Facilitation of human empathy through self-disclosure of anthropomorphic agents

Takahiro Tsumura¹²*, Seiji Yamada¹²

¹ Department of Informatics, The Graduate University for Advanced Studies, SOKENDAI, Tokyo, Japan
² National Institute of Informatics, Tokyo, Japan

* takahiro-gs@nii.ac.jp

Abstract

As AI technologies progress, social acceptance of AI agents including intelligent virtual agents and robots is getting to be even more important for more applications of AI in human society. One way to improve the relationship between humans and anthropomorphic agents is to have humans empathize with the agents. By empathizing, humans take positive and kind actions toward agents, and emphasizing makes it easier for humans to accept agents. In this study, we focused on self-disclosure from agents to humans in order to realize anthropomorphic agents that elicit empathy from humans. Then, we experimentally investigated the possibility that an agent’s self-disclosure facilitates human empathy. We formulate hypotheses and experimentally analyze and discuss the conditions in which humans have more empathy for agents. This experiment was conducted with a three-way mixed plan, and the factors were the agents’ appearance (human, robot), self-disclosure (high-relevance self-disclosure, low-relevance self-disclosure, no self-disclosure), and empathy before and after a video stimulus. An analysis of variance was performed using data from 576 participants. 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. These results support our hypotheses.

Introduction

Humans live in society and use various tools. We sometimes treat artifacts like human beings. It is known that humans tend to treat artificial objects like humans in media equations [1]. At the same time, we often empathize with artificial objects. Artificial objects that we empathize with include cleaning robots, pet-type robots, and anthropomorphic agents that provide services in online shopping and at help desks. These are already used in human society and coexist with humans. In addition, the appearance of these agents varies depending on the application and usage environment. However, some humans cannot accept agents [2-4]. when agents permeate society, they should have elements that are acceptable to humans.

We consider empathy to be a significant element for gaining human acceptance as a social member. For humans to get along with each other, it is important to empathize with the other [3-6]. In the field of psychology, research has focused on empathy. Omdahl [7] roughly classifies empathy into three types: (1) affective empathy, which is
an emotional response to the emotional state of others, (2) cognitive understanding of the emotional state of others, which is defined as cognitive empathy, and (3) empathy including the above two. Preston and De Waal [8] suggested that at the heart of the empathic response was a mechanism that allowed the observer to access the subjective emotional state of the target. They defined the perception-action model (PAM) and unified the differences in empathy. They defined empathy as three types: (a) sharing or being influenced by the emotional state of others, (b) assessing the reasons for the emotional state, and (c) the ability to identify and incorporate other perspectives. Focusing on whether they act in consideration of the interests of others from the perspective of others, a change of perspective was considered from the perspective of time discounting. Olderbak et al. [9] described theoretical and empirical support for the emotional specificity of empathy and presented an emotion-specific empathy questionnaire that assesses affective and cognitive empathy for six basic emotions.

Various questionnaires are used as a measure of empathy, but we examined two famous ones. The Ten Item Personality Inventory (TIPI) is used to investigate human personality [10]. In this study, empathy may be biased by human personality, so TIPI may be used for questionnaire surveys. The Interpersonal Reactivity Index (IRI), also used in the field of psychology, is used to investigate the characteristics of empathy [11].

Empathy has been studied not only in the field of psychology but also in the field of engineering. For example, empathy has received a lot of attention in the field of virtual reality. Bertrand et al. [12] proposed a theoretical analysis of various mechanisms of empathy practice to define a possible framework for the design of empathy training in virtual reality. Curran et al. [13] investigated empathy by showing a video from the visual perspective of a person watching a virtual reality movie. Empathy has also been studied in the online environment. Pfeil et al. [14] performed a qualitative content analysis on 400 messages from a bulletin board on depression to investigate how empathy was expressed and facilitated in online communication. Empathy is also attracting attention in product design, and Bennett et al. [15] discussed and investigated a human-centered design process (promise of empathy) in which designers try to understand the target user in order to inform technology development.

As another field of studying empathy, there is one in which humans empathize with artificial objects. In the fields of human-agent interaction (HAI) and human-robot interaction (HRI), empathy between humans and agents and robots is studied. The following research has been conducted in the field of HRI. Leite et al. [16] conducted a long-term survey in an elementary school to present and evaluate a social robot empathy model aimed at interacting with children over a long period of time. They measured children’s perceptions of social presence, engagement, and social support. Mollahosseini et al. [17] incorporated a deep neural network (DNN)-based automated facial expression recognition (FER) system into 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.

In addition, the following research has been done in the field of HAI. McQuiggan et al. [18] proposed a unified inductive framework for modeling parallel and reactive empathy, which empathy models by choosing appropriate parallel or reactive empathy expressions. It was used to facilitate empathic behavior suitable for the run-time situation. Chen et al. [19] hypothesized that empathy and anti-empathy were closely related to a creature’s inertial impression of coexistence and competition within a group, and they 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.

Paiva defined the relationship between humans and empathetic agents, which are called empathy agents in HAI and HRI studies. As a definition of empathy between
agents and robots and humans, Paiva represents empathy agents in two different ways: targeting empathy and empathizing with observers [20] [22]. In this study, we use an empathy target agent to facilitate human empathy.

In previous research, we focused on “(human) empathy” as an attribute that an agent should have in order to be accepted by human society, and we aimed to realize an empathy agent. In this research, as basic research, we focused on how humans improve their relationships with agents. One method is to have humans empathize with agents. By empathizing, humans take positive actions toward agents and are more likely to accept them. Various studies have been done on linguistic information [23–25], nonverbal information [26–28], situations [29–31] and relationships [32–34] as factors that cause empathy.

In the field of psychology, research has focused on self-disclosure. Jourard [35] 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 et al. [36] measured self-presentation intimacy and internality, Berg [37] focused on the association between responsiveness and self-disclosure, and Laurencea et al. [38] suggested that both self-disclosure and partner responsiveness contribute to the experience of intimacy in interaction. Lockwood et al. [39] 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. In a few studies, the actual self-disclosure reflected in interpersonal relationships has been investigated. Therefore, Kreiner et al. [40] designed new objective and dynamic measurements to evaluate self-disclosure and stable self-disclosure characteristics. Oh Kruzic et al. [41] focused on individual and collaborative contributions of face and upper-body nonverbal channels in virtual environments via avatars. Lee et al. [42] found the effect of improving participants’ perceptions of intimacy and enjoyment by including self-disclosure during human interaction with chatbots.

From 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. We focused on self-disclosure and experimentally examined what characteristics of self-disclosure affect 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. As for the appearance, human and robot were prepared and tested as symbols of the HAI and HRI fields.

In the remaining sections, we propose our empathy agents for facilitating human empathy. Next, we cover our experiment and the results. Finally, we discuss the results and describe our future work.

**Our empathy agents for facilitating human empathy**

Feeling the relevance of previous studies on self-disclosure and empathy, we focused on both empathy and self-disclosure. In this study, 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 the characteristics of self-disclosure that are effective in promoting empathy. At the same time, we will investigate the relationship between anthropomorphic agents with different appearances and self-disclosure.

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 situation in an interaction with a human. For the above purposes, we considered two hypotheses. Experiments were conducted to investigate these hypotheses.
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 is most likely to facilitate empathy, and no self-disclosure suppresses empathy.

H2: When interacting with agents, appearance factors do not affect the promotion of empathy through self-disclosure.

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 and after watching a 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). Although there were 12 levels in total, participants were asked to join in only one of six different experiments due to the within-participant factor. 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. These results support our hypotheses. From the results of the analysis, our goal of eliciting more human empathy when empathy agents make self-disclosures in relation to a situation in an interaction with humans was achieved by our hypothesis.

Materials and methods

The protocol was approved by the ethics committee of the National Institute of Informatics (13, April, 2020, No. 1). All studies were carried out in accordance with the recommendations of the Ethical Guidelines for Medical and Health Research Involving Human Subjects provided by the Ministry of Education, Culture, Sports, Science and Technology and Ministry of Health, Labour and Welfare in Japan. Written informed consent was provided by choosing one option on an online form: “I am indicating that I have read the information in the instructions for participating in this research. I consent to participate in this research.” All participants gave informed consent. After that, they were debriefed about the experimental procedures.

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

We created a scenario to be used in the experiment for this study. There were both a common scenario and a self-disclosure scenario under this scenario. It was necessary to check that the type of self-disclosure we prepared for our research was what we had intended. By performing a manipulation check, we confirmed that the three types of self-disclosure (high-relevance self-disclosure, low-relevance self-disclosure, no self-disclosure) we prepared gave the intended impression of the scenario used in the study. The content of the scenario will be described in another subsection.

As a manipulation check, we conducted an experiment to investigate the relationship between the scenario and self-disclosure and the degree of self-disclosure. As the content of the check, we asked the participants to read only the text of the common scenario (scenario 1) and then read the scenario of each self-disclosure condition (scenario 2). Afterward, the participants 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 analysis of variance (ANOVA) among one-factor participants.

We used Yahoo! Crowdsourcing 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.

Before performing an ANOVA, we calculated the number of participants suitable for the ANOVA by using G*Power 3.1.9.7 [43]. A pre-analysis was performed by selecting ANOVA: fixed effects, omnibus, one-way in F tests. As parameter settings, the effect size was 0.25, the significance level was 0.05, and the detection power was 0.80. As a result of the calculation with G*Power, the number of participants, 159, was determined to be appropriate. Since the effect size was not known in advance, 0.25 was selected from the definition of Cohen [44]. Similarly, the significance level and detection power were also determined from Cohen’s definition.

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: \(F(2,151) = 76.70\), degree of self-disclosure: \(F(2,151) = 102.44\)]. The results of the analysis indicated that the main effect was significant, so the results of the multiple comparisons were investigated.

The case where the self-disclosure condition was highly relevant was found to be related to the most common scenario. In addition, since all of the self-disclosure conditions were significant, 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 scenario designed by us had a degree of relevance of self-disclosure.

Next, it was observed that the degree of self-disclosure under each self-disclosure condition seemed to be self-disclosure 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 was 7 points, so an average of 4 points or more for the two types of self-disclosure, and a result of less than 4 points meant no self-disclosure. From the above results, it was judged that the two types of disclosed content were self-disclosure.

Also, as a result of a post-hoc analysis done with G*power, 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. From the results of the post-hoc analysis, it was 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 of the scenario we created. Our study was conducted using this scenario.
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 experiment 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 total of 1011 participants; however, since there were 93 participants who gave inappropriate answers, their data was excluded as erroneous, so the final total was 918. 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 \(^48,49\). After that, as a result of using Cronbach’s \(\alpha\) coefficient for the reliability of the questionnaire, the coefficient was 0.7037 to 0.8202 under all conditions. Before performing a three-way analysis of variance (ANOVA), we calculated the number of participants suitable for the ANOVA by using G*Power 3.1.9.7 \(^43\). A pre-analysis was performed by selecting MANOVA: repeated measures and within-between interaction in F tests. As parameter settings, the effect size was 0.15, the significance level was 0.05, and the detection power was 0.80. Since the effect size was not known in advance, 0.15 was determined from the definition of Cohen \(^44\). Similarly, the significance level and detection power were also determined from Cohen’s definition. By performing a pre-analysis, a post-hoc analysis can be performed with an appropriate sample size. This is because, when the sample size is large, even if there is a significant difference in the results, answers may actually be judged by a very small difference.

As a result of the G*Power calculation, the number of participants, 576, was determined to be appropriate. For the analysis, 96 people were analyzed under each of six conditions in the order of participation in the experiment. Therefore, the total number of participants used in the analysis was 576. The average age was 45.35 years (S.D. = 11.11), with a minimum of 15 years and a maximum of 86 years. In addition, there were 312 males and 264 females.

However, since only BehaviorQ analysis uses a two-factor ANOVA, it was calculated with G*Power. A pre-analysis was performed by selecting ANOVA: a priori: compute required sample size – given \(\alpha\), power, and effect size in F tests. As parameter settings, the effect size was 0.25, the significance level was 0.05, and the detection power was 0.80. Since the effect size was not known in advance, 0.15 was selected from the definition of Cohen \(^44\). Similarly, the significance level and detection power were also determined from Cohen’s definition.

As a result of G*Power calculation, the number of participants, 158, was determined to be appropriate. For the analysis, 27 people were analyzed under each of six conditions in the order of participation in the experiment. Therefore, the total number of participants used in the analysis was 162. The average age was 43.49 years (S.D. = 10.58), with a minimum of 17 years and a maximum of 68 years. In addition, there were 94 males and 68 females.

Experimental procedure

The experiment was conducted in an online environment by using an MMD agent. The online environment used in this experiment has already been used as one experimental method \(^45,47\). 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. At this time, the participants were only to read while imagining the agent. After that, they read the abstract, and the empathy that they felt for the agent was aggregated in a
questionnaire survey. In this task, participants could not judge the appearance of the agent or self-disclosure.

In the second task, a three minute video that was made from the contents 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 6 conditions that combined two factors: appearance (humans, robots) 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 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, in this experiment, the independent variables were self-disclosure, agent’s appearance, and stimulation. The dependent variables were human empathy and human empathic response.

**Questionnaire**

Participants answered the questionnaire after completing the first task and the second task. 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, we used a modified 12-item questionnaire. 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 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 1 as BehaviorQ. This was an item for investigating the empathic response of the participants, and they answered this question with either yes or no.

**Table 1. Summary of questionnaire used in this experiment**

| Affective empathy | Personal distress | Q1 | The agent experienced an emergency, and you became anxious and uncomfortable. |
|-------------------|-------------------|----|--------------------------------------------------------------------------------|
|                   | Empathic concern  | 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. |
|                   |                   | 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 | BehaviorQ         |    | Finally, the agent asked you to lend it some money. What would you do?       |

**Appearance**

In this experiment, two types of agent appearances were prepared. Figures 1 and 2 show the 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.

![Fig 1. Scene of video when appearance was human](image1)
A part of the scene where the human-like agent and the participants interacted.

![Fig 2. Scene of video when appearance was human](image2)
A part of the scene where the robot agent and the participants interacted.

**Scenario**

The scenario was that the participants were chatting at a cafe during a lunch break as a colleague at the agent’s workplace. At this time, all the content spoken by the agent was set to be neutral by sentiment analysis. Sentiment analysis of each scenario was performed 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 agents’ self-disclosure was classified into three types: high-relevance, low-relevance, and no self-disclosure. To investigate self-disclosure suitable for promoting empathy, which is the purpose of this experiment, we shared most of the progress time and content of the scenario so that the scenario did not make a significant difference for any combination. Therefore, the difference in content was only the self-disclosure part of the agent. In this experiment, the agent spoke about his own work, so the content of the 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.
Analysis method

We employed an ANOVA for a three-factor mixed-plan. The between-participant factors were two levels of appearance and three levels of self-disclosure. There were two levels of empathy for the within-participant factor, before and after watching the video. From the results of the participants’ questionnaires, we investigated how self-disclosure and appearance affected the promotion of empathy as factors that elicit human empathy.

The value of empathy aggregated in the first task and the second task was used as the dependent variable. For the BehaviorQ, the Yes/No answer was replaced with a 1/0 dummy variable, and an ANOVA between two-factor participants was then performed. R (ver. 3.6.3), statistical software, was used for the ANOVA and multiple comparisons of all analysis in this work. The procedure for the three-factor ANOVA is shown as follows.

1. Test the three-way interaction \((A \times B \times C)\).
2. If the three-way interaction \((A \times B \times C)\) is significantly recognized, examine the simple interaction effect.
   - Example: Test a two-factor interaction \((B \times C)\) at a certain level of \(A\).
   - If the simple interaction is significantly recognized, the simple-simple main effect is tested.
     - When the simple-simple main effect is significant, multiple comparisons are performed if necessary.
     - Example: Test the main effect of \(C\) at some level of \(A\) and some level of \(B\).
3. If the three-way interaction is not significantly recognized and the two-way interaction \((A \times B, B \times C, A \times C)\) is significantly recognized, the simple main effect test is examined.
   - Example: Test the simple main effect of \(B\) at a certain level of \(A\).
   - If the simple main effect is significant, multiple comparisons are performed as needed.
4. If all interactions are not significantly recognized, the main effects of \(A, B,\) and \(C\) are tested.

Results

From the results of each questionnaire, a significant difference was found in the interaction between two factors, self-disclosure and before and after watching the video, other than for Q1-Q3, which were items related to personal distress. Therefore, we investigated the simple main effect except for personal distress.

We focused on the areas marked in green in Table 2 for analysis. Table 2 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. The results of the ANOVA were analyzed in more detail in terms of personal distress (Q1-Q3), empathic concern (Q4-Q6), perspective taking (Q7-Q9), and fantasy scale (Q10-Q12). Finally, 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. As a result, the self-disclosure after watching the video was found to be significant. Furthermore, as a result of analyzing behavior related to the empathic response after watching the video, no significant difference was found.
Table 2. Analysis results of multiple comparison

| Questionnaire                  | Condition                                    | Mean  | S.D.  | Factor                                      | F         | p       | η²   |
|-------------------------------|----------------------------------------------|-------|-------|---------------------------------------------|-----------|---------|------|
| Empathy (Q1-12)               | High-relevance self-disclosure-Before        | 38.87 | 6.298 | Self-disclosure before watching video       | 0.1080    | 0.8976  | ns   |
|                               | Low-relevance self-disclosure-Before         | 38.72 | 7.034 |                             | 15.61     | 0.0000 ***| 0.5194|
|                               | Low-relevance self-disclosure-After          | 36.72 | 6.701 | Before after high-relevance self-disclosure | 10.29     | 0.0016 ** | 0.0511|
|                               | Low-relevance self-disclosure-After          | 36.32 | 6.567 | Before after low-relevance self-disclosure  | 10.98     | 0.0012 ***| 0.0442|
|                               | No self-disclosure-Before                    | 35.88 | 6.407 | Before after no self-disclosure             | 10.24     | 0.0000 ***| 0.1527|
|                               | No self-disclosure-After                     | 34.01 | 7.507 |                             | 10.96     | 0.0012 ***| 0.0511|
|                               |                                             |       |       |                                             |           |         |      |
| Affective empathy (Q1-6)      | High-relevance self-disclosure-Before        | 17.23 | 3.802 | Self-disclosure before watching video       | 0.0661    | 0.9361  | ns   |
|                               | Low-relevance self-disclosure-Before         | 17.21 | 3.008 |                             | 9.03      | 0.0044 ***| 0.0437|
|                               | Low-relevance self-disclosure-After          | 17.16 | 3.701 | Before after high-relevance self-disclosure | 0.0312    | 0.9086  | ns   |
|                               | Low-relevance self-disclosure-After          | 16.20 | 4.102 | Before after low-relevance self-disclosure  | 18.32     | 0.0000 ***| 0.0884|
|                               | No self-disclosure-Before                    | 16.09 | 3.249 | Before after no self-disclosure             | 10.66     | 0.0009 ***| 0.1972|
|                               | No self-disclosure-After                     | 13.28 | 4.507 |                             | 10.01     | 0.0017 ** | 0.0505|
|                               |                                             |       |       |                                             |           |         |      |
| Cognitive empathy (Q7-12)     | High-relevance self-disclosure-Before        | 19.64 | 3.514 | Self-disclosure before watching video       | 0.0947    | 0.6956  | ns   |
|                               | Low-relevance self-disclosure-Before         | 20.81 | 4.275 |                             | 14.01     | 0.0000 ***| 0.0499|
|                               | Low-relevance self-disclosure-After          | 19.57 | 4.098 | Before after high-relevance self-disclosure | 29.79     | 0.0000 ***| 0.1355|
|                               | Low-relevance self-disclosure-After          | 19.42 | 3.508 | Before after low-relevance self-disclosure  | 0.6322    | 0.4725  | ns   |
|                               | No self-disclosure-Before                    | 19.49 | 3.528 | Before after no self-disclosure             | 10.10     | 0.0017 ** | 0.0505|
|                               | No self-disclosure-After                     | 18.52 | 4.945 |                             | 10.43     | 0.0007 ***| 0.1302|
|                               |                                             |       |       |                                             |           |         |      |
| Empathic concern (Q4-6)       | High-relevance self-disclosure-Before        | 9.870 | 2.414 | Self-disclosure before watching video       | 0.1261    | 0.8815  | ns   |
|                               | Low-relevance self-disclosure-Before         | 10.40 | 2.413 |                             | 17.44     | 0.0000 ***| 0.0584|
|                               | Low-relevance self-disclosure-After          | 9.755 | 1.954 | Before after high-relevance self-disclosure | 13.84     | 0.0003 ***| 0.0679|
|                               | Low-relevance self-disclosure-After          | 9.995 | 2.277 | Before after low-relevance self-disclosure  | 9.13      | 0.0758  | ns   |
|                               | No self-disclosure-Before                    | 9.807 | 2.483 | Before after no self-disclosure             | 28.43     | 0.0000 ***| 0.1302|
|                               | No self-disclosure-After                     | 9.198 | 2.657 |                             | 10.38     | 0.0012 ***| 0.0393|
|                               |                                             |       |       |                                             |           |         |      |
| Perspective taking (Q7-9)     | High-relevance self-disclosure-Before        | 10.36 | 1.955 | Self-disclosure before watching video       | 0.0437    | 0.9573  | ns   |
|                               | Low-relevance self-disclosure-Before         | 10.94 | 1.985 |                             | 6.63      | 0.0126 ** | 0.0307|
|                               | Low-relevance self-disclosure-After          | 10.38 | 1.700 | Before after high-relevance self-disclosure | 27.61     | 0.0000 ***| 0.0939|
|                               | Low-relevance self-disclosure-After          | 10.44 | 1.960 | Before after low-relevance self-disclosure  | 0.6449    | 0.4299  | ns   |
|                               | No self-disclosure-Before                    | 10.41 | 1.947 | Before after no self-disclosure             | 2.78      | 0.0943  | ns   |
|                               | No self-disclosure-After                     | 10.18 | 2.251 |                             | 2.82      | 0.0147  |      |
|                               |                                             |       |       |                                             |           |         |      |
| Fantasy scale (Q10-12)        | High-relevance self-disclosure-Before        | 9.276 | 2.166 | Self-disclosure before watching video       | 0.4404    | 0.6440  | ns   |
|                               | Low-relevance self-disclosure-Before         | 9.870 | 2.421 |                             | 16.36     | 0.0000 ***| 0.0549|
|                               | Low-relevance self-disclosure-After          | 9.274 | 2.715 | Before after high-relevance self-disclosure | 18.12     | 0.0000 ***| 0.0871|
|                               | Low-relevance self-disclosure-After          | 9.797 | 2.899 | Before after low-relevance self-disclosure  | 8.54      | 0.0091 ***| 0.0164|
|                               | No self-disclosure-Before                    | 9.083 | 2.076 | Before after no self-disclosure             | 11.86     | 0.0000 ***| 0.0587|
|                               | No self-disclosure-After                     | 8.347 | 2.365 |                             |           |         |      |

p: *p<.05 **p<.01 ***p<.001

First, there was no significant interaction between the appearance factor and the self-disclosure factor under all conditions. The results for empathy (Q1-12) showed an interaction between the self-disclosure factor and before and after watching the video [F(2,570) = 24.31]. 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 (high-relevance: mean = 38.02, S.D. = 7.033, low-relevance: mean = 35.62, S.D. = 6.567, no self-disclosure: mean = 34.01, S.D. = 7.557). In addition, the simple main effect before and after watching the video for each self-disclosure condition was significantly different before and after watching the 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. Also, as a result of a post-hoc analysis by G*power, the effect size of the empathy was 0.2921, and the power of the empathy was 1.000. From the results of the post-hoc analysis, it was found that self-disclosure was effective for empathy.

Similarly, the results for affective empathy (Q1-Q6) showed an interaction between the self-disclosure factor and before and after watching the video [F(2,570) = 13.91]. As a result of a multiple comparison, the simple main effect of the self-disclosure factor
Fig 3. Results of multiple comparison for self-disclosure after watching video for empathy. Error bars show standard deviation.

After watching the video showed a significant difference in the combination of all three levels as shown in Fig. 3 (high-relevance: mean = 17.21, S.D. = 4.008, low-relevance: mean = 16.20, S.D. = 4.162, no self-disclosure: mean = 15.28, S.D. = 4.497). However, the simple main effect after watching the video for each self-disclosure condition was not significantly different between before and after watching the video with high relevance. Under other self-disclosure conditions, a significant difference was observed before and after watching the 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. Also, as a result of a post-hoc analysis by G*power, the effect size of the affective empathy was 0.2208, and the power of the affective empathy was 0.9924. From the results of the post-hoc analysis, it was found that self-disclosure was effective for affective empathy.

Fig 4. Results of multiple comparison for self-disclosure after watching video for affective empathy. Error bars show standard deviation.

In addition, the results for cognitive empathy (Q7-Q12) showed an interaction between the self-disclosure factor and before and after watching the video \( F(2,570) = 21.32 \). As a result of a multiple comparison, the simple main effect of the self-disclosure factor after watching the video showed no significant difference between low-relevance
self-disclosure and no self-disclosure, as shown in Fig. 5 (high-relevance: mean = 20.81, S.D. = 4.275, low-relevance: mean = 19.42, S.D. = 3.398, no self-disclosure: mean = 18.72, S.D. = 4.043). There was a significant difference in the simple main effect before and after watching the video for each self-disclosure condition between high-relevance self-disclosure and no self-disclosure, and a significant difference was found between high-relevance self-disclosure and low-relevance self-disclosure. From the above, it was suggested that high-relevance self-disclosure facilitated cognitive empathy and that no self-disclosure suppressed cognitive empathy. Also, as a result of a post-hoc analysis by G*power, the effect size of the cognitive empathy was 0.2735, and the power of the cognitive empathy was 0.9999. From the results of the post-hoc analysis, it was found that self-disclosure was effective for cognitive empathy.

![Fig 5. Results of multiple comparison for self-disclosure after watching video for cognitive empathy. Error bars show standard deviation.](image)

**Analysis of affective empathy and cognitive empathy**

A more detailed investigation into the results of the above affective empathy and cognitive empathy was done for the four categories of personal distress, empathic concern, perspective taking, and fantasy scale. First, we investigated the results for personal distress and empathic concern, which are classifications of affective empathy. Regarding the results for personal distress for Q1-Q3, no interaction was observed between all the factors, and only the main effect before and after watching the video was significantly different. This suggests that personal distress was suppressed after watching the video, but it could not be said that appearance factors and self-disclosure factors were involved. The results for empathic concern for Q4-Q6 showed interactions between self-disclosure factors and before and after watching the video, similar to the results for affective empathy \[F(2,570) = 21.73\]. As a result of a multiple comparison, as shown in Fig. 6, the simple main effect of the self-disclosure factor after watching the video showed a significant difference in the combination of all three levels (high-relevance: mean = 10.40, S.D. = 2.413, low-relevance: mean = 9.495, S.D. = 2.277, no self-disclosure: mean = 8.917, S.D. = 2.657). It was suggested that self-disclosure facilitated empathic concern when the relevance was high and that low-relevance self-disclosure suppressed empathic concern. The same tendency was seen in the simple main effect before and after watching the video for each self-disclosure condition, and empathic concern was facilitated when self-disclosure was highly relevant. Empathic concern was suppressed when there was no self-disclosure. Also, as a result of
a post-hoc analysis by G*power, the effect size of the empathic concern was 0.2762, and the power of the empathic concern was 0.9999. From the results of the post-hoc analysis, it was found that self-disclosure was effective for empathic concern.

Next, we investigated the results for perspective taking and fantasy scale, which are classifications of cognitive empathy. The results regarding the perspective taking for Q7-Q9 showed an interaction between the self-disclosure factor and before and after watching the video, similar to the comprehensive analysis result and the analysis result for cognitive empathy \([F(2,570) = 10.64]\). However, as a result of a multiple comparison, as shown in Fig. 7, the simple main effect of the self-disclosure factor after watching the video showed a significant difference only between the high-relevance self-disclosure and no self-disclosure (high-relevance: mean = 10.94, S.D. = 2.296, low-relevance: mean = 10.44, S.D. = 1.960, no self-disclosure: mean = 10.18, S.D. = 2.251). The simple main effect before and after watching the video for each self-disclosure condition was significantly different before and after watching the video only in the case of high-relevance self-disclosure. This suggests that, in the case of high-relevance self-disclosure, perspective taking was facilitated. Also, as a result of a post-hoc analysis by G*power, the effect size of perspective taking was 0.1932, and the power of perspective taking was 0.9643. From the results of the post-hoc analysis, it was found that self-disclosure was effective for perspective taking.

The fantasy scale analysis results for Q10-Q12 showed interactions between self-disclosure factors and before and after watching the video, similar to the comprehensive analysis results and cognitive empathy analysis results \([F(2,570) = 17.08]\). The results of the multiple comparisons were also the same, as shown in Fig. 8 and the simple main effect of self-disclosure factors after watching the video showed that there was no significant difference between low-relevance self-disclosure and no self-disclosure (high-relevance: mean = 9.870, S.D. = 2.421, low-relevance: mean = 8.979, S.D. = 2.064, no self-disclosure: mean = 8.547, S.D. = 2.395). The simple main effect before and after watching the video for each self-disclosure condition was similar to cognitive empathy, and there was a significant difference between high-relevance self-disclosure and no self-disclosure. However, a significant difference between low-relevance self-disclosure and no self-disclosure was not found. This suggests that high-relevance self-disclosure facilitates the fantasy scale and that no self-disclosure suppresses it. Also, as a result of a post-hoc analysis by G*power, the effect size of the fantasy scale was 0.2447, and the power of the fantasy scale was 0.9985. From the results of the post-hoc
Fig 7. Results of multiple comparison for self-disclosure after watching video for perspective taking. Error bars show standard deviation.

In the analysis, it was found that self-disclosure was effective for the fantasy scale.

Fig 8. Results of multiple comparison for self-disclosure after watching video for fantasy scale. Error bars show standard deviation.

Also, only the results of the fantasy scale analysis showed an interaction between the appearance and before and after watching the video \( F(1,570) = 5.456 \). As a result of a multiple comparison, a significant difference was observed before and after watching the video when the appearance was a robot. From the above, it was suggested that the fantasy scale was suppressed when the appearance was a robot. However, as a result of a post-hoc analysis by G\(^*\)power, the effect size of the fantasy scale was 0.0979, and the power of the fantasy scale was 0.3960. From the results of the post-hoc analysis, it was found that self-disclosure was not effective for the fantasy scale.

Discussion

This experiment was conducted to investigate the conditions necessary for humans to empathize with anthropomorphic agents. For this purpose, we formulated two hypotheses and analyzed the data obtained from the experiments. The experimental results supported the first hypothesis in that, among the three types of self-disclosure
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. Hence, we need to discuss low-relevance self-disclosure. It was observed that the content of self-disclosure itself is self-disclosure from prior experiments. However, empathy is not always suppressed when relevance is low. In this experiment, empathy was suppressed, but empathy may be facilitated depending on the content, even with low-relevance self-disclosure. Therefore, it is possible that empathy may be facilitated by the type and content of self-disclosure, which is different from the low-relevance self-disclosure used in the experiment. For this reason, it is effective to have self-disclosure that is high in relevance when humans empathize with empathy agents and to not allow self-disclosure if humans do not want to empathize. When introducing empathy agents using self-disclosure that is low in relevance, it should not be important whether humans have empathy for the agents.

Next, the experiment supported the second hypothesis in that the appearance factor did not affect the promotion of empathy through self-disclosure when interacting with the agent. So far, appearance and self-disclosure have been studied for human empathy. There is a reason for the choice of appearance we used this time. In the scenario of this experiment, we adopted a body structure similar to that of humans on the premise that the agent is 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. However, there is a possibility that the result is possible only under this condition, so appearance must be further examined.

In addition, since the scenario of self-disclosure in this experiment was related to work, content related to work was added to the relevant self-disclosure, and hobby-related content was added to the nonrelevant self-disclosure. The video time was adjusted by adding weather stories and land stories without self-disclosure. However, in discussing whether or not such self-disclosure is related to the scenario used, the results may change depending on how much it is related. By conducting a preliminary experiment to verify the relevance to the self-disclosure scenario and verifying the degree of self-disclosure in the same way, it was judged that there was no problem in this experiment.

Furthermore, although we did not pay attention to it in this study and did not formulate a hypothesis, we will also discuss the analysis results for the classification of empathy. First of all, regarding affective empathy, the experimental results did not indicate that affective empathy was facilitated. The reason empathy was not facilitated can be understood from personal distress and empathic concern. Personal distress had been suppressed, but it was better to suppress it. Personal distress means discomfort or anxiety regarding the suffering of another person, and this result is as expected because the experiment did not suppose that the agent would be used in a situation where personal distress is facilitated. Empathic concern was facilitated by self-disclosure that was high in relevance and suppressed by no self-disclosure. Empathic concern means compassion and sympathy for another person, and it was a good result that these were facilitated according to the hypotheses. Since these are elements of affective empathy, it was not facilitated as a result, and the level of affective empathy was maintained before and after watching the video even in the case of high-relevance self-disclosure.

Moreover, cognitive empathy was facilitated by high-relevance self-disclosure, as shown in the experimental results, and suppressed by no self-disclosure. The experimental results for cognitive empathy can be understood in terms of perspective
taking and imagination. Perspective taking means thinking about things from the standpoint of the other person, and this was facilitated only by high-relevance self-disclosure. Although it was not facilitated under other conditions, it was not suppressed, and since the self-disclosure was high in relevance, it was easier to change the viewpoint of the participants and, thus, empathy was facilitated. Imagination means to be able to imagine the feelings of the other person, and it was facilitated by high-relevance self-disclosure and suppressed by self-disclosure. It is thought that self-disclosure of the other party is considered to be a necessary element for becoming the other party, and if the relationship is strong, creativity is facilitated, and thus, it becomes easier to imagine the feelings of the other party. This is suggested by the fact that it was suppressed in the case of no self-disclosure, and no significant difference was observed in the case of low-relevance self-disclosure.

Next, we discuss the results for behavior as an analysis of empathic response. In this experiment, participants played the role of observer for the target empathy agent. Observers responded empathetically to any information available from the target. In this experiment, 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, no significant difference was found. From these facts, appearance factors and self-disclosure factors were not observed to affect the empathetic response in this study. We think that this did not affect the behavior because the interaction time between the empathy agent and the participants was as short as 3 minutes.

### Strengths and limitations

The purpose of this study is to investigate whether humans can elicit more empathy when they make a self-disclosure related to a situation in an interaction with an empathy agent. As a result of the experiment, it was found that the self-disclosure factor promotes human empathy toward an anthropomorphic agent. In addition, interaction was not observed between the appearance factor and the self-disclosure factor in this study. 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 the experiment that it suppresses empathy. By properly using empathy depending on the situation, we think that the impact on humans can be adjusted when anthropomorphic agents are introduced into human society in the future.

However, as a limitation of this study, participants interacted with the empathy agent by watching a video. Thus, 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.

### 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.
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 and after watching a 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. These results support our hypotheses. Moreover, the empathic response was unaffected by appearance and self-disclosure factors. 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.

References

1. Reeves B, Nass C. The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places. 1st ed. Cambridge: Cambridge University Press; 1996.

2. Nomura T, Kanda T, Suzuki T. Experimental investigation into influence of attitudes toward robots on human–robot interaction. AI Soc. 2006 Mar; 20(2): 138–150. doi:10.1007/s00146-005-0012-7.

3. Nomura T, Kanda T, Suzuki T, Kato K. Prediction of Human Behavior in Human–Robot Interaction Using Psychological Scales for Anxiety and Negative Attitudes Toward Robots. IEEE Transactions on Robotics. 2008 Apr; 24(2): 442-451. doi:10.1109/tro.2007.914004.

4. Nomura T, Kanda T, Kidokoro H, Suehiro Y, Yamada S. Why do children abuse robots? Interaction Studies. 2016 Jan; 17(3): 347-369. doi:10.1075/is.17.3.02nom.

5. Gaesser B. Constructing Memory, Imagination, and Empathy: A Cognitive Neuroscience Perspective. Front Psychol. 2013 Jan; 3: 576. doi:10.3389/fpsyg.2012.00576.

6. Klinecki O, Mayer S, Jusyte A, et al. Empathy promotes altruistic behavior in economic interactions. Sci Rep. 2016 Aug; 6(1): 31961. doi:10.1038/srep31961.

7. Omdahl BL. Cognitive appraisal, emotion, and empathy. 1st ed. New York: Psychology Press; 1995. doi:10.4324/9781315806556.

8. Preston S, De Waal F. Empathy: Its ultimate and proximate bases. Behavioral and Brain Sciences. 2002 Feb; 25(1): 1-20. doi:10.1017/S0140525X02000018.

9. Olderbak S, Sassenrath C, Keller J, Wilhelm O. An emotion-differentiated perspective on empathy with the emotion specific empathy questionnaire. Front Psychol. 2014 Jul; 5: 653. doi:10.3389/fpsyg.2014.00653.

10. Gosling SD, Rentfrow PJ, Swann WB. A very brief measure of the Big-Five personality domains. J Res Pers. 2003 Jun; 37(6): 504-528. doi:10.1016/S0092-6566(03)00046-1.

11. Davis MH. A multidimensional approach to individual difference in empathy. JSAS Catalog of Selected Documents in Psychology. 1980 Jan; 10(85): 85-103.
12. Bertrand P, Guegan J, Robieux L, McCall CA, Zenasni F. Learning Empathy Through Virtual Reality: Multiple Strategies for Training Empathy-Related Abilities Using Body Ownership Illusions in Embodied Virtual Reality. Front Robot AI. 2018 Mar; 5: 26. doi:10.3389/frobt.2018.00026.

13. Curran MT, Gordon JR, Lin L, Sridhar PK, Chuang J. Understanding Digitally-Mediated Empathy: An Exploration of Visual, Narrative, and Biosensory Informational Cues. Proc SIGCHI Conf Hum Factor Comput Syst. 2019 May; 614: 1-13. doi:10.1145/3290605.3300844.

14. Pfeil U, Zaphiris P. Patterns of Empathy in Online Communication. Proc SIGCHI Conf Hum Factor Comput Syst. 2007 Apr; 919–928. doi:10.1145/1240624.1240763.

15. Bennett CL, Rosner DK. The Promise of Empathy: Design, Disability, and Knowing the "Other". Proc SIGCHI Conf Hum Factor Comput Syst. 2019 May; 298: 1–13. doi:10.1145/3290605.3300528.

16. Leite I, Castellano G, Pereira A, Martinho C, Paiva A. Empathic Robots for Long-term Interaction. Int J Soc Robot. 2014 Aug; 6(3): 329–341. doi:10.1007/s12369-014-0227-1.

17. Mollahosseini A, Abdollahi H, Mahoor MH. Studying Effects of Incorporating Automated Affect Perception with Spoken Dialog in Social Robots. RO-MAN '18. 2018 Aug: 783-789. doi:10.1109/ROMAN.2018.8525777.

18. McQuiggan SW, Robison JL, Phillips R, Lester JC. Modeling Parallel and Reactive Empathy in Virtual Agents: An Inductive Approach. Proc Int Joint Conf Auton Agents Multiagent Syst. 2008 May; 1: 167–174. doi:10.5555/1402383.1402411.

19. Chen J, Wang C. Reaching Cooperation Using Emerging Empathy and Counter-Empathy. Proc Int Joint Conf Auton Agents Multiagent Syst. 2019 May: 746–753. doi:10.5555/3306127.3331764.

20. Paiva A, Dias J, Sobral D, Aylett R, Sorel Perez P, Woods S, Zoll C, Hall L. Caring for Agents and Agents that Care: Building Empathic Relations with Synthetic Agents. Proc Int Joint Conf Auton Agents Multiagent Syst. 2004 Jul; 2: 194-201. doi:10.1109/AAMAS.2004.10047.

21. Paiva A. Empathy in Social Agents. Int J Virtual Real. 2011 Jan; 10(1): 1-4. doi:10.20870/IJVR.2011.10.1.2794.

22. Paiva A, Leite I, Boukricha H, Wachsmuth I. Empathy in Virtual Agents and Robots: A Survey. ACM Trans Interact Intell Syst. 2017 Sep; 7(3): 1-40. doi:10.1145/2912150.

23. Konrath S, Falk E, Fuhrer-Forbis A, Liu M, Swain J, et al. Can Text Messages Increase Empathy and Prosocial Behavior? The Development and Initial Validation of Text to Connect. PLoS One. 2015 Sep; 10(9): 1-27. doi:10.1371/journal.pone.0137585.

24. Shaffer VA, Bohanek J, Focella ES, Horstman H, Saffran L. Encouraging perspective taking: Using narrative writing to induce empathy for others engaging in negative health behaviors. PLoS One. 2019 Oct; 14(10): 1-16. doi:10.1371/journal.pone.0224046.
25. Tahara, Shunichi and Ikeda, Kazushi and Hoashi, Keiichiro. Empathic Dialogue System Based on Emotions Extracted from Tweets. IUI ’19. 2019 Mar: 52–56. doi:10.1145/3301275.3302281.

26. Tisseron S, Tordo F, Baddoura R. Testing Empathy with Robots: A Model in Four Dimensions and Sixteen Items. Int J of Soc Robotics. 2015 Feb; 7(1): 97-102. doi:10.1007/s12369-014-0268-5.

27. Yoshioka G, Takeuchi Y. Inferring Affective States by Involving Simple Robot Movements. Proceedings of the 3rd International Conference on Human-Agent Interaction. 2015 Oct: 73–78. doi:10.1145/2814940.2814959.

28. Okanda M, Taniguchi K, Itakura S. The Role of Animism Tendencies and Empathy in Adult Evaluations of Robot. Proceedings of the 7th International Conference on Human-Agent Interaction. 2019 Sep: 51–58. doi:10.1145/3349537.3351891.

29. O’Connell G, Christakou A, Haffey AT, Chakrabarti B. The role of empathy in choosing rewards from another’s perspective. Front Hum Neurosci. 2013 May; 7: 174. doi:10.3389/fnhum.2013.00174.

30. Tan, Xiang Zhi and Vázquez, Marynel and Carter, Elizabeth J. and Morales, Cecilia G. and Steinfeld, Aaron. Inducing Bystander Interventions During Robot Abuse with Social Mechanisms. Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction. 2018 Feb 26: 169–177. doi:10.1145/3171221.3171247.

31. Richards D, Bilgin AA, Ranjbartabar H. Users’ Perceptions of Empathic Dialogue Cues: A Data-Driven Approach to Provide Tailored Empathy. Intell Virtual Agents. 2018 Nov: 35–42. doi:10.1145/3267851.3267857.

32. Stephan A. Empathy for Artificial Agents. Int J of Soc Robotics. 2015 Feb; 7(1): 111–116. doi:10.1007/s12369-014-0260-0.

33. Hosseinpanah A, Krämer NC, Straßmann C. Empathy for Everyone? The Effect of Age When Evaluating a Virtual Agent. Proceedings of the 6th International Conference on Human-Agent Interaction. 2018 Dec: 184–190. doi:10.1145/3284432.3284442.

34. Giannopulu I, Etournaud A, Terada K et al. Ordered interpersonal synchronisation in ASD children via robots. Sci Rep. 2020 Oct; 10(1): 17380. doi:10.1038/s41598-020-74438-6.

35. Jourard SM. Self-disclosure: An experimental analysis of the transparent self. 1st ed. Oxford: John Wiley; 1971.

36. Carpenter JC, Freese JJ. Three Aspects of Self-Disclosure as They Relate to Quality of Adjustment. J Pers Assess. 1979; 43(1): 78-85. doi:10.1207/s15327752jpa4301_11.

37. Berg JH. Responsiveness and Self-Disclosure. In: Derlega VJ, Berg JH. editors. Self-Disclosure. Perspectives in Social Psychology. Boston: Springer US; 1987: 101–130. doi:10.1007/978-1-4899-3523-6_6.

38. Laurenceau JP, Barrett LF, Pietromonaco PR. Intimacy as an interpersonal process: the importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. J Pers Soc Psychol. 1998; 74(5): 1238–1251. doi:10.1037//0022-3514.74.5.1238.
39. Lockwood PL, Ang YS, Husain M et al. Individual differences in empathy are associated with apathy-motivation. Sci Rep. 2017 Dec; 7(1): 17293. doi:10.1038/s41598-017-17415-w.

40. Kreiner H, Levi-Belz Y. Self-Disclosure Here and Now: Combining Retrospective Perceived Assessment With Dynamic Behavioral Measures. Front Psychol. 2019 Mar; 10: 558. doi:10.3389/fpsyg.2019.00558.

41. Oh Kruzic, C, Kruzic D, Herrera F et al. Facial expressions contribute more than body movements to conversational outcomes in avatar-mediated virtual environments. Sci Rep. 2020 Nov; 10(1): 20626. doi:10.1038/s41598-020-76672-4.

42. Lee Y, Yamashita N, Huang Y, Fu W. "I Hear You, I Feel You": Encouraging Deep Self-Disclosure through a Chatbot. Proc SIGCHI Conf Hum Factor Comput Syst. 2020 Apr; 1-12. doi:10.1145/3313831.3376175.

43. Faul F, Erdfelder E, Lang AG et al. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007 May; 39(2): 175-191. doi:10.3758/BF03193146.

44. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers; 1988.

45. Davis RN. Web-based administration of a personality questionnaire: Comparison with traditional methods. Behav Res Methods. 1998 Dec; 31(4): 572–577. doi:10.3758/BF03200737.

46. Crump MJC, McDonnell JV, Gureckis TM. Evaluating Amazon’s Mechanical Turk as a Tool for Experimental Behavioral Research. PLoS One. 2013 Mar; 8(3): 1-18. doi:10.1371/journal.pone.0057410.

47. Okamura K, Yamada S. Adaptive trust calibration for human-AI collaboration. PLoS One. 2020 Feb; 15(2): 1-20. doi:10.1371/journal.pone.0229132.

48. Schonlau M, Toepoel V. Straightlining in Web survey panels over time. Survey Research Methods. 2015 Jul; 9(2): 125-137. doi:10.18148/srm/2015.v9i2.6128.

49. Leiner DJ. Too Fast, too Straight, too Weird: Non-Reactive Indicators for Meaningless Data in Internet Surveys. Survey Research Methods. 2019 Dec; 13(3): 229-248. doi:10.18148/srm/2019.v13i3.7403.