Radboud University Nijmegen

AI Bachelor Thesis

The Base of Trust in Human-Robot Interaction

Author:
M.L.E. de Bruijn (s4042026)

Supervisors:
R. van den Brule
Dr. W.F.G. Haselager

SOW-BKI300 AI Bachelor Thesis
Artificial Intelligence
Faculty of Social Sciences
Radboud University Nijmegen
July 19, 2013

Abstract

Trust in Human-Robot Interaction becomes more important as robots are emerging in the domestic environment. This thesis is focused on the influence of a previous encountered robot on how a second robot is perceived and especially how it influences the trustworthiness of the second robot. Both the influence of how well the first robot performs the task and the fluency of movements during the task of the first robot are being examined. Fur this purpose, an experiment showing two movie clips of a robot to the participants is conducted, followed by a questionnaire to obtain the opinion of the participants about the robot shown second. The fluency of movement of the first robot was contrasted in judgement of trustworthiness of the second robot. In other words, if the first robot did not move fluently, the second robot was judged to be more trustworthy than if the first robot moved fluently. Contrary, how well the task is performed by the first robot was unexpectedly assimilated in judgement of trustworthiness of the second robot. Put differently, if the first robot performed badly, the second robot was judged to be less trustworthy than if the first robot performed well.

1 Introduction

1.1 Trust in Human-Robot Interaction

The possibilities and capabilities of robots are increasing. Because of this, collaboration between human and robot is possible, in which a robot no longer has to be fully manipulated by a human. There are for instance autonomous robots in military operations[8], or robots used for medical applications[15]. Moreover, robots are introduced into the domestic environment[17], in which robots can affect our social daily life and environment, being in the same space as humans. A domestic robot can either work by itself or collaborate with humans. In the first case, a robot can keep out of the way of humans when needed. For instance, iRobot's Roomba[1] can vacuum clean without necessarily interacting with anybody. This is in contrast to the latter, in which a robot has to interact with humans, such as the human symbiotic robot TWENDY-one[5] or an intelligent home assistant as Care-O-Bot[3, 9].

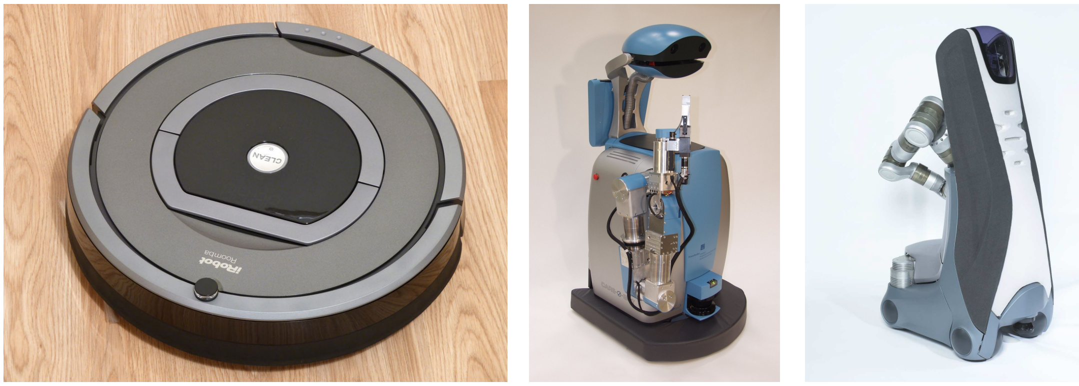
Domestic robots are fundamentally different from non-domestic robotics and other common domestic applications of advanced technologies[4, 17]. Whereas our homes are filled with advanced technologies, such as refrigerators, coffee machines, and vacuum cleaners, a robot has a more autonomous and physical presence[17]. This has a few consequences. Firstly, a domestic robot seems less stable and controllable than other domestic applications of advanced technologies, as domestic robots physically move around in, interact with and change the same surroundings as we do. This can even override our feeling of having control and decrease our feeling of safety. This depends on the size and capability of the robot and experience of the human[17].

For example, the iRobot's vacuum cleaning Roomba (Figure 1a). Since Roomba is small, not very dangerous-looking and can be turned off by pushing a button, there is a feeling of control. However, an intelligent home assistant, such as Care-o-Bot, is not as unambiguous as Roomba. It has multiple tasks and is equipped with multiple features to complete these tasks. For instance, Care-O-Bot II (Figure 1b) has the tasks of being a servant, walking assistant and communication tool[3]. Such a robot shows not only more complex behaviour than Roomba, but is also physically much bigger and more difficult to control. It is important for humans to feel safe around and to trust such a robot for them to use it beneficially.

Additionally, due to its autonomous and physical presence, a robot seems to have a mind of its own like a living entity[17]. Humans tend to anthropomorphize (and zoomorphize) robots[9, 12]. In this process, cognitive and emotional states are attributed to the robot and behaviour is interpreted as guided by rational choices and desires. The more capable and intelligent a robot seems to be, the more social understanding is expected of the robot. Robots can even be stereotyped based on their human-like face or voice[6].

For instance, iRobot's Roomba is clearly designed for one purpose (vacuuming), does not look human-like

[1] www.irobot.com



(a) iRobot's Roomba (b) Care-o-Bot II[3] (c) Care-o-Bot 3[9]

Figure 1: Overview of different domestic robots

and nobody is surprised if he does not understand whole monologues. However, there are plenty of families with the tendency to treat their Roomba as a pet, naming him, ascribing him a personality and building an intimate relationship with him[2, 14]. Care-o-Bot 3[9] (Figure 1c) is the successor of Care-o-Bot II. For the Care-o-Bot 3, it was deliberately decided to emphasize technomorphic perceptions, in order to reduce the anthropomorphic projections and interpretations of the robot.

Trust impacts the attitudes, behaviours and perceptions of humans with respect to robots and it affects the willingness of people to accept robot-produced information[4]. Therefore, trust is an important factor in human-robot interaction is domestic robotics. The current thesis focuses on the judgement of trustworthiness of a robot. As making a judgement of trustworthiness is a social judgement, the next section elaborates that topic.

1.2 Social Judgements

A social judgement of a target stimulus (the stimulus to be judged) depends on the context in which this stimulus is judged[11]. In order to form a social judgement of the target stimulus, an individual needs to create a cognitive representation of this target stimulus and needs to determine a standard of comparison to evaluate this target stimulus. Both the representation of the target and the standard of comparison are based on permanent as well as temporarily accessible information. Whereas the first is always accessible and thus independent of the context, the latter depends on information relevant for the judgement that is most accessible when making the judgement, and is thus context-dependent.

For a judgement, it matters whether the context information that comes to mind is sub- or superordinate to the target category[11]. Imagine you have to evaluate an strawberry ice lolly, the target category is then specifically this type of strawberry ice lolly. This is shown to you among different kinds and flavours of ice cream, which provides the context information. In this case, the context information is superordinate to the target category. However, if you would have to judge how you feel about ice cream in general (in this case the target category) and a specific kind of ice cream (e.g. a strawberry ice lolly) comes to mind (the context information), the context information is subordinate to the target category.

The relevant context information that comes to mind can either be assimilated of contrasted to the target. An assimilation effect occurs if the target is included in the superordinate category of the context information or if the context information is included in the superordinate category of the target. For the first case (inclusion of the target in the category of the context information), the features of this category and the category membership of the target are then included in the cognitive representation of the target. For the example of the ice creams, if you are asked to judge a strawberry ice lolly and this is shown among other types of ice creams (e.g. other ice lollies, scoop ice cream, soft ice cream) and you include the ice lolly in this category of ice creams in general, an assimilation effect occurs. The features of all kinds of ice creams can then be used to represent the ice lolly. For the second case (inclusion of the context information in the category of the target), your opinion of the context information influences your opinion of the target category. For the example of the ice creams, if you are asked how much you like ice cream in general (the target category) and you see your favorite ice lolly (the context information) and this is included in the target category, an assimilation effect occurs, making your judgement of ice creams more positive.

On the other hand, a contrast effect occurs if the relevant information that comes to mind is excluded from the target category. For the ice creams, this means a feature of the target is excluded from the representation of the target category . For instance, you are asked to judge a specific ice lolly but you notice that your favorite flavour is not present in the specific ice lolly to judge. This feature could then be excluded from the representation of the target, which results in a less positive judgement of the ice lolly than if the flavour would not have been excluded from the representation. For a contrast effect, the feature that is subtracted has to be valenced more extreme than the overall representation of the target. The valence of a feature is the degree to which an individual is attracted to or averted from the feature.

Furthermore, the information that is not included in the representation of the target may be included when constructing the standard of comparison. For this, the valence of the excluded feature has to be more extreme than the valence of the other information included in the standard of comparison. Due to this effect, the standard of comparison will be more extreme, leading to a more evident contrast effect. If you think of the missing flavour while determining the standard of comparison, this feature (the flavour) can be included in the standard of comparison, leading to more evident contrast effects than if it would not have been included. In this case the judgement of the ice lolly will be more negative if the missing flavour is included in the standard of comparison than if it is not.

If the context information is relevant for the judgement, it is either assimilated or contrasted to the target. There are a number of factors influencing whether a contrast or an assimilation effect emerges[11].

First, whether the context infomation belongs to the target category. This is more likely if, among other things, the information is representative for the target category, the target category is wide and if the stimuli are not perceived as distinct units of information. If the information does not belong to the target category, the information is excluded. Second, whether the context information comes to mind due to irrelevant influences. This could be if the participants are aware of the manipulation. If this is the case, the information is deliberately excluded. Third, whether the participant is intended to use the context information. Due to conversational norms, it can be that information that is used before is not used again, leading to a deliberate contrast effect. An example of this is that people exclude their happiness with their marriage in their happiness with life in general when they were already asked about their happiness with their marriage[11]. The happiness with their marriage will not be excluded from their happiness in general if they are not asked about the first beforehand. If all three of these factors do not lead to a contrast effect, the context information will be included in the representation of the target category, resulting in an assimilation effect.

An assimilation effect (inclusion of information that comes to mind) is assumed to be the default operation[11]. This has a few consequences. First, an assimilation effect is more likely to be obtained than a contrast effect. Second, in order for a contrast effect to emerge, more processing steps and effort are needed, thus more cognitive processing is necessary. Therefore it is possible for an assimilation effect to occur due to insufficient cognitive processing caused by for instance a cognitive load or lack of motivation.

Finally, if more information is included in a cognitive representation, each piece of information has less impact on the judgement[11]. For the example of evaluating a certain ice cream, if you base your judgement solely on whether the ice cream has a strawberry flavour, this feature of the ice cream has more impact on judgement than if the judgement is also based on whether the ice cream contains cream. Furthermore, the same information cannot be included in both the cognitive representation of the target and the cognitive representation of the standard of comparison[11].

1.3 Robot-Related Factors in Trustworthiness

Robot-related factors have a large influence on trust in human-robot interaction[4]. The extent to which behavioural style and level of task performance of a robot influence the social judgement of the trustworthiness of this robot is examined in an experiment by Van den Brule et al.[1]. This experiment made use of videos showing a robot performing the Van Halen Task, which will be explained in Section 2.2.

First, participants saw the baseline video, which was the same for all participants in order for them to get familiar with the task and appearance of the robot. Next, one of the sixteen test videos was shown, in which task performance and various behaviours were manipulated. A more extensive explanation of the manipulations relevant for the current experiment can be found in Section 2.1. To obtain the trustworthiness as judged by the participants, the experiment ended with a questionnaire.

The study showed that both level of task performance and fluency of movement of the robot influenced its judged trustworthiness significantly. The judged trustworthiness for a bad performing robot is lower than for a good performing robot, and the judged trustworthiness for a robot with trembling movements is lower than for a robot with smooth movements.

The difference in effect of motion fluency is larger for a good performing robot than for a bad performing robot. A possible explanation for this is that the level of task performance is probably not used in the representation of the good performing test robot. This is because the same information cannot be used in the cognitive representation of both the robot and its context, for which the good performing baseline robot is temporarily accessible. Here, only motion fluency can be used in judgement. When less information is used in making a judgement, each piece of information has more impact. For a bad performing robot, both task performance and motion fluency can be used in judgement, decreasing the impact of motion fluency relatively to its impact when judging a good performing robot.

1.4 The Current Experiment

The current experiment is focused on the influence of one robot, the baseline robot, on the trustworthiness of another robot, the test robot. As mentioned above, a baseline robot is used in order for participants to get familiar with the robot, followed by a test robot that has to be judged by the participants. The current experiment examines the influence of both the level of task performance and the fluency of motion of the baseline robot on the judged trustworthiness of the test robot. Next to trustworthiness, also the influence on the judgement of calibration, performance, motion fluency and expectations of mistakes made by a robot will be examined. For this purpose, whereas the baseline robot is manipulated, the test robot is fixed.

1.4.1 Research Questions

The main questions of the current experiment are whether the motion fluency and whether the level of task performance of a previous encountered robot is assimilated or contrasted to the trustworthiness of a robot. Next to trustworthiness, there will also be looked into judgement of calibration of behaviour and performance of, performance of, motion fluency of, and expectations of mistakes made by the second robot.

1.4.2 Expectations

For both motion fluency and task performance, a contrast effect on the judgement of trustworthiness of the test robot is expected. A contrast effect occurs, as explained in Section 1.2, if context information is excluded from the cognitive representation of the target stimulus and possibly is included in the cognitive representation of the standard of comparison. For motion fluency, this means that the test robot will be judged more trustworthy if the baseline robot trembles than if the baseline robot moves smoothly. For level of task performance, this means that the judged trustworthiness of the test robot is expected to be higher if the performance of the baseline robot is bad than if the baseline robot performs well. Moreover, the difference in judged trustworthiness of the test robot between seeing it after a smooth or trembling baseline robot is expected to be larger after a good performing baseline robot than after a bad performing baseline robot.

There are a number of reasons why a contrast effect is expected. As mentioned above, there are a number of factors involved whether information that comes to mind is included or excluded. One of these factors is whether the information belongs to the target category. When the information from the features of the baseline robot comes to mind, we expect that this will be judged not to belong to the target category. This is expected because first of all, it will be mentioned multiple times during the experiment that the baseline and test robot are different robots. On top of that, the robots have a different appearance. This is both to prevent that participants will think that the two robots are the same robot, thus the stimuli are presented as distinct units. Moreover, the target category is specifically the test robot, which is a small category relative to robots in general.

As mentioned, the difference in judged trustworthiness of the test robot between seeing it after a smooth or trembling baseline robot is expected to be larger after a good performing baseline robot than after a bad performing baseline robot. The reason for this is firstly that the same piece of information cannot be used to represent both the target stimulus and the standard of comparison. The target stimulus is in this case the good performing test robot, and the standard of comparison can be influenced by the baseline. Secondly, as more information is included in judgement, each piece of information has relatively less impact on judgement. Applied to the current experiment, if the baseline robot performs well, the level of task performance cannot be used in the representation of both the target stimulus and the standard. On the other hand, if the baseline robot performs badly, level of task performance differs for the baseline and the test robot and can thus be used in both representations. Therefore, if the baseline robot performs badly, relatively more information can be included in judgement, decreasing the impact of motion fluency. This decreases the difference in judged trustworthiness of the test robot between a trembling and a smooth baseline robot when the baseline robot performs badly, relatively to a good performing baseline robot.

2 Method

As mentioned, the current experiment is focused upon the influence of the level of task performance and fluency of movement of a baseline robot on the judged trustworthiness of the test robot.

2.1 Manipulations

Both level of task performance and motion fluency were manipulated in the baseline robot. A good performing robot performed the task without making any mistakes, while a bad performing robot removed two of the brown balls and twice a different coloured ball instead of removing the brown ball before or after. A smooth movement style involved a fluent movement while reaching for the balls, whereas a trembling movement style was created by a rapidly shaking arm of the robot. The baseline robot always looked straight ahead and did not hesitate in removing the balls. The test robot always performed well, removed the balls smoothly, looked ahead and did not hesitate in removing the balls from the conveyor belt. Thus, the experiment consisted of four different conditions. These conditions resulted in a 2 (level of task performance: good or bad) x 2 (motion fluency: smooth or trembling) between subject design.

2.2 Procedure

In order to test the influence of the baseline robot on the test robot, different videos of a robot performing the Van Halen Task were used. The videos used in the experiment by Van den Brule et al.[1] are reused in this experiment. The robot in the videos is based on TWENDY-one[5], with some small adaptions, as described by Van den Brule et al.[1]. Each video had a duration of approximately 40 seconds.

In the Van Halen Task, a robot is placed behind a conveyor belt. On this conveyor belt, fourteen coloured balls move along from right to left (from the participant's point of view). It is the task of the robot to remove all four of the brown balls, while leaving the balls with a different colour move along.

First, one of the baseline videos is shown to the participants, followed by the video of the test robot (Figure 2). The robot in the test video has a green colour, whereas the baseline robot is coloured red, in order to show the participants it involved a different robot. Furthermore, it was explicitly mentioned between the videos that the test robot was not the same robot as the baseline robot. Last, the participants were asked to fill out a questionnaire, to obtain their opinion of the test robot. The total experiment took approximately three minutes.

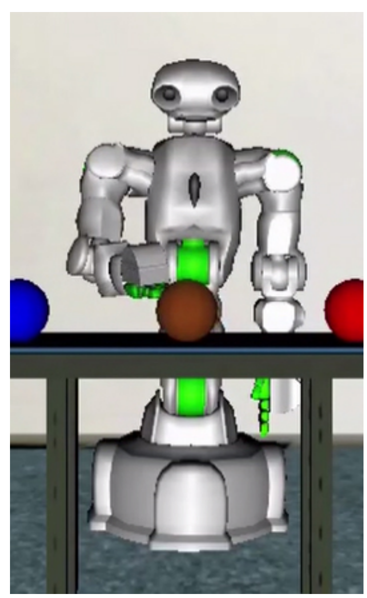


Figure 2: The test robot

2.3 Questionnaire

After watching both videos, the participants were asked to fill out a questionnaire. Here, it was explicitly mentioned that the questions involved the test robot and again that the baseline robot was a different robot. All rating were recorded on 7-point Likert scales, with a score of 1 meaning complete disagreement and 7 meaning complete agreement. The questionnaire was designed to obtain the opinion of the participants about the test robot, which included the following components.

2.3.1 Manipulation Checks

The participants were asked a few manipulation checks, in order to obtain how well they thought the robot performed his task, moved fluently doing so and to what extent they expected the robot to make mistakes. Furthermore, both whether the participant had seen the robot before in another experiment and whether the participant could explain the task of the robot were checked. The last was an open ended question. In order to explain the task of the robot sufficient, a participant had to mention that the robot had to sort or select balls. If the participant did not succeed in the explanation, he or she was a priori excluded from the analysis, and a replacement of this participant was run.

2.3.2 Trustworthiness

Four questions were asked to obtain the opinion of the participants about the trustworthiness of the test robot. The scale of trust was formed by the extent to which participants thought the robot was reliable, the extent to which the participants would trust the robot to sort balls without human supervision, and both the extent to which they wanted to judge the robot positively and negatively. Unfortunately, the reliability, measured by Cronbach's alpha, of the trust scale was low ($\alpha = 0.69$). However, as the same scale has been used before in a couple of other experiments (i.a. [1, 7, 10]), and the reliability of the scale in these experiments was high enough, this scale will still be used for analysis.

2.3.3 Calibration

How well the level of task performance aligned with the behaviour of the robot was investigated by two questions. The first asked to which extent the performance of the robot surprised the participants, based on the impression he had made, and the second asked to which extent the behaviour of the robot suited his performance. These questions constituted the calibration scale. However, since the reliability, again measured by Cronbach's alpha, of this calibration scale appeared to be too low to use ($\alpha = 0.39$), the questions constituting this scale are processed as separate items; item 1 refers to the extent to which the performance of the robot was surprising, based on the impression the robot made, and item 2 refers to the extent to which the behaviour of the robot suited his performance.

2.3.4 Demographics

Afterwards, both the age and gender of the participants were asked.

2.4 Participants

82 participants were recruited at the faculty of Social Sciences and could, in return for their participation, choose among different kinds of candy. Since two participants were excluded, a total of 80 participants remained. The first participant was excluded because he indicated to have prior knowledge about the experiment and the second because she could not tell what task the robot in the experiment was supposed to perform after the experiment was over, which was asked in the questionnaire, as mentioned above. This left 80 participants for analysis, of which 62 are female, ages 17-37 (median 22).

It should be noted that 19 of the participants indicated in the questionnaire that they had seen this robot before in another experiment. A 2 x 2 between-subject ANCOVA conducted over the manipulations checks, trustworthiness and the two items of calibration, with whether participants recognized the robot as covariate, showed no significant effect, $ps > .163$, of recognition.

2.5 Analysis

A 2 x 2 between-subject ANOVA is conducted on the items of the questionnaire. If the item was not normally distributed, still the 2 x 2 between-subject ANOVA was conducted. However, if this is marginally significant, the item is also analysed using a non-parametric analysis (bootstrap ANOVA).

3 Results

3.1 Manipulation Checks

3.1.1 Performance

All participants rated the performance of the test robot very high (all 6 or 7), which was not significant influenced by task performance or fluency of movement of the baseline robot (all $Fs < 1$, all $ps > .3$). Thus, the performance rating of the test robot was not influenced by the baseline robot.

3.1.2 Motion Fluency

The rating of motion fluency of the test robot is significantly influenced by the motion fluency of the baseline robot, $F(1,76) = 4.95$, $p = .029$, $\eta_p^2 = .06$. The test robot is judged to be more fluent when the baseline robot moves trembling ($M = 4.35$, $SD = 1.27$), than when the baseline robot moves smoothly ($M = 5.08$, $SD = 1.59$). This indicates a contrast effect of motion fluency of the baseline robot on the judged fluency of the test robot. Level of task performance of the baseline robot showed no significant effect and also no significant motion fluency x performance interaction effect was found (all $Fs < 0.2$, all $ps > .7$).

3.1.3 Expectation of Mistakes

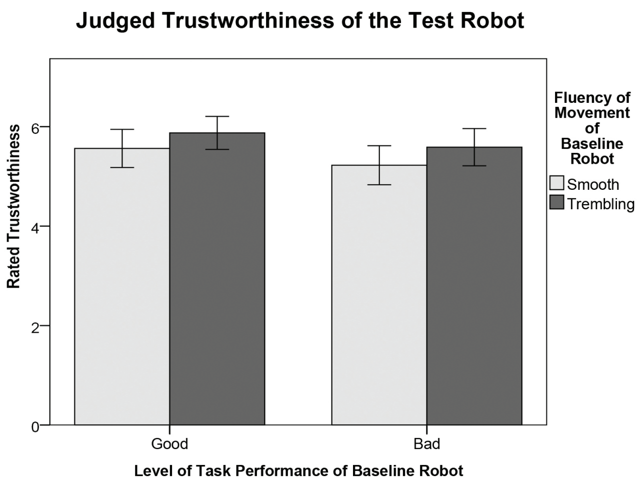
Both fluency of movement of the baseline robot and task performance of the baseline robot showed a significant effect on the rated expectation for the test robot to make mistakes.

The effect of motion fluency was significant, $F(1,76) = 5.71$, $p = 0.019$, $\eta_p^2 = 0.07$, and indicated that the test robot was expected to make more mistakes if this robot was preceded by a baseline robot with smooth movements ($M = 4.35$, $SD = 2.01$), than when preceded by a baseline robot with trembling movement ($M = 3.42$, $SD = 1.78$). This indicated the expected contrast effect of motion fluency on the expectation of mistakes.

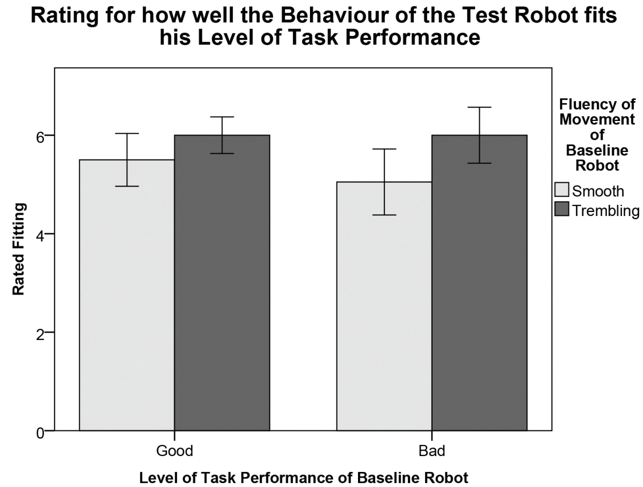
Furthermore, the significant effect of task performance, $F(1,76) = 17.63$, $p < .001$, $\eta_p^2 = .19$, indicated that the test robot was expected to make more mistakes if this robot was preceded by a bad performing baseline robot ($M = 4.70$, $SD = 1.70$), than when the baseline robot performed well ($M = 3.07$, $SD = 1.85$). Not in line with the expectations, this appears to be an assimilation effect of level of task performance on the expectation of mistakes. No significant motion fluency x performance interaction effect was found, $F(1,76) = 0.1$, $p > .7$.

3.2 Judgment of Trustworthiness

The main effect of level of task performance of the baseline robot and motion fluency of the baseline robot were both marginal significant on judgement of the trustworthiness of the test robot (Figure 3a).



(a) Rated trustworthiness of the test robot over different levels of task performance of the baseline robot(x-axis) and motion fluency of the baseline robot(colours). The error bars represent 95% confidence intervals.



(b) Rated fitting of the behaviour of the test robot to its level of task performance over different levels of task performance of the baseline robot(x-axis) and motion fluency(colours) of the baseline robot. The error bars represent 95% confidence intervals.

Figure 3: Rated (a) trustworthiness and (b) fitting of the behaviour to the level of task performance of the test robot

The main effect of motion fluency, $F(1{,}76) = 3.62$, $p = .061$, $\eta_p^2 = .05$, indicated that a robot is rated as more trustworthy when this robot is preceded by a trembling robot ($M = 5.73$, $SD = 0.76$) than when the robot is preceded by a smooth moving robot ($M = 5.39$, $SD = 0.84$). This is in line with the contrast effect that was expected to be found. The main effect of task performance, $F(1{,}76) = 3.10$, $p = .082$, $\eta_p^2 = .04$, indicated that a robot is rated more trustworthy when preceded by a good performing robot ($M = 5.72$, $SD = 0.77$) than when preceded by a bad performing robot ($M = 5.41$, $SD = 0.83$). This is not in line with the expected contrast effect, but seems to be an assimilation effect. No significant motion fluency x task performance interaction effect was obtained, $F(1{,}76) = 0.02$, $p > .8$.

Since the data of the trust scale is not normally distributed and has a relatively low reliability and shows to be marginally significant using 2 x 2 between-subject ANOVA, also a bootstrap ANOVA using 2000 bootstrap samples was conducted. This showed, too, a marginally significant effect for task performance of the baseline robot ($p = 0.062$). However, the effect of motion fluency of the baseline robot appeared to be significant ($p = 0.022$).

3.3 Calibration

There were no significant effects of task performance and motion fluency of the baseline robot on the first item, the surprise of the performance, based on the made impression (all $Fs < 2$, all $ps > .160$).

Furthermore, there is no significant effect of level of task performance of the baseline robot on item 2, the extent to which the behaviour fits the performance, nor a significant interaction effect is found (both *Fs* = 0.74, all $ps > 0.3$). However, there is a significant effect of motion fluency of the baseline robot on this item (Figure 3b). This effect, $F(1{,}76) = 7.69$, $p = .007$, $\eta_p^2 = .09$, indicates that the behaviour of the test robot is considered more fitting to its level of task performance when this robot is preceded by a trembling baseline robot ($M = 6.00$, $SD = 1.01$), than when the baseline robot moved smoothly ($M = 5.28$, $SD = 1.30$). Thus, motion fluency showed a contrast effect on the second item of calibration.

Since no marginal significant results were obtained for calibration, no non-parametric analysis was conducted, even though the data of this scale is not normally distributed and has a low reliability.

4 Discussion

4.1 Summary of Results

The main question of the experiment is whether the level of task performance and the whether fluency of movement of the baseline robot influence the judged trustworthiness of the test robot. The information of motion fluency of the baseline robot is contrasted to the test robot. Hence, the test robot is judged to be more trustworthy if the baseline robot trembled than when the baseline robot moved smooth. However, the information of level of task performance of the baseline robot is not contrasted, as was expected, but rather seems to be assimilated to the test robot. This means that the test robot is judged to be less trustworthy after a bad performing than after a good performing baseline robot.

These effects were not only found for judged trustworthiness, but also for some other variables. After a trembling baseline robot, the test robot is judged to be more fluent, is expected to make less mistakes, and its behaviour is considered to be more fitting to its level of task performance, than after a smooth baseline robot, showing the contrast effect for motion fluency. Additionally, for the assimilation effect of level of task performance, after a bad performing baseline robot, the test robot is expected to make more mistakes than after a good performing baseline robot.

Apparently, the information of motion fluency of the baseline robot is excluded from the cognitive representation of the test robot, whereas the information of task performance of the baseline robot is included in the cognitive representation of the test robot. More generally, when people have met a robot, both its behavioural style and task performance leave an impact on how they perceive a next robot. First of all, when a robot behaves in a trustworthy manner, its trustworthiness is not negatively influenced by (previously encountered) robots that behaved untrustworthy. Secondly, when a robot performs well, but you have experience with bad performing robots, this negatively influences the trustworthiness of the well performing robot.

4.2 Assimilation Effect of Task Performance

This leaves us with the question of why the level of task performance is assimilated even though this was not expected and even while the motion fluency does show the expected contrast effect.

Firstly, for a contrast effect to occur, it has to be decided by the participants making the judgement that the information that comes to mind does not belong to the target category and/or comes to mind due to irrelevant influences and/or that he/she is not intended to use that information. It was expected for the participants to see the robots independent and separate from each other, thus that the relevant information that came to mind was judged not to belong to the target category. The unexpected assimilation effect for task performance could indicate that the robots were not seen as separate and independent as intended, even though it was clearly shown and mentioned multiple times during the experiment. This could have been caused by the fact that the robots in the current experiment performed the same task and were based on the same robot. Moreover, the data from the experiment appeared to be binomial distributed. A possible cause for this could be that the participants interpreted the questions in two different ways. This hypothesis is based on the experiences of the experimenter during the experiment. Some of the participants indicated to be unsure in how to interpret some of the questions, and more specifically, which robot to involve in their judgement. For example, for the question (translated from the original in Dutch) 'To what extent did the performance of the robot surprise you, based on the impression he made on you?', some participants doubted whether the 'he' was the test robot, or if they had to think of the baseline robot as well. Since not all participants would indicate doubt, it is reasonable to assume more participants were unsure. Therefore, it could be that the participants were divided in the way to answer this and possibly other questions, and in which robots they involved in their judgement. However, this does not provide an explanation for the difference in effect between task performance and motion fluency.

It could also be the case that motion fluency was decided to belong to the target category, as task performance. Motion fluency would still be excluded if the participants decided for motion fluency, but not for task performance, that it came to mind due to irrelevant influences and/or that they were not intended to use that information. Motion fluency could have been deliberately excluded, maybe because the manipulation was more obvious or maybe because of the order of the questions asked.

Moreover, an assimilation effect can also occur as default mode. As mentioned in Section 1.2, more cognitive processing is necessary for a contrast effect. If participants are not motivated to process the information sufficiently, an assimilation effect is more likely to emerge. Motivation is higher for higher relevance of judgement for an individual's personal goals, higher need for closure and higher fear of invalidity[11]. In situations where robots and humans rely on each other, as is the case for domestic robotics, it matters for the individual how the robot behaves, performs and if he can be trusted. However, in the current experiment, it did not affect the participants how the robot behaved or performed, or whether they could trust him. It could have been that participants were not enough motivated to process the information sufficiently, resulting in an assimilation effect. But this statement is inconsistent with the contrast effect of motion fluency.

When participants get to know about the task of the robot, they can easily form an impression of how the robot would perform the task well or badly. Level of task performance is an explicit source. Contrary, the way of performing the task (motion fluency) is more subjective. According to dual-process models[13], more subjective information sources are processed automatically and unconsciously, whereas explicit sources are processed more deliberately and consciously[1]. Subjective sources, such as behavioural style and appearance, have shown to be processed automatically in human-human interaction. For instance, the judgement of, among other things, trustworthiness or likeability of an human face has shown to be made after 100 ms. of exposure[16], which was highly correlated with the judgements of the same traits without time limit. As mentioned above, insufficient processing leads more likely to an assimilation effect. Contrary to deliberative processes, automatic processes are not influenced by insufficient processing. This could explain why a lack of motivation could have resulted in an assimilation effect for task performance, but not for motion fluency. Unfortunately it should be mentioned that processing of motion fluency in human-robot interaction could not be shown to be automatic and unconscious by Van den Brule et al.[1].

Even if there is no difference in processing between level of task performance and motion fluency, this does not change that forming an impression of how the robot would perform the task well or badly is easily done. For participants to judge the performance of the test robot, the baseline robot (the context information) had no effect. On the other hand, judgement of motion fluency did get significantly influenced by the motion fluency of the baseline robot. When participants saw a trembling baseline robot, they used this information to judge the test robot more fluency.

Furthermore, there could be another difference between the manipulations of task performance and motion fluency. The robot trembled, or not, for all four balls to be removed, whereas task performance only differed for two balls, which could have made motion fluency more notable than task performance. This leaves the possibility of insufficient processing for task performance, but not for motion fluency as possible cause for the unexpected assimilation effect of task performance even while motion fluency was contrasted.

Besides the unexpected assimilation effect of level of task performance, another expectation was also not fulfilled. The difference in judged trustworthiness between trembling or smooth movements was expected to be smaller when the baseline robot performed badly than when the baseline robot performed well. This difference was expected to emerge from the fact that in case of a bad performing baseline robot, more pieces of information had to be included in the cognitive representation of the standard, having each less impact on judgement. However, no such difference emerged. Since level of task performance showed an assimilation effect on judgement, only motion fluency could have influenced the standard to which the test robot was compared. Hence, for both good and bad performance of the baseline robot, the same number of pieces of information can actually be contrasted, namely only motion fluency.

4.3 Appearance

In order to emphasize the difference between the baseline and test robot, they were given a different colour. The baseline robot is red, whereas the test robot is green. However, this could have affected the attitude of the participants towards the robots, for instance by having a preference for either red or green, or by triggering a good (green) vs. bad (red) difference. Although, according to Hancock et al.[4], the attitude towards a robot is a human-related factor for which only limited evidence was found.

4.4 Future Research

Following the speculations above, we will propose a number of approaches for future research.

Firstly, in this experiment, participants were shown the baseline robot directly before the test robot to judge. A possible approach for future research could be to extend the period between the first and second robot. This could for example be accomplished by giving participants some short, not-related tasks in between the videos of the two robots. If it is indeed the case that in the current experiment the robots were not seen as independent and separated as intended, this extend between the robots is expected to separate the robots more, which is expected to result in a contrast effect for both motion fluency and level of task performance. However, it should also be considered that if the time between the robots is extended, the baseline robot could be less temporarily available for constructing a cognitive representation, maybe decreasing the magnitude of the effect.

Furthermore, it could be interesting to provide the participants with different kinds of robots performing the same task to see if the assimilation effect of level of task performance only emerges for one kind of robot, or if the task performance of one kind of robot also is assimilated in judgement of another kind of robot. Reversed, it could also be interesting to provide the participants with the same kind of robot, but now they perform different tasks. This will show whether the effects also transfer over different tasks.

Thirdly, by presenting participants a cognitive load during the video of the test robot it could be determined whether motion fluency is processed deliberatively or automatically. If this is a deliberative process, participants in the cognitive load condition will not take motion fluency into account while making the judgement, and thus judge different than the control group (participants without cognitive load).

Fourthly, the robot in the current experiment did not affect the participants, as it did not personally matter to them how the robot performed or behaved or whether they could trust the robot. If an experiment can be created in which the judgement of trust is more important for the participant, the participants will be more motivated to sufficiently process all information in order to make a correct judgement. The judgement of trust would be more important for participants when they gain from making the correct judgement and/or when an incorrect judgement harms them. For instance if the correct judgement of Roomba saves you a lot of time cleaning or misjudging the Roomba wrecks your furniture. This is expected to lead to an increased magnitude of the effects and for the results of that experiment to show whether the found assimilation effect of task performance is really an assimilation effect or if it only occurred due to lack of motivation or due to distraction.

An important aspect of domestic robotics is that these robots are present in our social daily lives. The current experiment involved a robot placed behind a conveyor belt. For this purpose, the experiment could be redesigned towards a more domestic setting and task. Furthermore, the robot performed his task and the participants watched the robot perform the task, but the experiment did not require any interaction. in addition, interaction could be another way to motivate participants to pay attention to the robot.

Finally, to examine the influence of the colour of the robot, an experiment showing two robots could be conducted for which the colours are reversed for half of the participants. If the judgement of the second robot does not significantly differ from whether this robot is red or green, an effect by colour on judgement can be excluded.

The current thesis has shown that the influence of previously encountered robots should be taken into consideration when being concerned with and has given propositions for the future of trust in human-robot interaction.

References

[1] R. van den Brule, R. Dotsch, G. Bijlstra, D. H. J. Wigboldus, and W. F. G. Haselager. Human-Robot Trust - A Multi-Method Approach. *In press*, 2013.

[2] J. Forlizzi and C. DiSalvo. Service Robots in the Domestic Environment: a Study of the Roomba Vacuum in the Home. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 258–265. ACM, 2006.

[3] B. Graf, M. Hans, and R. D. Schraft. Care-O-bot II - Development of a next generation robotic home assistant. *Autonomous robots*, 16(2):193–205, 2004.

[4] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. C. Chen, E. J. De Visser, and R. Parasuraman. A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53(5):517–527, 2011.

[5] H. Iwata and S. Sugano. Design of Human Symbiotic Robot TWENDY-ONE. In *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*, pages 580–586. IEEE, 2009.

[6] S. Kiesler, A. Powers, S. R. Fussell, and C. Torrey. Anthropomorphic Interactions With a Robot and Robot-Like Agent. *Social Cognition*, 26(2):169–181, 2008.

[7] M. E. U. Ligthart, R. van den Brule, and W. F. G. Haselager. Human-Robot Trust: Is Motion Fluency an Effective Behavioral Style for Regulating Robot Trustworthiness. *In preparation*, 2013.

[8] P. Lin, G. Bekey, and K. Abney. Autonomous Mlitary Robotics: Risk, Ethics, and Design. Technical report, DTIC Document, 2008.

[9] C. Parlitz, M. Hägele, P. Klein, J. Seifert, and K. Dautenhahn. Care-O-bot 3- Rationale for human-robot interaction design. In *Proceedings of 39th International Symposium on Robotics (ISR), Seul, Korea*, pages 275–280, 2008.

[10] L. van de Ree, R. van den Brule, and W. F. G. Haselager. The Influences of Looking Behavior on the Trustworthiness of a NAO Robot. *In preparation*, 2013.

[11] N. Schwarz and H. Bless. *Constructing Reality and its Alternatives: An Inclusion/Exclusion Model of Assimilation and Contrast Effects in Social Judgment.* Lawrence Erlbaum Associates, Inc, 1992.

[12] M. Scopelliti, M. V. Giuliani, and F. Fornara. Robots in a Domestic Setting: a Psychological Approach. *Universal Access in the Information Society*, 4(2):146–155, 2005.

[13] E. R. Smith and J. DeCoster. Dual-Process Models in Social and Cognitive Psychology: Conceptual Integration and Links to Underlying Memory Systems. *Personality and Social Psychology Review*, 4(2):108–131, 2000.

[14] J. Sung, L. Guo, R. E. Grinter, and H. I. Christensen. "My Roomba Is Rambo": Intimate Home Appliances. In *UbiComp 2007: Ubiquitous Computing*, pages 145–162. Springer, 2007.

[15] K. M. Tsui and H. A. Yanco. Assistive, Rehabilitation, and Surgical Robots from the Perspective of Medical and Healthcare Professionals. In *Proceedings of AAAI Workshop on Human Implications of Human-Robot Interaction*, pages 34–39, 2007.

[16] J. Willis and A. Todorov. First Impressions: Making Up Your Mind After a 100-ms Exposure to a Fface. *Psychological science*, 17(7):592–598, 2006.

[17] J. E. Young, R. Hawkins, E. Sharlin, and T. Igarashi. Toward Acceptable Domestic Robots: Applying Insights from Social Psychology. *International Journal of Social Robotics*, 1(1):95–108, 2009.