The Perception of a Robot Partner's Effort Elicits a Sense of Commitment to Human-Robot Interaction

Forthcoming in *Interaction Studies*

Marcell Székely

Affiliation: Central European University, Department of Cognitive Science

Address: Oktober 6 Utca 7, Budapest 1051, Hungary

Email: szekelymarcell@gmail.com

Henry Powell

Affiliation: University of Warwick

Address: Social Sciences Building, Coventry CV4 7AL, UK

Email: henry.powell87@gmail.com

Fabio Vannucci

Affiliation: Università di Genova and Istituto Italiano di Tecnologia

Address: Via Enrico Melen, 83, 16152 Genova, Italy

Email: Fabio.Vannucci@iit.it

Francesco Rea

Affiliation: Istituto Italiano di Tecnologia

Address: Via Enrico Melen, 83, 16152 Genova, Italy

Email: Francesco.Rea@iit.it

Alessandra Sciutti

Affiliation: Istituto Italiano di Tecnologia

Address: Via Enrico Melen, 83, 16152 Genova, Italy

Email: Alessandra.Sciutti@iit.it

John Michael

Affiliation: University of Warwick

Address: Department of Philosophy, Social Sciences Building, Coventry CV4 7AL, UK

Email: j.michael.2@warwick.ac.uk

Abstract: Previous research has shown that the perception that one's partner is investing effort in a joint action can generate a sense of commitment, leading participants to persist longer despite increasing boredom. The current research extends this finding to human-robot interaction. We implemented a 2-player version of the classic snake game which became increasingly boring over the course of each round, and operationalized commitment in terms of how long participants persisted before pressing a 'finish' button to conclude each round. Participants were informed that they would be linked via internet with their partner, a humanoid robot. Our results reveal that participants persisted longer when they perceived what they believed to be cues of their robot partner's effortful contribution to the joint action. This provides evidence that the perception of a robot partner's effort can elicit a sense of commitment to human-robot interaction.

Keywords: commitment, effort, human-robot interaction, joint action

1. Introduction

There is a vast potential for robots to assist humans in joint actions in many different domains, from disaster relief to health care, education, and manufacturing (Breazeal et al., 2004; Lenz et al., 2008; Clodic et al., 2009; Sciutti et al., 2012; Grigore et al., 2013).[1] In optimizing human-robot interactions in order to tap this potential, one challenge is to minimize the risk of human interactants becoming frustrated or impatient when the joint action is not going well, for example because their robot partner makes mistakes or is slow in making its contribution. This may be a particularly important challenge in the case of humanoid robots since, on the one hand, their human-like appearance tends to raise people's expectations about their abilities, their reliability, and their user-friendliness (Ferrari & Eyssel, 2016), and, on the other hand, they will increasingly be expected to perform a wide range of tasks flexibly and autonomously, and may therefore be slow or error-prone in some situations. Indeed, this latter issue may be all the more acute for robots designed to engage in autonomous trial-and-error learning (Cully et al., 2015).

Previous research has shown that humans' patience towards a robot that performs suboptimally can be increased if the robot employs a mitigation strategy such as seeking human assistance and/or adapting its approach (Lee et al., 2010; Brooks et al., 2016; Mirnig et al., 2017), or expresses a negative emotional reaction and attempts to rectify its mistake (Hamacher et al., 2016). The present study aimed to build upon this research by addressing the more general question of how to sustain human interactants' willingness to persist in interacting with a robot partner despite increasing boredom or frustration -- irrespective of whether that boredom or frustration arises from errors on the part of the robot or from the nature of the interaction itself. Specifically, the study was designed to test the hypothesis that

[1] This potential is reflected in the ambitious aims specified by SPARC (the Partnership for Robotics in Europe) in the "Strategic Research Agenda for Robotics in Europe 2014-2020," 2014. [Online]. Available: http://www.eurobotics-project.eu.

human interactants' willingness to persist may be boosted by exposing the human to cues that the robot has invested effort in the joint action. If this is the case, it may provide roboticist with a useful -- and low-cost -- tool in designing robots that look and /or behave in such a way as to elicit patience and persistence on the part of human interactants.

Intriguingly, recent research on joint action in humans (i.e. human-human interaction) provides strong reasons to suspect that this may be the case. This research takes its starting point in Michael, Sebanz & Knoblich's (2015) hypothesis that the willingness to remain engaged in joint actions and to resist tempting alternative options and distractions is governed by a *sense of commitment*. The concept of a sense of commitment is closely related to the concept of trust, insofar as both refer to psychological mechanisms that help to sustain agents' motivation to contribute to joint action. Trust, however, helps to sustain agents' motivation to contribute to joint action only *indirectly* -- by stabilizing one agent's expectation that her partner will continue contributing to the joint action in a cooperative manner, it at least reduces a source of uncertainty which could undermine the first agent's motivation to contribute. But it does not *directly* explain why that first agent would then herself contribute to the joint action in a cooperative manner. Indeed, she might be distracted or tempted to disengage irrespective of her level of trust in her partner. Research on the sense of commitment aims to fill this gap: the sense of commitment is hypothesized as a mechanism which stabilizes agents' motivation to contribute to joint actions (and more generally to others' goals) and to persist in the face of tempting alternative options and distractions. More specifically, it is hypothesized as a mechanism which stabilizes this motivation in response to cues that one's partner values the joint action and may be relying on one to make one's contribution.

In support of this hypothesis, Michael, Sebanz & Knoblich (2016) reported evidence that a high degree of spatiotemporal coordination within joint action may function as such a cue, leading agents to remain engaged in the joint action for a longer time and making them more likely to persist until the goal is achieved, while Chennells et al. (under review) demonstrated that repeatedly coordinating with the same partner in a decision-making context is sufficient to elicit a sense of commitment, leading agents to resist tempting alternatives and thereby sustaining cooperation through fluctuations in individuals’ interests. In a similar vein, Székely & Michael (2018) probed the hypothesis that one's perception of a partner's investment of effort in a joint action may provide such a cue -- i.e. if one perceives one's partner to be investing a high degree of effort, or to have invested a high degree of effort, this may lead one to feel committed to making one's contribution to the joint action. To illustrate, Székely & Michael sketch the following example: 'Imagine that you have agreed to attend a cocktail party at your colleague’s apartment but, on the occasion, find yourself tired or otherwise tempted to leave after only a short time. If your colleague has obviously invested a great deal of effort in preparing the hors d’oeuvres and decorations, you might find that a sense of commitment leads you to stick around for a few hours after all.' (2018, p. 38). In support of this, their results showed that participants persisted longer at a boring joint action when they perceived cues of a (human) partner’s effortful contribution.

In the present study, we adapted Székely & Michael (2018)'s paradigm to the context of human-robot interaction in order to investigate whether participants’ persistence would be similarly reinforced by the perception of cues that a humanoid robot partner was making an effortful contribution (Michael & Salice, 2017). In particular, we chose to focus on the perception of cognitive effort rather than physical effort (as in Székely & Michael, 2018), since we surmised that many people's a priori conception of robots may lead them to be disinclined to perceive the actions of a robot as physically effortful (but see Feltz et al., 2014).

The paradigm implements a two-player version of the classic ‘snake game’, in which the participant controls the left-right axis while their partner controls the up-down axis. In our study, as in Székely & Michael (2018), participants were in fact paired with a virtual partner (i.e. an algorithm controls the up-down axis), but were led to believe that their partner was the iCub, a humanoid robot. To bolster this belief, we first exposed participants to a pair of videos of the iCub practicing the tasks that would be performed during the experiment, and then informed them that their game controls during the experiment would be linked with those of the iCub via internet. In fact, however, participants were paired with a virtual partner. To manipulate the perception of the partner's effort, we told participants that, before each round of the snake game, their partner would have to perform a cognitive task in order to ‘unlock’ the round. The cognitive task consisted in deciphering a captcha, which was either difficult (High Effort condition) or easy (Low Effort condition). In fact, however, there were no captchas to be solved, and the visual display indicating the partner's progress in solving each captcha was pre-programmed. Next, the participant and the partner had the shared goal of retrieving as many apples as possible by jointly controlling the snake. Since the apples appeared at an ever-slowing rate, each round became increasingly boring, generating an incentive to disengage. To reinforce this, we determined that participants would not receive points or any other extrinsic rewards for the collection of apples. Participants had been instructed to press a ‘finish’ button whenever they determined that it was time to move on to the next round. This enabled us to operationalize commitment in terms of how long participants persisted in each round.

Our minimal interaction setup had two key virtues. The first key virtue was that it enabled us to maintain a high degree of experimental control. By restricting the possibilities for interaction and communication with the robot, we were able to ensure that participants' experiences were as similar as possible, and thereby to focus narrowly on the effect of our manipulation. Secondly, it enabled us to compare our results with those of a control experiment (Experiment 2) in Székely & Michael's (2018) study, in which participants were correctly informed that their partner was a disembodied virtual agent (i.e. 'an algorithm'). The current experiment differed from that experiment only with respect to participants' beliefs as to the nature of their partner. This comparison makes it possible to isolate the impact of the participants' *belief* that the partner who was investing effort in the joint action was a humanoid robot. And indeed, previous research provides good reason to expect that this difference in beliefs may lead to very different behavior. For example, Stenzel and colleagues (2012) found that participants who were informed that a robot partner had been designed to function in a human-like manner spontaneously 'co-represented' the robot partner's task -- i.e., they exhibited the 'Social Simon effect' (Sebanz, Knoblich & Prinz, 2003) -- whereas participants who were informed that the robot functioned in a deterministic manner did not. This finding supports the conjecture that the belief that one's interaction partner is a human-like robot partner can increase the extent to which one interacts with the partner as though it were human. We therefore hypothesized that if participants believed that their partner were an humanoid robot such as the iCub, the perception of the partner's apparently effortful contribution could elicit a sense of commitment to the interaction. This hypothesis is also supported by research showing that participants were more motivated while performing an exercising task with an embodied robot partner than with a virtual partner (Fasola & Matarić, 2013), and by research indicating that participants were more motivated when exercising jointly with a robot partner than when exercising individually according to the instructions of a robot teacher (Schneider & Kümmert, 2016).

Thus, while the results of Székely & Michael's (2018) control experiment indicated that participants' persistence was unaffected by the perceived effort of a disembodied virtual agent, we predicted that they would persist longer in the High Effort condition than in the Low Effort condition in the current experiment. Indeed, we predicted that our results would closely resemble those of Experiment 1 in Székely & Michael's (2018) study, in which participants were led to believe that they were interacting with a person (although they were in fact paired with the same disembodied virtual agent used in Experiment 2 of that study and in the current study), and exhibited greater commitment as a function of their partner's perceived effort (i.e. longer persistence in the High Effort condition than in the Low Effort condition).

2. Method

2.1 Participants

Using G*Power 3.1 (Faul et al., 2009) we determined that a sample size of 26 would provide 80% statistical power for detecting a medium-sized effect equivalent to what we observed in a pilot study (*d*= .58), assuming a two-tailed t-test and an alpha level of .05. We therefore recruited twenty-six students (17 females; age range: 18-28, M = 22.05, SD = 2.58), using the participant database at the University of Warwick (UK), where the experiment was conducted. Our stopping rule was therefore as follows: we continued recruitment until twenty-six participants had completed the number of trials which we determined a priori to mark the minimum threshold (as explained below).

Six additional participants were excluded prior to analysis because they did not finish the minimum number of trials, as explained below. Thus, 32 participants in total were tested. All participants were naïve to the purpose of the study, reported normal or corrected to normal vision, and signed informed consent prior to the experiment. The experiment was conducted in accordance with the Declaration of Helsinki and was approved by the Humanities & Social Sciences Research Ethics Sub-committee (HSSREC) at the University of Warwick. Each participant received £6 for participating.

2.2 Material

The experiment was displayed on a 13-inch computer screen (resolution: 2560 x 1600 pixels, refresh rate: 60 Hz). The program for the experiment was written in Python (Peirce, 2007), with a framerate of 17 frames per second. The easy captchas (Low Effort condition) consisted of 3 characters; the animation of the partner 'deciphering' them was programmed to take 4-8 seconds. The difficult captchas (High Effort condition) consisted of 12 characters; the animation of the partner 'deciphering' them was programed to take 16-20 seconds. The videos which participants viewed of captchas being deciphered can be found in the Supplementary Material (See S2). The captcha before the practice round was of intermediate length (8 characters), taking 12 seconds to decipher. The examples of easy and difficult captchas that were presented in the instruction phase are depicted in Figure 1.

Fig. 1. Sample Captchas. In the instruction phase, participants were presented with examples of easy and difficult captchas.

The algorithm for the partner, which controlled the up-down axis, was programmed to behave in a human-like manner: it follows the shortest path to the apple, but sometimes (randomly) makes mistakes, reacting too late or turning in the wrong direction.

2.3 Procedure

Participants signed up for slots of one hour. On arrival at the lab, they were first informed that they would be playing 20 rounds of the snake game together with a partner, with the participant controlling the left-right axis, and the partner controlling the up-down axis, and that their joint task would be to collect as many apples as possible in each round by jointly maneuvering the snake. At the end of each round of the snake game, they were given feedback about how many apples they had collected in that round. They were not provided with a running total of the number of apples collected overall. Participants did not receive points or any other rewards for the collection of apples. This was because we did not want to provide external incentives for the collection of apples; instead, our focus was on the motivation arising out of a sense of commitment to the partner.

In addition, they were informed that they and their partner had each been assigned an additional task. Their partner would have the additional task of solving a captcha before each round in order to unlock the round. The captchas would sometimes be easy (Low Effort condition) and sometimes difficult (High Effort condition), as depicted in Figure 1. The participant would have the task of determining when it was time to conclude each round of the snake game, and move on to the next round, by pressing the spacebar.

They were then told that they would be linked with their partner via the internet, and that their partner was the iCub, a humanoid robot located at our partner lab (the Cognitive Robotics and Interaction Lab, based at the Italian Institute of Technology) in Genoa, Italy. They were informed that the iCub had been practicing solving captchas and playing the snake game earlier in the day. Next, they viewed a video (See Figure 2 and S1 in the Supplementary Material) of the iCub robot practicing the snake game, followed by a video of the iCub practicing captchas.

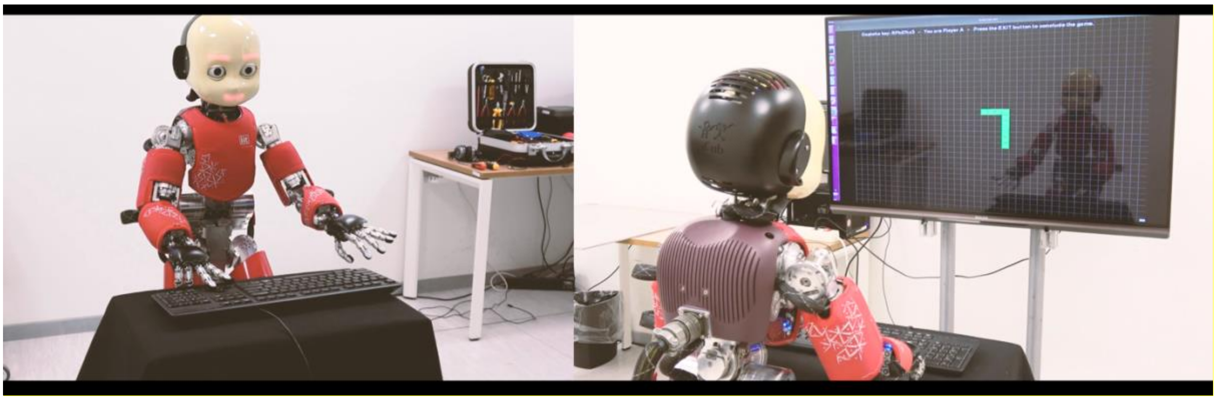

Fig. 2: Frames from the videos of the iCub. In the instruction phase, participants viewed videos of their robot partner practicing the snake game and practicing deciphering captchas. For the video of the snake game, the robot was programmed to move the index and middle fingers of its right hand to alternately press the left and right arrows while looking at the screen, and thereby to navigate the snake. For the video of the captcha task, the robot was programmed to reach the keyboard with its two hands, as if going to press the different keys to solve the captcha. The sounds of keyboard clicks were added to the video a posteriori.

In a within-subject design, the experiment consisted of 20 trials in total, 10 in the High Effort condition and 10 in the Low Effort condition. Each trial consisted of a captcha phase, followed by a round of the snake game. In the captcha phase, a video was presented in which stars progressively appeared to indicate that the partner was solving a captcha, and finally the completed captcha key was displayed (See Figure 3 and S2 in the Supplementary Material). The trials were presented in pseudorandomized order.

If a participant had not completed the full experiment (i.e. 20 trials) after a maximum of one hour, we interrupted them at the end of whatever trial they were currently performing, and then debriefed them and sent them home. This was because there was no upper limit on how long participants could persist (persistence being the dependent measure), and previous research (Székely & Michael, 2018) led us to anticipate that a small number of participants would tend to persist for so long that they would only complete a few trials, making it unfeasible to collect sufficient data from them to conduct meaningful analyses. As a consequence of this procedure, 6 participants were interrupted before completing a minimum of 8 trials in each condition (16 in total); the data from these participants was excluded prior to analysis.

To make the joint action increasingly boring, apples were programmed to appear at an ever-slowing rate within each round. In the first 10 seconds, each new apple appeared immediately. After 10 seconds, new apples appeared with a delay of 40 frames; this delay was doubled every ten seconds. Participants were instructed to press the ‘finish’ button when they determined that it was time to end each round and move on to the next round.

The experiment was preceded by one practice trial. The captcha for the practice trial was of intermediate length between the captcha for the High and Low Effort conditions (8 characters), and took 12 seconds to decipher.

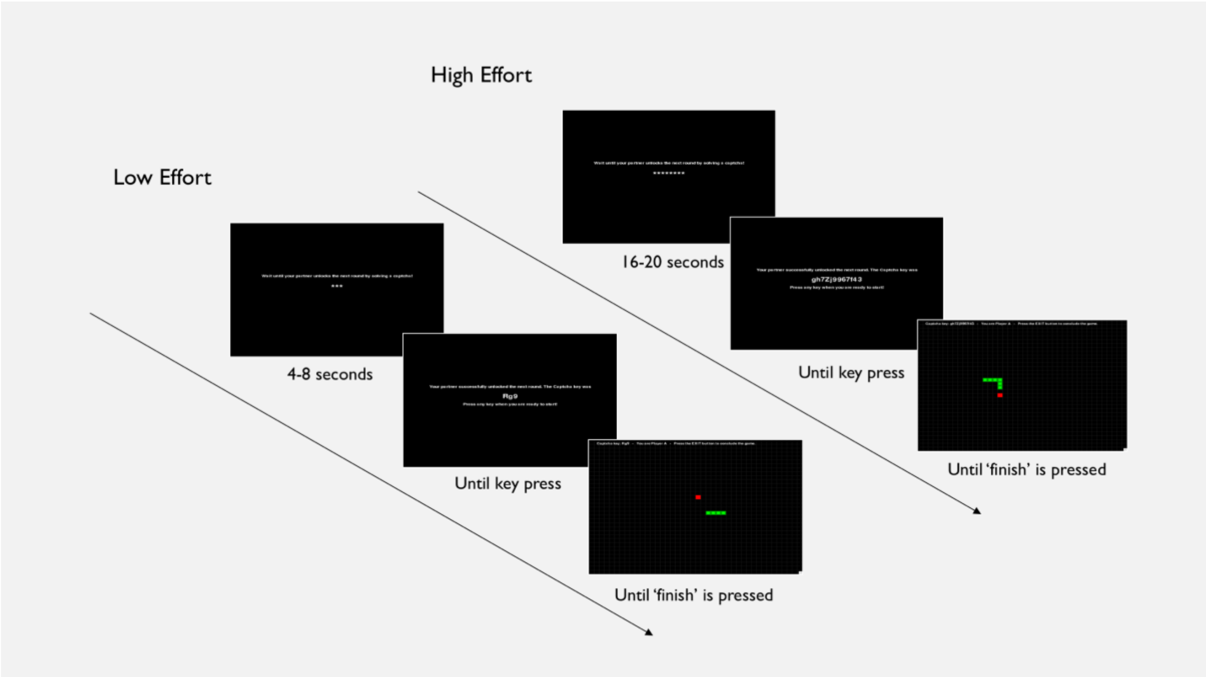

Fig. 3: Trial Structure. Each trial consisted of a captcha phase, followed by a round of the snake game. In the captcha phase, a video was presented in which stars progressively appeared to indicate that the partner was solving a captcha. The text on the screen read: 'Please wait until your partner unlocks the next round by solving a captcha!' Next, the completed captcha key was displayed (See S2 in the Supplementary Material). The text on the screen read: 'Your partner successfully unlocked the next round. The Catptcha key was: [Captcha key here]. Press any key when you are ready to start!' Finally, the participant played a round of the snake game. During this round, the text on the top of the screen read: 'Captcha key: [Captcha key here] – Your Player A – Press the EXIT button to conclude the game.' Each round continued until the participant pressed the spacebar to 'conclude' the round.

3. Results

3.1 Main Analysis

For the analysis, we excluded the data from 6 participants who did not complete at least 16 trials (8 in each condition) within the scheduled time slot of one hour. This left us with a sample size of 26 participants, as we had planned. We also excluded the data from trials on which participants collected 0 apples (6 trials; 2.3% of the data). There was a high degree of variability in persistence times across trials, with participants persisting as long as 25 minutes on some trials. Since we had no a priori basis for setting any particular upper bound, we did not exclude any of these longer trials. Instead, we elected to use individual participants' median persistence times in seconds as the basis for our analyses. In order to test the data for normality and homogeneity of variance we conducted a Shapiro-Wilk test, which revealed a significant deviation from normality, p = .003. We therefore performed a log10 transformation on the data to meet the assumption of normality. We then conducted a paired-samples t-test, which revealed significant difference between conditions, with participants persisting longer in the High Effort condition (logtransformed M = 1.85, SD = 0.28) than in the Low Effort condition (logtransformed M = 1.80, SD = 0.26), $t(25)$ = 2.76, p = .011, d = 0.54 (See **Fig. 4**, **Fig. 5,** and **Table 1**).

We also conducted a Wilcoxon Signed-Ranks test on the non-normalized data, which indicated that participants persisting longer in the High Effort condition (M = 83.5, SD = 43.83) than in the Low Effort condition (M = 73.97, SD = 36.96;), Z = 263, p=.025, r = 0.50.

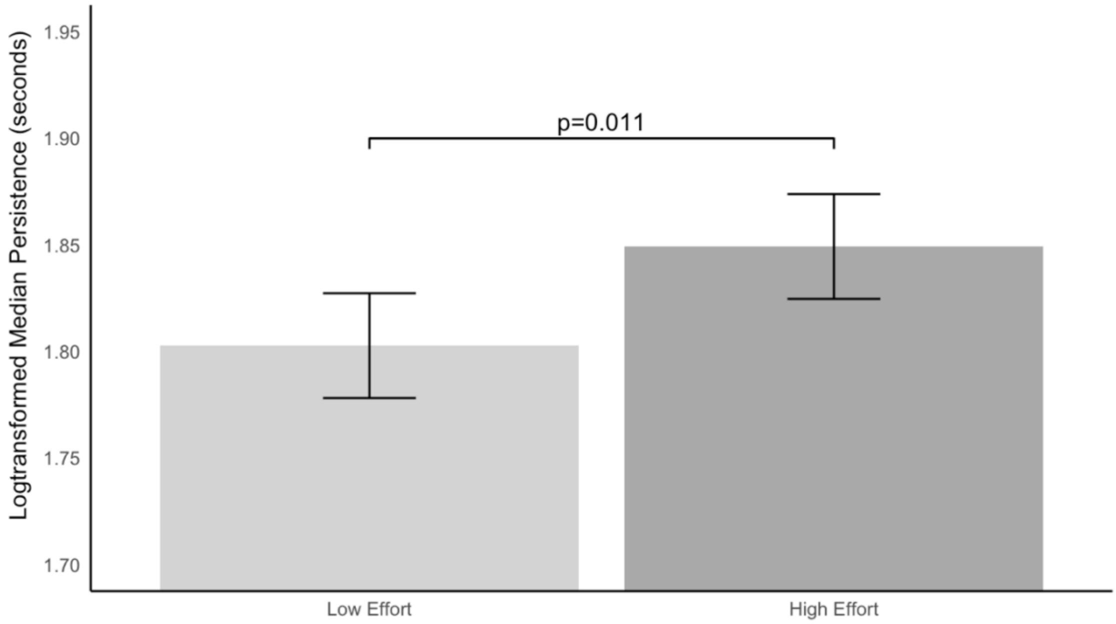

Figure 4: Results. Persistence for High and Low Effort conditions. Error bars represent the within-subject confidence intervals (following the method proposed by Cousineau, 2005; cf. Loftus & Masson, 1994)

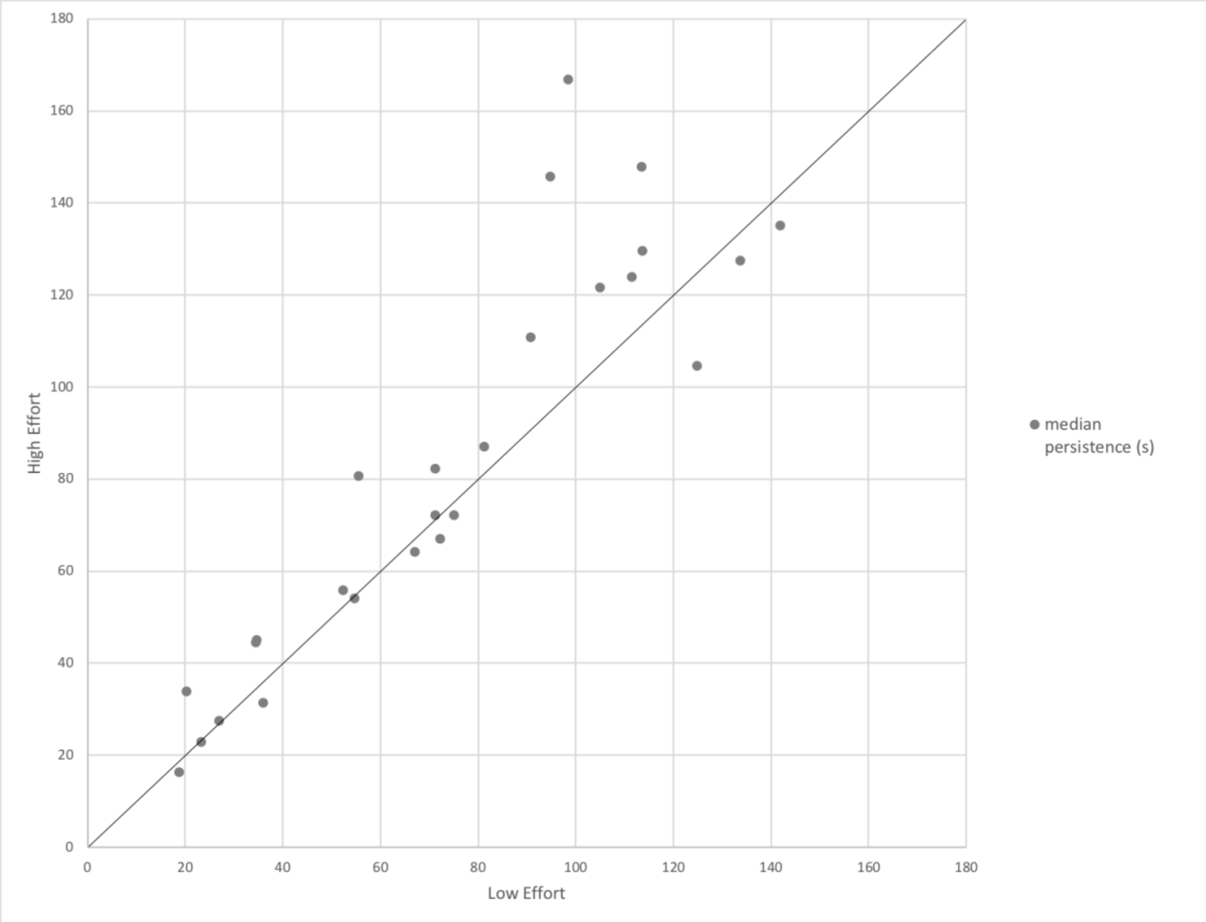

Figure 5: Individual Data. Each dark circle represents one participant's median persistence in seconds for each of the two conditions: the median persistence (in seconds) for the High Effort condition lies on the Y axis, while the corresponding median persistence for the same participant in the Low Effort condition (in seconds) lies on the X axis. The identity line indicates where each participant's dot would lie if her or his median persistence did not differ between conditions.

Table 1: Descriptive statistics for the High and Low conditions in the First Half and Second Half of the experiment. All values are given in seconds.

	Low Effort First Half	High Effort First Half	Low Effort Second Half	High Effort Second Half
mean	88.14	97.85	66.69	75.70
sd	54.41	54.37	39.22	46.05
min	22.24	26.55	12.23	13.44
max	221.97	220.77	137.25	169.78

3.2 Secondary Analyses

3.2.1 ANOVA

We were also interested in whether participants' persistence varied over the course of the experiment, and whether the effect of the manipulation was constant over the course of the experiment. To probe these two issues, we performed an ANOVA with Effort (High/Low) and Time (First Half of the experiment/Second Half of the experiment) as separate factors. The results showed a significant main effect of Time, with participants persisting significantly longer in the first half of the experiment (logtransformed M=1.89, SD= 0.26) than in the second half (logtransformed M=1.76, SD= 0.32), F (1,25) = 6.42, p=0.018, $\eta_p 2$ = 0.2. There was also a significant main effect of Effort, with participants persisting significantly longer in the High Effort condition (logtransformed M= 1.85, SD= 0.28) than in the Low Effort condition (logtransformed M=1.80, SD=0.26), F (1,25) = 7.5, p=0.013, $\eta_p 2$ = 0.22. There was no statistically significant interaction between Time and Effort, F (1,25) = .024, p=0.879, $\eta_p 2$ = 0.001, (See **Fig. 6**, **Fig. 7,** and **Table 1**).

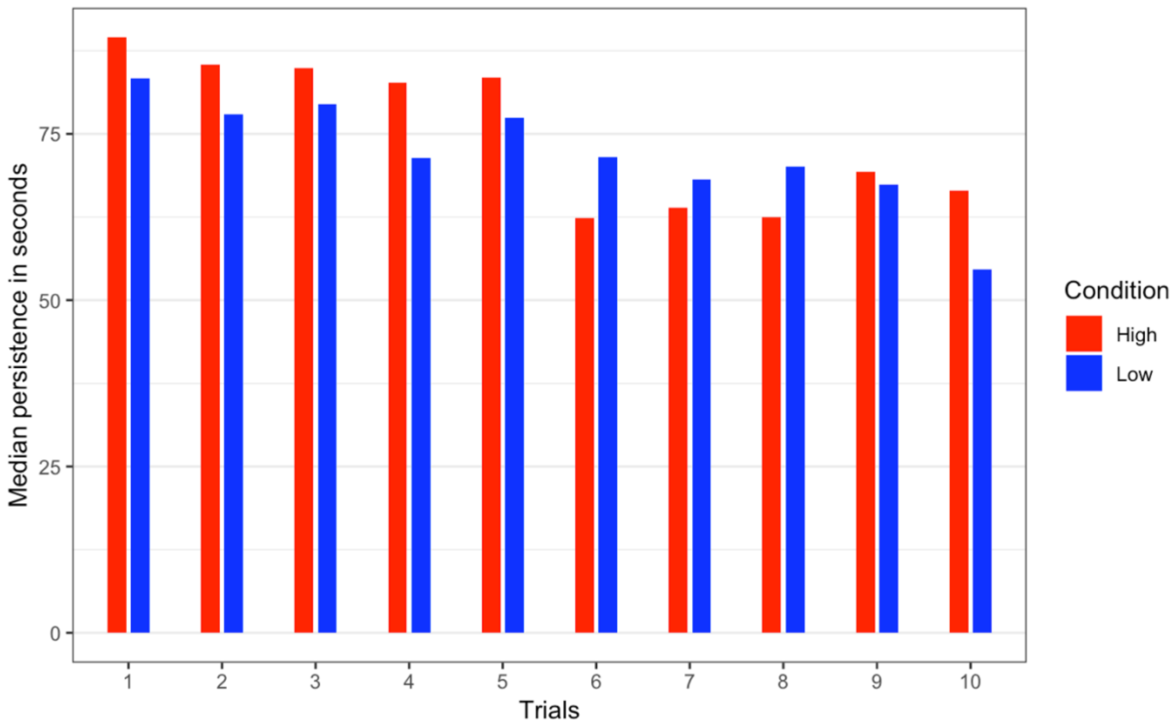

Figure 6: Descriptive data. The X axis represents the time course of the experiment, i.e. the ten rounds of each condition. Each red bar represents the median persistence in seconds for one trial in the High Effort condition; each blue bar represents the median persistence in seconds for one trial in the Low Effort condition. For ease of comparison, the first High Effort trial appears adjacent to the first Low Effort trial, the second High Effort trial adjacent to the second Low Effort trial, and so on.

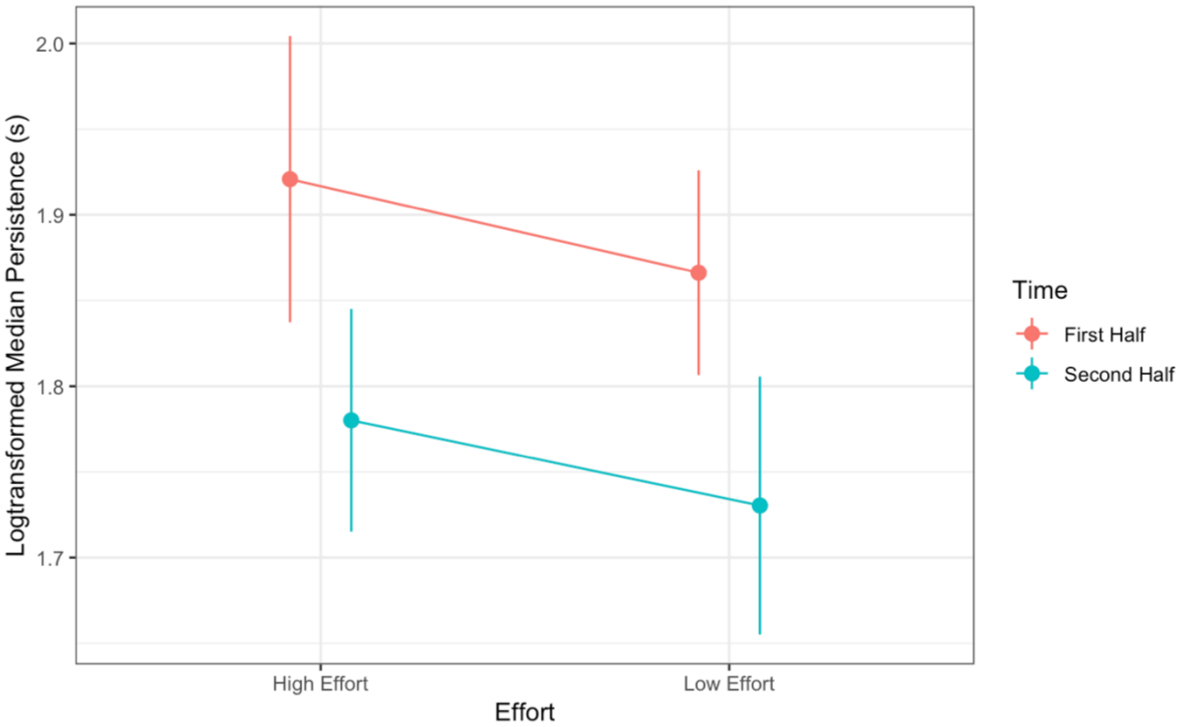

Figure 7: Results. Persistence for High and Low Effort conditions for the First Half and the Second Half of the experiment. Error bars represent the within-subject confidence intervals (following the method proposed by Cousineau, 2005; cf. Loftus & Masson, 1994).

3.2.2 Linear Regression

Some participants persisted longer in general than other participants. One possible reason for this is that some participants felt more committed to the experimenter than others, and therefore persisted longer. If so, then we may expect the same participants who persisted longer in general to also exhibit a larger difference between the High and Low Effort conditions than participants who persisted less in general, i.e. because they have a more acute sense of commitment. To explore this question, we first confirmed that the assumptions for regression analysis were not violated, and then conducted a linear regression to predict the difference between the High and Low Effort conditions based on the sum of the medians in the High and Low effort conditions. The sum of the medians in the High and Low Effort conditions did not significantly predict differences between the High and Low Effort conditions, $b = 0.162$, $t(24) = 0.805$, $p = 0.428$, (See **Fig. 8**).

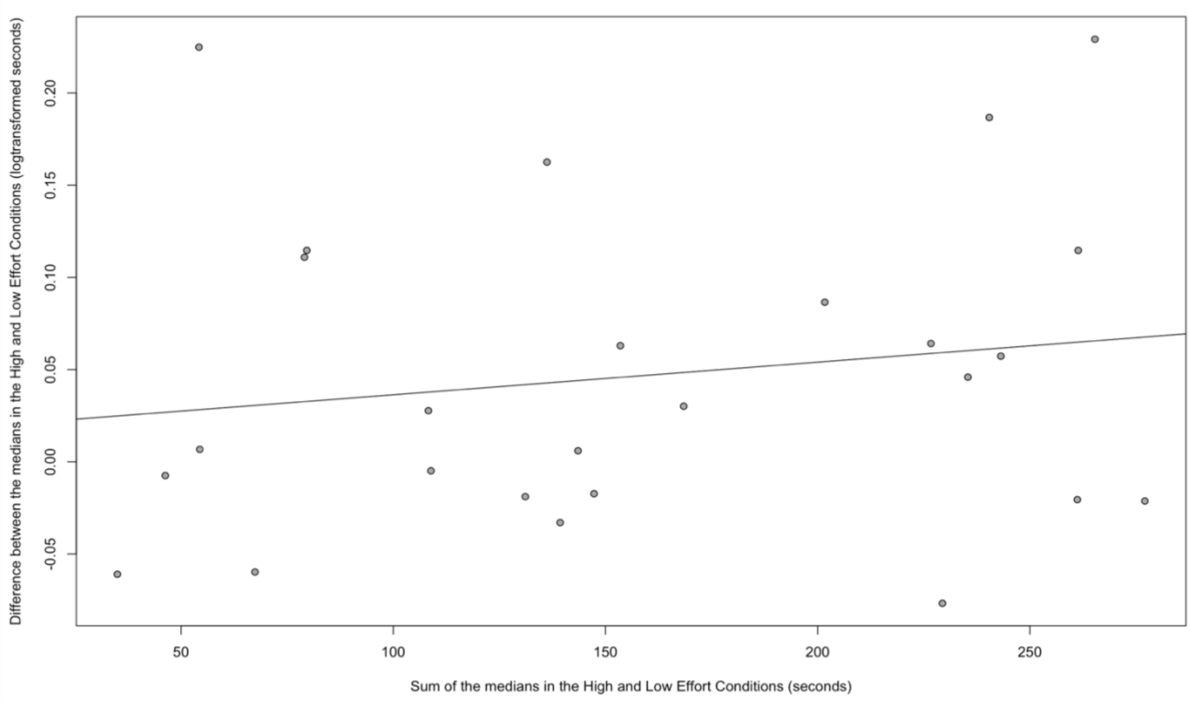

Figure 8: Results. The effect of general persistence, expressed as the effect of the sum of the medians in the High and Low Effort conditions (x axis) upon the difference between the High and Low Effort conditions (y axis).

4. Discussion

We implemented a 2-player version of the classic snake game which became increasingly boring over the course of each round. Before each round of the game, participants perceived what they believed to be cues that their partner, a humanoid robot, was deciphering a captcha to unlock the round. In the High Effort condition, those cues indicated a high degree of effort; in the Low Effort condition, they indicated a low degree of effort. To operationalize participants' commitment, we measured how long they persisted in each round of the snake game despite increasing boredom. In line with our prediction, the results revealed that participants persisted significantly longer in the High Effort condition than in the Low Effort condition. This supports the hypothesis that the perception of a humanoid robot's effortful contribution to a joint action is sufficient to elicit a sense of commitment, leading to increased persistence in the face of increasing boredom and/or frustration.

We also found that participants' persistence decreased over the course of the experiment, with median persistence per trial bring shorter in the second half of the experiment than in the first half. This finding confirms that participants did indeed find the task increasingly boring and/or frustrating. With this in mind, it bears emphasizing that the ANOVA revealed no interaction between Time (First Half / Second Half) and Effort (High / Low), indicating that the effect of the manipulation did not decrease over the course of the experiment. In other words, participants continued to persist longer in the High Effort condition than in the Low Effort condition despite increasing boredom and/or frustration.

To explore the question whether the effect of our manipulation may have been driven by a subset of participants who were particularly committed in general (i.e. to the task, the experimenter and/or the partner), we also we conducted a linear regression to discern whether persistence in general (i.e. the sum of the medians in the High and Low Effort conditions) predicted the size of the difference in persistence between the High and Low Effort conditions. We did observe a quantitative difference that was consistent with this possibility (i.e. the participants who persisted longer also exhibited larger differences between conditions, as illustrated in **Figure 8**), but since the difference was not statistically significant, we cannot draw any conclusions about any differences there may have been in how the manipulation affected different participants.

Our findings build upon and extend previous research showing that humans evaluate humanoid robots with respect to trustworthiness (DeSteno et al., 2012; Lee et al., 2013). While trust is an important stabilizing force in joint action, it does not directly explain why one agent would persist in contributing to a joint action despite increasing boredom or frustration. A sense of commitment, in contrast, can play this functional role: a sense of commitment can boost an agent's motivation to contribute to joint action as a function of cues indicating that her partner values the joint action and may be relying on her contribution (Michael, Sebanz & Knoblich, 2015). If a human interactant's sense of commitment can be triggered in interaction with a robot, as Michael & Salice (2017) have hypothesized, this could provide roboticists with effective, low-cost tools for designing robots that elicit patience and persistence on the part of human interactants. The findings reported here provide the first direct evidence that this is the case.

The suggestion that a human interactant's sense of commitment can be elicited by a robot partner also provides a new perspective on previous research. For example, Clodic and colleagues (2006) designed an interactive robot guide, ‘Rackham’, who made explicit agreements with museum visitors to guide them through an exhibition, and who was designed to maintain the joint commitment by monitoring whether his ‘clients’ were following him, waiting for them and adjusting his pace to theirs. Clodic and colleagues' work may be interpreted as providing a starting point for roboticists aiming to design robots that can elicit human interactants' commitment to interaction. More recently, Kahn et al. (2012) observed how children reacted when an adult came to collect a robot with whom they had been playing, and announced the intention to put it away in the closet despite the robot's pleas to continue playing. In many cases, the children defended the robot and exhibited moral outrage as the adult ignored it's pleas. This can be interpreted as evidence that by interacting with the robot, the children developed a sense of commitment to the robot, which led them to feel concern for its well-being and to judge that it was entitled to be treated fairly.

The current research also complements previous results showing that humans’ patience toward robots that perform suboptimally can be increased if the robot appears to invest effort in adapting its strategy to mitigate the consequences of its mistake (Lee et al., 2010; Brooks et al., 2016; Mirnig et al., 2017), or expresses a negative emotional reaction and attempts to rectify its mistake (Hamacher et al., 2016). Our study complements these previous results by addressing the more general question of how to sustain human interactants' willingness to persist in interacting with a robot partner despite increasing boredom or frustration -- irrespective of whether that boredom or frustration arises from errors on the part of the robot or from the nature of the task.

One important question for future research will be to investigate to what extent it may also be possible for *non-humanoid* robots to elicit a sense of commitment on the part of a human. In this context, it is worth noting that Székely & Michael (2018), using the same paradigm as was employed in the present study, observed no significant effect of the partner's apparent effort upon participants' persistence in a control experiment (Experiment 2 of their study) in which the partner was described as a disembodied virtual agent (i.e., an 'algorithm'). Thus, the differences which we observed between the High Effort and Low Effort conditions in the present study were comparable to those which Székely & Michael (2018) observed between the High and Low Conditions of their Experiment 1 (in which participants believed they were interacting with a human) rather than to those which they observed in Experiment 3 of their study (in which participants believed they were interacting with a disembodied algorithm (see **Figure 9**).

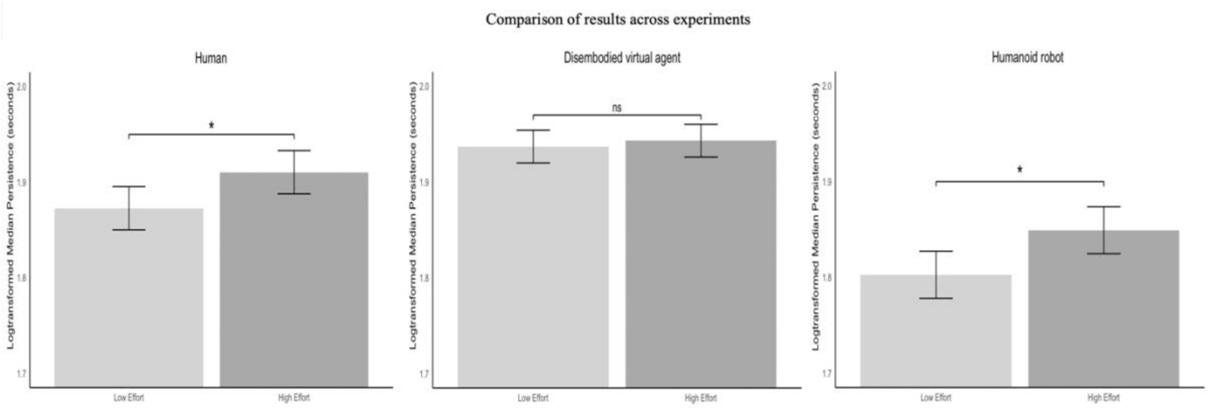

Figure 9: Comparison of results across experiments in which participants' beliefs about their partner were varied. Persistence for High and Low Effort conditions three experiments. In Székely & Michael, 2018, Experiment 1, participants believed they were interacting with a human partner. In Székely & Michael, 2018, Experiment 2, participants believed they were

interacting with a disembodied virtual partner. In the current experiment, they believed they were interacting with a humanoid robot partner. Error bars represent the within-subject confidence intervals (following the method proposed by Cousineau, 2005; cf. Loftus & Masson, 1994).

Of course, there are many differences between a disembodied virtual agent and the robot used in the present study, which not only had a body but was indeed highly human-like in appearance and in movement. It will therefore be important to probe the relative importance of such physical features as a body, a face, and a human-like appearance, as well as behavioral features such as gaze detection (Sciutti et al., 2015; Palinko et al., 2016), anticipatory gaze (Sciutti et al., 2012), a human-like movement profile (Sciutti et al., 2013), and the capacity to adapt movements to increase their legibility for a human partner (Dragan et al., 2013; Stulp et al., 2015). With respect to the notion of legibility, for example, the present study motivates the hypothesis that a robot's willingness to choose an action that is not optimal for itself (e.g., in terms of energy), but which maximizes legibility for a human partner, may be perceived as effortful and thereby boost a human partner's commitment to the interaction. If so, this may be a useful means of increasing the human's patience towards the robot in the event that the robot makes an error, performs a task slowly, or misunderstands an instruction.

Moreover, it will be important to investigate to what extent a human's sense of commitment to a robot partner may be increased by the robot's physical presence. In our study, participants' only exposure to the robot was through a pair of pre-recorded videos at the beginning of the experiment. In general, however, the physical presence of a robot partner may be expected to increase a human's sense of commitment to interacting with it. Indeed, this conjecture is motivated by the results of a recent study showing that people were more likely to comply with a robot's odd request (to throw some books into a rubbish bin) when the robot was physically present than when the interaction was mediated by video (Bainbridge et al., 2008), and also Wainer and colleagues' (2006) finding that participants enjoyed interacting with a physically present robot more than with a remote telepresent robot. Similarly, it is also likely that a human's sense of commitment to an interaction with a robot partner could be strengthened by enabling them to communicate: research has shown that people treat a computer agent more like a human when there is an initial verbal interaction between them (Lee, Kiesler & Forlizzi 2010), and that verbal communication can help guide a human interactant in forming appropriate expectations about a robot's capabilities (Fischer, 2011).

5. Conclusion

The findings reported here have important implications, indicating that human interactants' willingness to persist in interacting with a robot partner despite increasing boredom or frustration may be enhanced by implementing cues that the robot is investing effort. Further research is needed to investigate what other cues of a robot partner's effort contribution, such as physical effort or time, may increase a human user's commitment to remain patiently engaged.

Acknowledgements

This research was supported by a Starting Grant from the European Research Council (n 679092, SENSE OF COMMITMENT) and by the European Project CODEFROR (FP7-

PIRSES-2013-612555).

References

Bainbridge, W. A., Hart, J., Kim, E. S., & Scassellati, B. (2008, August). The effect of presence on human-robot interaction. In *Robot and Human Interactive Communication*, 2008. RO-MAN 2008. The 17th IEEE International Symposium on (pp. 701-706). IEEE.

Breazeal, C., Brooks, A., Gray, J., Hoffman, G., Kidd, C., Lee, H., ... & Mulanda, D. (2004). Humanoid robots as cooperative partners for people. *Int. Journal of Humanoid Robots*, *1*(2), 1-34.

Brooks, D. J., Begum, M., and Yanco, H. A. (2016). “Analysis of reactions towards failures and recovery strategies for autonomous robots,” in *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2016)* (New York, NY: IEEE), 487–492.

Chennells, M., Michael, J. (2018). Effort and performance in a cooperative activity are boosted by the perception of a partner’s effort, *Nature: Scientific Reports* (2018) *8:15692* | DOI:10.1038/s41598-018-34096-1

Clodic, A., Cao, H., Alili, S., Montreuil, V., Alami, R., & Chatila, R. (2009). Shary: a supervision system adapted to human-robot interaction. In *Experimental Robotics* (pp. 229-238). Springer Berlin/Heidelberg.

Clodic, A., Fleury S., Alami, R., Chatila, R., Bailly, G., Brethes, L., …others (2006). Rackham: An interactive robot-guide. In *ROMAN 2006-the 15th IEEE International Symposium on Robot and Human Interactive Communication*, Hatfield (pp. 502–509).

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method. *Tutorials in Quantitative Methods for Psychology*, 1(1), 42–45.

Cully, A., Clune, J., Tarapore, D., & Mouret, J. B. (2015). Robots that can adapt like animals. *Nature*, *521*(7553), 503-507.

DeSteno, D., Breazeal, C., Frank, R. H., Pizarro, D., Baumann, J., Dickens, L., & Lee, J. J. (2012). Detecting the trustworthiness of novel partners in economic exchange. *Psychological science*, *23*(12), 1549-1556.

Dragan, A. D., Lee, K. C., & Srinivasa, S. S. (2013, March). Legibility and predictability of robot motion. In *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on* (pp. 301-308). IEEE.

Fasola, J., and Matarić, M.J. A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction* 2.2 (2013): 3-32.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160.

Feltz, D. L., Forlenza, S. T., Winn, B., & Kerr, N. L. (2014). Cyber buddy is better than no buddy: A test of the Köhler motivation effect in exergames. *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, 3(2), 98-105

Ferrari F, Eyssel F. 2016 Toward a Hybrid Society. In: *Springer Handbook of Robotics*. Springer International Publishing, p. 909–918.

Fischer, K. (2011). How people talk with robots: Designing dialog to reduce user uncertainty. AI Magazine, 32(4), 31-38.

Grigore, E. C., Eder, K., Pipe, A. G., Melhuish, C., & Leonards, U. (2013, November). Joint action understanding improves robot-to-human object handover. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on* (pp. 4622-4629). IEEE.

Hamacher, A., Bianchi-Berthouze, N., Pipe, A. G., & Eder, K. (2016, August). Believing in BERT: Using expressive communication to enhance trust and counteract operational error in physical Human-Robot Interaction. In *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on* (pp. 493-500). IEEE.

Kahn Jr, P. H., Kanda, T., Ishiguro, H., Freier, N. G., Severson, R. L., Gill, B. T., ... & Shen, S. (2012). “Robovie, you'll have to go into the closet now”: Children's social and moral relationships with a humanoid robot. *Developmental psychology*, *48*(2), 303.

Lee MK, Kiesler S, Forlizzi J. Receptionist or information kiosk: How do people talk with a robot? Presented at the Conference on Computer-Supported Cooperative Work, New York, 2010.

Lee, J. J., Knox, B., Baumann, J., Breazeal, C., & DeSteno, D. (2013). Computationally modeling interpersonal trust. *Frontiers in psychology*, *4*, 893.

Lenz, C., Nair, S., Rickert, M., Knoll, A., Rosel, W., Gast, J., ... & Wallhoff, F. (2008, August). Joint-action for humans and industrial robots for assembly tasks. In *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on* (pp. 130-135). IEEE.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic bulletin & review*, *1*(4), 476-490.

Michael, J., & Salice, A. (2017). The Sense of Commitment in Human–Robot Interaction. *International journal of social robotics*, 9(5), 755-763.

Michael, J., Sebanz, N., & Knoblich, G. (2015). The sense of commitment: A minimal approach. *Frontiers in psychology*, *6*, 1968.

Michael, J., Sebanz, N., & Knoblich, G. (2016). Observing joint action: Coordination creates commitment. *Cognition*, *157*, 106-113.

Mirnig, N., Stollnberger, G., Miksch, M., Stadler, S., Giuliani, M., & Tscheligi, M. (2017). To err is robot: How humans assess and act toward an erroneous social robot. *Frontiers in Robotics and AI*, *4*, 21.

Palinko, O., Sciutti, A., Wakita, Y., Matsumoto, Y., & Sandini, G. (2016, November). If looks could kill: Humanoid robots play a gaze-based social game with humans. In *Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on* (pp. 905-910). IEEE.

Peirce, J. W. (2007). PsychoPy – Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13.

Schneider, S. & Kümmert, F. Exercising with a humanoid companion are more effective than exercising alone, *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, Cancun, 2016, pp. 495-501.

Sciutti, A., Bisio, A., Nori, F., Metta, G., Fadiga, L., & Sandini, G. (2012). Anticipatory gaze in human-robot interactions. In *Gaze in HRI from modeling to communication" workshop at the 7th ACM/IEEE international conference on human-robot interaction, Boston, Massachusetts, USA*.

Sciutti, A., Bisio, A., Nori, F., Metta, G., Fadiga, L., & Sandini, G. (2013). Robots can be perceived as goal-oriented agents. *Interaction Studies*, *14*(3), 329-350.

Sciutti, A., Bisio, A., Nori, F., Metta, G., Fadiga, L., Pozzo, T., & Sandini, G. (2012). Measuring human-robot interaction through motor resonance. *International Journal of Social Robotics*, *4*(3), 223-234.

Sciutti, A., Schillingmann, L., Palinko, O., Nagai, Y., & Sandini, G. (2015, March). A gaze-contingent dictating robot to study turn-taking. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts* (pp. 137-138). ACM.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition*, *88*, B11–B21. doi:10.1016/S0010- 0277(03)00043-X

Stenzel, A., Chinellato, E., Bou, M. A. T., del Pobil, Á. P., Lappe, M., & Liepelt, R. (2012). When humanoid robots become human-like interaction partners: corepresentation of robotic actions. *Journal of Experimental Psychology: Human Perception and Performance*, 38(5), 1073.

Stulp, F., Grizou, J., Busch, B., & Lopes, M. (2015, September). Facilitating intention prediction for humans by optimizing robot motions. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on* (pp. 1249-1255). IEEE.

Székely, M., & Michael, J. (2018). Investing in commitment: Persistence in a joint action is enhanced by the perception of a partner's effort. *Cognition*, *174*, 37-42.

Wainer, J., Feil-Seifer, D. J., Shell, D. A., & Mataric, M. J. (2006, September). The role of physical embodiment in human-robot interaction. In *Robot and Human Interactive Communication*, 2006. ROMAN 2006. The 15th IEEE International Symposium on(pp. 117-122). IEEE.