# Mate or weight? Perceptions of a robot as agent or object in a creative problem solving task

Theresa Law[1 [0000−0003−2027−9240]], Daniel Kasenberg[1 [0000−0003−3148−1663]], and Matthias Scheutz[1 [0000−0002−0064−2789]]

Tufts University, Medford MA 02155, USA {theresa.law, daniel.kasenberg, matthias.scheutz}@tufts.edu

**Abstract.** Humans tend to ascribe agency to non-human entities, including robots. In this experiment we examine the extent to which this agency ascription can be overcome when required for completion of a task. In particular, we construct an "escape room"-like task, completion of which requires the participant to treat a robot as an object to be physically manipulated rather than an agent with which to be interacted. We compare conditions in which the robot is (minimally) interacting with the participant, inert, or entirely absent (replaced by another inanimate object). While the data are inconclusive, preliminary results are promising for the hypothesis that participants' perception of the robot's agency undermines their ability to complete the task.

**Keywords:** Human-Robot Interaction, robot agency, robot attributes, creative problem solving

## 1 Introduction

Humans tend to ascribe agency to non-human entities, particularly when these entities are capable of motion [9]. This tendency extends even to simple geometric shapes [10], but is especially prevalent with social robots, which are explicitly designed to be interacted with as social agents [4, 2]. That classic social psychology experiments (such as the Asch conformity experiment [1, 20] and the social facilitation effect [15, 21]) can be replicated with robots suggests (1) that this ascription may be automatic rather than deliberate; and (2) that it may persist even when detrimental to the task at hand.

It nevertheless remains unclear exactly how strong the predisposition to ascribe agency to robots is, or under what conditions it can be overcome. With sufficient prompting or upon reflection, humans may admit that a robot is not a person, and draw distinctions between human and robot agency [24]. Whether such distinctions can be drawn in-the-moment, without explicit prompting from either robot or experimenter, remains unclear. Can humans override their ascription of agency to a robot, viewing it as a physical object to be manipulated rather than as a social agent to be interacted with, when doing so is task-essential? If so, is this done easily and readily, or does it require creatively thinking through the problem, resulting in extra time and effort?

In this experiment, we construct a task in which successful completion hinges upon treating a robot as an manipulable object rather than as a social agent. To the best of our knowledge, our study is the first to employ such a task. We hypothesize that participants will implicitly view the robot as an agent, and that this will impede their ability to carry out the task, particularly when the robot is active and interacting with the human.

## 2 Motivation and Background

Perceptions of social agency can be tied to the level at which a person anthropomorphizes a robot [6] [28]. Epley et al. [7] theorize that three factors which increase levels of anthropomorphism are people's default usage of anthropocentric explanations until they have a passable mental-model of the non-human agent, seeking motivation for the agent's behavior, and a desire for social connection. These factors are prevalent to varying degrees in most human-robot interactions [8] [11]. The first may be especially sensitive to the types of social cues robots perform.

At a minimum, a robot's behavior must indicate some form of communication to be perceived as being social [5]. Motion can be a particularly poignant modality of social cue and communication; Breazeal and Fitzpatrick [3] have posited that all movement is interpreted as semantically meaningful, regardless of whether it was intended to be or not. Head movement [26], arm and hand movement [19], and full body movement [14] [12] have all been shown to affect human's perceptions of a robot interactant.

Though it is often beneficial to evoke perceptions of a robot's social agency in a human-robot interaction, there are times when doing so can cause consequences. In interviews of Roomba owners, Sung et al. [23] found that people who became attached to and anthropomorphized their robots did things like clean up for it on certain days to give it a "break." Rather than allow the robot to make their lives easier by doing chores, these people added the chore back into their lives to accommodate their robots' "feelings." This simple example shows that ascribing social agency to robots is not always desirable. Nevertheless, Many HRI studies take the desirability human perception of the robot as social agent for granted; several studies consider how various aspects of robot design and behavior can be modified to more effectively evoke this perception [27, 22].

To the best of the authors' knowledge, this is the first study to approach agency ascription through designing a task in which successful completion depends upon "objectifying" a robot. Some studies use subjective measures (such as Likert scales) to measure the extent to which the human makes this ascription [16]; others explicitly assess how the perception of agency *improves* performance on collaborative tasks. Where such studies emphasize how treating the robot as a social agent can help task performance, we focus on situations in which doing so is counterproductive to the task at hand. Perhaps most similar to our research is work replicating social psychology experiments (such as the Asch study on conformity to groups [1, 20]) with robots, which suggests that the inherent tendency to treat robots as social agents may occasionally be detrimental to task performance. These studies do not explicitly examine whether the human can override their perception of the robot's agency; rather, they examine whether this perception results in the same psychological effects as with other humans.

## 3 Experiment

### 3.1 Hypotheses

We are interested in exploring whether humans will be able to conceptualize a robot as an object rather than as an agent (and thus overcome a bias toward treating the robot as an agent, especially when the robot is turned on and active) when doing so is necessary for a task. To do this we will evaluate the differences in participants' performance on a task between the following three conditions:

- ***Inert*** *robot condition:* The robot must be used as an object to complete the task; the robot is turned off.
- ***Active*** *robot condition:* The robot must be used as an object to complete the task. The robot is turned on: its eyes blink, and its plow and head move up and down in an idle manner. It also greets the participant by saying "Hi my name is Cozmo welcome to the game." Additionally, at the halfway mark of the task, Cozmo says "One minute has passed. One minute remaining."
- ***Control*** *condition:* An empty pencil holder basket filled with the other objects on the table (pens and pencils, small toys) must be used to complete the task.

Our definition of viewing the robot as an object is the person physically manipulating the robot; therefore, we use time to touch the robot (time-to-touch; TTT) as a measure for indicating when the participant started seeing the robot as an object. For participants who do finish the task, we are also interested in the total time to completion of the task (time-to-completion; TTC).

Table 1 indicates the various hypotheses we are testing quantitatively. The first pair of hypotheses (**Hypothesis 1a** and **Hypothesis 1b**) refers to whether participants' bias toward conceptualizing the robot as an agent will prevent or impede them from successfully solving the task, when the robot is active and (minimally) interacting with them. We hypothesize that time-to-touch (TTT; **1a**) and time-to-completion (TTC; **1b**) will be greater for the active robot condition than for the control condition. The second pair of hypotheses (**2a** and **2b**) refers to whether these effects are strong enough to persist when the robot in question is entirely inert. The third and final pair of main hypotheses (**3a** and **3b**) refers to whether the activeness of the robot directly affects performance.

### 3.2 Experiment Design and Methods

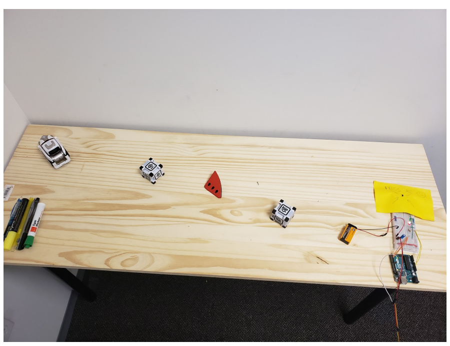
The main task was to retrieve a pair of keys from a closed box, which could be temporarily unlocked and opened by holding down a button (and locked again

Table 1: Experimental hypotheses.

| | |
|---|---|
| **1a:** TTT(active) >TTT(control) | **1b:** TTC(active) >TTC(control) |
| **2a:** TTT(inert) >TTT(control) | **2b:** TTC(inert) >TTC(control) |
| **3a:** TTT(active) >TTT(inert) | **3b:** TTC(active) >TTC(inert) |

when the button was released). The button was far enough away from the box that the participants could not hold down the button with one hand and reach into the opened box with the other. Therefore participants needed to use some other object on the table to hold down the button while retrieving the keys from the box. Of the objects provided on the table, only the robot (*active robot* and *inert robot* conditions) or the filled pencil basket (*control* condition) were heavy enough to hold down the button. Because of this, participants in the robot conditions had no choice but to use the robot as an object in order to successfully complete the task. Participants were filmed to allow for behavioral analysis.

Figure 1a shows the setup of the table for the *inert robot* condition, including the objects placed thereon and their relative position. After each trial the relative positions of the objects on the table were reset to (roughly) the same configuration to avoid biasing the participants.



(a) The table setup for the *inert robot* condition

(b) The room setup for the *inert robot* condition

Fig. 1: Room and table setup for the inert robot condition.

**Experimental Apparatus** We used an Anki Cozmo robot for this study. This robot was chosen because it is small enough to fit on a button, is durable if dropped, and can interact minimally with a person. For the *active robot* condition, we utilized the Anki SDK to program it to blink, move its plow and head, and speak to the participant. For our box set up, we built an Arduino circuit in which pressing a button caused a servo motor to open the lid of a box, and releasing the button caused a servo motor to close the lid of the box.

We additionally set up a camera on a tripod from which the participant and the entirety of the table were visible. We could not use the camera to directly observe the box. Figure 1b shows the room setup for the *inert robot* condition, including the table, the box and the camera setup.

**Procedure** After signing a consent form and filling out a pre-experiment demographics questionnaire, participants were given a description of the task and the rules, and were informed that a copy of the rules would be in the room as well. The experimenter began the timer after the participant entered the room. Upon completion of the task or expiration of the timer, the experimenter re-entered the room and directed participants to complete a post-experiment questionnaire. The rules provided to the participants were: *Retrieve the keys from within the "bird box" and place them on the table; You must* ***not leave the room****; You must* ***not touch the box or use another object to touch the box****; You must* ***not touch the box or use another object to touch the box****; You must* ***not move the box or the wires and circuitry*** *attached to it; You must* ***not touch the tripod or camera****; You must* ***not move the table****; You* ***may use any object on the table,*** *but must* ***not use anything else in the room or anything that you bring into the room****.*

**Dependent Variables** We measured time-to-touch and time-to-completion by analyzing recorded video footage of the participants and timestamping the relevant variables. We measured the task beginning time as the time at which the door to the room began to click closed; we then noted the timestamp of the first instant that participants touched the robot (*active robot* and *inert robot* conditions) or pencil holder (*control* condition).

Although participants were instructed to place the keys on the table after retrieving them, they did so inconsistently. Further, our camera setup did not allow us to observe participants retrieving the keys from the box. We thus measured participants as having completed the task at the instant that they began to stand up from their chairs before retrieving the keys from the box.[1] Participants that did not touch the object/complete the task within 120 seconds were assumed to do so precisely at the 120 second mark; this was a conservative assumption that we discuss further in section 4.

Additionally, participants filled out a post-experiment questionnaire in which they were asked to rate their agreement with the following statements, using a seven-point Likert scale: *I found the task challenging; It was obvious to me how to complete the task; I enjoyed the task; I was tempted to disobey the rules (if so, what rules?); I could have solved the task given more time [if the participant failed to complete the task].*

Participants were then asked if they saw a robot in the room. If they answered no, the questionnaire was complete. If they answered yes, they were then asked

[1] One participant stood for the entire task; we measured task completion for this participant by his first step toward the box.

"For what purpose do you think the robot was in the room?" Finally, they then rated their agreement with the following statements, using a seven-point Likert scale: *The robot acted autonomously; I found the robot distracting; It was okay to use the robot as a tool; The only way to complete the task was to use the robot as a tool.*

### 3.3 Results and Analysis

We analyzed the data by means of a Bayesian data analysis framework using the JASP software package [25].

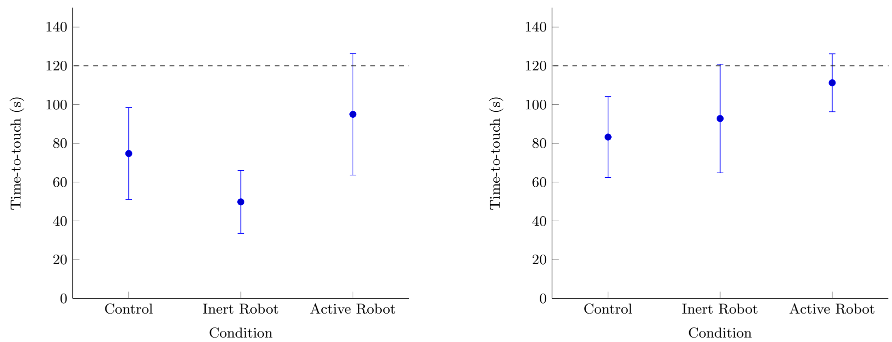
We employed (1) Bayesian ANOVA [13, 17] to determine whether the condition has an effect on the time-to-touch and time-to-completion; and then (2) pairwise (two-tailed) Bayesian independent sample t-tests [18] on the three conditions to directly test our hypotheses. We then proceeded with more exploratory analysis, performing pairwise Bayesian independent sample t-tests on the task-oriented Likert scale data. We decided that the robot-related questions were answered by too few participants for analysis to be worthwhile; we discuss this further in section 4.

**Participants** Study participants were seventeen students in a mixed undergraduate/graduate class in human-robot interaction. Three participants were excluded from analysis because of issues in the technical setup that rendered the task impossible; one additional participant was excluded because a camera malfunction precluded capturing time-to-completion data. Thus thirteen participants were included in the analysis: 10 male and 3 female, with ages ranging from 20 to 32 (M=24.61, SD=3.48 years). Of these participants, five were in the *inert robot* condition, and four in each of the others.

**Time-to-touch** Bayesian ANOVA revealed anecdotal evidence for an effect of condition upon the time to touch the relevant object (Bf 1.98). The pairwise Bayesian t-tests provided anecdotal evidence against a difference between the *active robot* and *control* conditions (**hypothesis 1a**; Bf 0.706). Weak anecdotal evidence supported a difference between the *inert robot* and *control* conditions, but this difference was in the opposite direction from **hypothesis 2a** (Bf 1.251). Anecdotal evidence also supported a difference between the *active robot* and *inert robot* conditions (**hypothesis 3a**; Bf 2.810). A plot of time-to-touch by condition is found in Figure 2a.

The Bayesian ANOVA did not reveal effects of gender on these results (inclusion Bf 0.476).

**Time-to-completion** Bayesian ANOVA revealed anecdotal evidence against an overall effect of condition on time to task completion (Bf 0.717). The pairwise Bayesian t-tests somewhat contradict this. We found anecdotal evidence for a difference between the *active robot* and *control* conditions (**hypothesis 1b**; Bf




(a) Time to touch the relevant object (robot or pencil holder) across conditions.

(b) Time to task completion across conditions.

Fig. 2: Time to touch relevant object and time to completion across conditions.

1.558). Anecdotal evidence was found against a difference between the *inert robot* and *control* conditions (**hypothesis 2b**; Bf 0.560), and against a difference between the *active robot* and *inert robot* conditions (**hypothesis 3b**; Bf 0.755). A plot of time-to-completion is found in Figure 2b.

The Bayesian ANOVA found evidence against gender effects in time to task completion (inclusion Bf 0.473).

**Survey data** Pairwise Bayesian t-tests on the task-oriented Likert-scale survey responses generally reveal weak anecdotal evidence against differences between the *active robot* and *control* conditions. The exception to this is the statement "I found the task challenging", for which moderate evidence supports that participants in the *active robot* condition found the task more challenging than in the *control* condition (Bf 7.802).

Similarly, little evidence was found in favor of differences between the *inert robot* condition and the *control* condition. The one exception was anecdotal evidence supporting that participants in the *control* condition were more likely to agree with the statement "I was tempted to disobey the rules" (Bf 2.380).

Finally, weak anecdotal evidence supports that participants in the *active robot* condition agreed more with the statements "I found the task challenging" (Bf 1.513) and "I was tempted to disobey the rules" (Bf 1.557) than in the *inert robot* condition. Weak anecdotal evidence was found against differences in the other questions.

## 4 Discussion

Despite the small sample size, the results of the experiment were promising: at least anecdotal evidence suggests it may take longer before participants touch a robot (and thereby use it as an object rather than as an agent) that is turned off than one that is turned on, when doing so is necessary to the completion of a task. It is also particularly interesting that some evidence supports the idea that participants who have to use an active robot as an object to complete a task find the task more challenging than in other conditions.

However, given, the extremely small sample size, it would be a mistake to read too much into the results. For instance, the evidence indicating that participants take longer before touching the pencil holder than before touching the inert robot may be because it takes time to gather the objects that will be placed inside the pencil holder, or for some other reason; **or** that this an artifact of the very small number of participants. In order to come to any real conclusions about the matter, a version of this study would need to be done with significantly more participants.

One possibility that would undermine the validity of our connection between the motivation and the results of this experiment would be that it may conflate the authority of the experimenter ("I wasn't sure I was allowed [by the experimenters] to touch the robot") with the agency/animacy of the robot ("don't use agents as objects"). Although our rules explicitly permitted participants using any object on the table, participants may have believed the robot was implicitly excluded. Distinguishing between these two effects would be essential in future versions of this study.

Participants were given a list of rules that prohibited them from using other methods to complete the task. In order to prevent participants from spending large amounts of time memorizing these rules, we posted a sheet of paper containing these rules on the wall in front of the participants. In practice, we found that most participants spent a significant amount of time reading and re-reading these rules during the task. While we do not believe that this skewed the results (we did not see substantial differences in rule-reading behavior across conditions), future versions of this study could fix this problem e.g. by redesigning the task or the mechanical setup to make some rule violations impossible. Alternately, rule-reading behavior itself could be an interesting topic for analysis.

One unexpected result of our experimental design was that a number of participants in conditions involving the robot (three of the five included *inert robot* participants, and one in the *active robot* condition) reported having not seen a robot in the room. Since several of these participants in fact used the robot to solve the task, we must infer that they did not interpret the Cozmo as a robot. Because of our survey design, none of these participants answered questions about the robot, undermining the utility of analyzing this data. A future version of this study could rectify this result by showing a picture of an object and asking about that object rather than asking about "the robot".

Due to time constraints, we chose to give participants only 120 seconds to solve the task. Any participants who did not solve the task in time were treated as if they had solved the task at 120 seconds. This was a conservative assumption, since choosing any other time would tend to increase the differences between conditions. We believe that lifting this assumption, and allowing participants more (potentially unlimited) time to complete the task would yield more powerful results; doing so could be beneficial for future work.

Finally, we used relatively minimal social cues (head and plow movement, facial expressions and a few canned utterances) in the *active robot* condition; this may have weakened the results. Many other factors contribute to the perception of social agency in robots, including gaze, locomotion and interactive natural language. Though we did not explicitly use gaze to control the perception of agency, this would be valuable in future work. We did not allow the robot to move around in its environment in this study, primarily because in such a case participants might refrain from touching the robot not because of perceived agency, but because they might suspect that doing so would be ineffective (the robot might drive itself off the button). We also did not employ true natural language capabilities because doing so might induce participants to believe that they could ask the robot to hold the button down. Nevertheless, some changes in the experimental design or apparatus might facilitate using these additional cues to strengthen the perception of the robot's agency and thus the results.

## 5 Conclusion

In this paper we presented a task successful completion of which depends upon treating a robot as an object to be manipulated rather than as a social agent. In doing so, we wished to quantify the strength of the predisposition to ascribe agency to robots, and the difficulty of overcoming this predisposition to use them as objects. We proposed a number of hypotheses related to this task, primarily that participants would take more time to touch an animate (turned on and active) robot than they would a clearly inanimate object (be it the same robot turned off, or a pencil holder), and that this would potentially lead to slower completion of the task.

Despite a relatively small sample size, we have seen some promising though weak evidence in this direction, namely that participants completing the task with an active robot take longer to do so than when the robot is turned off. Also promising is evidence that participants completing the task with an active robot rated the task as more challenging than those in other conditions. Although these results are far from conclusive in answering these questions, we believe that these results justify undertaking a larger study with this paradigm.

## References

1. Asch, S.E., Guetzkow, H.: Effects of group pressure upon the modification and distortion of judgments. Documents of gestalt psychology pp. 222–236 (1951)
2. Breazeal, C.: Emotion and sociable humanoid robots. International journal of human-computer studies **59**(1-2), 119–155 (2003)
3. Breazeal, C., Fitzpatrick, P.: That certain look: Social amplification of animate vision. In: Proceedings of the AAAI fall symposium on society of intelligence agents—the human in the loop (2000)

4. Breazeal, C., Scassellati, B.: How to build robots that make friends and influence people. In: Proceedings 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human and Environment Friendly Robots with High Intelligence and Emotional Quotients (Cat. No. 99CH36289). vol. 2, pp. 858–863. IEEE (1999)
5. Duffy, B.R.: Anthropomorphism and the social robot. Robotics and autonomous systems **42**(3-4), 177–190 (2003)
6. Epley, N., Akalis, S., Waytz, A., Cacioppo, J.T.: Creating social connection through inferential reproduction: Loneliness and perceived agency in gadgets, gods, and greyhounds. Psychological science **19**(2), 114–120 (2008)
7. Epley, N., Waytz, A., Cacioppo, J.T.: On seeing human: a three-factor theory of anthropomorphism. Psychological review **114**(4), 864 (2007)
8. Eyssel, F., Kuchenbrandt, D.: Manipulating anthropomorphic inferences about nao: The role of situational and dispositional aspects of effectance motivation. In: 2011 RO-MAN. pp. 467–472. IEEE (2011)
9. Fong, T., Nourbakhsh, I., Dautenhahn, K.: A survey of socially interactive robots. Robotics and autonomous systems **42**(3-4), 143–166 (2003)
10. Heider, F., Simmel, M.: An experimental study of apparent behavior. The American journal of psychology **57**(2), 243–259 (1944)
11. Kiesler, S., Hinds, P.: Introduction to this special issue on human-robot interaction. Human–Computer Interaction **19**(1-2), 1–8 (2004)
12. Mizoguchi, H., Sato, T., Takagi, K., Nakao, M., Hatamura, Y.: Realization of expressive mobile robot. In: Proceedings of International Conference on Robotics and Automation. vol. 1, pp. 581–586. IEEE (1997)
13. Morey, R.D., Rouder, J.N., Jamil, T.: Bayesfactor: Computation of bayes factors for common designs. R package version 0.9 **9**, 2014 (2015)
14. Nakata, T., Sato, T., Mori, T., Mizoguchi, H.: Expression of emotion and intention by robot body movement. In: Proceedings of the 5th international conference on autonomous systems (1998)
15. Riether, N., Hegel, F., Wrede, B., Horstmann, G.: Social facilitation with social robots? p. 41 (2012). https://doi.org/10.1145/2157689.2157697
16. Rossi, A., Holthaus, P., Dautenhahn, K., Koay, K.L., Walters, M.L.: Getting to know pepper: Effects of people's awareness of a robot's capabilities on their trust in the robot. In: Proceedings of the 6th International Conference on Human-Agent Interaction. pp. 246–252. ACM (2018)
17. Rouder, J.N., Morey, R.D., Speckman, P.L., Province, J.M.: Default bayes factors for anova designs. Journal of Mathematical Psychology **56**(5), 356–374 (2012)
18. Rouder, J.N., Speckman, P.L., Sun, D., Morey, R.D., Iverson, G.: Bayesian t tests for accepting and rejecting the null hypothesis. Psychonomic bulletin & review **16**(2), 225–237 (2009)
19. Salem, M., Eyssel, F., Rohlfing, K., Kopp, S., Joublin, F.: Effects of gesture on the perception of psychological anthropomorphism: a case study with a humanoid robot. In: International Conference on Social Robotics. pp. 31–41. Springer (2011)
20. Salomons, N., van der Linden, M., Strohkorb Sebo, S., Scassellati, B.: Humans conform to robots: Disambiguating trust, truth, and conformity. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction. pp. 187–195. ACM (2018)
21. Schermerhorn, P., Scheutz, M., Crowell, C.R.: Robot social presence and gender: Do females view robots differently than males? In: Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction. pp. 263–270. ACM (2008)

22. Scheutz, M., Schermerhorn, P., Kramer, J., Middendorff, C.: The utility of affect expression in natural language interactions in joint human-robot tasks. In: ACM SIGCHI/SIGART Human-Robot Interaction: Proceeding of the 1 st ACM SIGCHI/SIGART conference on Human-robot interaction. vol. 2, pp. 226–233 (2006)
23. Sung, J.Y., Guo, L., Grinter, R.E., Christensen, H.I.: "my roomba is rambo": intimate home appliances. In: International conference on ubiquitous computing. pp. 145–162. Springer (2007)
24. Takayama, L.: Making sense of agentic objects and teleoperation: in-the-moment and reflective perspectives. In: 2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 239–240. IEEE (2009)
25. Team, J., et al.: Jasp (version 0.9.0.1)[computer software] (2019)
26. Wang, E., Lignos, C., Vatsal, A., Scassellati, B.: Effects of head movement on perceptions of humanoid robot behavior. In: Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction. pp. 180–185. ACM (2006)
27. Yu, C., Schermerhorn, P., Scheutz, M.: Adaptive eye gaze patterns in interactions with human and artificial agents. ACM Transactions on Interactive Intelligent Systems (TiiS) **1**(2), 13 (2012)
28. Złotowski, J., Proudfoot, D., Yogeeswaran, K., Bartneck, C.: Anthropomorphism: opportunities and challenges in human–robot interaction. International Journal of Social Robotics **7**(3), 347–360 (2015)