Deceitful Robots: The effects of Non-verbal Empathy and Competence on perceptions of a social robot

Miguel Filipe Ribeiro Pais
Instituto Superior Técnico, No. 62466
miguelpais@ist.utl.pt

Abstract

This work presents a pre-study of behavior design and a controlled experiment into the effects of competence and non-verbal empathic behaviors on the attitudes towards robots in a guidance scenario using the NAO robot. The goal of the pre-study was to gain input for the design of the non-verbal behaviors of the robot in the controlled experiment, by observing the behaviors executed by 5 participants. Participants were restricted to non-verbal expression and wore a mask that diminished facial expression, in this way matching NAO’s expressive capabilities. The controlled experiment was run with 46 participants using a between-subjects design featuring 3 conditions of robotic behavior: high competence empathic alignment, high competence empathic misalignment and low competence empathic alignment. We analyzed the effects of these different conditions on the perceived trustworthiness, friendship and social presence of the robot. With this study, we identified the positive effect of competence on the perceived trustworthiness of a robotic guide and extracted lessons on the interpretation of robotic non-verbal behavior.

1. Introduction

Some areas of research on robotics have been led by the vision of the usage of robots as true social actors capable of interacting with humans and potentiating scenarios of partnership, cooperation, teaching, guiding, entertainment, aiding and helping.

Reeves and Nass [1996] have shown that people respond socially to computers similarly to the way they might respond to other people. Humanoid robots, with a similar morphology, moving and expressing capabilities are, in this way, the most prominent entities for the creation of the most natural way of interfacing with humans: the social interface [Breazeal, 2002].

However, as Breazeal and Scassellati [1999] argue, in order to potentiate these social scenarios, a robot must be believable and life-like, must have behavioral consistency, ways of expressing its internal states, be prepared for changes in environmental conditions, be flexible in dealing with unexpected events and quick enough to respond to situations in an appropriate manner. A robot must thus convey intentionality, that is, suggest the presence of internal beliefs, desires and intentions [Breazeal and Scassellati, 1999].

However, recent studies have identified that simple and unexpected behaviors, such as cheating and deception, might also be effective ways of conveying intentionality, leading humans to make greater attributions of robotic mental state [Short et al., 2010]. These behaviors contradict the initial expectations of humans, built upon biases that lead them to regard automatous systems as well-intentioned entities [Dzindolet et al., 2003]. This area of automated systems has, over the years, covered the relationships that arise in partnerships between humans and aid systems, and confirmed not only these biases [Dzindolet et al., 2003], but also identified the effects of social and affective factors on the perceived trustworthiness of these systems. These affective effects surface also in the area of agent systems, given agent systems’ greater capabilities to elicit emotional responses by conveying empathy [Brave et al., 2005]. Featuring human-like embodiments and the capability for the usage of facial expressions, body movement, posture, gestures, gaze and voice, social robots are a step above agent systems as appropriate entities to convey empathy and, in this way, elicit social and affective responses from humans.

In this work we will thus connect ideas from the area of automated systems, that identified how affective factors could affect the perceived trustworthiness of an aid system, with research on the different ways to convey empathy in social robots and their effects. This will allow us to study and compare how a robot’s different empathic behaviors and competence levels might lead to different trustworthiness perceptions, on a guiding scenario of human-robot interaction with the NAO robot.
2. Related Work

Our role with this work was to explore the idea of a deceptive robot. Terada and Ito [2010] defines deception as the act to propagate beliefs that are not true and characterizes deceptive acts as intended acts to inflict a loss on others. Given this, deception makes sense in contexts in which a human either competes, cooperates or depends in part on a system to achieve something. In this context, deception means an intended act from a system to inflict a loss on a human that trusted and cooperated with it. A well studied area in which humans depend or cooperate with artificial systems is the area of Automated Systems. This area encompasses all the systems that either automatically carry out a task for a user or provide instructions or advice on how it should be performed.

The assumption behind automation systems lies in the idea that a human-computer "team" will always be more productive than either the human or the automated aid would be working alone. However, it’s not always the case that these human-robot teams work optimally, as the use of the automation is guided by the trust that is perceived by the operator of such systems.

Josang and Presti [2004] defined trust as the user's willingness to depend on a system and rely on the information or capabilities it provides even though negative consequences are possible and the system might make mistakes.

This trustworthiness perception influences whether appropriate or inappropriate automation reliance occurs. Appropriate automation reliance occurs whenever an operator trusts and relies on an automated aid that is more reliable than manual operation, or conversely, distrusts and chooses not to rely on an automated aid that is less reliable than manual operation [Dzindolet et al., 2003]. Inappropriate use of automation, however, occurs whenever an operator doesn’t trust and doesn’t rely on an aid that is more reliable than manual operation (disuse) or when it trusts and relies on an aid that is less reliable than manual operation (misuse) [Dzindolet et al., 2003], potentially leading to disastrous consequences [Parasuraman and Riley, 1997].

Cramer [2010], reviewing literature from Castelfranchi and Falcone [2000] and McAllister [1995] identified that the concept of trust can be divided in two different subdimensions: Cognition-based trust and Affective-based trust. Cognition-based trust arises from the assessment of a system’s competence and reliability through the user’s understanding of the system’s inner workings and perceptions of control over system outcomes. Affective-based trust, on the other hand, focuses on social and affective factors and, in this way, involves faith in the motives of a system, bonding and the perception that the user and the system can successfully collaborate.

Apart from this, the initial trust placed upon an automated system is also influenced by biases that may surface from the user’s initial mental model of the robot. The positive bias assumption from social psychology defines that, in the absence of other information, people believe in goodness of others, and positive rather than neutral evaluations are given [Dzindolet et al., 2003].

Dzindolet et al. [2003] went then to explore whether this positive bias was valid in the context of automated systems, by conducting a study in which participants rated their trust on an automated image recognition aid before an history of its reliability had been exposed. This study did confirm that participants with little information on the reliability of an automated aid believed that it would perform well and better than themselves, even before a solid proof of its reliability had been shown.

After this initial impression, the user starts to build a more solid view of the system’s trustworthiness based on the interaction progresses. However, and since trust is not affect solely by cognitive factors, affective factors can also influence the this perception.

Parasuraman and Miller [2004] studied the effects of etiquette in automated systems, arguing that it could profoundly affect the trust the user would place upon them. They defined etiquette as the the set of behaviors that mark individuals as members of trustworthy groups or behaviors as pleasing, according to cultural norms. Since people often respond socially to computer technology in ways that are similar to social interaction with humans [Reeves and Nass, 2003], by ascribing technology with human traits, we might expect to increase the social responses from people to that technology and, in this way, affect the calibration of human trust in the usage of such automated systems. In particular, they were interested in the effect that resulted from the interaction of etiquette with a cognitive factor, such as reliability, in order to be able to answer the question: Does good etiquette compensate for poor reliability and result in increased usage decisions? Does poor etiquette wipe out the benefits of good reliability?.

From this assumption, they manipulated these two factors and measured their combined effects on trust in the context of an airplane on-board computer monitoring system that gave instructions about the aircraft. They found out that user’s performance on the system was significantly better when the reliability of the system was high rather than low. However, good automation etiquette significantly enhanced performance, regardless of automation reliability. More interestingly, the effects of etiquette were powerful enough to overcome low reliability, with performance and trust in the low reliability/good etiquette condition not significantly differing from that in the high reliability/poor etiquette condition.

As we saw, the decision to rely or not on automated systems comes mainly from the perceived trustworthiness that...
they inspire, which is influenced not only by its perceived cognitive factors, such as reliability, but also by social and affective factors, such as etiquette. These social and affective effects are consistent with the idea that people are able to respond socially to technology, even when it doesn’t feature human-like methods of expressiveness. These social perceptions from technology can greatly affect the cooperation relationship that arises between a human and a system, to the point of compensating for poor reliability.

Transporting these results to the area of social robots, Cramer et al. [2010] explored the effects of providing robots with different empathic accuracy in different situation contexts. In their study, it was identified that trust decreased in general when empathic responses from robots were incongruent with the affective state of the user. However, in negative valence conditions, that is, in situations in which the outcome had been negative, robots that reacted as if the situation had been positive led to greater perceived empathic abilities.

3. Objectives

With this work we address the study of the effects evoked from a robot’s different levels of competence and empathic abilities in a scenario where it is offering guidance to humans. Empathy will be manipulated through its alignment with the user’s emotional state, while competence will differ between a high and low level. We are particularly interested in identifying if any of those combinations can affect trust to such an extent that a perception of a robot arises as a deceptive entity and, if so, what will be its effects. By considering solely empathy expressed through non-verbal behavior, we are also disregarding any form of verbal communication, focusing on a more restrict form of expression that is particularly relevant for the development of social robots.

A study will thus be conducted, built around a scenario with a robot serving as a guide, in which its non-verbal empathic ability and its competence rate will be manipulated. We will then analyze the effects of these different combinations. Our hypothesis are:

- **H1.1**: Robotic competence should positively influence the trust placed upon the robot, and the compliance to its instructions;
- **H1.2**: A robotic empathic alignment should positively influence the trust placed upon the robot, the compliance to its instructions and its perceived social skills, here conceptualized as friendship and social presence.

In this way, we assume that as a user perceives a robot as more competent, so does increase its perceived trustworthiness. On the other hand, a robot’s emotional alignment to the subject’s emotional state should affect not only its perceptions of friendship and social presence, but also its perceived trustworthiness.

4. Scenario

The testbed for this study was the well-know Memory Game. Instead of using actual cards, a computer game version was created with the screen showing N randomly dispersed pairs of cards facing down, thus hiding their symbol. Each card in a given pair features exactly the same symbol, but no two distinct pairs of cards feature the same symbol. The game then takes place in rounds with the player selecting a first card, revealing its symbol and then selecting another of the remaining facing down cards so that its hidden symbol matches the one of the first card. If both cards show the same symbol, that pair is considered found and is not able to be selected again in the future. Otherwise, if the card’s symbols don’t match, the selected cards are faced down again. The game progresses in this fashion until the user finds out every pair of cards. In this study’s version of the game, the screen presented an area with 36 cards facing down, with a cursor over one of them, which players could manipulate using the arrow keys. Upon pressing the return key, the card under the current cursor would be turned, showing its particular symbol. The cursor would then be moved to the desired second card, which would then be selected in a similar fashion, finally revealing either a match or not.

The objective given to the user was to try to match all the pairs of cards in the fewest number of rounds possible. However, during the game, the NAO robot was present nearby, and manipulated the gameplay by always expressing directional gestures that revealed to the test subject which second card the robot thought the subject should select. After every two cards were selected, the robot also appraised the situation by expressing some emotional reaction.

5 Design

In order to design a guidance robot for this particular game scenario, two main questions arose. The first question dealt with how to design the guiding communication protocol between the human and the robot, and the second dealt with how to design robot behaviors able to express emotions such as happiness, anger and sadness. To solve these two questions, an empirical approach was taken and a preliminary study was conducted with 5 pairs of participants involving a 16-card version of the same memory game. In it, participants were asked either to play the game itself or to act as the robot in one of two opposite scenarios:
helping or misleading the playing partner. In both of them, subjects acting as the robot had in front of them a screen showing the solution to the game being played by the partner. In the helping condition, the “robot” was supposed to help the partner achieve the best possible score, by guiding him correctly. In the misleading condition however, the guidance done by the “robot” should mislead the player in order to make him have the worst performance possible. In both of the conditions, nothing was said to the participants acting as the robot on how they should guide the player, but two restrictions were imposed: they were not allowed to verbally guide the player, so that only non-verbal communication would be possible, nor should they display any facial expressions that might influence the information being communicated. In order to achieve this last restriction, the participant acting as the robot was asked to wear a mask that only left the eyes uncovered. With these two restrictions enforced, we were thus able to observe and analyze the non-verbal facial-expressionless protocol that emerged between the “robot” and the player.

![Figure 1. A supervisor subject of the pre-study performing a deictic gesture](image)

By analyzing the actions and strategies employed by the humans playing the role of the robot when given goals similar to the ones of the robot in the main study being developed, we got a sense for how the robot behaviors of that scenario should be designed. From this preliminary study, a total of 49 unique behaviors were identified. 28 of these guided the design of the communication protocol between the robot and the human in the study being developed, as they were gestures that leaked to the playing subject information on which card they should select. Most of these consisted of deictic gestures, that is, directional pointing behaviors that expressed in which direction the player should move the game’s cursor. The other 21 behaviors consisted of emotional displays for basic emotions, head nods and playful behaviors in general. We also got a sense for the strategies that humans employed when given the task to mislead another subject. We observed how participants which were given the task of misleading the player used emotional expressions to be able to get away with the amount of incorrect instructions given and how they, in this way, turned the potential impression of deception into a general impression of lack of competence.

The communication protocol between the human and the robot was thus designed in our main study using the same ways the humans acting as the robot chose to communicate with their partner in this preliminary study, that is, using mostly deictic gestures. All the behaviors displayed by the robot in the main study were developed by mimicking as accurately as possible the gestures performed by the humans in the study and the emotional expressions were produced in a similar fashion.

6. Implementation

In order to develop the proposed scenario, several components were used, developed and integrated.

First of all, a game module was developed, responsible for dealing with the player input and providing the virtual game area. It managed everything related to the state of the game and, given this, it was also responsible for generating the perceptions that provided information to the other modules on what was happening in the game, so that the necessary actions could be conducted.

Then, the NAO Mind module was developed. The NAO Mind was then the component that modeled the mind of the robot, kept its state and changed it accordingly based on the low-level perceptions received from the game component. When the mind reached particular states, it would be also its job to order the execution of the necessary reactions by the physical robot.

However, the concern of modeling the emotional state of the robot was kept outside the NAO mind module, since another component, the FAtiMA agent architecture, had been built specifically for this purpose. FAtiMA expects some perceptions to be fed into its appraisal cycle that then shift the internally modeled emotional state, according to some rules scripted into the system, which also define which actions should be triggered when certain emotional states are reached. This component was thus connected to the NAO mind component, so that both the Mind could feed FAtiMA’s appraisal cycle with the necessary perceptions, and also so that the Mind could receive back the information on which actions should be triggered based on the current emotional state, so that the correct behaviors could be ordered for execution on the physical robot.

To connect the NAOMind with FAtiMA, a component called ION was used, while to connect the Mind with the robot, the library NAOMind was developed.
Figure 2 shows the general picture of the architecture involved in the scenario developed. In summary, the MemoryGame component continuously generated raw, low-level perceptions that required a change in the modeled state of the robot. They were thus sent to the NAO Mind component which interpreted them in order to shift the modeled state of the robot. Based on this state, the mind could then either trigger the necessary behaviors to be executed by the robot or feed higher-level perceptions to the Emotional Model, abstracted by the FAiTiMA agent architecture. This component continuously appraised the received perceptions and, upon certain emotional states, triggered reactions. These reactions were then sent back to the Mind, potentially launching the display of other behaviors on the physical robot.

7 Manipulations

In the experiment two variables were manipulated:

- Competence: the rate at which the robot’s instructions were incorrect, that is, how many times the robot’s instruction led the human to the wrong card, between:
  - High Competence: 3 out of 10 robot instructions were incorrect;
  - Low Competence: 5 out of 10 robot instructions were incorrect;

- Empathy: how the robot reacted after either a match or a non-match have occurred, between:
  - Empathically Aligned with the Subject’s emotional state: the robot reacted empathically in regards to the action performed by the subject, being generally Happy upon a match and Sad upon the failure to do so;
  - Empathically Mis-aligned with the Subject’s emotional state: the robot reacted in an empathically inverted manner, expressing Happiness upon a failure to match a card and Indifference or Anger upon a match;

From these two independent variables, 3 test conditions arose in which the study was based on:

- High Competence Robot and Empathically on the Subject’s Side;
- High Competence Robot and Empathically Against the Subject;
- Low Competence Robot and Empathically on the Subject’s Side;

8 Method

46 participants were gathered and a between-groups study was conducted. In all of them, participants were asked to play a full game of Memory supervised and manipulated by the robot in a way corresponding to the study condition.

Prior to that, participants filled out a pre-questionnaire which elicited information about their age and gender and general attitudes towards robots.

Participants were then briefed about the game they would be playing and were shown a video that explained how the robot would interact with them, in order to minimize potential misunderstandings in the communication protocol. In that video, the complete process of how the robot would guide the user was shown, from the moment that the first card was selected, until the confirmation that the user had reached the card that the robot thought was the right one. In an effort to not bias test subjects into over-trusting the robot, it was emphasized that it was not necessarily true that the robot would help the test subject or that it was its intention to do so.

During the game, and in each round, the robot would thus move in a idle way until the user selected the first card. At that moment, the robot would then point in one of four possible directions (up, down, left and right), indicating to the subject to which adjacent place he should move the cursor to. When finally the subject reached the desired card, the
robot would confirm that it should be selected by expressing a head nod. Upon the revelation of the second card, the robot would then appraise the situation by expressing an emotional reaction, depending if the user matched or not and if he did so by following or not a correct or incorrect robotic instruction.

Participants’ perceptions were captured by analyzing both their actual actions during the game as well as their answers to a post questionnaire. In this way, it was possible to distinguish and differentiate between what subjects perceived or stated and what they actually did. Upon the completion of the game, participants were debriefed, some of them were interviewed and one cinema ticket was offered as a reward for the participation.

9. Results

When analyzing the two factors manipulated in the study, a successful manipulation was identified for the independent variable of competence through the dependent variable of Perceived Competence (U = 82.000, p < .001, $M_{HC} = 5.3$, $STD_{HC} = 0.651$, $M_{LC} = 4.20$, $STD_{LC} = 0.941$), but not for the manipulation of empathy, through the dependent variable of Perceived Goal of the Robot (U = 224.500, p = 0.990). This means that participants truly perceived the less competent robot as such, but were not able to perceive the robot that reacted in an empathically mis-aligned manner as an entity that was against them during the game.

9.1 Manipulation of Competence

Statistically significant results for the manipulation of competence were identified on the scales of Actual Task Performance ($U = 31.500$, $p < .001$), Compliance ($U = 106.000$, $p = 0.004$), Trust in Provided Information ($U = 104.500$, $p = 0.003$), Dependability ($U = 132.000$, $p = 0.025$), Usefulness ($U = 144.000$, $p = 0.045$) and Trust ($F(1,43) = 5.194$, $p = 0.028$).

For the variable of Actual Task performance, it was found that the Low Competence group took significantly more significantly more rounds to finish the game than the High Competence group ($M_{HC} = 25.77$, $STD_{HC} = 1.794$, $M_{LC} = 32.53$, $STD_{LC} = 6.802$).

For the variable of Compliance, it was found that the Low Competence group presented significantly less following of the robot’s instructions when compared to the High Competence group ($M_{HC} = 71.82$, $STD_{HC} = 15.696$, $M_{LC} = 56.20$, $STD_{HC} = 21.153$, $U = 106.000$, $p = 0.004$).

Participants rated their Trust on the robot’s instructions (Trust in Provided Information) has higher in the High Competence condition when compared to the Low Competence one ($M_{HC} = 5.33$, $STD_{HC} = 1.184$, $M_{LC} = 4.75$, $STD_{LC} = 0.644$, $U = 104.500$, $p = 0.003$).

The dependability trust ratings on the robot were also higher in the High Competence condition when compared to the Low Competence one ($M_{HC} = 5.33$, $STD_{HC} = 1.184$, $M_{LC} = 4.75$, $STD_{LC} = 0.644$, $U = 132.000$, $p = 0.025$).

Participants rated the highly competent robot as significantly more useful than the low competent one ($M_{HC} = 6.36$, $STD_{HC} = 0.773$, $M_{LC} = 5.69$, $STD_{LC} = 1.109$, $U = 144.000$, $p = 0.045$).

At last, when analyzing the general variable of Trust, a statistically significant result was found between the Low Competence condition and the High Competence one ($F(1,43) = 5.194$, $p = .028$), with subjects rating significantly higher the Trust placed upon the Highly competent robot ($M_{HC} = 5.37$, $STD_{HC} = 0.932$, $M_{LC} = 4.73$, $STD_{LC} = 0.790$).

9.2 Manipulation of Empathy

For the manipulation of Empathy, and even though this manipulation was deemed as unsuccessful, a statistically significant result was found on the dimension of Utility of Winning ($U = 90.500$, $p = .001$). Subjects in the Emphatically Aligned Robot condition rated their own desire to perform well in the game as higher than in the Emphatically Misaligned condition ($M_{HE} = 6.30$, $STD_{HE} = 0.750$, $M_{LE} = 5.27$, $STD_{LE} = 0.884$).
10. Discussion

With this work, we tried to advance the knowledge on the area of social robots on the potential effects of competence and empathic abilities on the perceived trust, friendship and social presence of a guide robot. We conducted a study built around a memory game scenario in which a robot provided guiding instructions to the player on which card he should select. This guidance was provided at two different levels of competence, either high and low, and with featuring emotional responses either aligned to the subject’s emotional state or mis-aligned with it.

The results from our study showed that subjects perceived the difference between the two levels of competence present in our study. We found that participants interacting with a highly competent robot performed higher on the memory game they played, and presented higher compliance to the robot’s instructions. They also rated the perceived Trust, Usefulness and Dependability of the robot as higher in the highly competence conditions. With this, we found support for the first hypothesis of this study.

Competence can thus positively affect the trust and the compliance to the instructions of a robotic guide. This occurred because, as the players noticed that the guidance provided by a robot could fail, and did so quite frequently, more attention was paid to the task, and with it, a greater awareness emerged that enabled the players to more easily detect when the robot was giving a wrong instruction. The humans didn’t stop following the robot, but they would end up not complying to its instructions when they realized that they were wrong.

Users perceived the lack of competence of the robot, and thus trusted less in it and started executing the task on their own to a greater extent, instead of blindly following the robot. From the fact that the robot was less competent, its perceived Usefulness and Dependability levels were also compromised. However, lack of competence alone was not enough for the robot to be perceived as a deceptive entity, with some participants specifically highlighting in the interviews the fact that, even though they did perceive the robot as a not so competent entity, the fact that it seemed emotionally on the same plane as them, led them to excuse its failures.

However, the empathically mis-alignment to the subject’s emotional state, that made the robot react with emotions such as happiness in negative contexts and emotions such as anger in positive ones, didn’t lead to a perception of the robot as against the user. We didn’t thus find support for this second hypothesis.

What we did find out was that the interpretation of robotic non-verbal and facial-expressionless behavior is closely linked to the context to which those behaviors are expressed. Context is thus one of the most prominent factors in the interpretation of emotional non-verbal and facial-expressionless behavior. This means that, positive behaviors such as happiness when seen on the context of a positive event, such as the match of a card, are interpreted correctly, as shown by the interviews performed with the subjects in those conditions. However, exactly the same displays, when seen in a negative context, such as the mis-match of two cards, stop being interpreted as happiness, and the player tries instead to interpret them in a way consistent with the current context or with the mental model of the robot’s character, which may give rise to the expectation that the robot is there to help. According to the interviews, only few subjects observing a display of happiness upon a mis-match interpreted it as happiness. Instead, subjects ended up interpreting the robot as being angry at himself, surprised, or asking for forgiveness.

Surprisingly, and even though the users didn’t perceive the non-empathic robot as an entity against them in the game they were playing, they did perceive a difference, which led them to rate the Utility of Winning dimension as higher in the empathic condition. This means that, the manipulation introduced for empathy didn’t give away the idea to the player that the robot was against them, but they did find a different encouragement behavior from the empathically aligned robot when compared to the empathically mis-aligned one. This made the players that interacted with an empathic robot rate considerably higher their desire to win the game, as they felt more support coming from the robot on the task they were performing. With this, we can say that designers of human-robot interaction experiments should align emotional responses from robots with the events happening, in order to elicit the idea of the robot as a partner, so that the whole interaction can be performed as if the human and the robot are a team.

11 Conclusions

We explored whether the combination of non-verbal behavior and empathy could compromise its perceived trust-worthiness, to the point of being perceived as a deceptive entity.

What we found instead were important lessons for the area of the interpretation of robotic non-verbal behavior. We found out that, without the presence of facial expressions, the interpretation of robotic non-verbal behavior is closely tied to the task context, with the user’s trying to interpret the displays of the robot in a way consistent with it. This highlighted the fact that having a robot express its emotional state solely through non-verbal facial-expressionless behaviors carries a higher degree of ambiguity and allows for a much stronger role of context in the interpretation of the behaviors. Designers of interaction experiments should choose robots and embodiments that allow them to express
their emotional expressions, not only through body posture and movements, but also through facial-expressions so that the possible ambiguity in the interpretation of behaviors and the role of context in their interpretation can be reduced.

We didn’t thus find support for the idea that the manipulated values of non-verbal empathy and competence could give away the idea of the robot as an entity against the player during the game. A non-verbal empathic mis-alignment to the subject’s emotional state was also not perceived in a way that could compromise the robot’s trustworthiness.

Competence did however affect the perceived trustworthiness of the robot, leading to higher performance and compliance to the robot’s instructions. However, this effect was not strong enough to make the less competent robot seem as a deceptive entity. However, when analyzing the specific sub-dimensions of trust that were affected by the competence manipulation, we realized that the perception of the character of the robot was not compromised by it, but only its perception of reliability. This was also indicated by the fact that the perceived friendship and social presence of the robot were not influenced by its perceived competence. This means that ascribing emotional responses to robots may allow for a not so competent robot to still be regarded as having a trustworthy and friendly character, enabling users to forgive its poor performance. Since providing robots with emotional capabilities might be an easier task than increasing their reliability, this lesson becomes relevant for the designers of interaction experiments. In this way, instead of going through the effort of providing higher degrees of reliability to a robotic system in order to guarantee acceptability from user, we might achieve the same result by ascribing robots with empathic responses.

References


D. J. Mcallister. Affect and cognition-based trust as foundations for interpersonal cooperation in organizations. 1995.


