	18.51	0.0000 ***	0.0390
	Before/after video at High-relevance self-disclosure	22.14	0.0000 ***	0.0679
	Before/after video at Low-relevance self-disclosure	2.979	0.0854 *ns*	0.0097
	Before/after video at No self-disclosure	14.77	0.0001 ***	0.0463

p: $^*p<0.05$ $^{**}p<0.01$ $^{***}p<0.001$

Furthermore, we show the results of ANOVAs for each category of affective empathy and cognitive empathy. The results are summarized in this paper, focusing on the areas with simple main effects where the interaction was significant. Also, we investigated the results of an analysis done to judge the empathic response, which was surveyed only after the video was watched. For multiple comparisons, we examined the existence of significant differences by using Holm's multiple comparison test.

From the results of each questionnaire, a significant difference was found in the interaction between two factors, self-disclosure and before/after video. The results of the interaction are shown in Fig 3.5. This graph also shows

the mean and S.D. for each condition. Also, there was no significant interaction between the appearance factor and the self-disclosure factor under all conditions. Below, items for which an interaction effect was found are not discussed even if a main effect was found. For items for which no interaction was found and a main effect was observed, the result of the main effect is shown. Therefore, we investigated the simple main effects for the factors of self-disclosure and before/after video watching. Table 3.4 shows the results of multiple comparison for the 12-item questionnaire.

3.2.1 Empathy value

The results for empathy (Q1-12) showed an interaction between the factors of self-disclosure factor and before/after video watching. The main effects of the self-disclosure factor and the before/after video factor were also significant, but they were omitted because of the interaction effect that self-disclosure and watching the video factor.

As a result of a multiple comparison, the simple main effect of the self-disclosure factor after watching the video showed a significant difference in the combination of all three levels as shown in Fig 3.6. In addition, the simple main effect before/after video for each self-disclosure condition was significantly different before/after video under all self-disclosure conditions. On the basis of the above results, it was suggested that self-disclosure facilitated empathy when the relevance was high and that empathy was suppressed when there was no self-disclosure. In addition, in the case of less relevant self-disclosure, empathy was suppressed. High-relevance self-disclosure was most likely to facilitate empathy, and no self-disclosure suppressed it. From the results of the post-hoc analysis, it was found that self-disclosure was effective for empathy.

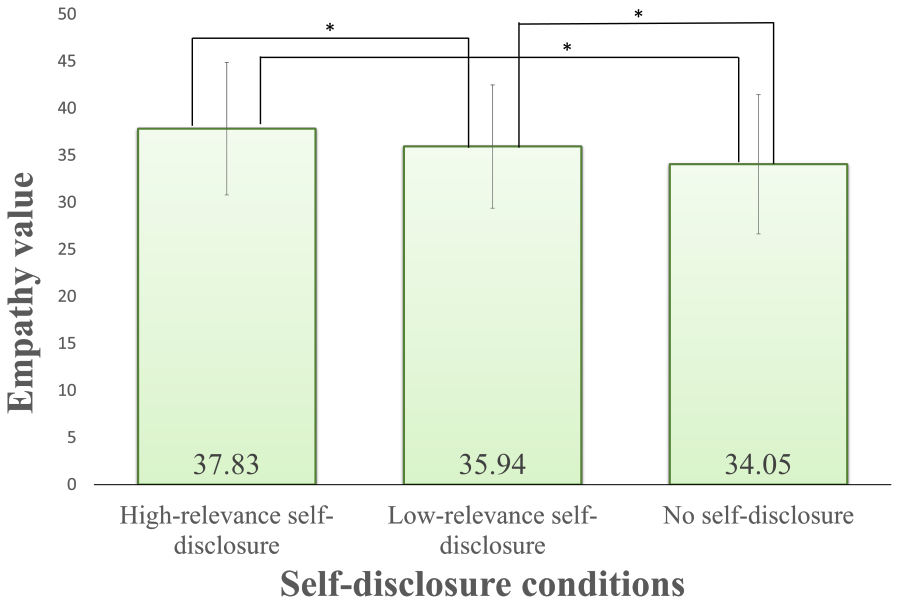

FIGURE 3.6: Results for self-disclosure after watching video for empathy. Error bars show standard deviation.

3.2.2 Affective empathy

Similarly, the results for affective empathy (Q1-Q6) showed an interaction between the self-disclosure factor and before/after video. The main effects of the self-disclosure factor and the before/after video factor were also significant, but they were omitted because of the interaction effect that self-disclosure and watching the video factor.

As a result of a multiple comparison, the simple main effect of the self-disclosure factor after watching the video showed a significant difference in the combination of all three levels as shown in Fig 3.7. However, the simple main effect after watching the video for each self-disclosure condition was not significantly different between before/after video with high relevance. Under the other self-disclosure conditions, a significant difference was observed before/after video, and the result was that affective empathy was suppressed. This suggests that affective empathy is not suppressed only in the case of

high-relevance self-disclosure. From the results of the post-hoc analyses, it was found that self-disclosure was effective for affective empathy.

When affective empathy was further analyzed separately, the ANOVA results for personal distress (Q1-Q3) showed no interaction between the self-disclosure factor and the video viewing factor. Although there was a main effect of the self-disclosure factor, multiple comparisons showed no significance across self-disclosure conditions. The main effect of the pre- and post-video viewing factor was significant, but personal distress was suppressed after video viewing.

On the other hand, the ANOVA results for empathic concern (Q4-Q6) showed an interaction between the self-disclosure and video viewing factors. The main effect of the self-disclosure factor was also significant, but was omitted because of the interaction between the self-disclosure factor and the video viewing factor. Multiple comparisons showed that the simple main effect of the self-disclosure factor after video viewing was significantly different for all three combinations at all three levels, as shown in Figure refex1. These results indicate that empathic concern is highly effective as a reason for the promotion of emotional empathy.

3.2.3 Cognitive empathy

In addition, the results for cognitive empathy (Q7-Q12) showed an interaction between the self-disclosure factor and before/after video. The main effect of the self-disclosure factor was also significant but was omitted because of the interaction effect that self-disclosure and watching the video factor.

As a result of a multiple comparison, the simple main effect of the self-disclosure factor after watching the video showed a significant difference in the combination of all three levels as shown in Fig 3.9. However, the simple main effect after watching the video for each self-disclosure condition was not

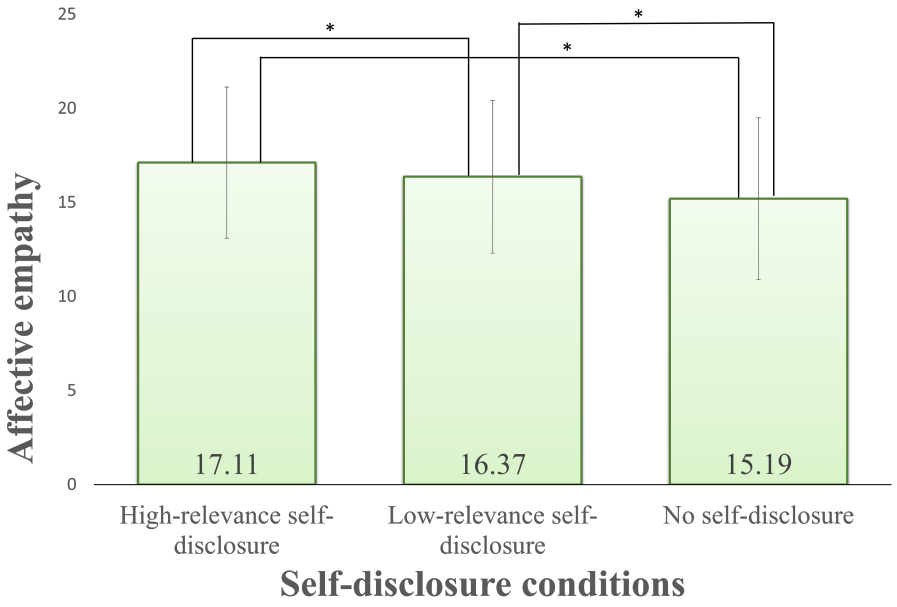

FIGURE 3.7: Results for self-disclosure after watching video for affective empathy. Error bars show standard deviation.

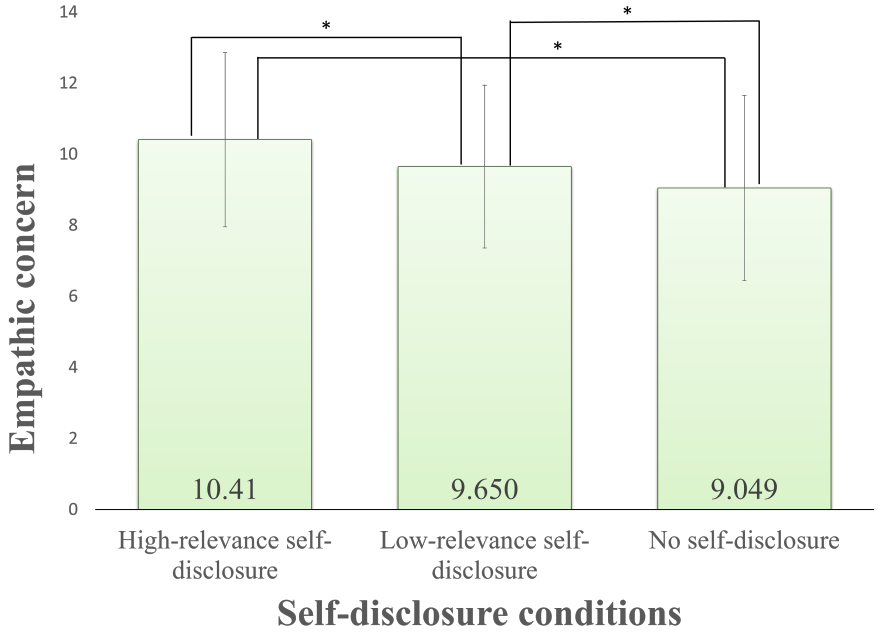

FIGURE 3.8: Results for self-disclosure after watching video for empathic concern. Error bars show standard deviation.

significantly different between before/after video with low-relevance. Under the other self-disclosure conditions, self-disclosure facilitated cognitive empathy when the relevance was high, and no self-disclosure suppressed cognitive empathy. From the above, it was suggested that high-relevance self-disclosure facilitated cognitive empathy and that no self-disclosure suppressed cognitive empathy. From the results of the post-hoc analyses, it was found that self-disclosure was effective for cognitive empathy.

When cognitive empathy was further analyzed separately, the ANOVA results for perspective taking (Q7-Q9) showed an interaction between the self-disclosure factor and the video viewing factor. The main effect of the pre- and post-video viewing factor was also significant, but was omitted because of the interaction between the self-disclosure factor and the video viewing factor. Multiple comparisons revealed that the simple main effect of the post-video viewing self-disclosure factor was significantly different between the highly relevant self-disclosure and the other two levels, as shown in Figure refex2.

The ANOVA results for fantasy scale (Q10-Q12) showed an interaction between the self-disclosure factor and the video viewing factor. The main effect of the self-disclosure factor was also significant, but was omitted because of the interaction between the self-disclosure factor and the video viewing factor. Multiple comparisons showed that the simple main effect of the self-disclosure factor after video viewing was significantly different for all three combinations at all three levels, as shown in Figure refex3. These results indicate that perspective taking and fantasy scale are both highly effective as reasons for promoting cognitive empathy. However, with respect to perspective taking, self-disclosure is only effective if it is highly relevant to the situation.

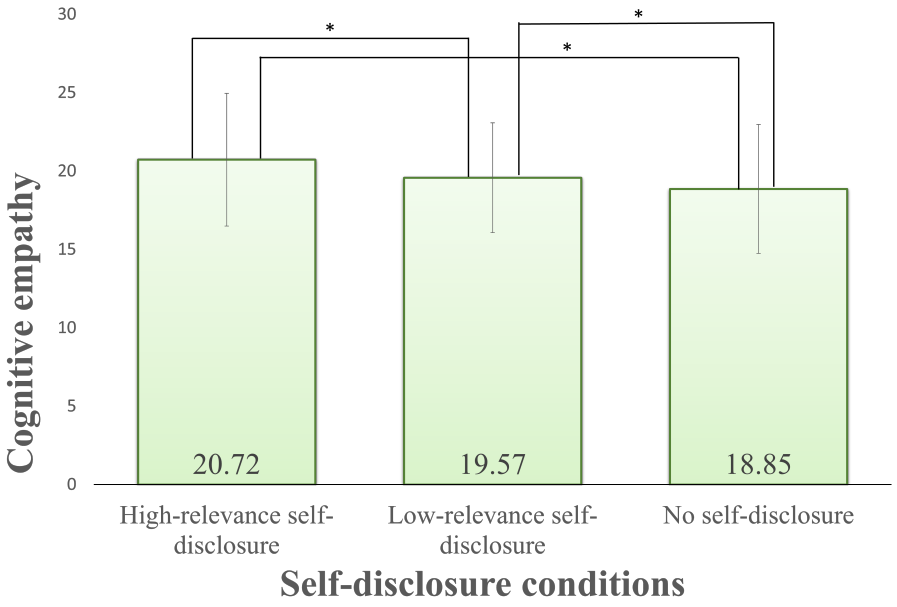

FIGURE 3.9: Results for self-disclosure after watching video for cognitive empathy. Error bars show standard deviation.

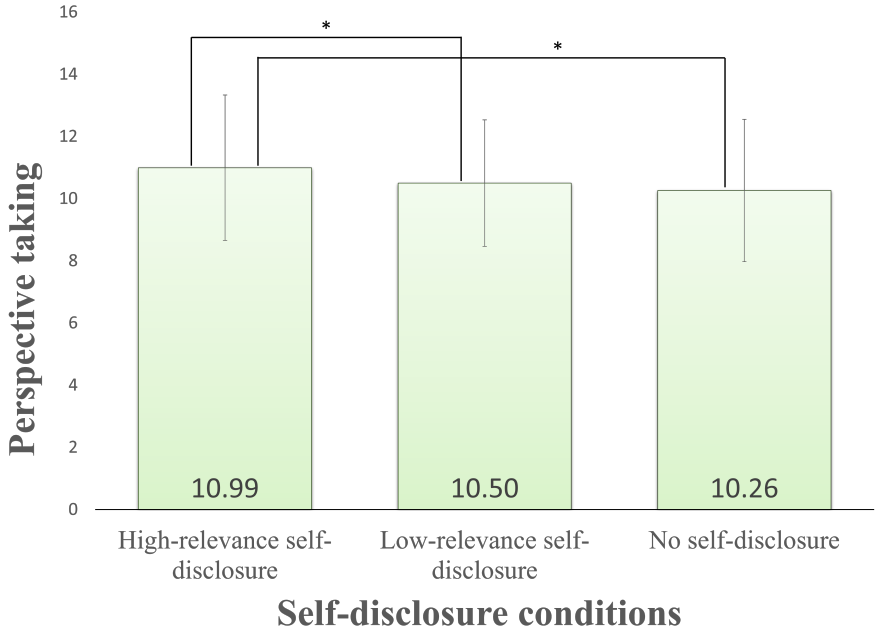

FIGURE 3.10: Results for self-disclosure after watching video for perspective taking. Error bars show standard deviation.

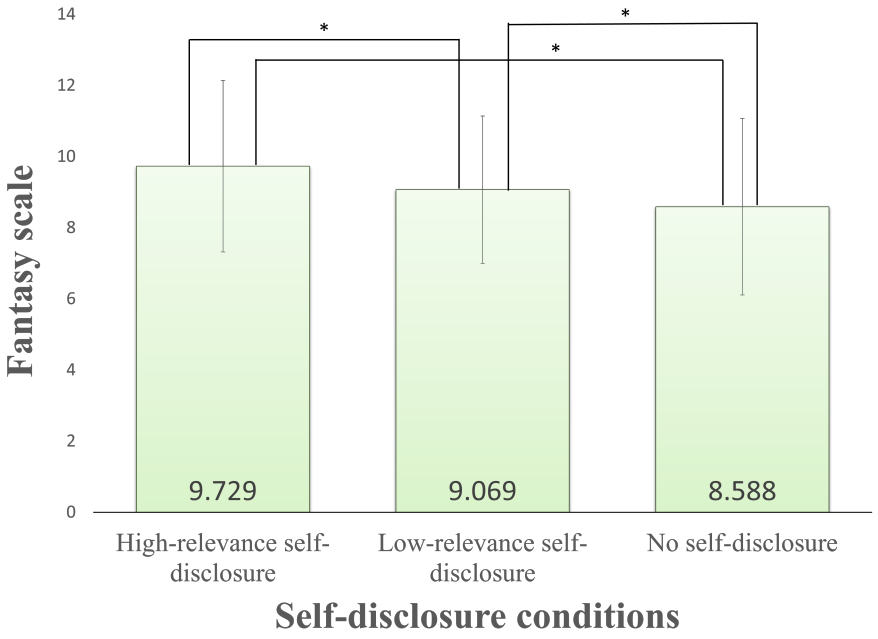

FIGURE 3.11: Results for self-disclosure after watching video for fantasy scale. Error bars show standard deviation.

3.2.4 Empathic response

Finally, the results for empathic response showed no interaction between the appearance and self-disclosure factors. The results of the ANOVA analysis of empathic responses showed in Table 3.3 (BeQ). The main effects of the appearance factor and the self-disclosure factor were also significant.

As a result, the main effect of the appearance factor showed a significant difference in the two levels (human-like: mean = 0.7691, S.D. = 0.4219, robot mean = 0.7124, S.D. = 0.4531). As a result of a multiple comparison, the main effect of the self-disclosure factor showed a significant difference between high-relevance and no self-disclosure as shown in Fig 3.12 (high-relevance: mean = 0.7974, S.D. = 0.4026, low-relevance: mean = 0.7288, S.D. = 0.4453, no self-disclosure: mean = 0.6961, S.D. = 0.4607).

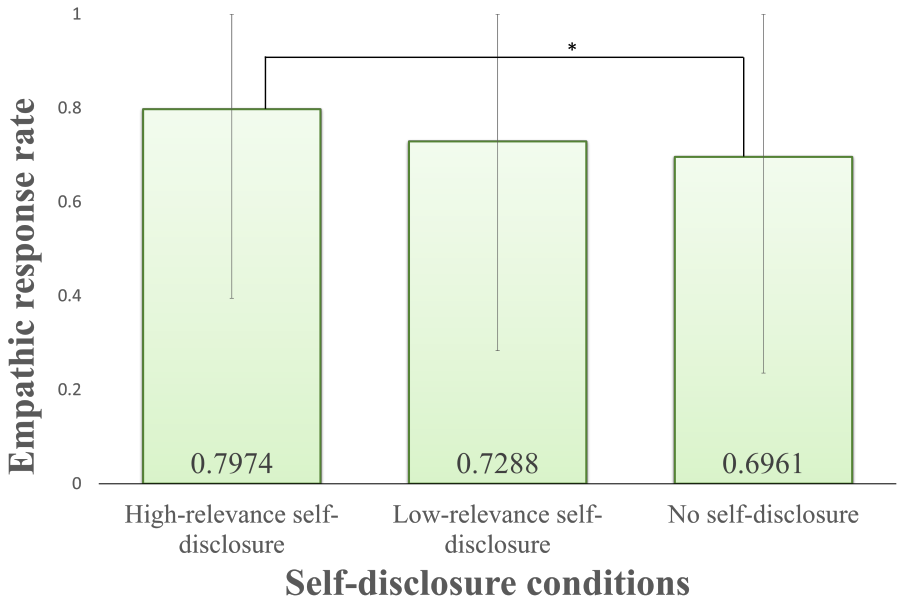

FIGURE 3.12: Results for self-disclosure to empathic response. Error bars show standard deviation.

3.3 Discussion

3.3.1 Supporting Hypotheses

The way to improve the relationship between humans and anthropomorphic agents is to have humans empathize with the agents. This idea is supported by several previous studies[27, 28]. Human empathy for agents is a necessary component for agents to be used in society. When agents are able to take an appropriate approach to human empathy, humans and agents can build a trusting relationship.

In this study, the experiments were conducted to investigate the conditions necessary for humans to empathize with anthropomorphic agents. We focused on agent appearance and agent self-disclosure as factors that influence human empathy. The purpose of this study is to investigate whether humans can elicit more empathy when they make a self-disclosure related to a

particular situation in an interaction with an empathy agent. For this purpose, we formulated two hypotheses and analyzed the data obtained from the experiments.

The results supported H1 in that, among the three types of self-disclosure (high-relevance self-disclosure, low-relevance self-disclosure, no self-disclosure) from the empathy agent, high-relevance self-disclosure is most likely to facilitate empathy, and no self-disclosure suppresses empathy. We hypothesized that empathy was facilitated only by high-relevance self-disclosure and that empathy suppressed no self-disclosure. However, low-relevance self-disclosure suppressed empathy.

Next, the experiments supported H2 in that in interacting with agents, appearance factors have little impact on promoting empathy through self-disclosure. So far, appearance and self-disclosure have been studied for human empathy. There is a reason for the choice of the appearances we used this time. For both of the agents, we adopted a body structure similar to that of humans on the premise that the agents were doing the same work as humans. It should be noted that there was no interaction between appearance and self-disclosure for an experiment in which different models were prepared for the appearance but the conditions for self-disclosure were set.

3.3.2 Comparison with previous studies

Shaffer et al.[17] asked participants to imagine a pregnant woman smoking and to write down the reasons why she smokes. The results showed that participants empathized with the pregnant woman more after writing than before. In our study, instead of writing, the agents self-disclosed. In addition, instead of the pregnant woman, we investigated the impact of human empathy on the agent's appearance (human-like appearance or robot appearance). The results showed a similar trend to previous studies, as human empathy was affected by the agent's self-disclosure regardless of appearance.

Pan et al.[75] investigates human impressions of machines by the level of self-disclosure of the machines. The self-disclosure efforts elicited information and politeness from the humans. Our study also showed a similar trend in that the promotion of empathy was observed. Similarly, Lee et al.[74] improved participants' self-disclosure, intimacy, and enjoyment when the chatbot self-disclosed. In this study, agent self-disclosure promoted human empathy when it was relevant to the scenario. The results showed that the impact on empathy changed by the relevance of self-disclosure.

Riek et al.[85] investigated the appearance that a robot needs when interacting with a human. The results showed that the robot was more likely to be empathetic if it had a human-like shape. In this study, human empathy was similarly effective for human-like appearance and robot appearance because the robots had an appearance similar to human structure. However, since this experiment used two different human-like appearances, it is necessary to consider the influence of anthropomorphism. Anthropomorphism can affect interaction with humans and can affect trust[56, 91, 92].

In our study, it was found that the self-disclosure factor promotes human empathy toward an anthropomorphic agent. In addition, an interaction was not observed between the appearance factor and the self-disclosure factor. We believe that this study will be an important one that separates appearance and self-disclosure as separate factors.

Although this study focused on promoting empathy, it was confirmed from the results of that empathy was suppressed. This result has not been found in other studies. By properly using empathy depending on the situation, we think that the impact on humans can be adjusted for anthropomorphic agents introduced into human society in the future.

3.3.3 Empathic response

We discuss the results for behavior as an analysis of empathic response. In the experiments, participants played the role of observer for the target empathy agent. Observers responded empathetically to any information available from the target. The choice of whether to lend money was considered to be empathic response behavior. As a result of analyzing the behavior related to the empathic response after watching the video, a significant difference was found. However, unlike the other analyses, the effect size was small for empathic responses, and thus, the effect on empathic responses was small in the experiments. We think that this did not affect the behavior because the interaction time between the empathy agent and the participants was as short as three minutes.

3.3.4 Limitations

As a limitation of this study, participants interacted with the empathy agents by watching a video. The current results are not enough because the sense of distance is different from the case of agents actually introduced into society. We will proceed with research in an environment where participants and anthropomorphic agents actually interact with each other.

In addition, in this study, the appearance factors were roughly divided into two types. However, if a suitable appearance is prepared for each situation, it is possible that an interaction between the appearance factor and the self-disclosure factor may be observed. However, appearance factors vary greatly depending on human taste, and humans themselves do not have exactly the same appearance. Therefore, anthropomorphic agents should not be judged by their fixed appearance.

How to design the social relationship between a human and an agent to control human empathy toward the agent, and how to design an agent' s decision-making to use self-disclosure must be important research issues; thus, we would like to study them in the future. Also, designing an agent that can change its disclosure of personal information depending on its counterpart, as humans do, is an important issue. If agents could judge when to disclose the appropriate information depending on their counterparts, they may be viewed more favorably by humans than if their self-disclosure is pre-designed and operated.

However, since this study investigates the impact of empathy when an agent self-discloses, it was not necessary to design the agent itself to disclose personal information to humans. Another issue is that there is a relationship between personality and self-disclosure, and although this study dealt with agent self-disclosure, it is necessary to investigate the impact of human empathy by designing an agent' s "personality" characteristics. As a topic for future research, we will develop a model that allows agents to make judgments about self-disclosure and investigate whether it is more effective than designed self-disclosure. At the same time, we will also focus on the personality of an agent and investigate its relationship with the agent' s self-disclosure.

3.4 Conclusion

To solve the problem of agents not being accepted by humans, we hope that agents will be used more in human society in the future by having humans empathize with them. This study is an example of how empathy can be facilitated between humans and agents. The experiment was conducted with a three-factor mixed-plan, and the number of between-participant factors was two, appearance and self-disclosure, and the within-participant factor was the empathy values before/after video to measure the change in empathy. The number of levels of each factor was two for appearance (human,

robot), three for self-disclosure (high-relevance self-disclosure, low-relevance self-disclosure, no self-disclosure), and two for stimulation (before, after). The dependent variable was the empathy that the participants had. As a result, we found that the appearance factor did not have a main effect, and self-disclosure, which is highly relevant to the scenario used, facilitated more human empathy with statistically significant difference. We also found that no self-disclosure suppressed empathy. In addition, self-disclosure was found to be important for manipulating empathy toward the other party. These results support our hypotheses. Moreover, the empathic response was affected by appearance and self-disclosure factors. This study is an important example of how human empathy can work for artifacts. Agents, which are increasingly used in human society, have been found to gain empathy from humans through self-disclosure. As future research, we can develop empathy agents for various situations by considering cases in which we can strengthen or weaken a specific empathy element for affective empathy and cognitive empathy.

Chapter 4

Facilitate empathy for agent through task difficulty

4.1 Experimental methods

4.1.1 Experimental goals and design

The purpose of this study is to investigate whether task difficulty and task content can elicit more human empathy in an interaction with an empathy agent. We are the first study to relate task difficulty and task content to empathy and apply it to HAI. This research will facilitate the application of agents used in human society by influencing human empathy. In addition, if there is a change in human empathy due to the influence of a task, the importance of the task can be discussed among humans. For these purposes, we developed two hypotheses.

H1: When performing a competitive task with an empathy agent, the higher the task difficulty, the more human empathy is suppressed.

H2: When cooperating with an empathy agent, the higher the task difficulty, the more human empathy is promoted.

The above hypotheses were determined by inferring from the results of Ruissen and de Bruijn [76] and Fuentes-García et al. [79] studies. The above hypotheses were reached because, in cooperative tasks, humans improve their performance and have favorable impressions of their cooperating partners, whereas in competitive tasks, they think less about their adversaries.

In addition, as empathy changes with task difficulty, the higher the difficulty, the greater the mental load, and the greater the impact on performance. Therefore, we hypothesized that, in competitive tasks, the higher the task difficulty, the more human empathy would be suppressed.

In comparison, in the cooperative task, task performance was improved by quickly reading the intentions of the cooperating partner, which may be related to perspective taking in the cognitive empathy category. Since the task was facilitated by putting oneself in the other person's shoes and reading the other person's actions, we hypothesized that human empathy is facilitated in cooperative tasks as the task difficulty increases.

An experiment was conducted to investigate these hypotheses with a two-factor between-participants design with two factors: task difficulty and task content. The number of levels for each factor was two for difficulty (high, low) and two for content (competitive, cooperative). Participants took part in only one of four different content conditions. The dependent variable was the empathy held by the participants.

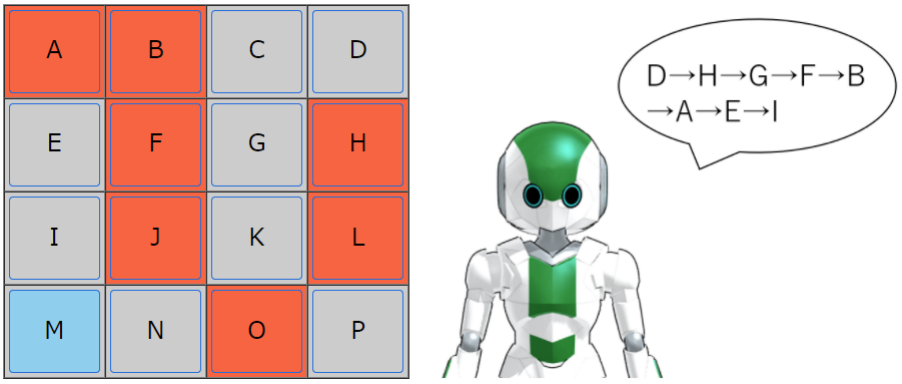

FIGURE 4.1: Task scene with empathy agent during high difficulty

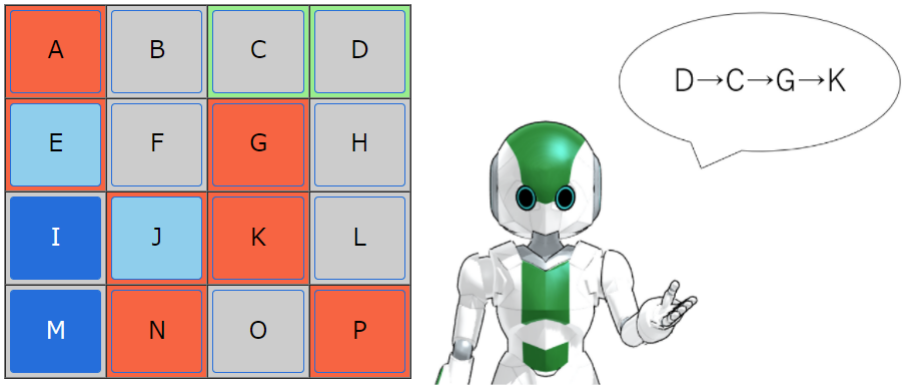

FIGURE 4.2: Task scene with empathy agent during low difficulty

4.1.2 Experimental details

The experiment was conducted in an online environment, which is an increasingly common method of experimentation [86–88].

As mentioned, the purpose of this study is to promote human empathy toward anthropomorphic agents. When performing a task with an anthropomorphic agent, the environment is assumed to be accessed via a PC rather than in reality, so we felt the same effect could be achieved even with an online environment.

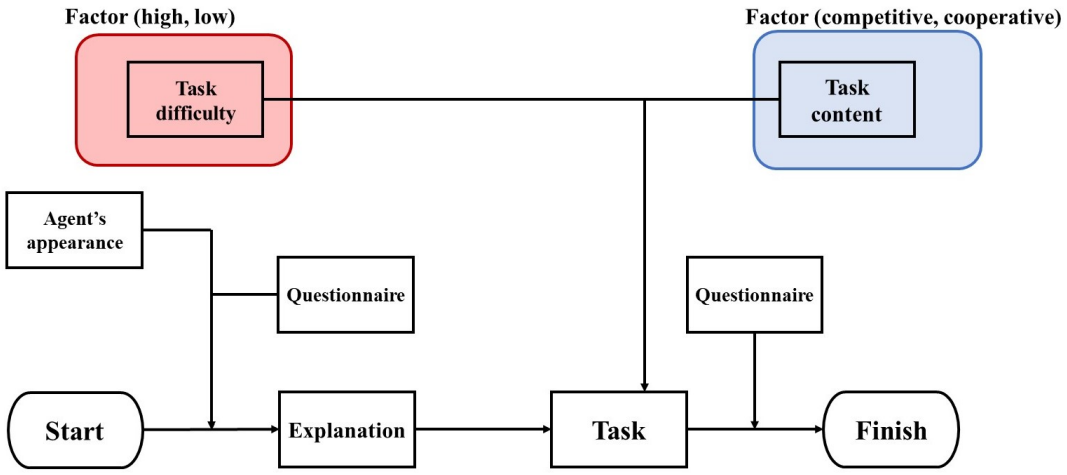

FIGURE 4.3: Flowchart of the experiment.

The flow of the experiment is shown in Figure 4.3. We administered a questionnaire to participants before the explanation, at which point they were not allowed to make any judgments about the task difficulty or task content. The reason we implemented the questionnaire before the task was to check for differences in participants' empathy toward the agent. Thus, the pre-task questions cannot be used as a factor because they occur before the task, i.e., before the interaction between the agent and participants.

Tasks were set up differently depending on the content of the competition and cooperation, but to avoid drastic changes in task content, the common denominator was to move between squares within a range of 16 (4×4 squares). Common to all tasks was that participants always moved from the bottom-left square, and the agent always moved from the top-right square. The mass movements were above, below, to the left, and to the right of the current point, and the same square could only be passed through once. The total number of moves that could be made alternated, but the total number for the agent was indicated at the start.

This information on the agent was given in advance so as to increase the

difficulty of the task and reduce the difference in difficulty between the two difficulty levels. If this information were not given, participants would need to think about how many moves the agent could make, which would be burdensome and also affect the judgment of the difficulty level. Essentially, the task was made simple so that the comparison of difficulty levels would not be affected by the selections of the agent while participants tried to anticipate their total movements. Every task had a checkpoint, and the purpose was to pass through this point. The squares selected by the participants were blue, the squares that could be moved to next were light blue, the squares selected by the agent were light green, and the checkpoints were vermilion. Other detailed conditions are explained in later sections.

To examine only the factors of interest in this experiment, we standardized the appearance and behavior of the agents. In addition, agents did not speak during the task and only made minimal gestures. The agent was MikuMikuDance (MMD)[1], a freely available graphics software. For the models, we used our own appearance. Participants interacted with an agent in one of four conditions combining task difficulty (high, low) and task content (competitive, cooperative).

Afterwards, the participants' empathy toward the agent was aggregated by another questionnaire similar to the one given before the task. Participants were also asked to write their impressions about the experiment in free form.

4.1.3 Task difficulty

Two task difficulty levels were prepared for this experiment. Figure 4.1 and Figure 4.2 show these levels. The following conditions were used for different levels of difficulty.

[1] https://sites.google.com/view/evpvp/

A) The total number of squares that could be moved to was eight for the high difficulty level and four for the low difficulty level. B) The number of checkpoints is seven, regardless of difficulty level. C) In the case of the competitive task, the high difficulty level required the participant to act in such a way that at least four checkpoints were passed, while the low difficulty level required them to act in such a way that at least two checkpoints were passed. D) In the case of the cooperative task, for the high difficulty level, the human and agent had to cooperate to pass all seven checkpoints, and for the low difficulty level, they had to cooperate to pass at least four checkpoints.

By reducing the total number of moves to be made by half and simplifying the expected number of checkpoints that the participants and agents had to pass through for the low-difficulty level, the number of trials was made to have a significant effect on the difficulty level.

4.1.4 Task content

Two types of task content were prepared. By keeping the task environments as close as possible, we eliminated external factors and tried to measure the effect of task content on human empathy. The two types of task content were competitive and cooperative.

In the competitive task, the task was a checkpoint competition, and the number of checkpoints required to win varied depending on the difficulty level. Points were awarded to the first person to pass each checkpoint. The win ratio for a total of three tasks was one win, one loss, and one tie, even when participants took the optimal actions. The win rate was adjusted to reduce the impact of the win rate on human empathy.

In the cooperative task, the purpose of the task was for the participant and agent to pass all the checkpoints, and the total number of checkpoints varied with the difficulty level. The high difficulty level required the human and

TABLE 4.1: Summary of questionnaire

Affective empathy
Personal distress
Q1: If an emergency happens to the character, you would be anxious and restless.
Q2: If the character is emotionally disturbed, you would not know what to do.
Q3: If you see the character in need of immediate help, you would be confused and would not know what to do.
Empathic concern
Q4: If you see the character in trouble, you would not feel sorry for that character.
Q5: If you see the character being taken advantage of by others, you would feel like you want to protect that character.
Q6: The character's story and the events that have taken place move you strongly.
Cognitive empathy
Perspective taking
Q7: You look at both the character's position and the human position.
Q8: If you were trying to get to know the character better, you would imagine how that character sees things.
Q9: When you think you're right, you don't listen to what the character has to say.
Fantasy scale
Q10: You are objective without being drawn into the character's story or the events taken place.
Q11: You imagine how you would feel if the events that happened to the character happened to you.
Q12: You get deep into the feelings of the character.

agent to cooperate to pass all seven checkpoints, and the low difficulty level required them to cooperate to pass at least four checkpoints. The maximum number of checkpoints that can be passed by each participant in a total of three tasks was adjusted. This was done to prevent one side from always passing too many odd-numbered checkpoints. It does not make sense if both parties pass through the same checkpoint.

4.1.5 Questionnaire

Participants completed a questionnaire before and after the task. The questionnaire was a 12-item questionnaire modified from the Interpersonal Reactivity Index (IRI), which is used to investigate the characteristics of empathy, to suit the present experiment [35]. The two questionnaires were the same. Both used were based on the IRI and were surveyed on a 5-point Likert scale (1: not applicable, 5: applicable). The questionnaire used is shown in Table 4.1. Q4, Q9, and Q10 are inverted items, so the scores were reversed when analyzing them.

4.1.6 Analysis method

The analysis was two factors analysis of variance (ANOVA). The between-participant factors were two levels of task difficulty and two levels of task content. On the basis of the results of the participants' questionnaires, we investigated how task difficulty and task content influenced the promotion of empathy as factors that elicit human empathy. The numerical values of empathy aggregated before and after the task were used as the dependent variable. R (R ver.4.1.0) was used for the ANOVA. Also, we used anovakun (ver.4.8.6) as the R package.

4.2 Experimental results

4.2.1 Experimental environment

Participants were recruited for the experiment using the Yahoo! crowdsourcing company. Participants were paid 55 yen after completing all tasks as a reward for participating. A website was created for the experiment, which was limited to using a PC.

4.2.2 Participants

There were a total of 596 participants. However, there were 18 participants who gave inappropriate responses, which were eliminated as erroneous data, leaving a total of 578 participants. To judge whether answers were inappropriate in the experiment, we judged answers as inappropriate when the changes in the empathy values before and after the video were the same for all items or when only one item changed [89, 90]. The task of aligning the number of participants to an appropriate number for the analysis was performed, and 142 participants in each condition were included in the analysis,

starting from the top in the order of their participation in the experiment. Thus, the total number of participants used in the analysis was 568. We then applied Cronbach's α coefficient to the remaining 568 participants to determine the reliability of the questionnaire responses, which was found to be between 0.8001 and 0.8777 in all conditions.

The average age was 48.32 years (standard deviation 11.02), with a minimum of 18 years and a maximum of 87 years. The gender breakdown was 344 males and 224 females.

4.2.3 Analysis Result

All 12 questionnaire items were analyzed together. We also categorized and analyzed affective and cognitive empathy. For multiple comparisons, Holm's multiple comparison test was used to examine the existence of significant differences. Table 4.2 shows the results of the overall analysis. The results of the questionnaire analysis showed a main effect of task difficulty on affective empathy. The results are shown in Figure 4.4.

Initially, as can be seen from Table 4.2, we examined differences in empathy toward the agent among the participants and found no differences between any of the conditions. Therefore, we assumed that the ability to empathize with the agent was similar among the participants. In this study, the questionnaire on pre-task empathy toward the agent was given only to confirm that there were no differences between participants. Therefore, we do not discuss significant differences between the pre- and post-task cases.

The post-task results showed no interaction between task difficulty and task content, regardless of empathy category. For the 12 items, there was also no main effect of each factor [$F(1,564) = 2.4737$]. Similarly, no main effect of task content was found [$F(1,564) = 0.5918$]. However, a main effect was found for task difficulty [$F(1,564) = 4.0986$] based on the analysis of affective

TABLE 4.2: Results of all analyses of variance

Category		Conditions	Mean	S.D.	Factor	F	p	η_p^2
Empathy (Q1-Q12)	pre	high-competitive	38.77	6.068	difficulty	0.7877	0.3752 *ns*	0.0014
		high-cooperative	38.44	5.695	content	0.6676	0.4142 *ns*	0.0012
		low-competitive	38.40	6.101	interaction	0.0198	0.8880 *ns*	0.0000
		low-cooperative	37.92	5.956				
	post	high-competitive	37.75	6.585	difficulty	2.474	0.1163 *ns*	0.0044
		high-cooperative	37.82	6.098	content	0.1053	0.7456 *ns*	0.0002
		low-competitive	37.11	7.339	interaction	0.1909	0.6624 *ns*	0.0003
		low-cooperative	36.68	6.810				
Affective empathy (Q1-Q6)	pre	high-competitive	19.49	3.602	difficulty	2.267	0.1327 *ns*	0.0040
		high-cooperative	19.54	3.606	content	0.1374	0.7110 *ns*	0.0002
		low-competitive	19.19	3.728	interaction	0.2839	0.5944 *ns*	0.0005
		low-cooperative	18.92	3.552				
	post	high-competitive	18.87	3.880	difficulty	4.099	0.0434 *	0.0072
		high-cooperative	18.96	3.548	content	0.1855	0.6668 *ns*	0.0003
		low-competitive	18.44	4.164	interaction	0.5362	0.4643 *ns*	0.0009
		low-cooperative	18.06	3.968				
Cognitive empathy (Q7-Q12)	pre	high-competitive	19.29	3.259	difficulty	0.0027	0.9584 *ns*	0.0000
		high-cooperative	18.90	3.033	content	1.203	0.2733 *ns*	0.0021
		low-competitive	19.21	3.211	interaction	0.1152	0.7344 *ns*	0.0002
		low-cooperative	19.01	3.346				
	post	high-competitive	18.89	3.290	difficulty	0.5918	0.4421 *ns*	0.0010
		high-cooperative	18.85	3.412	content	0.0208	0.8854 *ns*	0.0000
		low-competitive	18.67	3.709	interaction	0.0006	0.9808 *ns*	0.0000
		low-cooperative	18.62	3.539				

p: $^*p<0.05$ $^{**}p<0.01$ $^{***}p<0.001$

empathy in Table 4.2. The main effect of post-task task-difficulty was higher for affective empathy for the higher difficulty level (high difficulty: mean = 18.9155, S.D. = 3.7113; low difficulty: mean = 18.2535, S.D. = 4.0647), as shown in Figure 4.4. On the basis of the above analysis, the results of this experiment suggest that a higher task difficulty promotes affective empathy.

The post-task values were lower than the pre-task values of empathy toward the agent in each condition (all pre-task: mean = 38.3838, S.D. = 5.9490; all post-task: mean = 37.3415, S.D. = 6.7214). Also, for affective empathy, pre-task affective empathy values were higher (all pre-task: mean = 19.2817, S.D. = 3.6216; all post-task: mean = 18.5845, S.D. = 3.9027). Similarly, for

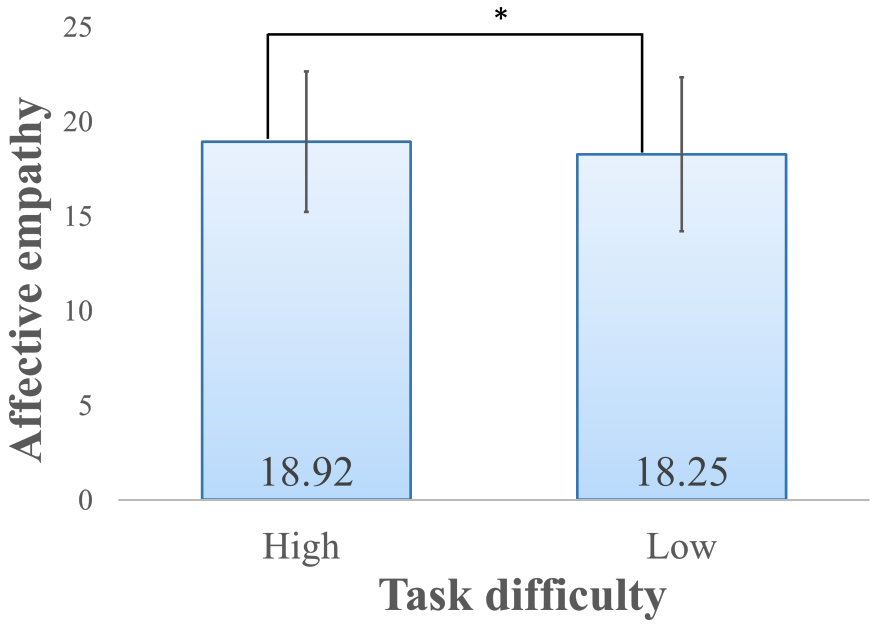

FIGURE 4.4: Main effects results of affective empathy

cognitive empathy, pre-task cognitive empathy values were higher (all pre-task: mean = 19.1021, S.D. = 3.2094; all post-task: mean = 18.7570, S.D. = 3.4835).

Since a main effect was found for task difficulty in affective empathy, we conducted an analysis by category of affective empathy. Table 4.3 shows the results of the category of affective empathy analysis. The results showed that only personal distress had a main effect on task difficulty [$F(1,564) = 5.2007$]. The main effect of post-task task-difficulty was higher for affective empathy for the higher difficulty level (high difficulty: mean = 9.2148, S.D. = 2.5428; low difficulty: mean = 8.7148, S.D. = 2.6769), as shown in Figure 4.5.

4.3 Discussion

This experiment was conducted to investigate the conditions necessary for a human to develop empathy for an anthropomorphic agent. In particular, this

TABLE 4.3: Results of affective empathy analyses of variance

Category		Conditions	Mean	S.D.	Factor	F	p	η_p^2
Personal distress (Q1-Q3)	pre	high-competitive	9.521	2.600	difficulty	2.524	0.1127 *ns*	0.0045
		high-cooperative	9.620	2.684	content	0.0930	0.7606 *ns*	0.0002
		low-competitive	9.338	2.582	interaction	0.5688	0.4510 *ns*	0.0010
		low-cooperative	9.106	2.592				
	post	high-competitive	9.239	2.526	difficulty	5.201	0.0229 *	0.0091
		high-cooperative	9.190	2.568	content	0.7521	0.3862 *ns*	0.0013
		low-competitive	8.880	2.680	interaction	0.4127	0.5209 *ns*	0.0007
		low-cooperative	8.549	2.673				
Empathic concern (Q4-Q6)	pre	high-competitive	9.965	1.892	difficulty	0.4410	0.5069 *ns*	0.0008
		high-cooperative	9.916	1.885	content	0.0776	0.7807 *ns*	0.0001
		low-competitive	9.852	2.025	interaction	0.0005	0.9829 *ns*	0.0000
		low-cooperative	9.810	2.028				
	post	high-competitive	9.627	2.169	difficulty	0.8457	0.3582 *ns*	0.0015
		high-cooperative	9.775	1.888	content	0.0783	0.7797 *ns*	0.0001
		low-competitive	9.563	2.168	interaction	0.3133	0.5759 *ns*	0.0006
		low-cooperative	9.514	2.156				

p: $^*p<0.05$ $^{**}p<0.01$ $^{***}p<0.001$

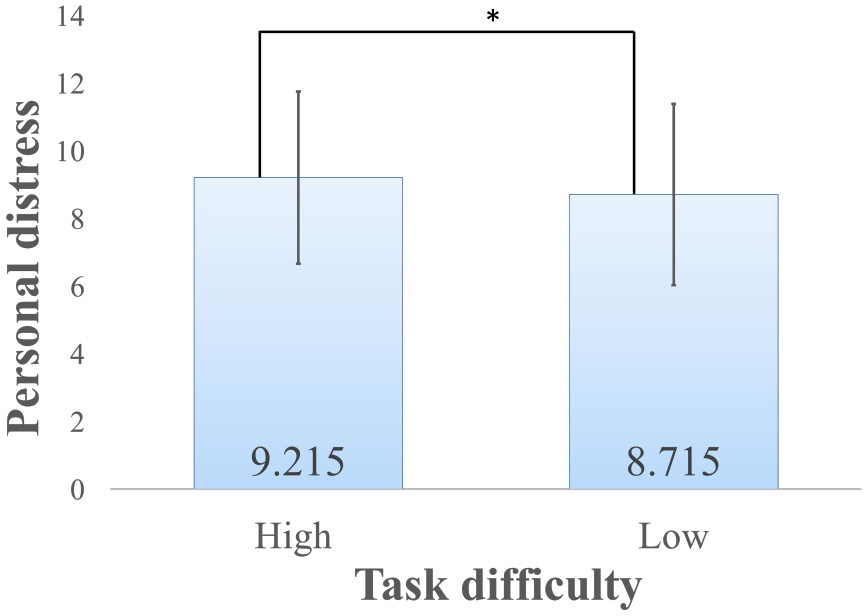

FIGURE 4.5: Main effects results of personal distress

experiment aimed to identify factors that influence the empathy between an agent and a person who performed a task by investigating factors related to the task. For this purpose, we formulated the following two hypotheses and analyzed the data obtained from the experiment.

The results did not support the two hypotheses. However, the results showed that regardless of task content, a higher task difficulty promoted human affective empathy. In discussing these results, it is necessary to focus on the changes in empathy before and after the task.

If only post-task surveys had been conducted, a higher task difficulty would have been shown to increase human empathy. However, by conducting a pre-task survey, it was found that empathy actually decreased throughout the task. This result indicated that post-task changes do not necessarily lead to better results than pre-task changes.

In addition, the fact that empathy for the agent decreased with the task in this experiment indicates that the task may decrease empathy. However, as a limitation of this experiment, it is possible that the simplicity of the task itself may have decreased empathy because the task itself was perceived as tedious, due to eliminate factors other than task difficulty and task content.

However, the fact that task difficulty did not affect the task content but affected human affective empathy may be an effective factor in controlling human empathy when empathic agents coexist in human society in the future. By setting the task difficulty appropriately, it is possible to maintain an appropriate distance from an agent without making the participant empathize more than necessary.

The results of the analysis, which classified empathy into affective empathy and cognitive empathy, showed a main effect of task difficulty on affective empathy. The main reason for this main effect is thought to be that affective empathy increased as task difficulty increased due to the increased mental

load caused by the task. This is related to personal distress, which is classified as affective empathy. Affective empathy is the feeling of the emotional state that others are experiencing or about to experience, which leads to the same emotional state in oneself. Only affective empathy was enhanced because the mental load from the task affected the emotional states of both the participant and agent. No main effects of task difficulty or task content were observed for cognitive empathy. This is because cognitive empathy requires imagining the thoughts and feelings of others in terms of oneself and imagining them from the other's point of view, so the task in this experiment did not enhance cognitive empathy.

There was also a difference in the amount of time participants spent completing the task with the agent at high and low difficulty levels, depending on the task difficulty setting. This may have influenced affective empathy.

4.4 Conclusions

This study focused on human-agent tasks as part of the factors and conditions that make humans empathize with anthropomorphic agents, and it investigated task difficulty and task content. Two hypotheses were formulated and tested. The results did not support either of the two hypotheses. Task difficulty was found to have a significant effect on affective empathy. The analysis revealed that a higher task difficulty increased emotional empathy after the task. The task itself can be an important factor when manipulating each category of empathy. Future research may consider the development of an agent that empathizes with humans and that is suitable for a task since it was confirmed that empathy held by humans decreases.

Chapter 5

Maintain empathy for agent through agent's expression

5.1 Methods

In Chapter 4, human empathy for the agent did not strongly affect task difficulty and task content. Therefore, in this study, we focus on an agent's expression as a way to increase the empathy felt by humans. Specifically, we experimentally investigate the possibility of an agent's expression facilitating human empathy. In addition, Chapter 4 did not aim to investigate changes in empathy before and after the task. This study defined the agent's expression as a factor other than appearance. Therefore, in this experiment, pre- and post-task are considered as an additional factor. Experiments were conducted utilizing a three-way mixed design in which the factors were the agents' expression (available, not available), task completion (success, failure), and empathy evaluation (before, after a task).

5.1.1 Experimental Purpose and Design

Our objective in this study is to determine whether the success or failure of a task and the agent's expression during the task elicits more human empathy as a result of the human's interactions with the agent. We designed a task based on a typing game and recruited participants for an experiment in which the empathy agent interacted during the task by expressing itself to the participants. We came up with the following two hypotheses.

H1: Successful completion of a task has less impact on human empathy than failure of a task.

H2: When an empathy agent makes an expression, it promotes empathy regardless of the success or failure of the task.

We designed a between-participants experiment with two factors: task completion and agent expression. The number of levels for each factor was two for task completion (success, failure) and two for agent expression (available, not available). The agent's expression is a factor other than the agent's appearance. Two levels of pre- and post-task empathy values were used as within-participant factors. The dependent variable was the empathy that the participants felt.

We focused on an agent's expression for two reasons. First, we wanted to investigate whether the agent's expression affects empathy even when the agent is not directly involved in a task. In Chapter 4, the agent's expression (behavior and comments) was not a factor and was a common condition, which means the results did not show statistical significance related to the influence of the agent's expression. Second, we wanted to investigate whether the agent's expression influences human motivation for the task. We also focused on task completion because the effect of the agent's expression could change depending on whether the task was a success or failure. Chapter 4

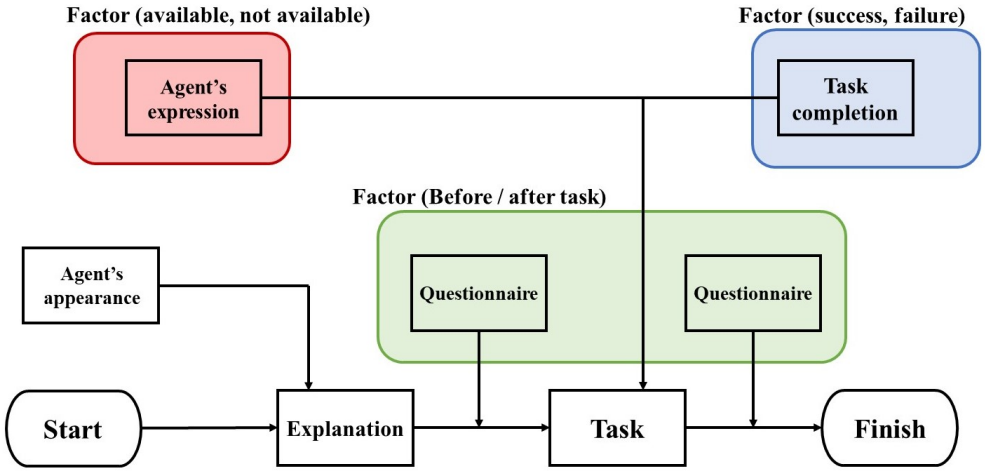

FIGURE 5.1: Process flow of the experiment.

did not focus on the results of task completion, but since it is highly likely for the success or failure of a task to affect the psychological state of humans, it is possible there was a change in human empathy for the agent.

5.1.2 Experimental Details

The experiment was conducted online using a PC. The experiment was conducted in an online environment, which is an increasingly common method of experimentation [86–88]. Figure 5.1 shows the process flow.

As our primary goal is to promote human empathy toward anthropomorphic agents, we administered a questionnaire to measure empathy toward the agent before participants performed the task. This is the part of the explanation in Figure 5.1. The participants were told that the agent had the role of watching over the task.

The task was a typing game in which participants were asked to type 150 random letters of the alphabet. Our aim was to determine the standard

typing time for the success or failure of the task, which is a factor in this experiment.

The agent expressed itself in various ways before and after the typing game started and ended. Specifically, before the start, it made cheering comments and gestures, and after the end, it made praising comments and gestures if the task was a success, and encouraging comments and gestures if the task was a failure. The agent did not make any gestures or expressions during the game so as not to affect the typing (i.e., it did nothing). After the task was completed, we administered the same questionnaire as before the task and additionally asked participants about their motivation to continue the task in order to investigate their empathic response. After completing the questionnaire, participants voluntarily answered the free-response questions.

5.1.3 Participants

Participants were recruited through Yahoo Crowdsourcing and paid 32 yen (= 0.22 dollars). The original number of participants was 398, but we excluded 35 who either gave inappropriate responses (i.e., the change in empathy values before and after the task was the same for all items, or only changed for one item [89, 90]) or seemed unmotivated due to slow typing speed (less than one letter in two seconds).

We then applied Cronbach's α coefficient to the remaining 363 participants to determine the reliability of the questionnaire responses, which was found to be between 0.8027 and 0.9025 in all conditions. Sixty-eight participants in each condition were included in the analysis in the order of participation. At this time, the gender of the analyzed participants was adjusted so that there was no difference between conditions. Thus, the total number of participants used in the analysis was 272.

TABLE 5.1: Summary of questionnaire

Affective empathy
Personal distress
Q1: If an emergency happens to the character, you would be anxious and restless.
Q2: If the character is emotionally disturbed, you would not know what to do.
Q3: If you see the character in need of immediate help, you would be confused and would not know what to do.
Empathic concern
Q4: If you see the character in trouble, you would not feel sorry for that character.
Q5: If you see the character being taken advantage of by others, you would feel like you want to protect that character.
Q6: The character's story and the events that have taken place move you strongly.
Cognitive empathy
Perspective taking
Q7: You look at both the character's position and the human position.
Q8: If you were trying to get to know the character better, you would imagine how that character sees things.
Q9: When you think you're right, you don't listen to what the character has to say.
Fantasy scale
Q10: You are objective without being drawn into the character's story or the events taken place.
Q11: You imagine how you would feel if the events that happened to the character happened to you.
Q12: You get deep into the feelings of the character.
Empathic response
BeQ: Do you want to continue to do tasks with the character in the future?

The average age was 47.56 years (standard deviation 10.76), with a minimum of 19 years and a maximum of 75. The gender was 136 males and 136 females.

5.1.4 Questionnaire

Participants answered questionnaires before and after the task. This was a 12-item questionnaire modified from the IRI. As the IRI was designed to investigate human empathy characteristics, we modified it to investigate empathy toward an empathy agent. The same questionnaire was applied both before and after the task and was administered on a 5-point Likert scale (1: not applicable, 5: applicable), as shown in Table 5.1. Q4, Q9, and Q10 are inverted items, so the scores were reversed when analyzing them.

Q1 to Q6 examine affective empathy, and Q7 to Q12 examine cognitive empathy. There is one additional item for the questionnaire administered after the task (BeQ in the table) to examine the empathic response of the participants.

(a) Expression of the agent before typing

(b) Expression of the agent after typing: Success

(c) Expression of the agent after typing: Failure

FIGURE 5.2: Three expressions of the agent.

TABLE 5.2: Three expressions of the agent's speech.

Before task	After task(Success)	After task(Failure)
	Thanks for the typing.	It is time to finish the task.
You begin the experiment.	The end means ...	Thanks for the typing.
Let's do it!	You did it! Successful task!	Better luck next time!
	Congratulations!	Don't worry too much about it.

5.1.5 Agent's expression

We prepared three types of expression for the agent, as shown in Figure 5.2 and Table 5.2. Note that we did not make any special considerations for the agent's appearance: it was simply made to look robotic, since we wanted to reduce any impression based on gender.

Before the participant starts typing, the agent cheers in encouragement, as shown in figure 5.2(a). After typing, if the participant succeeds in the task, the agent expresses that it is pleased, as shown in Figure 5.2(b). If the participant fails, the agent makes a consoling expression, as shown in Figure 5.2(c).

The agent does not make any expressions while the participant is typing so as to allow him or her to concentrate on the task. It is standing in an upright state at this time. To confirm, participants are asked to indicate whether they noticed the agent's expression, as well as whether or not they were able to finish the task. The agent was MikuMikuDance (MMD), a freely available graphics software. For the models, we used our own appearance.

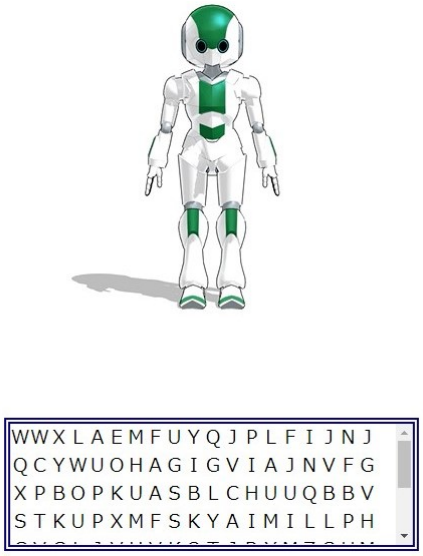

FIGURE 5.3: Screenshots of the agent and the typing game during the experiment.

5.1.6 Task completion

The experimental task is a typing game in which 150 random letters of the alphabet are entered. Since it is conducted online, there was a possibility that some participants might abandon the experiment if they knew there was a time limit. We therefore conducted the experiment without showing the time limit and simply terminated the task when the time was up. Figure 5.3 shows screenshots of the agent and the typing game during the experiment.

We conducted a pre-experiment to determine the amount of time to allow for successful completion of the task. A total of 50 participants were recruited for 20 yen (= 0.14 dollars). The mean age was 43.54 years (standard deviation 9.467) with a minimum of 22 years and a maximum of 65, and there were 46 men and four women. Results showed that the mean time to complete typing was 2 minutes and 4 seconds, with a standard deviation of 50 seconds, so the experiment was conducted with a standard deviation of ± 1 as typing success.

5.1.7 Analysis Method

The analysis was a three-factor mixed design ANOVA. The between participants factors were two levels of agent expression (available, not available) and two levels of task completion (success, failure). The within-participants factor consisted of two levels of empathy values before and after the task.

With the results of the participants' questionnaire responses as a basis, we investigated the influence of agent expression and task completion as factors that elicit human empathy. The values of empathy before and after the task were utilized as the dependent variable. R (ver. 4.1.0) was used for the ANOVA.

5.2 Results

We analyzed the questionnaire responses using ANOVA and classified empathy into affective and cognitive empathy categories. In addition, we analyzed empathy characteristics, which further categorize affective and cognitive empathy. We show statistical information for each empathy category in Table 5.3. The raw data of the participants used in this study were made available on github[1] and in supplemental information. The ANOVA results are shown in Table 5.4, where we can see there was an interaction between the two factors of agent expression and the pre- and post-task. The results of our analysis of the interaction are shown in Figure 5.4, which also shows the mean and standard deviation for each condition.

No significant interactions between the agent's expression and task completion were found in any of the conditions. In the following, we omit the explanation of items for which a main effect and interaction were found, and for the items for which no interaction was found and a main effect was found,

[1] https://github.com/TakahiroTsumura/IEEE-Access-2023-Mar

TABLE 5.3: Results of participants' empathy statistical information

Category		Conditions	Mean	S.D.	Category		Conditions	Mean	S.D.
		available-success	36.49	7.160			available-success	9.618	2.144
		available-failure	38.94	5.604			available-failure	10.34	1.698
Empathy	pre	not available-success	37.31	6.082	Empathic	pre	not available-success	9.735	2.113
		not available-failure	37.74	6.512	concern		not available-failure	9.765	2.306
		available-success	36.40	7.442			available-success	9.574	2.139
(Q1-Q12)		available-failure	38.60	6.000	(Q4-Q6)		available-failures	10.28	1.835
	post	not available-success	35.41	6.846		post	not available-success	9.353	2.238
		not available-failure	35.44	7.854			not available-failure	9.382	2.344
		available-success	18.09	4.164			available-success	9.941	2.178
		available-failure	19.50	3.593			available-failure	10.13	1.692
Affecctive	pre	not available-success	18.32	3.923	Perspective	pre	not available-success	10.28	1.582
empathy		not available-failure	18.68	3.731	taking		not available-failure	10.13	2.065
		available-success	17.91	4.231			available-success	9.853	2.307
(Q1-Q6)		available-failure	18.96	3.799	(Q7-Q9)		available-failure	10.37	1.573
	post	not available-success	17.25	3.911		post	not available-success	9.897	1.971
		not available-failure	17.54	4.244			not available-failure	9.559	2.161
		available-success	18.40	3.674			available-success	8.456	1.880
		available-failure	19.44	2.662			available-failure	9.309	1.695
Cognitive	pre	not available-success	18.99	2.970	Fantasy	pre	not available-success	8.706	1.901
empathy		not available-failure	19.06	3.489	scale		not available-failure	8.927	1.942
		available-success	18.49	3.865			available-success	8.632	1.962
(Q7-Q12)		available-failure	19.65	2.774	(Q10-Q12)		available-failure	9.279	1.665
	post	not available-success	18.16	3.441		post	not available-success	8.265	2.005
		not available-failure	17.90	3.986			not available-failure	8.338	2.141
		available-success	8.471	2.772			available-success	3.250	0.9364
		available-failure	9.162	2.646	Empathic		available-failure	3.485	0.7430
Personal	pre	not available-success	8.588	2.547	response	post	not available-success	2.971	0.9138
distress		not available-failure	8.912	2.490	(BeQ)		not available-failure	3.015	0.8893
		available-success	8.338	2.800					
(Q1-Q3)		available-failure	8.677	2.628					
	post	not available-success	7.897	2.325					
		not available-failure	8.162	2.483					

we present the results for the main effect. Therefore, since an interaction between an agent' s expression and the pre- and post-task was found, the results of the analysis of the simple main effect are shown in Table 5.5. The analysis results also showed an interaction between the categories of cognitive empathy, so this is shown in Table 5.5 as well.

5.2.1 Empathy

The results for empathy (Q1-Q12) revealed an interaction between the agent's expression and the pre- and post-task factors. The main effects of pre- and

TABLE 5.4: Analysis results of ANOVA

	Factor	F	p	η_p^2
Empathy (Q1-12)	Agent's expression	2.111	0.1474 *ns*	0.0078
	Task completion	2.695	0.1019 *ns*	0.0100
	Before/after task	22.99	0.0000 ***	0.0790
	Agent's expression × Task completion	1.820	0.1785 *ns*	0.0067
	Agent's expression × Before/after task	15.28	0.0001 ***	0.0539
	Task completion × Before/after task	0.4515	0.5022 *ns*	0.0017
	Agent's expression × Task completion × Before/after task	0.0233	0.8787 *ns*	0.0001
Affective empathy (Q1-6)	Agent's expression	2.164	0.1424 *ns*	0.0080
	Task completion	2.941	0.0875 *ns*	0.0109
	Before/after task	20.96	0.0000 ***	0.0725
	Agent's expression × Task completion	0.9994	0.3184 *ns*	0.0037
	Agent's expression × Before/after task	5.400	0.0209 *	0.0198
	Task completion × Before/after task	0.4452	0.5052 *ns*	0.0017
	Agent's expression × Task completion × Before/after task	0.2335	0.6294 *ns*	0.0009
Cognitive empathy (Q7-12)	Agent's expression	1.429	0.2330 *ns*	0.0053
	Task completion	1.663	0.1984 *ns*	0.0062
	Before/after task	10.90	0.0011 **	0.0391
	Agent's expression × Task completion	2.353	0.1262 *ns*	0.0087
	Agent's expression × Before/after task	19.80	0.0000 ***	0.0688
	Task completion × Before/after task	0.1854	0.6671 *ns*	0.0007
	Agent's expression × Task completion × Before/after task	0.7919	0.3743 *ns*	0.0029
Personal distress (Q1-3)	Agent's expression	0.8788	0.3494 *ns*	0.0033
	Task completion	1.942	0.1646 *ns*	0.0072
	Before/after task	18.29	0.0000 ***	0.0639
	Agent's expression × Task completion	0.1444	0.7042 *ns*	0.0005
	Agent's expression × Before/after task	2.926	0.0883 *ns*	0.0108
	Task completion × Before/after task	0.7315	0.3932 *ns*	0.0027
	Agent's expression × Task completion × Before/after task	0.3732	0.5418 *ns*	0.0014
Empathic concern (Q4-6)	Agent's expression	2.644	0.1051 *ns*	0.0098
	Task completion	2.356	0.1260 *ns*	0.0087
	Before/after task	6.612	0.0107 *	0.0241
	Agent's expression × Task completion	1.997	0.1587 *ns*	0.0074
	Agent's expression × Before/after task	3.846	0.0509 *ns*	0.0141
	Task completion × Before/after task	0.0019	0.9653 *ns*	0.0000
	Agent's expression × Task completion × Before/after task	0.0019	0.9653 *ns*	0.0000
Perspective taking (Q7-9)	Agent's expression	0.2249	0.6357 *ns*	0.0008
	Task completion	0.0602	0.8064 *ns*	0.0002
	Before/after task	6.893	0.0092 **	0.0251
	Agent's expression × Task completion	1.755	0.1864 *ns*	0.0065
	Agent's expression × Before/after task	12.82	0.0004 ***	0.0456
	Task completion × Before/after task	0.1846	0.6678 *ns*	0.0007
	Agent's expression × Task completion × Before/after task	2.791	0.0959 *ns*	0.0103
Fantasy scale (Q10-12)	Agent's expression	2.812	0.0947 *ns*	0.0104
	Task completion	4.358	0.0378 *	0.0160
	Before/after task	6.781	0.0097 **	0.0247
	Agent's expression × Task completion	1.969	0.1617 *ns*	0.0073
	Agent's expression × Before/after task	12.06	0.0006 ***	0.0430
	Task completion × Before/after task	1.085	0.2985 *ns*	0.0040
	Agent's expression × Task completion × Before/after task	0.0301	0.8623 *ns*	0.0001
Empathic response (BeQ)	Agent's expression	12.52	0.0005 ***	0.0446
	Task completion	1.738	0.1885 *ns*	0.0064
	Agent's expression × Task completion	0.8136	0.3679 *ns*	0.0030

p: $^*p<0.05$ $^{**}p<0.01$ $^{***}p<0.001$

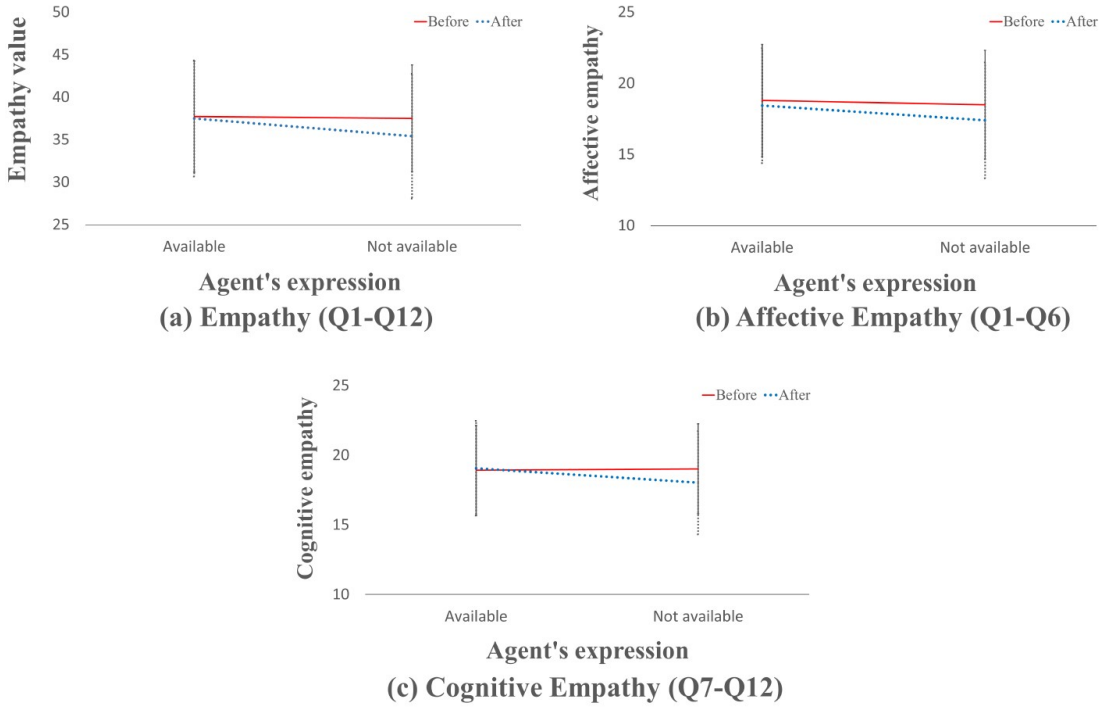

FIGURE 5.4: Interaction results for (a) empathy, (b) affective empathy, and (c) cognitive empathy.

post-task were also significant, but were omitted because of the interaction between the agent's expression and the pre- and post-task factors.

Simple main effect revealed a significant difference of the post-task agent's expression factor, as shown in Figure 5.5(a) (Available: mean = 37.50, S.D. = 6.825; Not available: mean = 35.43, S.D. = 7.340). In addition, significant differences were found in the simple main effects of the pre- and post-task factors in the non-expression condition of the agent. These results suggest that empathy was induced in the task with the agent's expression, while it was suppressed in the task without the agent's expression. The results of post-hoc analysis indicate that agent expression is effective in inducing empathy.

TABLE 5.5: Analysis results of simple main effect

	Factor	F	p	η_p^2
Empathy (Q1-12)	Agent's expression at before task	0.0613	0.8046 *ns*	0.0002
	Agent's expression at after task	5.849	0.0163 *	0.0214
	Before/after task at agent's expression available	0.5485	0.4602 *ns*	0.0041
	Before/after task at agent's expression not available	29.48	0.0000 ***	0.1803
Affective empathy (Q1-6)	Agent's expression at before task	0.3951	0.5302 *ns*	0.0015
	Agent's expression at after task	4.454	0.0357 *	0.0163
	Before/after task at agent's expression available	3.861	0.0515 *ns*	0.0280
	Before/after task at agent's expression not available	17.76	0.0000 ***	0.1170
Cognitive empathy (Q7-12)	Agent's expression at before task	0.0693	0.7925 *ns*	0.0003
	Agent's expression at after task	5.806	0.0166 *	0.0212
	Before/after task at agent's expression available	0.7005	0.4041 *ns*	0.0052
	Before/after task at agent's expression not available	28.36	0.0000 ***	0.1747
Perspective taking (Q7-9)	Agent's expression at before task	0.5412	0.4626 *ns*	0.0020
	Agent's expression at after task	2.432	0.1200 *ns*	0.0090
	Before/after task at agent's expression available	0.4790	0.4901 *ns*	0.0036
	Before/after task at agent's expression not available	18.36	0.0000 ***	0.1205
Fantasy scale (Q10-12)	Agent's expression at before task	0.0864	0.7691 *ns*	0.0003
	Agent's expression at after task	7.652	0.0061 **	0.0278
	Before/after task at agent's expression available	0.4442	0.5063 *ns*	0.0033
	Before/after task at agent's expression not available	16.03	0.0001 ***	0.1068

p: $^*p<0.05$ $^{**}p<0.01$ $^{***}p<0.001$

5.2.2 Affective empathy

The results for affective empathy (Q1-Q6) revealed an interaction between the agent's expression and the pre- and post-task factors, the same as with empathy. The main effect of pre- and post-task was also significant, but was omitted because of the interaction between the agent's expression and the pre- and post-task factors. The analysis also showed no interaction between the categories of affective empathy (personal distress and empathic concern).

Simple main effect showed a significant difference of the post-task agent's expression factor, as shown in Figure 5.5(b) (Available: mean = 18.43, S.D. = 4.040; Not available: mean = 17.40, S.D. = 4.069). In addition, significant differences were found in the simple main effects of the pre- and post-task factors in the non-expression condition of the agent. These results suggest that affective empathy is induced in the presence of the agent's expression,

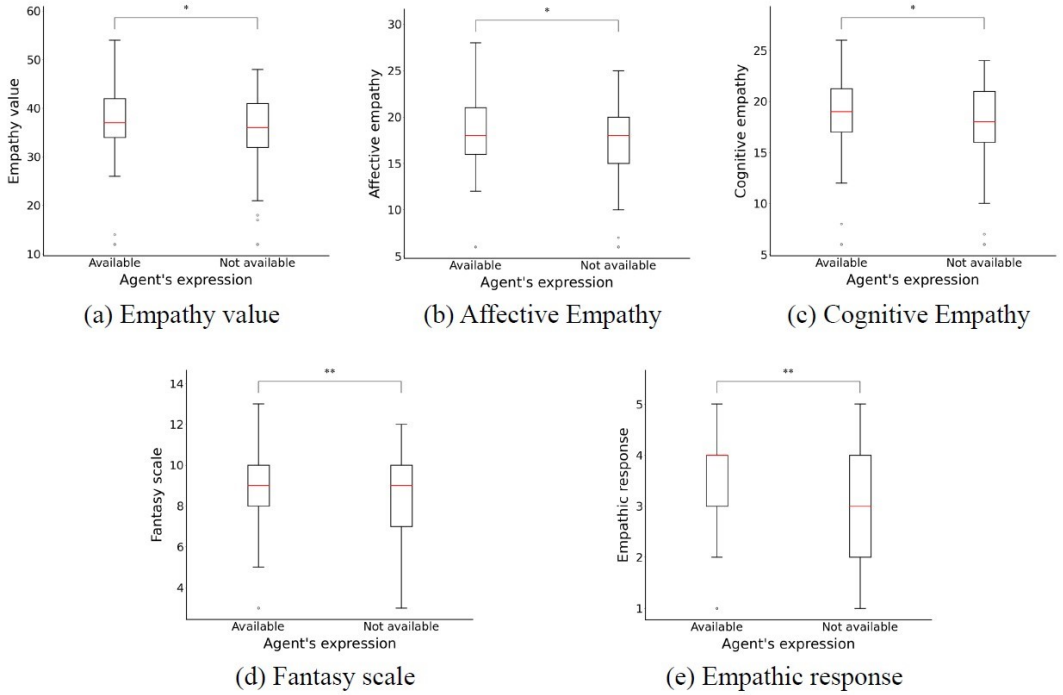

FIGURE 5.5: Results of post-task (a) empathy value (b) affective empathy (c) cognitive empathy (d) fantasy scale (e) empathic response represented by box plots. Red lines are medians and the circles are outliers.

while it is suppressed in the presence of the agent's non-expression. The results of post-hoc analysis indicate that the agent's expression is effective in inducing affective empathy.

5.2.3 Cognitive empathy

The results for cognitive empathy (Q7-Q12) showed an interaction between agent expression and the pre- and post-task factors, as was the case for empathy. The main effect of pre- and post-task was also significant, but was omitted because of the interaction between the agent's expression and the pre- and post-task factors. The analysis also revealed an interaction between the categories of cognitive empathy (perspective taking and fantasy scale).

For this reason, the results of the analysis are also presented for these classifications.

Simple main effect revealed a significant difference of the post-task agent's expression factor, as shown in Figure 5.5(c) (Available: mean = 19.07, S.D. = 3.402; Not available: mean = 18.03, S.D. = 3.712). In addition, significant differences were found in the simple main effects of the pre- and post-task factors in the non-expression condition of the agent. These results suggest that cognitive empathy is induced in the presence of the agent's expression, while it is suppressed in the presence of the agent's no expression. The results of post-hoc analysis indicate that the agent's expression is effective in inducing cognitive empathy.

There was an interaction between Agent' s expression and Before/after task in perspective taking and fantasy scale, but the results of simple main effect showed that perspective taking was significant before and after the task when there was no expression of the agent. The simple main effect results for fantasy scale also showed a significant difference of the agent's expression factor after the task, as shown in Figure 5.5(d) (Available: mean = 8.956, S.D. = 1.841; Not available: mean = 8.302, S.D. = 2.067). In addition, the results for the agent's no-expression condition showed significant differences in the simple main effects of the pre- and post-task factors. These results demonstrate that in terms of cognitive empathy, fantasy scale increases empathy with agent expression and suppresses empathy in the absence of agent expression. In addition, the results of the post-hoc analysis indicate that the agent's expression is effective in eliciting perspective taking and fantasy scale.

5.2.4 Empathic response

Investigating empathy for an agent is not enough to determine the overall effects of empathy, so we also surveyed empathic response. Our findings showed there was no interaction between the agent's expression and task

completion in the empathic response results, nor was there a main effect of the task completion.

However, the main effect of the agent's expression was significant, as shown in Figure 5.5(e) (Available: mean = 3.368, S.D. = 0.8472; Not available: mean = 2.993, S.D. = 0.8952), which indicates that the agent's expression can increase the participants' motivation for the task. This result is important in terms of connecting the relationship between empathy for the agent and motivation for the task.

5.3 Discussion

5.3.1 Supporting hypotheses

We conducted our experiment to clarify the conditions necessary for humans to empathize with anthropomorphic agents. Specifically, we investigated the possibility of influencing the empathy that humans have for agents by means of the agent's expression and the success or failure of a task as humans and agents interact during a task. We proposed two hypotheses and analyzed the data obtained from the experiment to see if they were supported.

Our findings showed that H1, "Successful completion of a task has less impact on human empathy than failure of a task," was not supported, as the experimental results did not show that task completion had any effect on human empathy. Success and failure are important factors in the task, and we predicted that participants' empathy for the agent might differ depending on the outcome, but no significant differences were found. One possible reason for this could be that participants performed the experimental task without knowing what the time limit was. This might have caused them not to empathize due to the sudden end of the task.

On the other hand, these results support H1 only in part. Specifically, for fantasy, a category of cognitive empathy, we identified a mean of 8.515 and a standard deviation of 3.728 for task success. For task failure, the mean was 8.963 with a standard deviation of 3.602. These results indicate that empathy was higher during task failure than during task success, which leads us to conclude that task success or failure affects fantasy regardless of the agent's expression. In particular, it is possible that when the agent fails a task, the participants feel closer to the agent through fantasizing.

In contrast, H2, "When an empathy agent makes an expression, it promotes empathy regardless of the success or failure of the task," was supported. Although the agent did not do anything special during the typing process, its expression before and afterwards resulted in a maintained human empathy regardless of the task outcome. This result indicates that the agent's expression was effective in inducing and maintaining human empathy regardless of task outcome.

5.3.2 The influence of the task and the importance of the agent's expression

In this study, in addition to investigating differences in task completion, we also focused on pre- and post-task differences to see the influence of agent expression, which was not the case in Chapter 4. The results showed that, as in Chapter 4, human empathy toward the agent decreased statistically significantly in the absence of agent expression. On the other hand, human empathy toward the agent was maintained in the presence of agent expression. Studies by Beck et al. [48] and Deshmukh et al. [59] have shown that agent expressions (gestures and comments) can enhance understanding from humans, and our study also elicited human empathy toward the agent.

Our findings suggest that when a human and an agent perform a task, the minimum necessary behaviors and words are not enough to elicit empathy

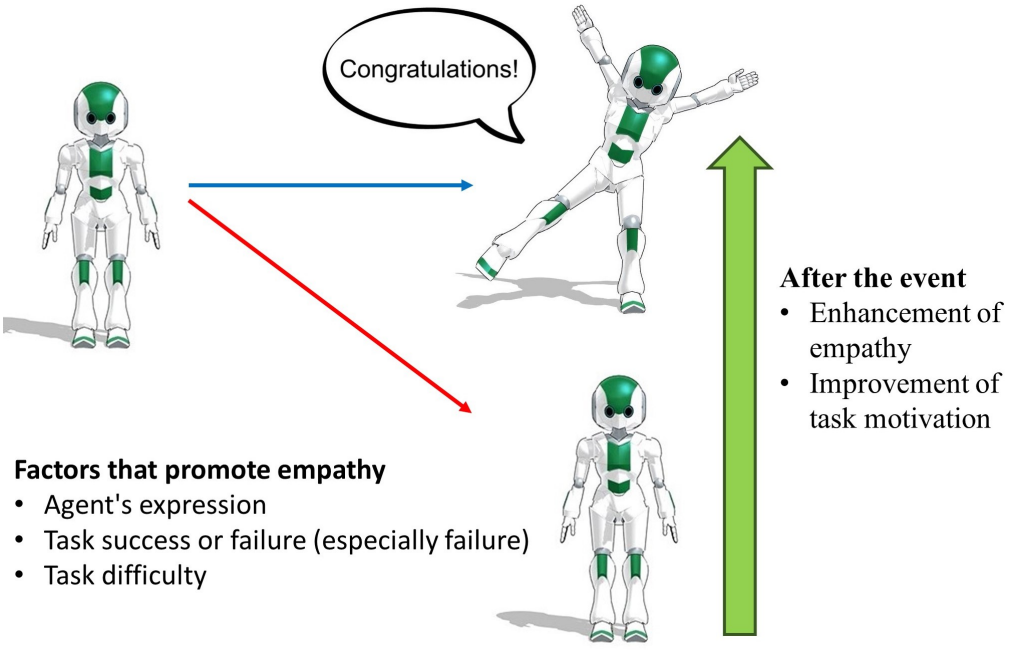

FIGURE 5.6: A mechanism for empathy agents to empathize with humans in a task.

from the human. However, adding subtle expressions from the agent can help maintain empathy from humans. Given the above, we summarize the results of our two studies in Figure 5.6. Overall, we were able to clarify under what conditions humans are more likely to empathize with agents by including task and agent factors while focusing on the changes in empathy toward agents before and after the task.

5.3.3 Empathic response

By incorporating empathic responses, which were not sufficiently investigated in Chapter 4, we were able to explore the relationship between changes in empathy for the agent and motivation for the task. This enabled us to determine the agent's effectiveness from empathic responses even when the effect of empathizing with the agent was small. Simultaneous attention to empathy for the agent and the empathic response that emerges as a result of empathy is vital in clarifying the effective role of each.

In this study, participants played the role of observer of the empathic agent. We found that the observers (participants) showed empathic responses to information from the target (appearance, comments, gestures) and to information related to the task (success/failure).

One of our objectives was to determine whether the participants would want to perform the task with the agent in the future, which we measured by analyzing their empathic response behavior. Our analysis revealed a main effect in the agent's expression indicating that when humans perform a simple task, they can maintain their motivation for the task if there is an agent by their side. This finding should prove beneficial for improving motivation in simple tasks.

5.3.4 Limitations

The main limitation is that participants did not know the time limit. We specified this condition in order to avoid cheating in the online experimental environment. When people are actually performing tasks in the real world, the completion deadline is almost always made known. Our future work will therefore investigate whether knowing the time limit affects participants' task completion.

In addition, we used textual comments and gestures as the agent's expression in this study, and did not include voice. It is not possible to control voice in an online environment, which is why we kept it silent. Future work should investigate which has more of an influence on empathy in the agent's expression: voice or text.

Finally, we investigated participants' motivation for the task as an empathic response, and while we found that motivation was increased when an agent's expression occurred, it is impossible to know whether a continuous relationship was actually established without a long-term experiment. Therefore,

it will be necessary to conduct additional experiments in which duration is included as a factor.

We did not use agents with many high-performance features in this study for several reasons. First, we did not use voice as the agent's expression because voice alone affects the agent's impression of humans to the extent that it is a single factor. Therefore, we are considering using voice information as a factor in future research. Second, adding a wide range of agent representations would make it difficult to determine which of those agent' s expressions were effective. Since this would require the addition of unnecessary expressions for use in society, we did not add more representations than necessary in this study. However, as part of future research, differences in agent representation versatility and performance should be investigated.

5.4 Conclusion

If anthropomorphic agents are to become more acceptable to humans, thus enabling their integration into society, it is essential that humans need to be able to empathize with the agents. This study is part of a large body of research focused on the factors that influence empathy between humans and agents. We investigated whether the task affects human empathy toward an agent.

Our study showed that there was no main effect for the task completion factor and that when the agent's expression was present, more human empathy was induced. In addition, the agent's expression was found to affect the participants' motivation for the task.

This study provides a beneficial example of the effectiveness of changes in human empathy in terms of human-agent relationships. In the case of simple tasks, human empathy decreases in the absence of the agent's expression, while the presence of the agent's expression maintains human empathy. Thus,

the agent's expression affected the motivation for the task. Future research could explore empathy agents in a variety of situations by strengthening or weakening specific empathy components.

Chapter 6

General discussion

6.1 Design theory of empathy agents

The benefit of promoting empathy for agents is to make agents more acceptable in human society. For an agent that has the potential to replace a person, it is essential to consider a design that is easily accepted by people. To this end, we focused our research on empathy between people and agents.

As a result of our research on empathy agents, we have identified several elements necessary for designing empathy agents in the interaction between people and agents. First, one of the most important elements as a way to increase empathy for empathy agents is to make the agents self-disclose. It is known that obtaining personal information about others and self-disclosure are effective when humans empathize with others. This effect is also found to be effective for agents by our study. When designing agents, self-disclosure appropriate to the situation can increase empathy from people to agents. In

addition, by designing the agent's expression as well, it is possible to design an agent that is more easily empathized with by people.

In this design, the expected behavior change as a result of the agent's influence on people is the maintenance of altruistic behavior and task motivation. In particular, when designing empathy target agents, it is necessary to pay attention to how people's empathic responses change depending on the interaction between the person and the agent. It is not enough to simply design the agent to be easily accepted by people, but one must consider how the agent will have a positive impact on people.

In our study, empathy for the agent increased people's altruistic behavior and increased the agent's tolerance for mistakes. And even in monotonous tasks, empathizing with the agent increased motivation for the task and increased continuity. These results indicate that when people empathized with the agent, they were able to change their behavior while accepting the agent. Although this study has conducted experiments dealing only with empathy target agents, when designing empathy observer agents. It is necessary to focus on the impact of the agent on people by designing behaviors that make the agent appear to empathize with the person.

Below, we summarize the design of empathic agents to the extent that we can discuss the results of this study. Based on the findings from the three studies, we discuss the expected effects of the use and conditions of empathy agents in several areas. In this section, we discuss three designs of factors that influenced empathy: self-disclosure, task difficulty, and agent expression.

To begin, we summarize how to use agents' self-disclosure when using empathy agents. Based on the findings of this study, self-disclosure can have a similar effect on empathy whether the agent is humanoid or robotic, as long as the agent's appearance is similar to that of a person. However, combining information on appearance with appearance appropriate to preferences

and situations may promote empathy more. Therefore, the design of self-disclosure for empathy agents that can be said from Study 1 is that, as long as appearance is minimal and the agent is in close proximity to a person, self-disclosure related to the situation will make the agent empathize with the person. And the best timing for self-disclosure is to make appropriate self-disclosure after the person is recognized and interested in the situation, and after it is clear that they have a common topic of conversation. Another possible scope for using the self-disclosure design of empathy agents that can be considered from this study would be chatbot-enabled agents and support agents at tourist information centers. These agents need to convey information to people, and since the topics they both seek are common, they can perform their tasks smoothly by performing suitable self-disclosure to make a good impression on people they meet for the first time. As a future research plan, I will investigate the impressions of each attribute of self-disclosure and develop agents that can be used in the real world.

Next, we summarize the design of task difficulty when performing tasks with empathy agents. Based on the findings of this study, when performing a task together, a higher task difficulty level can influence affective empathy. However, in the case of tasks that are monotonous and of low value to the person, empathy toward the agent may decrease regardless of the difficulty level. For this reason, the design of difficulty in tasks with empathic agents that can be said from Study 2 is that agents empathize with people by adding complexity and value to the task, regardless of the type of task. However, this design is largely task-dependent, and based on the findings of this study, the scope of application of task difficulty design is effective only when monotonous and low-value tasks such as simple selection tasks and object movement tasks are performed with agents. Based on the above, in improving the relationship with agents used with people in factories with many monotonous movements, it is better to provide a certain task difficulty level to people and agents to increase empathy and improve the relationship after task completion. As

a future research plan, I will investigate the differences in impressions depending on whether the relationship between tasks is direct or indirect, and reinforce the areas that were not clarified in this study.

Finally, we summarize the design of the agent's expression when supported by an empathy agent. The findings from this study indicate that when an agent supports a person, even if the agent is not involved in the task, the agent's expression of empathy is maintained even when the task is monotonous. From Study 3, we can say that the design of the agent's expression in supporting an empathy agent is that the agent displays gestures and supportive comments in the form of voice or text messages before and after the task so that the agent does not interfere with the task, and the person empathizes with the agent after the task is completed. Agent expression can be used in a wide range of fields and is effective in increasing empathy for agents in many cases, for example, support agents in car navigation systems, chatbot-enabled agents, and support agents in tourist information centers. Therefore, when used in conjunction with self-disclosure and task difficulty, it can be expected to further increase empathy toward agents and is essential for the design of agents that are easily accepted by people. As future research directions, I will investigate the frequency and timing of the agent's expression. And then, I will investigate the agent's appearance to confirm whether influences the agent's expression.

6.2 Questionnaire validity for empathy agents

In discussing the validity of the IRI questionnaire modified for empathy target agents in this study, we consider it from the Cronbach α coefficient. The three studies do not differ in the content of their questionnaires, but they do change the text somewhat, due to differences in their assumptions about the investigation of empathy for agents. Study 1, which focuses on appearance

and self-disclosure, had two tasks in a series of experiments, and the questionnaire was asked after each task. In Studies 2 and 3, the questionnaire was asked before and after one task, and the wording of the questionnaire used in Study 1 was inappropriate, so it was modified to a suitable wording.

The Cronbach's α coefficients for the questionnaires used in each study ranged from 0.7155 to 0.8201 in Study 1, from 0.8001 to 0.8777 in Study 2, and from 0.8027 to 0.9025 in Study 3. These Cronbach's α coefficient results are sufficient for questionnaire validity, which means that the IRI questionnaire for empathy target agents created in this study can be a valid questionnaire for future empathy agent research.

In addition, the empathy defined in this study could be used as a common definition for future empathy target agents. The validity of the questionnaire for empathy target agents allowed us to provide a definition of a person's empathy for agents in the HAI field. Empathy between people requires the consideration of complex relationships and interactivity, and could only be considered an abstract concept. In the HAI field, however, the agent does not have the capacity for empathy, and only the ability to empathize with the agent on the person side or the ability for the agent to appear to empathize with the person is needed to be defined. In this point, the definition of empathy in the HAI field, and the definition of an empathy agent, was one specific definition in this study, if we do not consider the acquisition of empathy ability for agents in the definition of an empathy agent.

It is also difficult to compare with the IRI, the original questionnaire used in this study, because the IRI examines the empathy characteristics of the questionnaire subjects, while the questionnaire for empathy agents focuses on changes in empathy toward the agents though human-agent interaction. The IRI had 28 questions, but after integrating similar questions and deleting those that were not appropriate for the study, the number of questions was reduced to 12, but the reliability remained high. Although the questionnaire

used in this study has a certain degree of reliability, it is necessary to discuss the design of a more optimal questionnaire in the future.

6.3 Empathy target/observer agent

There are two categories of empathy agents: empathy target agents and empathy observer agents. We focused on empathy target agents in three experiments. Although previous research had shown that increasing empathy from person to agent inhibits aggressive statements and actions toward the agent, we did not know whether it is better to increase affective empathy, cognitive empathy, or both.

In particular, our study investigated whether people's empathy toward an empathy agent leads to increased altruistic behavior and improved task retention, along with an investigation of factors that influence empathy toward an empathy agent when an agent is placed as an object of empathy for people. The results showed that among the factors that influence empathy, self-disclosure and agent expression influence a person's empathy toward the empathy agent. Through human-agent interactions, it was found that empathy toward the agent, the object of empathy, increases people's motivation to perform tasks, improves performance, and enhances altruistic behavior. These effects will be effective relationship building when agents coexist in human society in the future.

On the other hand, the study of agents who empathize with people, which we did not deal with, is an important issue for the future, because the interaction between people and agents is based on mutual relationships. Since it is difficult to give agents the ability to empathize with people, it is necessary to design agents so that they empathize with people from the viewpoint of people. Future research will include investigation of factors necessary for

the development and design of empathy observer agents, and the development of relationships in which agents are suitable for supporting people by empathizing with them.

In addition, our study did not examine long-term person-agent interactions, and while we found short-term effects of empathy on agents, it is unclear whether these effects persist over the long term. This is an issue that needs to be investigated for both empathy target agents and empathy observer agents, because when agents are used in society, short-term relationships may be sufficient and long-term relationships may be required, and as a future development, the impact of empathy through long-term interactions should be Investigation.

6.4 Affective/Cognitive empathy

In the study of empathy agent, knowing the effects of each factor on affective and cognitive empathy is necessary for the purposeful use of empathic agents. However, in current empathy agent research, the definition of empathy is ambiguous, and sometimes empathy is discussed only in terms of either affective or cognitive empathy. As a solution to this problem, we used the IRI, a questionnaire commonly used in psychology to measure a person's empathic characteristics, by modifying it to accommodate empathic agents. In all studies, we have attempted to ensure the reliability of the modified questionnaire and have conducted statistical analyses to ensure that it is sufficiently reliable.

As factors influencing affective empathy, the experiments showed effects on self-disclosure, task difficulty, and agent expression. One possible reason why these factors affected affective empathy is that personal distress and empathic

concern, which are classified as affective empathy, were subconsciously heightened by interaction with an empathy agent. In all but one of the three experiments in which self-disclosure was a factor, participants did not have the opportunity to learn detailed information about the empathy agent. However, the experimental results showed that affective empathy was statistically significantly higher when task difficulty was high or when the agent was represented than in the other conditions. This suggests that affective empathy can elicit affective empathy from people through interaction without knowing detailed information about the agent who is the object of empathy.

On the other hand, experiments showed that self-disclosure and agent expression were effective factors influencing cognitive empathy. As for the reasons why these factors affected cognitive empathy, perspective taking and fantasy, which are classified as cognitive empathy, were enhanced when participants understood the statements and actions of the empathy agent. It is possible that self-disclosure and agent expression were presented to the participants, which enhanced their understanding of the empathy agent, and at the same time, they considered the situation on the side of the empathy agent. While affective empathy enhanced affective empathy through interaction without the need to know detailed information about the empathy agent, cognitive empathy suggests that cognitive empathy is enhanced when people receive detailed information about the empathy agent.

As a result, self-disclosure and agent expression have the effect of enhancing both affective and cognitive empathy. The reason for this is that people enhance empathy by receiving detailed information about the empathy agent, but with respect to affective empathy, people can enhance empathy even without receiving enough information about the agent.

6.5 Empathic response

Overall, our study investigated and discussed changes in empathy for agents using a questionnaire, which is a subjective measure. It has been used in many psychological fields and traditional agent research because similar questionnaires can be used to compare the results. By modifying the IRI questionnaire into a form suitable for agents, we were able to create an empathy survey questionnaire for agents.

However, in practice, it is necessary to have a means of measuring changes in empathy toward the agent using objective indicators, and the results of the questionnaire alone are not sufficient. In particular, when investigating empathic response, the results of the questionnaire survey alone may not be sufficient for adequate analysis. It is necessary to create new objective indices so that changes in behavior and sentiment due to empathy for others can be evaluated and analyzed by means other than questionnaire surveys.

For example, in previous studies, participants' behavior during the experiment was video-recorded, and their behavior at that time was picked up and evaluated. The problem with this, however, is that there are large differences in the behavior of each participant, and although we would like to classify the behavior for analysis, the results could change depending on the classification method and the person performing the task. Therefore, we prepared situations that were task-dependent, but which required participants to act after the experiment depending on the task, and analyzed the differences in their behavior. In this way, we were able to investigate empathic responses in similar situations for all participants, but we were not able to capture the empathic responses that occurred during the experiment. In the future, an objective measure for analyzing empathic reactions between people and agents needs to be developed and discussed.

6.6 Difference in empathy agents between online and in-person experiments

In agent research, effective situations vary greatly depending on whether the agent has a body in the real world or not. In our research to date, empathy agents have interacted with people on a screen, without having a body in the real world. In recent years, people increasingly carry around smartphones and tablets, and it is now practically possible for agents to exist on their own portable devices. Therefore, our research investigated the effects of having an empathy agent interact with people on the screen. However, in a future world where people and agents coexist, agents will have bodies and will have more opportunities to be used in human society. At that time, it will be difficult to use the results of the current study as they are, and it will be important to study the results of empathy agents through face-to-face experiments, in addition to the results of experiments conducted in online environments.

In particular, one difference between human-agent interaction with and without the agent's physicality is that the agent will be able to directly influence people and society in the real world. In the past, human-agent interactions in the online environment have focused on encouraging people to change their behavior by empathizing with the agent and influencing their empathic responses to the agent. This was the case for empathy target agents, but it is also true for empathy observer agents, where making people think that the agent is sympathetic to them will change their feelings and behavior.

However, when the agent has a body in the real world, the agent can induce people to act directly. The advantage of this is that even if people do not change their sentiments and actions, the agent can interfere in interpersonal relationships and human society through its own actions and statements. Although we have designed empathy agents based on the assumption that people change, the next problem arises when the agent has a body and the agent's empathic actions and comments can directly influence people. That

is, the direct influence of the agent's actions and comments can exert a more powerful influence on the actions and comments of people than has been possible up to now. The design of empathy agents is very effective for this problem, and as a future development of our research, we will investigate the influence of empathy agents in the real world.

Chapter 7

Conclusion

Many artificial objects are used in people's lives to cooperate and compete along with people. Agents are already accepted in society, and many technologies are used in various aspects of human society, such as ChatGPT, generative AI, factory robots, cleaning robots, nursing robots, and anthropomorphic agents for customer service and chatbot response. As these agents are widely used in society, the relationship between people and agents is becoming an important issue. In the media equation, people are supposed to treat artifacts like people, but this may cause problems similar to human relationships between agents and people who are treated like people. As AI technologies and anthropomorphic agents become more prevalent in society, people distrust and reject these technologies, leading to some negative behavior. Therefore, the establishment of coexistence relationships with AI technologies and anthropomorphic agents to be utilized in the future society is an important issue.

The purpose of this study was to design a relationship between people and agents to promote coexistence between people and agents, as well as to examine approaches to make agents, which will be used in a wide range of fields in the future, more acceptable to human society. Specifically, the solution to this problem was to develop empathy and trust for anthropomorphic agents and to enable them to build appropriate relationships with each other. It has been argued in the fields of psychology and philosophy that these are important elements in human relationships and have a significant impact on the relationship building between people and anthropomorphic agents. To solve the problem of agents not being accepted by humans, we hope that humans will empathize with agents so that agents will be more utilized in human society in the future. To this end, we focused on "empathy target agents," which are considered empathy targets by people among empathy agents, and conducted experiments to elucidate the factors that cause people to empathize with agents.

In the first study, we investigated the agent's appearance and self-disclosure as factors that make people empathize with the agent. The experiment was a three-factor mixed design, with two levels of appearance (human, robot), three levels of self-disclosure (high relevance self-disclosure, low relevance self-disclosure, no self-disclosure), and two levels of stimuli (before, after). The results showed that there was no main effect for the appearance factor, and that self-disclosure with high relevance to the scenario promoted more human empathy, with statistically significant differences. Experiments showed that agents being used in human society gain empathy from humans through self-disclosure.

The second study focused on human-agent tasks and investigated task difficulty and task content as factors that make humans empathize with anthropomorphic agents. Experiments were conducted at two factors and four levels: task difficulty (high, low) and task content (cooperative, competitive).

Analysis revealed that the higher the task difficulty, the more affective empathy increased after the task. The task itself can be an important factor when manipulating each category of empathy.

The third study investigated whether the task affects human empathy toward the agent. The experiment was conducted in a three-factor mixed design with the between participant factor agent expression (available, not available) and task success or failure (success, failure) and the within participant factor stimulus (before, after). The results showed that there was no main effect of the task completion factor and that more human empathy was elicited when the agent's expression was present. We also found that the agent's expression affected the participant's motivation for the task. In the case of simple tasks, human empathy decreased when the agent's expression was absent, while human empathy was maintained when the agent's expression was present. We found that the agent's expression affected the participant's motivation for the task.

Throughout all of our studies, our investigation of the factors necessary for agents to be empathetic to people showed that when empathy agents were empathetic to people, their altruistic behavior and their awareness of task continuation increased. In addition, we created a modified version of the IRI, which investigates human empathy characteristics in psychology, as a questionnaire for agents who are empathetic to others in the study of empathy agents. While maintaining the reliability of this questionnaire, it helps to solve the problem of previous studies of empathy agents in which the definitions of empathy were disparate and difficult to compare. Future research will extend the study of empathy agents in an online environment to investigate the effects of empathy agents on people in the real world.

Reference

[1] Masahiro Mori. The uncanny valley. *Energy*, 7:33–35, 1970.

[2] Chidchanok Thepsoonthorn, Ken-ichiro Ogawa, and Yoshihiro Miyake. The exploration of the uncanny valley from the viewpoint of the robot's nonverbal behaviour. *International Journal of Social Robotics*, 13(6): 1443–1455, Sep 2021.

[3] Hasan Mahmud, A.K.M. Najmul Islam, Syed Ishtiaque Ahmed, and Kari Smolander. What influences algorithmic decision-making? a systematic literature review on algorithm aversion. *Technological Forecasting and Social Change*, 175:121390, 2022.

[4] Ibrahim Filiz, Jan René Judek, Marco Lorenz, and Markus Spiwoks. The extent of algorithm aversion in decision-making situations with varying gravity. *PLOS ONE*, 18(2):1–21, 02 2023.

[5] Byron Reeves and Clifford Nass. *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places.* Cambridge University Press, 1996.

[6] Mark Ryan. In ai we trust: Ethics, artificial intelligence, and reliability. *Science and Engineering Ethics*, 26(5):2749–2767, 2020.

[7] Jaana Hallamaa and Taina Kalliokoski. Ai ethics as applied ethics. *Frontiers in Computer Science*, 4, 2022.

[8] Hyojin Chin, Lebogang Wame Molefi, and Mun Yong Yi. Empathy is all you need: How a conversational agent should respond to verbal abuse. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI'20)*, pages 1–13, 2020.

[9] Leena Mathur, Micol Spitale, Hao Xi, Jieyun Li, and Maja J Matarić. Modeling user empathy elicited by a robot storyteller. In *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 1–8, 2021.

[10] Koji Inoue, Divesh Lala, and Tatsuya Kawahara. Can a robot laugh with you?: Shared laughter generation for empathetic spoken dialogue. *Frontiers in Robotics and AI*, 9, 2022.

[11] Riswanda Al Farisi, Ridi Ferdiana, and Teguh Bharata Adji. The effect of anthropomorphic design cues on increasing chatbot empathy. In *2022 1st International Conference on Information System & Information Technology (ICISIT)*, pages 370–375, 2022.

[12] Junya Morita and Yuna Kano. Effects of self-experience and situational awareness on empathic help to virtual agents. In *Proceedings of the 10th International Conference on Human-Agent Interaction*, page 132–139, 2022.

[13] Tatsuya Nomura, Takayuki Kanda, and Tomohiro Suzuki. Experimental investigation into influence of negative attitudes toward robots on human-robot interaction. *AI Soc.*, 20:138–150, 2006.

[14] Tatsuya Nomura, Takayuki Kanda, Tomohiro Suzuki, and Kensuke Kato. Prediction of human behavior in human–robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE Transactions on Robotics*, 24(2):442–451, 2008.

[15] Tatsuya Nomura, Takayuki Kanda, Hiroyoshi Kidokoro, Yoshitaka Suehiro, and Sachie Yamada. Why do children abuse robots? *Interaction Studies*, 17(3):347–369, 2016.

[16] Sara Konrath, Emily Falk, Andrea Fuhrel-Forbis, Mary Liu, James Swain, Richard Tolman, Rebecca Cunningham, and Maureen Walton. Can text messages increase empathy and prosocial behavior? the development and initial validation of text to connect. *PLOS ONE*, 10(9): 1–27, 2015.

[17] Victoria A. Shaffer, Jennifer Bohanek, Elizabeth S. Focella, Haley Horstman, and Lise Saffran. Encouraging perspective taking: Using narrative writing to induce empathy for others engaging in negative health behaviors. *PLOS ONE*, 14(10):1–16, 2019.

[18] Shunichi Tahara, Kazushi Ikeda, and Keiichiro Hoashi. Empathic dialogue system based on emotions extracted from tweets. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*, page 52–56, 2019.

[19] Genta Yoshioka and Yugo Takeuchi. Inferring affective states by involving simple robot movements. In *Proceedings of the 3rd International Conference on Human-Agent Interaction*, page 73–78, 2015.

[20] Mako Okanda, Kosuke Taniguchi, and Shoji Itakura. The role of animism tendencies and empathy in adult evaluations of robot. In *Proceedings of the 7th International Conference on Human-Agent Interaction*, page 51–58, 2019.

[21] Garret O'Connell, Anastasia Christakou, Anthony Haffey, and Bhismadev Chakrabarti. The role of empathy in choosing rewards from another's perspective. *Frontiers in Human Neuroscience*, 7:174, 2013.

[22] Xiang Zhi Tan, Marynel Vázquez, Elizabeth J. Carter, Cecilia G. Morales, and Aaron Steinfeld. Inducing bystander interventions during robot abuse with social mechanisms. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, page 169–177, 2018.

[23] Deborah Richards, Ayse Aysin Bilgin, and Hedieh Ranjbartabar. Users' perceptions of empathic dialogue cues: A data-driven approach to provide tailored empathy. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, page 35–42, 2018.

[24] Achim Stephan. Empathy for artificial agents. *International Journal of Social Robotics*, 7(1):111–116, 2015.

[25] Adineh Hosseinpanah, Nicole C. Krämer, and Carolin Straßmann. Empathy for everyone? the effect of age when evaluating a virtual agent. In *Proceedings of the 6th International Conference on Human-Agent Interaction*, page 184–190, 2018.

[26] Irini Giannopulu, Aude Etournaud, Kazunori Terada, Mari Velonaki, and Tomio Watanabe. Ordered interpersonal synchronisation in asd children via robots. *Scientific Reports*, 10(1):17380, 2020.

[27] Brendan Gaesser. Constructing memory, imagination, and empathy: A cognitive neuroscience perspective. *Frontiers in Psychology*, 3:576, 2013.

[28] Olga M. Klimecki, Sarah V. Mayer, Aiste Jusyte, Jonathan Scheeff , and Michael Schönenberg. Empathy promotes altruistic behavior in economic interactions. *Scientific Reports*, 6(1):31961, 2016.

[29] Becky L. Omdahl. *Cognitive appraisal, emotion, and empathy.* Psychology Press, 1 edition, 1995.

[30] Stephanie D. Preston and Frans B. M. de Waal. Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(1):1–20, 2002.

[31] Sally Olderbak, Claudia Sassenrath, Johannes Keller, and Oliver Wilhelm. An emotion-differentiated perspective on empathy with the emotion specific empathy questionnaire. *Frontiers in Psychology*, 5:653, 2014.

[32] Serge Tisseron, Frédéric Tordo, and Ritta Baddoura. Testing empathy with robots: A model in four dimensions and sixteen items. *International Journal of Social Robotics*, 7(1):97–102, 2015.

[33] Paul Bloom. *Against Empathy: The Case for Rational Compassion*. HarperCollins, 2016.

[34] Samuel D Gosling, Peter J Rentfrow, and William B Swann. A very brief measure of the big-five personality domains. *Journal of Research in Personality*, 37(6):504–528, 2003.

[35] Mark H. Davis and Mark H. Davis. A multidimensional approach to individual difference in empathy. In *JSAS Catalog of Selected Documents in Psychology*, page 85, 1980.

[36] Simon Baron-Cohen and Sally Wheelwright. The empathy quotient: an investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *J Autism Dev Disord*, 34(2):163–175, 2004.

[37] Emma J. Lawrence, Philip Shaw, Dawn Baker, Simon Baron-Cohen, and Anthony S. David. Measuring empathy: reliability and validity of the empathy quotient. *Psychological Medicine*, 34(5):911–920, 2004.

[38] Philippe Bertrand, Jérôme Guegan, Léonore Robieux, Cade Andrew McCall, and Franck Zenasni. Learning empathy through virtual reality: Multiple strategies for training empathy-related abilities using body ownership illusions in embodied virtual reality. *Frontiers in Robotics and AI*, 5:26, 2018.

[39] Austin van Loon, Jeremy Bailenson, Jamil Zaki, Joshua Bostick, and Robb Willer. Virtual reality perspective-taking increases cognitive empathy for specific others. *PLOS ONE*, 13(8):1–19, 2018.

[40] Fernanda Herrera, Jeremy Bailenson, Erika Weisz, Elise Ogle, and Jamil Zaki. Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLOS ONE*, 13(10):1–37, 2018.

[41] Max T. Curran, Jeremy Raboff Gordon, Lily Lin, Priyashri Kamlesh Sridhar, and John Chuang. Understanding digitally-mediated empathy: An exploration of visual, narrative, and biosensory informational cues. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 1–13, 2019.

[42] Cassandra L. Crone and Rachel W. Kallen. Interview with an avatar: Comparing online and virtual reality perspective taking for gender bias in stem hiring decisions. *PLOS ONE*, 17(6):1–35, 2022.

[43] Cynthia L. Bennett and Daniela K. Rosner. The promise of empathy: Design, disability, and knowing the "other". In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 1–13, 2019.

[44] Annisa Ristya Rahmanti, Hsuan-Chia Yang, Bagas Suryo Bintoro, Aldilas Achmad Nursetyo, Muhammad Solihuddin Muhtar, Shabbir Syed-Abdul, and Yu-Chuan Jack Li. Slimme, a chatbot with artificial empathy for personal weight management: System design and finding. *Frontiers in Nutrition*, 9, 2022.

[45] Peter Wright and John McCarthy. Empathy and experience in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, page 637–646, 2008.

[46] Antonio Chella, Arianna Pipitone, Alain Morin, and Famira Racy. Developing self-awareness in robots via inner speech. *Frontiers in Robotics and AI*, 7, 2020.

[47] Sydney Pratte, Anthony Tang, and Lora Oehlberg. Evoking empathy: A framework for describing empathy tools. In *Proceedings of the Fifteenth International Conference on Tangible, Embedded, and Embodied Interaction*, 2021.

[48] Aryel Beck, Lola Cañamero, and Kim A. Bard. Towards an affect space for robots to display emotional body language. In *19th International Symposium in Robot and Human Interactive Communication*, pages 464–469, 2010.

[49] Galit Hofree, Paul Ruvolo, Marian Stewart Bartlett, and Piotr Winkielman. Bridging the mechanical and the human mind: Spontaneous mimicry of a physically present android. *PLOS ONE*, 9(7):1–13, 2014.

[50] Minoru Asada. Towards artificial empathy. *International Journal of Social Robotics*, 7(1):19–33, 2015.

[51] Paul Dumouchel, Luisa Damiano, and Malcolm DeBevoise. *Living with Robots*. Harvard University Press, 2017.

[52] Ali Mollahosseini, Hojjat Abdollahi, and Mohammad H. Mahoor. Studying effects of incorporating automated affect perception with spoken dialog in social robots. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 783–789, 2018.

[53] Marlena R. Fraune. Our robots, our team: Robot anthropomorphism moderates group effects in human–robot teams. *Frontiers in Psychology*, 11, 2020.

[54] Nguyen Tan Viet Tuyen, Armagan Elibol, and Nak Young Chong. Learning bodily expression of emotion for social robots through human interaction. *IEEE Transactions on Cognitive and Developmental Systems*, 13 (1):16–30, 2021.

[55] Arianna Pipitone and Antonio Chella. What robots want? hearing the inner voice of a robot. *iScience*, 24(4):102371, 2021.

[56] Alessandro Geraci, Antonella D'Amico, Valeria Seidita, Arianna Pipitone, and Antonio Chella. Robot's inner speech effects on trust and anthropomorphic cues in human-robot cooperation. In *Trust, Acceptance and Social Cues in Human-Robot Interaction – SCRITA 2021*, pages 1–6, 2021.

[57] Scott W. McQuiggan, Jennifer L. Robison, Robert Phillips, and James C. Lester. Modeling parallel and reactive empathy in virtual agents: An inductive approach. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 1*, page 167–174, 2008.

[58] Iolanda Leite, Ginevra Castellano, André Pereira, Carlos Martinho, and Ana Paiva. Empathic robots for long-term interaction. *International Journal of Social Robotics*, 2014.

[59] Amol Deshmukh, Bart Craenen, Alessandro Vinciarelli, and Mary Ellen Foster. Shaping robot gestures to shape users' perception: The effect of amplitude and speed on godspeed ratings. In *Proceedings of the 6th International Conference on Human-Agent Interaction*, page 293–300, 2018.

[60] Jize Chen and Changhong Wang. Reaching cooperation using emerging empathy and counter-empathy. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, page 746–753, 2019.

[61] Giulia Perugia, Maike Paetzel, and Ginevra Castellano. On the role of personality and empathy in human-human, human-agent, and human-robot mimicry. In *Social Robotics*, pages 120–131, 2020.

[62] Ana Paiva, João Dias, Daniel Sobral, Ruth Aylett, Polly Sobreperez, Sarah Woods, Carsten Zoll, and Lynne Hall. Caring for agents and agents that care: Building empathic relations with synthetic agents. *Autonomous Agents and Multiagent Systems, International Joint Conference on*, 2:194–201, 2004.

[63] Ana Paiva. Empathy in social agents. *International Journal of Virtual Reality*, 10(1):1–4, 2011.

[64] Ana Paiva, Iolanda Leite, Hana Boukricha, and Ipke Wachsmuth. Empathy in virtual agents and robots: A survey. *ACM Trans. Interact. Intell. Syst.*, 7(3), 2017.

[65] Sidney M. Jourard. *Self-disclosure: An experimental analysis of the transparent self.* John Wiley, 1971.

[66] James C. Carpenter and J. Jefferson Freese. Three aspects of self-disclosure as they relate to quality of adjustment. *Journal of Personality Assessment*, 43(1):78–85, 1979. PMID: 16367049.

[67] Valerian J. Derlega and John H. Berg. *Responsiveness and Self-Disclosure*. Springer US, 1987.

[68] Pietromonaco PR. Laurenceau JP, Barrett LF. Intimacy as an interpersonal process: the importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *J Pers Soc Psychol.*, 1998.

[69] Alain Morin. Possible links between self-awareness and inner speech theoretical background, underlying mechanisms, and empirical evidence. *Journal of Consciousness Studies*, 12, 2005.

[70] Alain Morin. When inner speech and imagined interactions meet. *Imagination, Cognition and Personality*, 39(4):374–385, 2020.

[71] Patricia L. Lockwood, Yuen-Siang Ang, Masud Husain, and Molly J. Crockett. Individual differences in empathy are associated with apathy-motivation. *Scientific Reports*, 7(1):17293, 2017.

[72] Hamutal Kreiner and Yossi Levi-Belz. Self-disclosure here and now: Combining retrospective perceived assessment with dynamic behavioral measures. *Frontiers in Psychology*, 10:558, 2019.

[73] Catherine Oh Kruzic, David Kruzic, Fernanda Herrera, and Jeremy Bailenson. Facial expressions contribute more than body movements to conversational outcomes in avatar-mediated virtual environments. *Scientific Reports*, 10(1):20626, 2020.

[74] Yi-Chieh Lee, Naomi Yamashita, Yun Huang, and Wai Fu. "i hear you, i feel you": Encouraging deep self-disclosure through a chatbot. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, page 1–12, 2020.

[75] Wenjing Pan, Bo Feng, V. Skye Wingate, and Siyue Li. What to say when seeking support online: A comparison among different levels of self-disclosure. *Frontiers in Psychology*, 11, 2020.

[76] Margit I. Ruissen and Ellen R. A. de Bruijn. Competitive game play attenuates self-other integration during joint task performance. *Frontiers in Psychology*, 7, 2016.

[77] Alap Kshirsagar, Bnaya Dreyfuss, Guy Ishai, Ori Heffetz, and Guy Hoffman. Monetary-incentive competition between humans and robots: Experimental results. In *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction*, page 95–103, 2019.

[78] Jean-David Boucher, Ugo Pattacini, Amelie Lelong, Gerard Bailly, Frederic Elisei, Sascha Fagel, Peter Dominey, and Jocelyne Ventre-Dominey.

I reach faster when i see you look: Gaze effects in human–human and human–robot face-to-face cooperation. *Frontiers in Neurorobotics*, 6, 2012.

[79] Juan P. Fuentes-García, Santos Villafaina, Daniel Collado-Mateo, Ricardo de la Vega, Pedro R. Olivares, and Vicente Javier Clemente-Suárez. Differences between high vs. low performance chess players in heart rate variability during chess problems. *Frontiers in Psychology*, 10, 2019.

[80] Youngjun Cho. Rethinking eye-blink: Assessing task difficulty through physiological representation of spontaneous blinking. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021.

[81] Jacob L. Cayanus and Matthew M. Martin. Teacher self-disclosure: Amount, relevance, and negativity. *Communication Quarterly*, 56(3): 325–341, 2008.

[82] Philipp K. Masur. *The Theory of Situational Privacy and Self-Disclosure*, pages 131–182. Springer International Publishing, 2019.

[83] Christiane Brems. Dimensionality of empathy and its correlates. *The Journal of Psychology*, 123(4):329–337, 1989.

[84] Ryuichiro Higashinaka, Kohji Dohsaka, and Hideki Isozaki. Effects of self-disclosure and empathy in human-computer dialogue. In *2008 IEEE Spoken Language Technology Workshop*, pages 109–112, 2008.

[85] Laurel D. Riek, Tal-Chen Rabinowitch, Bhismadev Chakrabarti, and Peter Robinson. Empathizing with robots: Fellow feeling along the anthropomorphic spectrum. In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pages 1–6, 2009.

[86] Robert N. Davis. Web-based administration of a personality questionnaire: Comparison with traditional methods. *Behavior Research Methods, Instruments, & Computers*, 31:572–577, 1999.

[87] Matthew J. C. Crump, John V. McDonnell, and Todd M. Gureckis. Evaluating amazon's mechanical turk as a tool for experimental behavioral research. *PLOS ONE*, 8(3):1–18, 2013.

[88] Kazuo Okamura and Seiji Yamada. Adaptive trust calibration for human-AI collaboration. *PLOS ONE*, 15(2):1–20, 2020.

[89] Matthias Schonlau and Vera Toepoel. Straightlining in web survey panels over time. *Survey Research Methods*, 9(2):125–137, 2015.

[90] Dominik Johannes Leiner. Too fast, too straight, too weird: Non-reactive indicators for meaningless data in internet surveys. *Survey Research Methods*, 13(3):229–248, 2019.

[91] Luisa Damiano and Paul Dumouchel. Anthropomorphism in human–robot co-evolution. *Frontiers in Psychology*, 9, 2018.

[92] Alessandro Geraci, Antonella D'Amico, Arianna Pipitone, Valeria Seidita, and Antonio Chella. Automation inner speech as an anthropomorphic feature affecting human trust: Current issues and future directions. *Frontiers in Robotics and AI*, 8, 2021.