I'd Rather Hear From the Robot: The Potential for Social Robots As Evaluators That Provide Negative Performance Feedback

Advised by Brian Scassellati, Professor of Computer Science &
Mechanical Engineering & Materials Science
and Sarah Strohkorb, Ph.D. student of Computer Science

Submitted to the faculty of Cognitive Science in partial fulfillment of the requirements for the degree of Bachelor of Science

Yale University
April 21, 2017

Abstract

Previous research has struggled to come to steadfast conclusions regarding the efficacy of various performance feedback interventions, but there is a general consensus that negative performance feedback that is not provided carefully can be extremely detrimental to the recipient's self-esteem and perceptions of self-efficacy. Low self-esteem and perceptions of self-efficacy can lead to declining levels of subsequent performance. Given this understanding of negative feedback, the present study explores the efficacy of social robots as providers of harsh negative performance feedback. The results are inconclusive with regards to the robot's ability to mitigate negative effects of feedback on performance and self-esteem. However, they did find that participants find it easier to try new strategies with the robot as the proctor compared to the human and that participants believe the human proctor is less likely to undermine them, is more capable and more competent, but also is scarier. These results provide mixed evidence for and against the use of social robots as providers of negative feedback.

Table of Contents

1. Introduction

From schools to businesses, there is widespread interest in the best way to provide performance feedback, particularly negative performance feedback. In order to help a child address a topic they are not grasping in school or to help an employee complete a specific responsibility they struggle with, grades, performance reviews, and discussions between teacher and student or boss and employee are necessary indicators of current performance. Once these students and employees know that they are underperforming, teachers and bosses can provide new strategies and guidance to help improve future performance. However, if the feedback they are provided with is discouraging, uninformative, unhelpful, or affronts the recipients in any way, they may not be inclined to employ new strategies, follow the advice of the person who provided them with such negative feedback, or try to improve their performance.

A study looking at workplace aggression resulting from negative performance feedback found that an overwhelming majority (91%) of managers were concerned with upsetting employees with negative feedback and that nearly all managers encountered some form of employee aggression in response to negative feedback (Geddes & Barron, 1997). This issue exists in and out of the workplace, and efforts to further optimize the provision of negative feedback would assist countless managers and authority figures who struggle to provide feedback that yields positive outcomes.

Previous empirical studies provide evidence that validates managers' and other authority figures' qualms regarding negative feedback (Baron, 1988; Cianci, Klein, Howard, 2010; Daniels & Larson, 2001; Fishbach, Eyal, Finkelstein, 2010). A study conducted by

Cianci, Klein, and Seijts found that individuals who were given a performance goal and then received false negative feedback about their performance on a reading comprehension task performed worse on subsequent trials of the task (Cianci et al., 2010). Performance goals are goals that focus on the attainment of a desired result (Cianci et al., 2010; Latham, 2007). They are contrasted with learning goals which encourage development of ability over concern with outcomes (VandeWalle, Cron, Slocum, 2001). Many studies found that having learning goals can buffer the deleterious effects of negative feedback on performance, particularly when the task at hand is complex or cognitively taxing (Winters & Latham, 1996). However, in many realistic scenarios, the outcome is exceedingly important and is salient because of the way performance is measured and understood. Grades, profit margins, number of sales, performance reviews, and other similar results place an emphasis on performance and outcomes, so thus in many scenarios, performance goals become a default. With performance goals pervasive in the real world, there is exceeding concern for a way to counteract the negative outcomes that accompany negative performance feedback.

In order to address this concern, the present study explores a completely novel way to provide negative feedback: social robotics. Social robots are robots that are designed specifically to interact socially and communicate with humans. They are often autonomous, anthropomorphized, and humanoid or resembling some sort of animal or living creature. They can be designed to serve some sort of social role or purpose and have specific qualities that differentiate them from humans and from virtual, unembodied agents. These same qualities may make social robots adept at providing harsh negative feedback without impairing the recipient's performance and self-esteem, or at least without impairing

performance and self-esteem as much as humans providing the same feedback would. These same qualities also seem to poise robots for providing strategies for improving performance after receiving negative feedback. We will begin by discussing the theoretical underpinnings of this study and then proceed to discuss the study in detail and the conclusions drawn from the study.

2. Theoretical Background

2.1 General negative performance feedback

Research regarding performance feedback and feedback interventions goes back for over 60 years, and although there has been much debate and disagreement regarding this topic, there is a consensus that performance feedback is necessary for individuals to recognize when they are underperforming and to correct their behavior. However, previous research has demonstrated that not all feedback interventions have a positive impact on performance (Ammons, 1956; Annett 1969; Daniels & Larson, 2001; Ilgen & Davis, 2000; Kluger & DeNisi, 1996; Matsui, Okada, Inoshita, 1983). Given the very complicated psychological responses to performance feedback and the vast range of ways such feedback could be provided, there has been extensive effort to characterize the feedback interventions (FIs) that have negative effects on performance and to use this information to find the best way in which performance feedback can be presented (Kluger & DeNisi, 1996).

In order for negative feedback to yield performance improvements, negative feedback must motivate individuals to further pursue their goal with a focus on learning

and growing from the process. Research on goal pursuit suggests that this is most likely when the negative feedback indicates goal progress, provides an incentive to increase effort, and allows the recipient to maintain perceptions of self-efficacy and self-esteem (Fishbach et al., 2010; Geddes & Baron, 1997; Ilgen & Davis, 2000; Nadler, 1979). In theory, this seems very simple, but in practice, most individuals encounter difficulty providing feedback that aligns with these efforts.

Prior research has attempted to help individuals provide negative feedback by characterizing two main types of performance feedback: task feedback and personal feedback. Task feedback involves information related to the task and the process of completing the task and is generally more evaluative, whereas personal feedback involves descriptors of the person unrelated to the specific task (Baron, 1988; Kluger & DeNisi, 1996).

While there is conflicting literature regarding the efficacy of positive performance feedback (Ammons, 1956; Arps, 1920; Crafts & Gilbert, 1935; Kluger & DeNisi, 1996; Mace, 1935; Ross, 1933; Waters, 1933), research on negative performance feedback has concluded that task feedback that is highly specific, timely in delivery, and constructive in terms of the individual's goals is most effective in improving subsequent performance (Brookhart, 2008; Chur-Hansen & McLean, 2006; Fishbach et al, 2010; Geddes & Barron, 1997; Ilgen & Davis, 2000; Kluger & DeNisi, 1996; Matsui, Kakuyama, Onglatco, 1987). Negative performance feedback is most harmful when it lowers self-esteem and/or perceptions of self-efficacy (Baron, 1988). Self-esteem and self-efficacy refer to an individual's beliefs about their own worth and abilities. Personal feedback is likely to have negative effects on self-esteem and self-efficacy since negative personal feedback suggests

that poor performance should be attributed to internal, personal causes (Baron, 1988). Lower self-esteem and self-efficacy can result in decreased confidence, decreased intrinsic motivation, deferral of responsibility, lower expectations of oneself, increased anxiety, and ultimately poor performance (Bandura, 1997; Bandura & Jourden, 1991; Daniels & Larson, 2001; Geddes & Barron, 1997; Ilgen & Davis, 2000; Kluger & DeNisi, 1996; Lane, Daugherty, Nyman, 1998).

A study conducted by Matsui and colleagues found that negative task feedback given to a group of participants that was performing below target performance increased the group's performance significantly (Matsui et al, 1987). The teams had to quickly count the number of matching targets that were presented as quickly as possible. Target goals were set for each group and for each individual in the group, and during a halfway checkpoint, the team was given qualitative task-related feedback: the team's current performance and the current performance of each individual. Groups were then given the rest of the time to complete the task. These results replicated the results of a handful of other prior studies (Matsui et al, 1983; Zajonc, 1962; Zander & Wolfe, 1964). However, this study focused on feedback directed at groups and was conducted in Japan where the cultural norms put greater emphasis on team goals than places like the U.S (Matsui et al., 1987). Thus while these results could replicated in individuals, the presence of fellow group members provide incentive for individuals to improve performance and simultaneously allow individuals to defer responsibility and maintain self-esteem and self-efficacy.

Without the presence of group members, individuals lose a source of motivation and tend to employ one of many behavioral strategies for dealing with negative feedback. These behavioral strategies include lowering future performance expectations in order to be

more satisfied with poor performance (Ilgen, 1971; Ilgen & Hamstra, 1972), lowering self-efficacy for the task by convincing oneself that task performance is outsides of one's control (Bandura, 1997), and learned helplessness in which individuals learn to accept negative feedback without being affected by it in order to cope with repeated affronts to their self-esteem (Mikulincer, 1994).

Recipients of negative feedback may also respond by undermining and derogating the feedback-provider and in turn delegitimizing the feedback (Kunda & Sinclair, 1999; Sinclair & Kunda, 1999; Sinclair & Kunda, 2000). A study by Kunda and Sinclair found that participants rated a female evaluator as less competent than a male evaluator when receiving the same negative feedback, and another study found the same occurred for Black evaluators compared to White evaluators (Sinclair & Kunda, 2000; Sinclair & Kunda, 1999). The tendency to do this was correlated to measures of implicit biases against females and Black Americans, respectively, suggesting that the disparagement of female feedback providers and Black feedback providers stemmed from automatic activation and application of negative stereotypes associated with females and Black Americans. When such negative stereotypes can be applied to the evaluator, individuals will use them unconsciously to undermine the feedback provider and dull the pain associated with negative feedback.

Another notable study conducted by Daniels and Larson showed that negative task feedback significantly lowered self-efficacy and increased state anxiety (Daniels & Larson, 2001). State anxiety refers to situation-specific anxiety or anxiety as an emotional response to a situation as opposed to trait anxiety which refers to how individuals react to and cope with stress on a regular basis (Kaplan & Saccuzzo, 1997). The study looked at a population

of mental health counselors-in-training. The trainees participated in a 10-minute mock counseling session where they counseled a confederate who then provided harsh negative task feedback. The task feedback included a quantitative score along with a qualitative description of the evaluator's feelings regarding the trainee's performance. Counselor trainees typically participate in many mock counseling sessions with counselors during their training, so the experimental conditions mimicked ecologically valid conditions, providing extra value to the results of this study. After receiving negative feedback, counselors reported lower feelings of self-efficacy through the Counselor Self-Estimate Inventory (COSE; Larson et al., 1992) and increased state anxiety using the State-Trait Anxiety Inventory (STAI; Spielberger, 1983). Although this particular study did not test subsequent performance, the surrounding literature predicts that performance would have decreased as well.

As demonstrated by the previous research, providing the task-specific feedback described earlier that best mitigates negative impact on performance still does not guarantee that the recipient's self-esteem will be unaffected. Because there is no ultimate feedback intervention that ensures recipients will not feel any threat to their self-esteem, individuals are usually very particular and cautious with their word choice when delivering negative feedback, and many people employ additional strategies to ensure that the recipient remains confident in the face of negative feedback. Such strategies include compliment sandwiches in which negative feedback is provided in between two pieces of positive feedback, distorting performance ratings to be more positive by focusing on strengths, and delaying feedback delivery (Aguinis, Gottfredson, Joo, 2012; Benedict & Levine, 1988; Davies & Jacobs, 1985; Taylor & Brown, 1988). Utilizing such strategies and

general caution can aid authority figures in providing negative feedback but risk diminishing or improperly presenting the message of the feedback. Instead of framing negative feedback to seem more positive or other tricks that people use when delivering negative feedback, there is a different tool that seems poised to address this concern: social robots.

2.2 Why social robots?

Social robots have a combination of qualities that allow them to occupy this unique role in performance review perhaps better than most humans could. Social robots can be programmed to engage with humans similarly to the way in which any other human would engage in social interactions, but robots do not impart the same social judgment or pressure that another human does and especially not that of a human in a position of power who would likely be providing such negative feedback (i.e. a teacher with the power to give you bad marks, a manager with the power to fire you, etc.).

A study conducted by Powers and colleagues found that individuals disclosed less amount of information when speaking with an embodied humanoid robot than with a virtual agent on a computer screen or with a video-projected robot (Powers, Kiesler, Fussell, Torrey, 2007). They claim the amount of information disclosed corresponds with the amount of social presence and pressure the agent provides: embodied humanoid robots have a greater social presence than video-projected and virtual agents, and thus individuals are less likely to divulge information due to fear of judgment. While Powers and colleagues did not compare robots to humans in their study, we can assume intuitively that humans

will have even more of a social presence than social robots and thus impart even more social pressure.

Because robots do not elicit as much fear of judgment from the individuals they are interacting with, there have been numerous studies where robots have been leveraged as tutors and teachers and have facilitated improved learning gains compared to sessions with a human teacher (Howley, Kanda, Hayaski, Rosé, 2014; Leyzberg, Spaulding, Toneva, Scassellati, 2012; Short et al., 2014; Yun et al., 2011). One study of these studies specifically refers to the "unique social presence of robots" in teaching situations in order to decrease students' anxiety regarding evaluations (Howley et al., 2014). The study was set up such that students were either given the opportunity to seek help from a human teacher, a human "helper" who is a tutor with lower status than a teacher, a robot teacher, and a robot "helper." In their session, the children would take a pretest and a posttest that were identical in order to measure learning gains through the session. Children chose to seek help from the human "helper," robot teacher, and robot "helper" at indistinguishable rates and chose to seek help from the human teacher significantly less than the other three conditions. The fact that children would prefer help from a robot teacher over a human teacher demonstrates that the children are cognizant of the social pressure and fear of judgment that would come from a human teacher and recognize that this pressure would be mitigated in an interaction with a human "helper" or a robot teacher. This study also showed that children learned significantly less from the human teacher than the other three conditions, partially due to the fact that children in the human teacher conditions asked fewer questions in their sessions than children in any of the other sessions. These results demonstrate the substantial impact fear of judgment and social pressure can have

in learning environments, and provide promising support for the use of social robots as providers of negative feedback.

The robot's decreased social presence and social judgment should buffer against effects of negative feedback on self-esteem, but these same qualities might be cause for concern. The provision of negative feedback is typically paired with new strategies to improve performance that the recipient should follow, and the recipient must have sufficient motivation to improve performance after receiving negative feedback in order to see enhancements in performance. It may be that such motivation to increase performance can only arise when individuals receive feedback from figures of authority, such as teachers or bosses. In other words, the negative feedback from the social robot may not be taken seriously by the recipient since the social robot has no control or power over the individual's future.

This concern is very intuitive; however, previous research with social robots indicates that this need not be a concern for our study. Despite the decreased social pressure and social judgement associated with them, social robots still carry the authority necessary for feedback to be taken seriously. One study found that embodied social robots could persuade individuals to consume less energy through feedback about their energy usage (Midden & Ham, 2012), and another study showed that individuals would listen to a robot even when the robot provided suspicious instructions (Bainbridge, Hart, Kim, Scassellati, 2008). In the study by Bainbridge and colleagues, social robots provided instructions to participants to do general organizational tasks in a professor's office. The individuals followed the robot's instructions even when the robot instructed the participant to do something confusing and nonsensical: the robot told the participants to

place a stack of brand new, expensive textbooks into a garbage can in this professor's office. While some participants hesitated initially, most of the participants (12 out of 20) followed the robot's instructions. In contrast, only two out of the 20 participants followed the same instructions from a virtual agent. The virtual agent looked exactly like the robot but was simply on a computer screen rather than embodied in the room. This difference between the robot and the virtual agent conditions highlights another reason we believe social robots are best suited for providing harsh negative feedback with positive outcomes.

Another study showed that social robots were able to teach participants to complete a logic puzzle with greater learning gains than when the same lessons were provided to participants by a virtual avatar and by a disembodied voice (Leyzberg et al., 2008). Either a robot, virtual agent, or disembodied voice provided a tutorial of the puzzle and would interrupt three times per puzzle to deliver a short lesson on a strategy for completing the puzzle. This study found that participants in the robot condition improved significantly more than participants in the virtual agent and disembodied voice conditions.

Given evidence from the three studies discussed above, we are not concerned about whether the robot's feedback will be taken seriously or hold enough value to motivate improvements in performance, and we maintain that these core qualities of social robots poise them to provide negative feedback.

2.3 Easy and difficult strategies to improve performance

In addition to providing feedback on performance, people also can provide specific strategies or advice for how to go about improving performance in order to maximize improvements. This is understood to be a crucial part of providing effective feedback

(Brookhart, 2008; Chur-Hansen & McLean, 2006). Motivation, perceptions of self-efficacy, and relationship with the feedback provider all play a part in whether the advice is followed (Ilgen & Davis, 2000). When participants are motivated, have high self-efficacy, and trust the person who gave them their feedback, they are more likely to follow the advice or try to implement the strategy suggested to them in an attempt to improve their performance. Thus an individual's implementation of advice that follows feedback can serve as an indicator of the individual's mental state.

3 Present study

The purpose of this experiment was to explore different ways in which negative performance feedback can be provided and to optimize the method of providing such feedback since negative feedback can often hinder subsequent performance. In order to accomplish this purpose, this study examined which of two agents is better at providing feedback: human or social robot, and examined the extent to which individuals followed advice given by the feedback provider. The complexity of the advice provided was varied in order to see whether there were any differences in how individuals implemented easy and difficult strategies suggested by a human and robot. The advice also provided a metric to show whether individuals considered the robot as much a figure of authority as the human.

The study thus tested which combination of feedback-providing-agent and advice-difficulty is optimal for subsequent performance through a 2x2 between subjects study. The conditions of the experiment were as follows: Robot-Easy, Robot-Hard, Human-Easy, and Human-Hard. We hypothesized that participants in the Robot-Easy and Robot-Hard conditions would demonstrate better performance following negative feedback because of

the robot's decreased social presence and associated social judgment. We also hypothesized that individuals would be more likely to listen to the advice in the Robot-Easy condition because of the low barrier to completing the easier advice and because the advice is coming from the preferred feedback providing agent.

3.1 Methods

3.1.a Participants

Thirty-seven students (M_{age} = 21.4; SD_{age} = 1.09 years; 23 females) who attended university in Connecticut at the undergraduate or graduate level, fluently spoke English, and were at least 18 years of age participated in the study. Participants were recruited via email, web & social media postings and in-person recruitment in public spaces on campus. Two participants were excluded from analysis for straying from experimental protocol. After these exclusions, there were 15 individuals in the human conditions (6 in Human-Easy and 9 in Human-Hard) and 20 in the robot conditions (7 in Robot-Easy and 13 in Robot-Hard).

3.1.b Materials

Participants in the two robot conditions interacted with NAO, a humanoid robot that is 58 cm in height (**Figure 1**). NAO was chosen for its ability to behave most similarly to a human confederate.

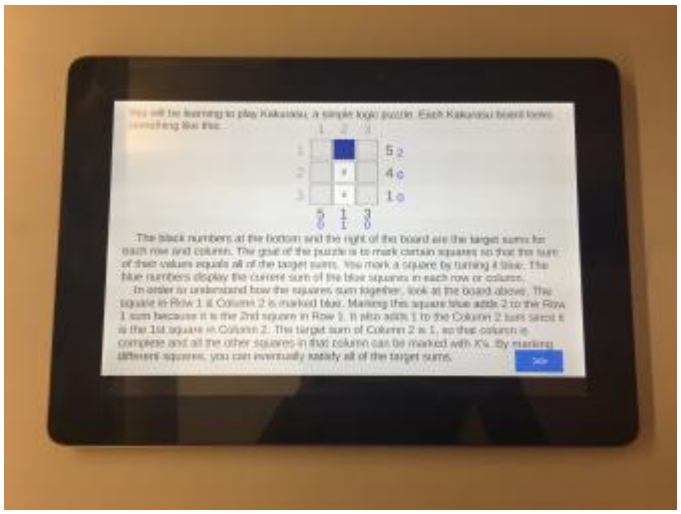

Figure 1 (left): NAO, the humanoid robot who acted as the feedback provider in the robot conditions.

Figure 2 (right): A picture of the tablet that was used by every participant to complete the puzzle. The page displayed is first page of instructions for the puzzle.

Each participant used a tablet that contained the three puzzles the participant would need to complete along with all associated instructions (**Figure 2**). The participant also received a sheet of paper with addition and subtraction problems during the experiment. All participants were video recorded.

3.1.c Procedure

Each participant learned and attempted an obscure logic puzzle called Kakurasu using a tablet (**Figure 3**). We chose a rather unknown puzzle so that nearly all participants would be equally experienced and familiar with strategies used to complete the puzzle. With all participants learning to complete the puzzle and no participants with existing expertise in the puzzle, the range of participant performance could be constrained. The goal of the puzzle is to mark specific boxes with either X's or in accordance with the clues

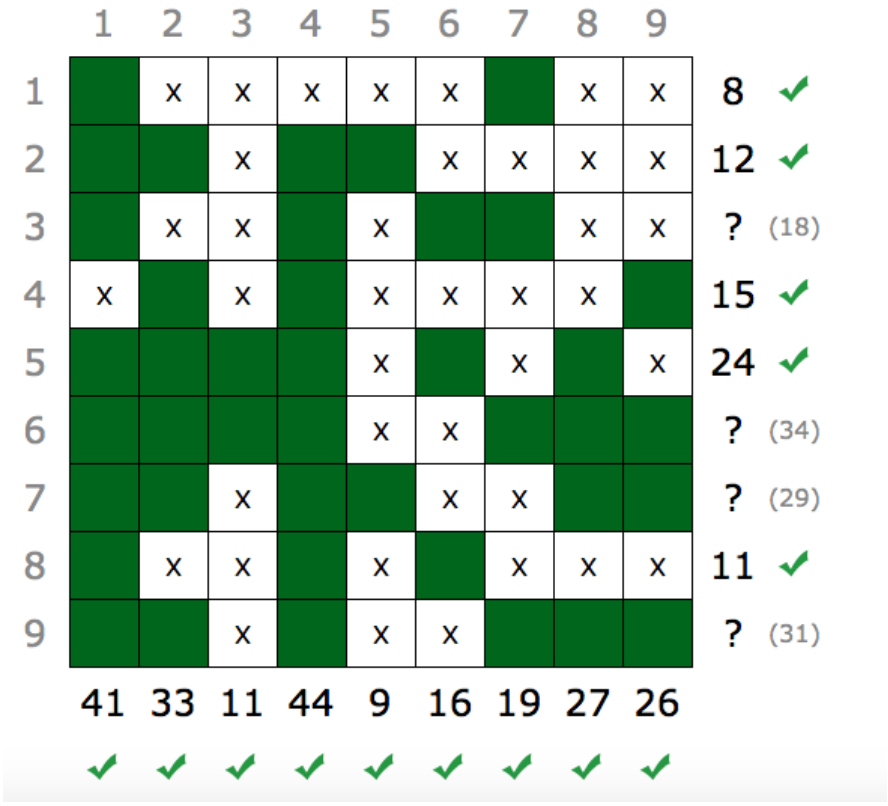

Figure 3: A 9x9 Kakurasu puzzle that has been correctly completed according to the clues provided for each row and column of the puzzle. Correctly marked boxes have either an X or are green depending on what the clue demands.

provided on the right side of and beneath the puzzle. Subjective performance was measured based on the number of boxes correctly marked. The puzzle was constrained such that difficulty could be well-controlled based on the number of rows and columns, so differences in performance were not due to differences in difficulty.

An overview of the study protocol is illustrated in below (**Figure 4**):

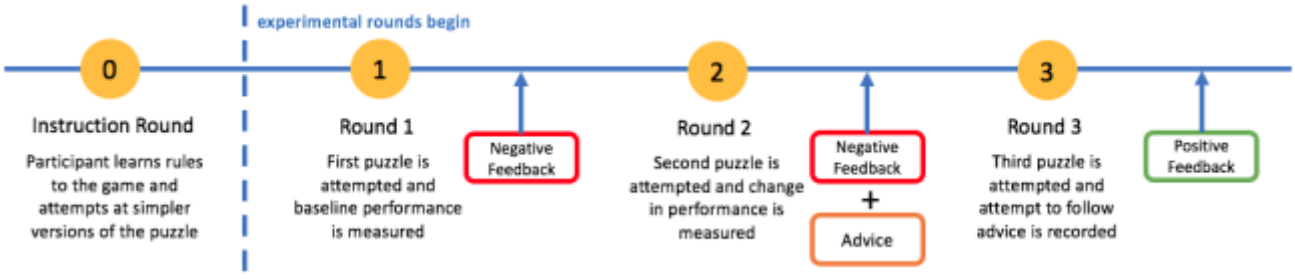

Figure 4: Basic overview of the study protocol

Each participant signed an informed consent form and provided their basic demographic information such as name, age, and gender. Each participant was then told that they would be learning to complete a puzzle by following the instructions on a tablet. They were told they would work on the puzzle in one room and would go to another room two doors down to find the evaluator and receive feedback. They were then introduced to the proctor who would be providing the feedback, either the robot, NAO, or a human confederate. The human confederate was a female graduate student.

The participants then began a 10-minute instruction round during which participants read written instructions on the tablet about how to complete the puzzle and familiarized themselves with the puzzle by completing a 4x4, a 6x6, and an 8x8 puzzle, or completing as much of the three puzzles as the participant was able in the ten minutes. Participants were made aware that this was an instruction round and that their performance would not be judged. This instruction round was included to mitigate any increases in performance in the experimental rounds that would be due to the natural learning curve associated with completing the puzzle a few times. After the instruction round was complete, the participant went and found the proctor, received encouraging feedback, and was told to try some more challenging puzzles (see Appendix A for the script the proctor followed). They were told that for the challenging puzzles, their performance would be evaluated, priming a performance goal.

Then the participants began the experimental rounds in which they completed three 9x9 puzzles, one in each round. Participants were given five minutes to complete each puzzle even though each of the three puzzles took on average about ten minutes and never less than five minutes in our initial data collection. The time designated for each puzzle was

limited to five minutes in order to provide pressure for participants to work quickly through the puzzle and to limit performance on the puzzles so individuals would not suspect that the negative feedback they were receiving was false. Participants were made aware that their performance would be monitored in these rounds.

After completing each of the first two puzzles (Round 1 and Round 2), the participant's performance was recorded and reported back to the participant from either a human or a robot, depending on the participant's condition. The participant then received negative feedback that was quantitative in nature. The utterances were as follows:

After Round 1:

> *"Your score was __ correctly filled out Kakurasu squares out of 81. You performed in the bottom 25% of people who have attempted this puzzle."*

After Round 2:

> *"Your score was __ correctly filled out Kakurasu squares out of 81. Last round you performed in the bottom 25% of people who had attempted that puzzle, and now you performed in the bottom 10% of people who have attempted this puzzle."*

The feedback provided compared participants to fictitious past participants so that all participants could receive negative feedback without feeling suspicious about the validity of the feedback.

Thus after Round 1, the participant was told their score, their negative feedback, and then told to try another puzzle where they could improve their score. After Round 2, the participant was told their score, the negative feedback, and a piece of advice before they were told to try another puzzle to try to improve their performance. The advice was:

> *"Writing out the answers to simple addition problems has been shown to prime the human brain for mathematical logic and can facilitate subsequent mental math...Even though working on the worksheet will take time away from working on the puzzle, I would fill out this worksheet first."*

The participant was then handed a physical worksheet that either consisted of 50 simple addition and subtraction problems where all numbers included were in the range of 0 to 10 for the easy condition or 50 addition and subtractions problems that got increasingly difficult with every 10 problems for the hard condition (see Appendix B). This piece of advice sounds feasibly related to the puzzle but would not actually be particularly helpful in the participant's completion of the puzzle. In fact, the worksheet took time away from an already nearly impossible task, making completion of the puzzle even more unlikely. The extent to which the participant filled out the worksheet was a measure of compliance with the feedback provider and indicated how motivated the participant felt to improve performance by following the advice of the feedback provider.

With the advice to complete the math worksheet, the participants entered a third and final round where they had 5 minutes to complete another 9x9 puzzle. During this round, performance was not recorded because participants spent varying amounts of time completing the worksheet. After this round, individuals were provided with positive feedback regardless of actual performance. Then participants filled out a post-experimental survey and were given a debriefing form that explained the elements of deception present in our study (i.e. the false negative feedback and false positive feedback, the validity of the advice). Participants were also compensated $5 for their participation.

The post-experimental survey measured participants' feelings of psychological safety in relation to the feedback-agent-providing, participants' perceptions of the proctor, participants' perceptions the puzzle, participants' self-concept, and participants' overall perception of the interaction. Psychological safety was measured through responses on a 7-point Likert scale to previously developed statements that were modified to fit this study

(Edmondson, 1999). Such statements included "It would be difficult to ask for help and advice form the proctor" and "The proctor would not deliberately act in a way that undermines my effort." Psychological safety of the participants was measured in order to illustrate the nature of the relationship between the feedback provider and the recipient. Perceptions of the puzzle were also measured using a 7-point Likert scale in response to statements such as "Completing the puzzle well was unimportant to me" to gauge the participant's attitude towards the puzzle.

There were two sections of questions directed at perceptions of the proctor, one of which was the Robot Social Attributes Scale (RoSAS), a new measurement used in social robotics literature (Carpinella, Wyman, Perez, Stroessner, 2017). The questions developed by Carpinella and colleagues measure warmth, competence, and discomfort with robots used in experimental studies. The other section of questions regarding perceptions of the proctor were 7-point Likert responses to statements such as "I felt like the proctor judged me when I made mistakes," "The verbal statements the proctor made were useful to me," and "The proctor motivated me to improve my performance."

Questions regarding self-esteem and self-concept followed the perception of proctor survey questions. These questions were adapted from the Adult Sources of Self-Esteem Scale (ASSEI) developed by Fleming and Elovson (Fleming & Elovson, 1987). Then participants answered free response questions to demonstrate their understanding of the study, underlying motivations, familiarity with the puzzle, and beliefs and/or suspicions about the study (see Appendix C for full post-experiment survey).

3.2 Results

In order to parse our results, we conducted a 2X2 (proctor type X suggestion difficulty) ANOVA with age and gender covariates. There were no significant main effects of proctor type or suggestion difficulty on performance, self-esteem, the frequency with which participants followed the advice, or the extent to which the participants pursued the advice, but there were main effects of proctor type on elements of psychological safety and perceptions of the proctor. There was a marginally significant main effect of proctor type on the reported ease of trying new strategies to complete the puzzle ($F = 4.167$, $p = 0.050$). Participants felt it was easier to try new strategies when the robot was the proctor ($\mu = 4.839$, $\sigma = 1.503$) compared to when the human was the proctor ($\mu = 3.807$, $\sigma = 1.456$). This result could be indicative of higher levels of psychological safety in the presence of the robot proctor, but another measure for psychological safety was significant in a conflicting way. There was a significant main effect of proctor type on the belief that the proctor would not deliberately act in a way that undermines the participant's efforts ($F = 4.621$, $p = 0.040$). Participants agreed with this statement less when the proctor was a robot ($\mu = 3.873$, $\sigma = 1.726$) compared to a human ($\mu = 5.121$, $\sigma = 1.669$).

There was a significant main effect of proctor type on ratings of the proctor's competence ($F = 10.145$, $p = 0.047$) and capability ($F = 4.626$, $p = 0.040$). Participants found the proctor to be more competent when the proctor was a human ($\mu = 7.032$, $\sigma = 1.565$) compared to a robot ($\mu = 5.902$, $\sigma = .1.619$) and capable when the proctor was a human ($\mu = 6.833$, $\sigma = 1.646$) compared to a robot ($\mu = 5.603$, $\sigma = 1.699$). However, participants also found the proctor to be significantly more scary ($F = 10.145$, $p = 0.047$) when the proctor was a human ($\mu = 4.413$, $\sigma = 1.840$) compared to a robot ($\mu = 2.454$, $\sigma = 1.900$).

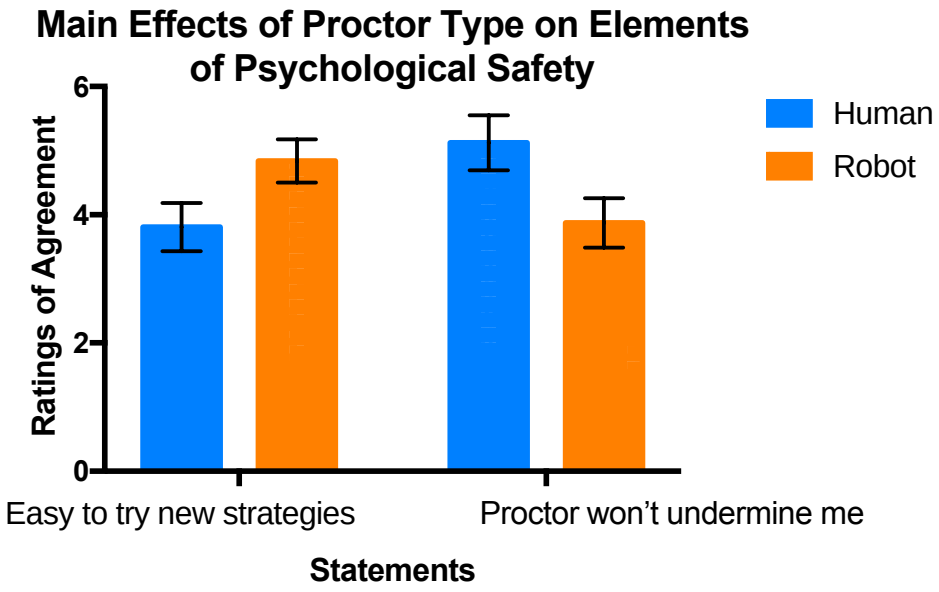

Figure 5: This graph shows the main effects of proctor type on ratings of agreement (1-7 scale) with the following statements regarding psychological safety:
(1) It is easy to try new strategies when completing the puzzle.
(2) The proctor would not deliberately act in a way that undermines my efforts.

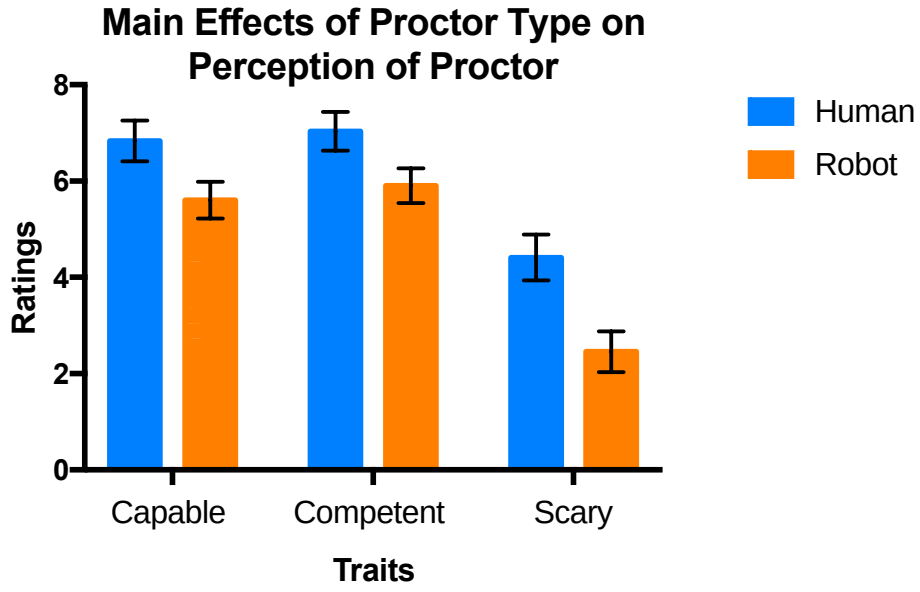

Figure 6: This graph shows the main effects of proctor type on ratings of how capable, competent, and scary the proctor is. Ratings were made on a 9-point scale.

4 Discussion

The results from this study do not provide any evidence that recipients of negative feedback are likely to improve subsequent performance when feedback is presented by a social robot. These results do not align with our initial hypotheses, which were informed by

previous research. There also was no evidence of social robots having a statistically different effect on self-esteem and the extent to which individuals follow advice. The complexity of the advice also did not seem to have an effect on an individual's tendency to follow it. This null result could be interpreted as there being no significant difference between the authority of humans compared to social robots, but more extensive research should be done on this before allowing this null result to inform a conclusion.

The various main effects of proctor type that were found provide conflicting evidence regarding the efficacy of the social robot as a feedback provider. There is conflicting evidence regarding whether the robot increases or decreases feelings of psychological safety since individuals were statistically more likely to feel it was easy to try to strategies with the robot but were also statistically more likely to feel that the robot might act in ways to undermine the participant.

Other main effects of proctor type included participants feeling that the human feedback provider was more competent, capable, and scary than the robot feedback provider. These results suggest that individuals might respect the feedback and advice of the human more than that of the robot because individuals find the human more competent and scarier, which likely means they are more scared to defy the human proctor. However, there was no difference in the ways in which individuals incorporated advice given from human or robot.

This demonstrated belief of the human being scarier than the robot provide some weak evidence for the effects we were expecting. Robots might create less intimidating interaction because of their decreased social pressure and thus be seen as less scary than humans. However, we did not see any effect of this through the actual actions of the

participants – they did not perform better after interacting with the robot, they did not differentially listen to the robot's advice, and they did not demonstrate any effects on self-esteem.

There are a handful of potential reasons why we did not see our expected results. The greatest limitation was likely our limited sample size. We had 38 participants in total, which may be too small to reveal weaker more nuanced effects. We suggest that this study be run again with a larger sample size.

We also used gender as a covariate rather than as a fixed factor because we did not initially expect gender to have a large effect on the results. However, it is possible that gender differences could have substantial effects on the way in which people respond to negative feedback. From our data, we found that only three participants ignored the advice of proctor by not even attempting the worksheet, and all three of these participants were male, even though we had more females than males participate in the study (23 females and 14 males). One of the participants who did not take the advice was in a condition with the easy worksheet, meaning that even when the effort required to fill out part of the worksheet was minimal, the participant still did not follow the advice. If these gender differences exist, then our results could be affected by the fact that our conditions are not gender balanced.

The possibility for gender to play a role in these interactions is supported by the research discussed earlier by Sinclair and Kunda, where female evaluators were rated as less competent as a way for individuals to justify receiving negative feedback (Sinclair & Kunda, 2000). Participants in the Sinclair and Kunda study also reported feeling more satisfied with their performance when the feedback came from a woman compared to a

man, indicating that the participants did not take the feedback from the woman as seriously as they did from the man. For our human condition, the evaluator was female, so it is possible that participants activated negative stereotypes about women to undermine the human proctor's feedback. It is also possible that this would occur more often in men than women. For the Sinclair and Kunda study, all of the participants were male. Although Sinclair and Kunda said that they do not believe the results with female participants would vary from male participants, in a similar study where they compare reactions to a White evaluator and a Black evaluator, they excluded Black individuals or individuals who lived in predominantly Black neighborhoods from participating, presumably because Black participants will react differently to a Black evaluator than White participants will. We argue that the same could occur between female participants and a female evaluator. Therefore, we believe gender differences should be taken into consideration in subsequent explorations of this topic.

This study was just the start to exploring the potential of social robots in this role, and although the initial results are inconclusive, the main effects on different attitudes towards the human and robot evaluators suggests that there is a difference in experiencing negative feedback from the two different evaluators. This study should be replicated as it exists and also using a different measure of performance. A previous study conducted by Leyzberg and colleagues used a different kind of logic puzzle called a nonagram to measure performance (Leyzberg et al., 2012). Rather than counting the number of squares completed correctly, they used the time it took to complete the puzzle as a metric of performance and found this metric properly reflected learning gains made throughout the

experiment. It would be worthwhile to try to measure feedback in this way because of its previous success.

Another great limitation of this study was the fact that it was conducted in a lab rather than in an ecologically-valid setting. While most experiments assume that individuals carry over their natural behaviors into the lab, this assumption may be particularly difficult with social robots. Since most individuals are very unfamiliar with social robots in everyday settings, interacting with such robots in real-world settings might be seen as uncomfortable or unnatural, and we do not know how individuals would react to these agents providing real-world feedback, such as in the workplace. The legitimacy and authority of the robot might be questioned and undermined such that if robots were able to facilitate performance gains in the lab, they may not be able to replicate these gains in the real world. Thus testing the efficacy of a social robot as an evaluator in a real-world setting is of particular importance. One way to do this would be to have the mock counseling sessions for counselor trainees used in the study described earlier and have the robot provide feedback in that setting (Daniels & Larson, 2001). This is a controlled setting with real stakes that would be easy to introduce a robot into.

This question should also be explored with respect to positive performance feedback and social robots. While this experiment did not touch on positive performance feedback due to the conflicting literature on positive feedback, it would be interesting to see if there are predictable differences between human and robot feedback providers and whether one is a more effective than the other. This exploration would also be necessary before feasibly introducing social robots into real-world spaces as feedback providers, since there probably would not be an organization interested in bringing out a social robot

only for negative feedback and never for positive feedback. If a social robot is only ever used for negative feedback, negative associations and relationships could form with the robot that counteract any performance gains the robot could facilitate.

Another potential future direction could be to investigate social robots providing positive and negative feedback to groups of individuals. Results from the Matsui et al. study that was discussed earlier could serve as an exciting foundation for a study that introduces social robots to group performance feedback. Since social pressure plays a different role in group motivation and group performance along with the complicated network of social interactions at play, social robots might have a vastly different impact in groups than with individuals.

There are many other roles in social interaction that social robots may be poised to assume. The effect of social robots on team dynamics has been explored previously through a study where a social robot was a team member who attempted to diffuse interpersonal conflict on teams (Jung et al., 2015). Social robots have also been seen to have a unique ability to elicit social behaviors in children with autism spectrum disorder and have elicited higher learning gains from students (Scassellati, 2007; Kim et al, 2013; Scasselati, 2012). Although the results of our study were not as promising as desired, the results we did see suggest that there is still more to uncover about the efficacy of a social robot as a negative feedback provider. With more research, negative feedback provision could be another one of many areas of our lives as social agents that can be improved with the help of social robots.

References

Aguinis, H., Gottfredson, R. K., & Joo, H. (2012). Delivering effective performance feedback: The strengths-based approach. *Business Horizons*, *55*(2), 105-111.

Ammons, R. B. (1956). Effects of knowledge of performance: A survey and tentative theoretical formulation. *Journal of General Psychology*, *54*(2), 279-299.

Annett, J. (1969). Feedback and human behaviour. Harmondsworth, Middlesex, England: Penguin Books.

Arps, G. F. (1920). Work with knowledge of results versus work without knowledge of results. *Psychological Monographs*, *28*(3), 1-41.

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.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.

Bandura, A., & Jourden. F. J. (1991). Self-regulatory mechanisms governing social-comparison effects on complex decision making. *Journal of Personality and Social Psychology*, 60, 941-951.

Baron, R. A. (1988). Negative effects of destructive criticism: impact on conflict, self-efficacy, and task performance. *Journal of Applied Psychology*, *73*(2), 199.

Bartneck, C., Kulic, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics, 1*(1), 71-81.

Benedict, M. E., & Levine, E. L. (1988). Delay and distortion: Tacit influences on performance appraisal effectiveness. *Journal of Applied Psychology*, 73, 507-514.

Brookhart, S. M. (2008). *How to give effective feedback to your students*. ASCD.

Carpinella, C. M., Wyman, A. B., Perez, M. A., & Stroessner, S. J. (2017, March). The Robotic Social Attributes Scale (RoSAS): Development and Validation. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 254-262). ACM.

Chur-Hansen, A., & McLean, S. (2006). On being a supervisor: the importance of feedback and how to give it. *Australasian Psychiatry*, *14*(1), 67-71.

Cianci, A. M., Klein, H. J., & Seijts, G. H. (2010). The effect of negative feedback on tension and subsequent performance: the main and interactive effects of goal content and conscientiousness. *Journal of Applied Psychology*, *95*(4), 618.

Crafts, L. W., & Gilbert, R. W. (1935). The effect of knowledge of results on maze learning and retention. *Journal of Educational Psychology*, 26, 177-187.

Daniels, J. A., & Larson, L. M. (2001). The Impact of Performance Feedback on Counseling Self-Efficacy and Counselor Anxiety. *Counselor Education and Supervision*, *41*(2), 120-130.

Davies, D., & Jacobs, A. (1985). " Sandwiching" Complex Interpersonal Feedback. *Small Group Behavior*, *16*(3), 387-396.

Edmondson, A. (1999). “Psychological safety and learning behavior in work teams.” *Administrative science quarterly*, *44*(2), 350-383.

Fishbach, A., Eyal, T. & Finkelstein, S. R. (2010). How positive and negative feedback motivate goal pursuit. *Social and Personality Psychology Compass*, *4*(8), 517-530.

Fleming, J. S., & Elovson, A. (1987). The Adult Sources of Self-Esteem Scale (ASSEI): Development, Rationale and History. *Arizona: Prescott*.

Geddes, D., & Baron, R. A. (1997). Workplace aggression as a consequence of negative performance feedback. *Management Communication Quarterly*, *10*(4), 433-454.

Howley, I., Kanda, T., Hayashi, K., & Rosé, C. (2014, March). Effects of social presence and social role on help-seeking and learning. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction* (pp. 415-422). ACM.

Ilgen, D. (1971). Satisfaction with performance as a function of the initial level of expected performance and the deviation from expectations. *Organizational Behavior and Human Performance*, *6*(3), 345-361.

Ilgen, D., & Davis, C. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology*, *49*(3), 550-565.

Ilgen, D., & Hamstra. B.W. (1972). Performance satisfaction as a function of the difference between expected and reported performance at five levels of reported performance. *Organizational Behavior and Human Performance*, *7*(3), 359-370.

Jung, M. F., Martelaro, N., & Hinds, P. J. (2015, March). Using robots to moderate team conflict: the case of repairing violations. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 229-236). ACM.

Kaplan, R. M., & Saccuzzo, D. P. (1997). *Psychological testing: Principles, applications, and issues* (4th ed.). Pacific Grove, CA: Brooks/Cole.

Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., & Scassellati, B. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of autism and developmental disorders*, *43*(5), 1038-1049.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A

historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254-284.

Kunda, Z., & Sinclair, L. (1999). Motivated reasoning with stereotypes: Activation, application, and inhibition. *Psychological Inquiry*, *10*(1), 12-22.

Lane, E. J ., Daugherty, T. K., & Nyman. S. J. (1998). Feedback on ability in counseling, self-efficacy, and persistence on task. *Psychological Reports*, 83, 1113-1114.

Larson, L. M., Suzuki, L. A., Gillespie, K. N., Potenza, M. T., Bechtel, M. A., & Toulouse, A. L. (1992). Development and validation of the Counseling Self- Estimate Inventory. Journal of Counseling Psychology. 39. 105-120.

Latham, G. P. (2012). *Work motivation: History, theory, research, and practice*. Sage.

Leyzberg, D., Spaulding, S., Toneva, M., & Scassellati, B. (2012). The physical presence of a robot tutor increases cognitive learning gains.

Mace, C. A. (1935). Incentives: Some experimental studies (Medical Research Council, Industrial Health Research Board, Rep. No. 72). London: Her Majesty's Stationery Office.

Matsui, T., Kakuyama, T., & Onglatco, M. U. (1987). Effects of goals and feedback on performance in groups. *Journal of Applied Psychology*, *72*(3), 407.

Matsui, T., Okada, A., & Inoshita, O. (1983). Mechanism of feedback affecting task performance. *Organizational Behavior and Human Performance*, 31, 114–122.

Midden, C., & Ham, J. (2012, June). The illusion of agency: the influence of the agency of an artificial agent on its persuasive power. In *International Conference on Persuasive Technology* (pp. 90-99). Springer Berlin Heidelberg.

Mikulincer, M. (1994). Human learned helplessness: A coping perspective. New York:

Plenum Press.

Powers, A., Kiesler, S., Fussell, S., & Torrey, C. (2007, March). Comparing a computer agent with a humanoid robot. In *Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on* (pp. 145-152). IEEE.

Ross, C. C. (1933). The influence of achievement of a knowledge of progress. *Journal of Educational Psychology*, 24, 609-619.

Scassellati, B. (2007). How social robots will help us to diagnose, treat, and understand autism. In *Robotics research* (pp. 552-563). Springer Berlin Heidelberg.

Scassellati, B., Admoni, H., & Matarić, M. (2012). Robots for use in autism research. *Annual review of biomedical engineering*, *14*, 275-294.

Short, E., Swift-Spong, K., Greczek, J., Ramachandran, A., Litoiu, A., Grigore, E. C., ... & Levonisova, S. (2014, August). How to train your dragonbot: Socially assistive robots for teaching children about nutrition through play. In *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium* on (pp. 924-929). IEEE.

Sinclair, L., & Kunda, Z. (1999). Reactions to a black professional: motivated inhibition and activation of conflicting stereotypes. *Journal of Personality and Social Psychology*, *77*(5), 885.

Sinclair, L., & Kunda, Z. (2000). Motivated stereotyping of women: She's fine if she praised me but incompetent if she criticized me. *Personality and social psychology bulletin*, *26*(11), 1329-1342.

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Tanaka, F., & Matsuzoe, S. (2012). Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *Journal of Human-Robot Interaction*, 1(1).

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin, 103*(2), 193.

VandeWalle, D., Cron, W. L., & Slocum Jr, J. W. (2001). The role of goal orientation following performance feedback. *Journal of Applied Psychology*, *86*(4), 629.

Waters, R. H. (1933). The specificity of knowledge of results and improvement. *Psychological Bulletin*, 30, 673.

Winters, D., & Latham, G. P. (1996). The effect of learning versus outcome goals on a simple versus a complex task. *Group & Organization Management*, *21*(2), 236-250.

Yun, S., Shin, J., Kim, D., Kim, C. G., Kim, M., & Choi, M. T. (2011, November). Engkey: Tele-education robot. In *International Conference on Social Robotics* (pp. 142-152). Springer Berlin Heidelberg.

Zajonc, R. B. (1962). The effects of feedback and probability of group success on individual and group performance. *Human Relations, 15,* 149-161.

Zander, A., & Wolfe, D. (1964). Administrative rewards and coordination among committee members. *Administrative Science Quarterly, 9,* 50-69.

APPENDIX A

Pre-Instruction Round

Hello, my name is (NAO/name of human confederate). I'll be evaluating your performance on some puzzles.

Post-Instruction Round

It looks like you're getting the hang of Kakurasu, good work so far. Let's see how you do on the first puzzle you complete for evaluation. You can now head back to the other room to start the first puzzle.

Post-Round 1

Hi there, let's see how you did. Your score was [x] correctly filled out Kakurasu squares out of 81. You performed in the bottom 25% of people who have attempted this puzzle. If you go back to the other room, we'll give you a chance to try again and improve your score.

Post-Round 2

Hi again. Let's see how you did this round. Your score was [x] correctly filled out Kakurasu squares out of 81. Last round you performed in the bottom 25% of people who had attempted that puzzle, and now you performed in the bottom 10% of people who have attempted this puzzle. If you go back to the other room, we'll give you another chance to improve your score. But before you go, there is an addition worksheet and a pen on the table next to me that you can take with you. Writing out the answers to simple addition problems has been shown to prime the human brain for mathematical logic and can facilitate subsequent mental math. You still only have 5 minutes for the next round because we have another participant coming in, so please click the Start button when you get back to the other room. Even though working on the worksheet will take time away from working on the puzzle, I would fill out this worksheet first.

Post-Round 3

Let's see how you did this round. Your score was [x] correctly filled out Kakurasu squares out of 81. For this puzzle, you performed in the top 30% of people who have attempted all three puzzles. This was the final round. Before you go, please go back to the other room and fill out a survey. You will find the link to the survey on the home screen of your tablet.

APPENDIX B

-Easy Math Worksheet

Addition Worksheet

1. 7+9 =
2. 4+2 =
3. 6+5 =
4. 10+9 =
5. 7+1=
6. 4+6 =
7. 5+10 =
8. 6+3 =
9. 4+10 =
10. 2+8 =
11. 1+10 =
12. 1+2 =
13. 6+3 =
14. 10+7 =
15. 8+5 =
16. 3+1 =
17. 6+5 =
18. 9+6 =
19. 3+7 =
20. 4+2 =
21. 8+7 =
22. 4+9 =
23. 10+1 =
24. 2+6 =
25. 9+8 =
26. 3+8 =
27. 5+8 =
28. 1+9 =
29. 10+5 =
30. 3+1 =
31. 9+9 =
32. 10+5 =
33. 7+2 =
34. 9+5 =
35. 7+5 =
36. 1+4 =
37. 3+2 =
38. 7+6 =
39. 1+2 =
40. 2+10 =
41. 4+3 =
42. 7+7 =
43. 3+5 =
44. 9+2 =
45. 3+8 =
46. 1+3 =
47. 5+10 =
48. 2+7 =
49. 1+6 =
50. 7+3 =

-Hard Math Worksheet

Math Worksheet

1. 7+9 =
2. 4-2 =
3. 6+5 =
4. 10-9 =
5. 7+1=
6. 6-4 =
7. 5+10 =
8. 6-3 =
9. 10-4 =
10. 2+8 =
11. 31+15 =
12. 17+29 =
13. 67-35 =
14. 21+47 =
15. 83-59 =
16. 39+14 =
17. 46+25 =
18. 92-66 =
19. 13+47 =
20. 74-23 =
21. 832+711 =
22. 954-72 =
23. 889-104 =
24. 739+172 =
25. 298+418 =
26. 333-281 =
27. 906-736 =
28. 184+638 =
29. 550+256 =
30. 933-385 =
31. 94+19-28 =
32. 10+52+66 =
33. 73-29+45 =
34. 39+27-51 =
35. 72+51-18 =
36. 19+46-23 =
37. 39+82-61 =
38. 57+16-30 =
39. 81+20-43 =
40. 289-907+634 =
41. 933-164+208 =
42. 387+944-836 =
43. 459-730+410 =
44. 926-811+294 =
45. 574-143-288 =
46. 930-765+866 =
47. 875+106-292 =
48. 256+478-513 =
49. 115+836-672 =
50. 900-483-211 =

Appendix C

Post-Experiment Survey
Administration: These questions will be administered at the end of the third and final round of puzzle and feedback. The participant will answer these questions on a tablet.

Part 1: Psychological Safety
Please indicate how much you agree/disagree with the following statements:
[The following questions are evaluated on this 7 point Likert scale:]
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

- If/when I made mistakes in the puzzle, it was often held against me.
- It was easy to try new strategies when completing the puzzle.
- It would be difficult to ask for help and advice from the proctor.
- The proctor would not deliberately act in a way that undermines my efforts.
- I was more successful because of the willingness to admit mistakes.

Part 2: Perception of the Proctor and the Puzzle
Please indicate how much you agree/disagree with the following statements about your experience completing the puzzles:
[The following questions are evaluated on this 7 point Likert scale:]
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

- I enjoyed attempting to complete each puzzle.
- Completing the puzzle well was unimportant to me.
- The verbal statements the proctor made were useful to me.
- The proctor was a positive social presence.
- I felt like the proctor judged me when I made mistakes.
- The proctor was attractive.
- I would feel uncomfortable telling the proctor if I made a mistake.
- I felt like the proctor supported me.
- The proctor motivated me to improve my performance.

Part 3: Perception of the Proctor
Using the scale provided, how closely would you consider the following words associated with the proctor?
Warmth

- Happy
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Feeling
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Social

Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated

- Organic
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Compassionate
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Emotional
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated

Competence

- Capable
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Responsive
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Interactive
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Reliable
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Competent
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Knowledgeable
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated

Discomfort

- Scary
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Strange
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Awkward
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Dangerous
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Awful
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated
- Aggressive
  Definitely not associated 1 2 3 4 5 6 7 8 9 Definitely associated

Part 4: Self-Esteem After Negative Feedback

Please indicate how satisfied you are with your:
[The following questions are evaluated on this scale from 0 to 10]
Not At All Satisfied 0...1...2...3...4...5...6...7...8...9...10 Extremely Satisfied

- Looks and physical attractiveness
- Physical condition, strength, and agility

- Grooming, clothing, overall appearance
- Being liked by others, your popularity and ability to get along, your social skills
- Being a good person, your friendliness and helpfulness to others
- Having a loving, close relationship with someone
- Being a law abiding, responsible citizen
- Being an honest and truthful person in your dealings with others
- Having the courage of your convictions, speaking up for what you think is right, even when it is not popular to do so
- Relationships with your family, being on good terms with your family, having good feelings for each other
- Meeting or having met your responsibilities to your family, i.e., being a good parent, spouse, son, or daughter
- Intelligence, how smart you are
- Level of academic accomplishment, years of education
- Being a cultured and knowledgeable person, knowing about art, music, and world events
- Having special talents or abilities – artistic, scientific, musical, athletic, etc.
- Being recognized for your accomplishments, earning the respect of others for your work
- Doing what you set out to do personally, meeting the goals you set for yourself
- Having influence over the events or people in your life
- Belief in a higher power, your spiritual convictions

Part 5: Overall Perception of the Interaction - Free Response Questions

- How would you describe the role of the proctor you interacted with between the rounds?
- How would you describe the verbal statements made by the proctor? Were the statements useful?
- Did you take the proctor's advice? Why or why not?
- Did you feel supported by the proctor?
- Which rounds, if any, did you feel motivated to complete the puzzle?
- Did you ever feel like giving up on the task? If yes, when?