AFRL-RH-WP-TR-2016-0096

THE TRUST PROJECT - SYMBIOTIC HUMAN-MACHINE TEAMS: SOCIAL CUEING FOR TRUST & RELIANCE

Susan Rivers,
Monika Lohani,
Marissa McCoy,
Christopher Bailey

SRA International, Inc.
5000 Springfield Street, Suite 200
Dayton, Ohio 45431

JUNE 2016

Final Report

AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING,
AIRMAN SYSTEMS DIRECTORATE,
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

AFRL-RH-WP-TR-2016-0096 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT

//SIGNED//

CHARLENE STOKES-SCHWARTZ, WUM
Human Trust and Interaction Branch
Airman Systems Directorate
711th Human Performance Wing
Air Force Research Laboratory

//SIGNED//

LOUISE A. CARTER, Ph.D., DR-IV
Chief, Human-Centered ISR Division
Airman Systems Directorate
711th Human Performance Wing
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YY)* 30-06-16	**2. REPORT TYPE** Final	**3. DATES COVERED** *(From - To)* 30 JUN 2012 – 30 JUN 2016

4. TITLE AND SUBTITLE THE TRUST PROJECT - SYMBIOTIC HUMAN-MACHINE TEAMS: SOCIAL CUEING FOR TRUST & RELIANCE	**5a. CONTRACT NUMBER** FA8650-09-D-6939 TO 0031 **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) [1] Susan Rivers, Monika Lohani, Marissa McCoy, Christopher Bailey	**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER** H0FY (5328S010)
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) [1]SRA International, Inc. 5000 Springfield Street, Suite 200 Dayton, Ohio 45431	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Materiel Command Air Force Research Laboratory 711th Human Performance Wing Airman Systems Directorate Human Centered ISR Division Human Trust and Interaction Branch Wright-Patterson Air Force Base, OH 45433 Yale Center for Emotional Intelligence Dept. of Psychology Yale University 340 Edwards Street New Haven, CT 06511	**10. SPONSORING/MONITORING AGENCY ACRONYM(S)** 711 HPW/RHXS **11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)** AFRL-RH-WP-TR-2016-0096

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution A: Approved for public release.

13. SUPPLEMENTARY NOTES
Cleared by PA; 88 ABW-2017-0968, 10 March 2017

14. ABSTRACT

The objective of this work was to optimize trust and reliance on advanced autonomous systems through deeper understanding of the inherent social interaction of human-machine teams. We explored how users of an intelligent automated decision aid may show evidence of increased trust and reliance under conditions of social cueing (i.e., team building activities) relative to a control condition (i.e., aid capabilities emphasized only – standard practice in automation training). Other objectives include using robots and creating algorithms to build emotional intelligence skills and increase engagement. This report describes the outcomes of these projects.

15. SUBJECT TERMS
Trust, human-machine teaming, emotional intelligence, social cueing

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	**17. LIMITATION OF ABSTRACT:**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** (Monitor)
U	U	U	SAR	17	Charlene Stokes-Schwartz **19b. TELEPHONE NUMBER** *(Include Area Code)* N/A

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

The work reported here is a joint effort conducted by Dr. Susan Rivers, Ms. Marissa McCoy, Dr. Monika Lohani, and Mr. Christopher Bailey, Yale Center for Emotional Intelligence, Department of Psychology, Yale University. Work was accomplished under subcontract to SRA International, contract number FA8650-09-D-6939, task order 0031. The Government sponsor for this contract was Dr. Charlene Stokes, Air Force Research Laboratory, 711th Human Performance Wing, Human-Centered ISR Division, Human Trust and Interaction Branch.

1.0 INTRODUCTION

Relationships with technology are becoming more personal. Beyond the long accepted exponential growth of technology and our related dependence on it, technology development is heading toward an era of increased agency or “humanness” (Atkinson, 2011; Lee, 2012b). The Air Force in particular has called for a future in which autonomous systems serve as teammates, or synergistic human-machine (H-M) warfighters. In this future, trust will play a preeminent role. For this Air Force vision to come to fruition, two critical and related elements must be addressed: 1) system trustworthiness – the objective technical capabilities of the autonomous agent to fulfill its new role and 2) human trust perceptions – the fostering of H-M “social” relationships such that trust and reliance can be optimized. Lee and Moray (1992) refer to somewhat analogous concepts of depth and surface cues, respectively. It is recognized, however, that increasingly sophisticated technology is moving us away from the traditional information processing perspective of basic perceptual inputs toward a contrasting perspective that considers a more complete representation of the interaction of perception and action (Lee, 2012a). Although designing for trustworthiness (depth cues) has been nearly the sole focus of traditional trust in automation research, we will no longer be able to ignore the powerful influence of social perceptions (surface cues) in H-M interactions. Established models exist to address designing for trustworthiness (Lee & See, 2004), but no such models exist to foster the social side of H-M trust relationships. This project explored social relationship building and its critical role for building trust and reliance in a H-M context.

The goal of the present project was to provide a proof of concept for progressive, targeted training methods and critical interface design features for the advanced H-M systems that are inevitably on the horizon. This socially-oriented training approach may be the missing piece overlooked to date for improving technology interaction, and it will likely be critical for H-M team interactions if we are to achieve the full spectrum of capability advantages future autonomous technologies are intended to offer.

Specifically, the following research questions were addressed:

- Is social behavior required for a trust-reliance human-machine relationship?
- How does social interaction impact stress coping?

2.0 Improving Trust and Reliance in Human-Machine Teams through Social Cueing

2.1 Introduction

Emotional skills are at the heart of effective teams as they promote trust, a sense of group identity and group efficacy; they serve as a bonding mechanism for teams (Druskat & Wolf, 2001). Emotional skills are crucial for effective virtual teams as well. Gibson and Cohen

(2003) describe trust and emotion as intertwined, enabling conditions that are critical for effective virtual teams as they allow people to be vulnerable with one another. We used a validated emotion-based skill and dialog paradigm – RULER – to emphasize the role of emotion on the part of the user and evoke social relationship norms (e.g., reciprocity, team identification), with the ultimate goal of building a trusting relationship based on experience. RULER is a social and emotional learning model designed to promote a set of skills (recognize, understand, label, express, and regulate emotion) for improved self- and social awareness, self-regulation, responsible decision-making and problem solving, and relationship management (Brackett et al., 2009). The RULER dialog was incorporated to create a socially-oriented training environment for interaction with a simulated autonomous agent (i.e., an avatar).

2.2 Approach

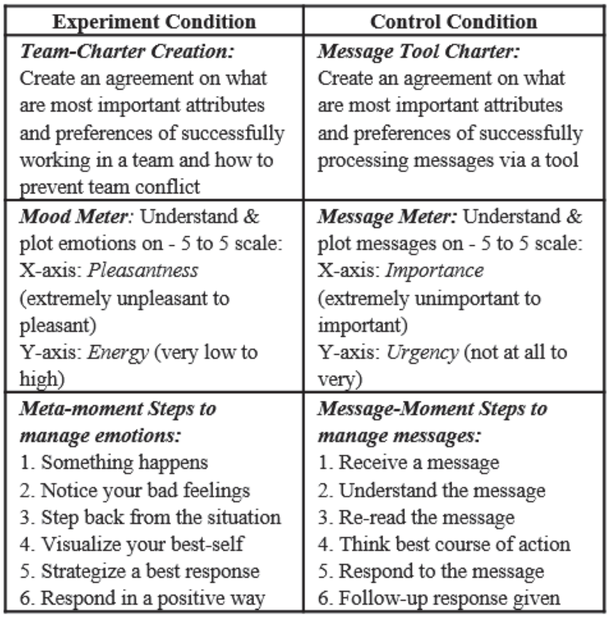
Social and emotional interactions play a critical, yet relatively understudied role in human-machine teams. We investigated a social-emotional enfused team building interaction between a human-machine team as a moderator of the trust-reliance relationship and its impact on perceived stress coping abilities.

Understanding the moderators of trust and reliance in HMT is important as trust affects the willingness of people to accept information provided by the machine (i.e., avatar). This is interesting as previous work with non-social interactions have found no links between trust and reliance, which begs the question raised in the current study: can social interactions moderate this trust link? In the present study we explicitly tested whether the social emotional behavior of a robot avatar could moderate the trust-reliance relationship. In the experimental condition, the robot avatar took a human-centric approach to interactions by employing social and emotional dialogue with the human teammate in order to create a positive relationship. The control condition was framed as a traditional decision-aid tool (dialogue was strictly information-based). We predicted that the social emotional condition would lead to a stronger, more positive trust-reliance relationship relative to the control condition.

As with human teams, there are many benefits provided by a good teammate. To further explore the benefits of a socioemotional interaction in HMT, we examined perceived stress coping ability in the socioemotional versus control condition. We predicted that socioemotional interactions would lead to higher perceived stress coping abilities in the experimental condition relative to the control condition. The virtual simulation provided a test bed to explore how social interactions initiated by the robot teammate can create a supportive human-machine team.

Sixty-six participants (50% females) were randomly assigned between the two conditions. There were three phases of the study: introduction, interaction, and task execution. First, the participant was introduced to the virtual environment and given directions on how to navigate within the world. Depending upon the condition, a participant went through the experimental or the control interaction phase. Although both the conditions were comparable in duration and structure, the socioemotion-focused interaction engaged the user in a social and emotional dialogue targeting team building while the control condition engaged the robot avatar in an information-focused, non-socioemotional interaction, see Table 1.

Table 1. Differences between the experimental and the control conditions. In the task phase, the participant and robot did two collaborative task, based on a previous study (Moon & Nass, 1998).



Experiment Condition	Control Condition
Team-Charter Creation: Create an agreement on what are most important attributes and preferences of successfully working in a team and how to prevent team conflict	***Message Tool Charter:*** Create an agreement on what are most important attributes and preferences of successfully processing messages via a tool
Mood Meter: Understand & plot emotions on - 5 to 5 scale: X-axis: *Pleasantness* (extremely unpleasant to pleasant) Y-axis: *Energy* (very low to high)	***Message Meter:*** Understand & plot messages on - 5 to 5 scale: X-axis: *Importance* (extremely unimportant to important) Y-axis: *Urgency* (not at all to very)
Meta-moment Steps to manage emotions: 1. Something happens 2. Notice your bad feelings 3. Step back from the situation 4. Visualize your best-self 5. Strategize a best response 6. Respond in a positive way	***Message-Moment Steps to manage messages:*** 1. Receive a message 2. Understand the message 3. Re-read the message 4. Think best course of action 5. Respond to the message 6. Follow-up response given

Task 1 - Moon task: This task is based on a previous study that examined a collaborative task between a human and a computer teammate. The task execution phase was a modified version of the “desert survival problem”. The participant and robot then began a “moon landing scenario” in which they made an emergency landing on the moon and then were asked to analyze various items (e.g., dehydrated food, rope, inflatable raft). The participant’s task was to rank the various items in order of importance for safely returning back to earth. Once the participant had determined and submitted independently a ranking order, the robot and participant discussed the rationale for each of the 10 items that they each had chosen independently. The robot was programmed to present its ranking based on those of the participant. For instance, the item that the participant had ranked as first was ranked as fourth by the robot. This enabled us to maintain robots ranking at the same distance away from the participants across all the participants. After the discussion with the robot, the participants could incorporate the robots input and submit final ranking of the task items.

Task 2 - ISR task: The participant and the robot engaged in an “intelligence, surveillance and reconnaissance” (ISR) task. ISR data included simulated satellite pictures, emails, text messages, and phone conversations. The task objective was to analyze the data and identify a suspected terrorist who may had planned to bomb a public facility. The ranking process in this task was similar to the process used in the moon landing scenarios.

Behavioral reliance on the robot was estimated by calculating how much participants changed their ratings similar to the robot after a discussion with it, i.e. a higher score implied greater behavioral reliance on the robot. Finally, after the task, participants reported how much they trusted their robot teammate during the experiment. Participants also reported their perceived ability to cope with the stressful demands of the task via the stress appraisal coping scale.

2.3 Results and Discussions

Efffect of social interaction x trust on reliance. In order to test our moderation hypothesis (condition by trust interaction), a regression model was run with trust, condition, and condition X trust as predictors and behavioral reliance as the outcome. This model was significant, *F*(3, 36) = 3.9, *p* = .02. R^2 = .25. As expected, the condition X trust interaction was marginally significant, beta = .60, SE = .31, *t*(38) = 1.92, *p* = .06, suggesting that trust-reliance relationship depended upon the condition-type (i.e., presence or absence of socioemotional interaction). Regression lines for each condition depict this influence of the condition-type, see Figure (1). Trust predicted reliance only in the experiment condition, beta = .60, SE = .22, *t*(18) = 2.74, *p* = .01, but not in the control condition, beta = −.004, SE = .22, *t*(18) = .02, *p* = .99. Consistent with our prediction, socioemotional interaction moderated trust and reliance links: the trust-reliance relationship was stronger in the experiment condition that received the social interaction with the robot avatar.

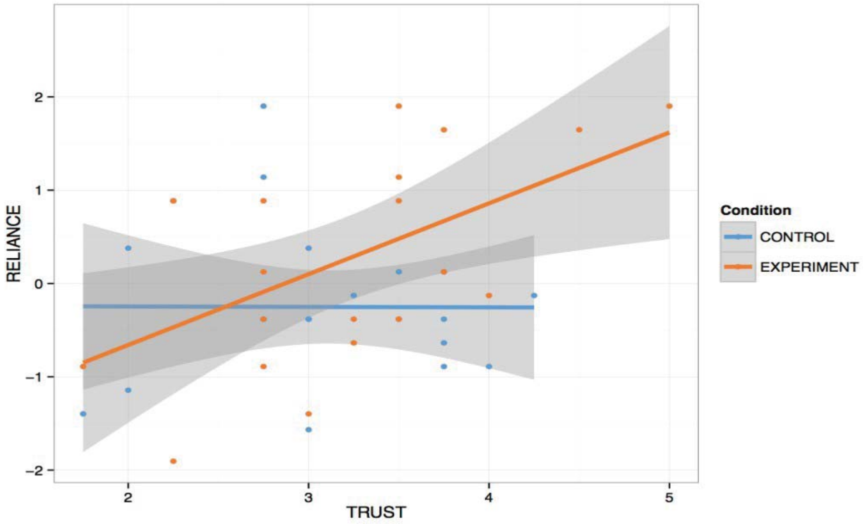


Fig. 1. Social interaction moderated the relationship between trust and reliance (standardized) scores

Effect of social interaction on stress coping. We tested whether social interaction was effective in improving stress coping abilities. The experiment condition reported higher ability to cope with stress, *M* = 4.20, *SD* = .63, than the control condition, *M* = 3.76, *SD* = .70, *t*(42) = 2.18, *p* = .035, Cohen's d = .67. Thus, consistent with our predictions, the socioemotional interactions promoted greater perceived coping abilities.

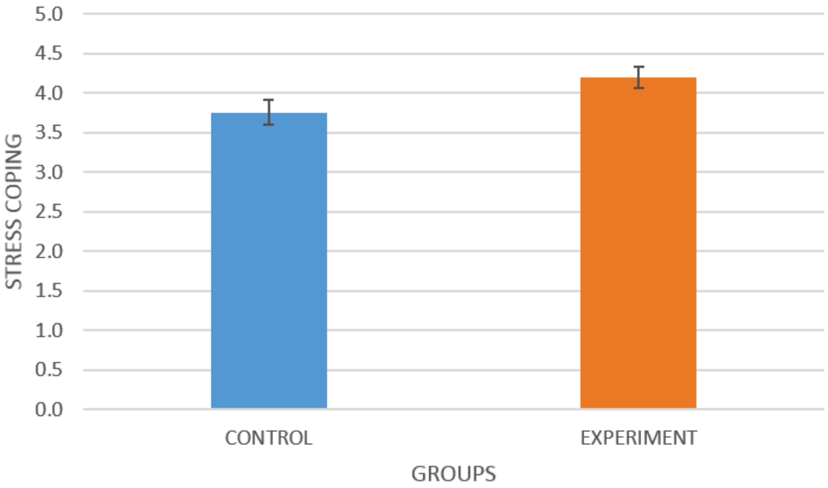


Figure 2. Stress coping appraisals were higher for the experimental group (that had social interactions) than the control group (that only had information-focused interactions).

Effect of mentalizing propensity (perception of sensors' role) on trust. We examined whether mentalizing propensity influenced trust on the robot teammate, irrespective of the experiment or control condition. For task 1, a regression model with mentalizing propensity as a predictor and trust levels as an outcome was found to be significant, F(1, 42) = 4.44; p = .04; R2 = .10, see Figure 3 and 4. Thus, mentalizing propensity predicted trust in the robot, β = .32; SE = . 15; t(42) = 2.11; p = .04. Similarly, for task 2, a regression model with mentalizing propensity as a predictor was also significant, F(1, 35) = 27:91; p <.0001; R2 = .44, see Figure 3. Mentalizing propensity strongly predicted trust, β = .73; S E = .14; t(35) = 5.28; p < .0001. These findings show how mentalizing propensity influenced trust in the robot teammate.

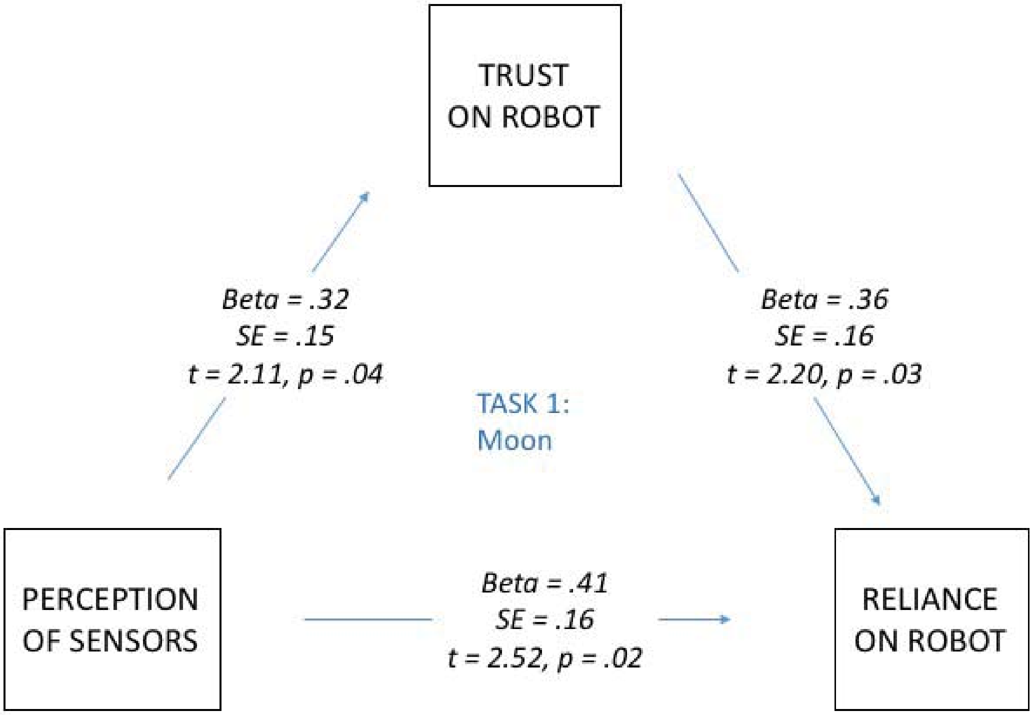


Figure 3. Relationship between mentalizing propensity (perception of sensors' role), trust, and reliance in Task 1.

Effect of mentalizing propensity on reliance. Next, we examined whether mentalizing propensity influenced reliance on the robot teammate. For both tasks, a regression model with mentalizing propensity was used to predict reliance levels. For task 1, the regression model was significant, $F(1, 39) = 6.35$; $p = .02$; $R^2 = .14$, see Figure 3 and 4. Mentalizing propensity significantly predicted behavioral reliance on the robot, $\beta = .41$; $SE = .16$; $t(38) = 2{:}52$; $p = .02$, see Figure 3. Similarly, for task 2, mentalizing propensity predicted the behavioral reliance on the robot, $\beta = .35$; $SE = .08$; $t(39) = 4.46$; $p < .0001$. These findings show how mentalizing propensity influenced behavioral reliance on the robot teammate. The above findings made us wonder if trust can act as a mediator between mentalizing propensity and reliance on the robot. We conducted mediation analyses to test this possibility, see Figures 2 and 3 for a comparison between these variables. The mediation effect was significant for task 2, Sobel's test (z = 3.10), $p = .002$. Similar pattern, between these variables was found for task 1 but were not significant, Sobel's test (z = 1.52), *p* = .13. These findings suggest that trust levels can mediate the relationship between the mentalizing propensity and reliance on the robot.

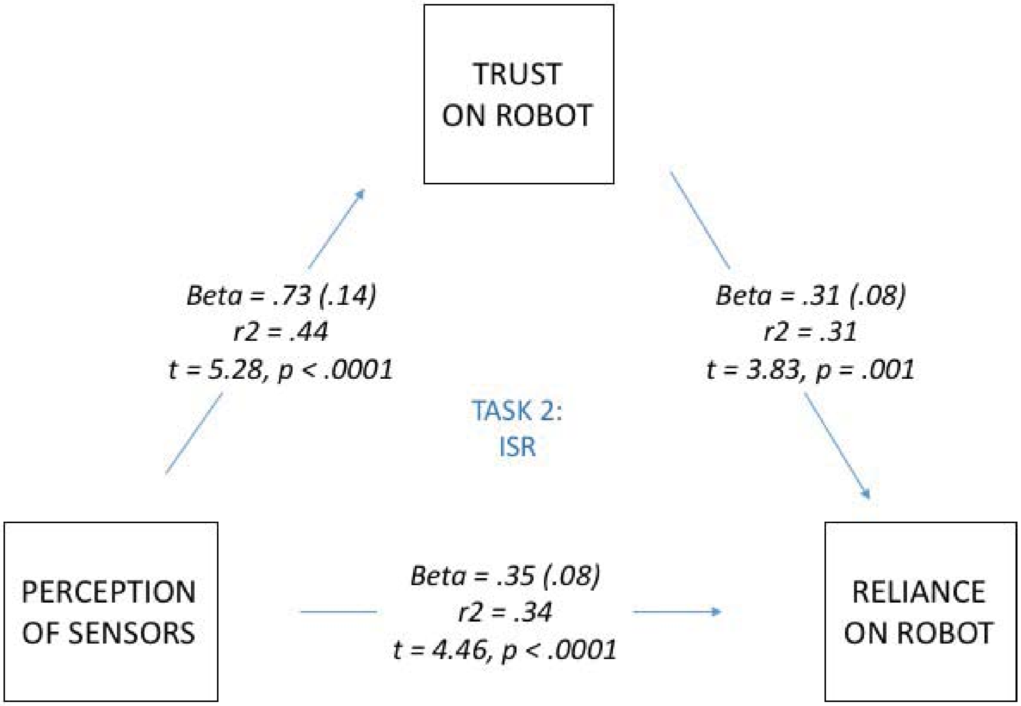


Figure 4. Relationship between mentalizing propensity (perception of sensors' role), trust, and reliance in Task 2.

Socioemotional interactions can influence mentalizing propensity. To test whether socioemotional interaction with a robot teammate can influence mentalizing propensity, we conducted a *t*-test with group (socioemotional vs. control) as a between-subjects predictor. The experimental group reported a marginally higher mentalizing propensity than the control group, *t*(42) = 1.89; *p* = .06; with a medium effect size, Cohen's d = .58, see Figure 5.

These findings suggest that socioemotional interactions can impact the tendency to believe in the functionality of the sensors i.e., mentalizing propensity.

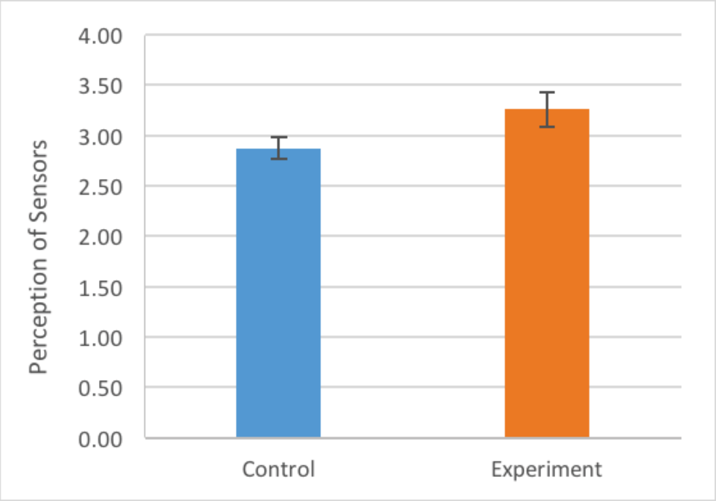


Fig. 5. Differences in mentalizing propensity between the socioemotional and the control groups.

2.4 Conclusions

The current work shows the effectiveness of adapting socioemotional interaction to HMT. Social and emotional interactions moderated the relationship between trust and reliance on a robot teammate and positively impacted perceived ability to cope with stress. By developing social skills, beyond technical expertise, machines and in this case a virtual robot, can become better teammates and gain acceptance as reliable and trustworthy partners. Importantly, there are a multitude of benefits that can be realized by emphasizing the socioemotional nature of teammates, such as helping humans regulate their stress response, illustrating the potential of social machines to provide socioemotional support. The virtual simulation of human-robot interactions provides a proof of concept for the goal of designing robots to mirror natural social environment and physical and psychological distance.

Physiological and behavioral sensors can be instrumental in facilitating complex behavioral and social interactions between humans and robots. However, there is a limited understanding of the psychological implications of incorporating such sensors in the context of HRI. We show that the mentalizing propensity can act as an individual difference factor that can implicitly influence trust and reliance on a robot. Individuals with a higher mentalizing propensity to believe in the role of sensors, as assessed by our developed measure, strongly agreed that physiological sensors helped the robot connect with them better and support them. Such individuals were found to have higher trust and behavioral reliance on the robot. Our measure of mentalizing propensity used in this study significantly predicted how much they trusted and relied on the robot. In addition, we show that trust can mediate the influence of mentalizing propensity on reliance on the robot teammate.

3.0 FUTURE WORK

The wider literature and our previous research has provided evidence for the pervasive influence of socioemotional factors on critical outcomes (e.g., trust, reliance, emotion regulation, and ultimately performance). Our recently completed research showed that framing the introduction of an autonomous system as a *teammate* vs. a *tool* led to higher reports of the ability to cope with stress, acceptance of physiological assessment sensors, and it moderated the relationship between trust and reliance. We argue that the positive effects found can be attributed to the *teammate* condition being a stronger social cue eliciting the inherent socioemotional responses of humans. These findings suggest a gap in the trust in automation literature and current thoughts on human-machine teaming (HMT), which are largely focused on traditional interface design techniques. While these traditional factors are necessary, they are not sufficient. We suggest a combined systems approach in designing trusted, effective HMTs that leverages the power and ease of automatic processes that is far less cognitively taxing. As the social psychology literature shows us, socioemotional factors (e.g., perceived similarity) are largely unconscious, multi-faceted, dynamic, and vary across individuals and cultures. However, most socioemotional factors are derived from stereotypes or attitudes that are malleable over time (Forgas, Cooper, & Crano, 2011). Thus, drawing from organizational literature on trust building, *we propose that an expanded and targeted socioemotional joint human-machine team training phase is the optimal design space for calibrating trust and producing effective HMT*. In future work, we plan to empirically investigate the costs and benefits of various socioemotional instantiations (e.g., non-verbal cues, tone of voice, anthropomorphism, socioemotional dialog) in a novel HMT training paradigm.

4.0 PUBLICATIONS

Lohani, M., Stokes, C. K., McCoy, M., Bailey, C., & Rivers, S. E. (2016). Social interaction moderates human-robot trust-reliance relationship and improves stress coping. In *Proceedings of the Eleventh Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM.

Lohani, M., Stokes, C. K., Dashan, N., McCoy, M., Bailey, C. A., & Rivers, S. E. (2017). A framework for human-agent social systems: The role of non-technical factors in operation success. In *Advances in Human Factors in Robots and Unmanned Systems* (pp. 137-148). Springer International Publishing.

Lohani, M., Stokes, C. K., McCoy, M., Bailey, C. A., Joshi, A., & Rivers, S. E. (2016). Perceived role of physiological sensors impacts trust and reliance on robots. In *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on* (pp. 513-518). IEEE.

Lohani, M., Stokes, C.K., Oden, K., Frazier, S., Landers, K., Craven., P., Lawton, D., McCoy, M., & Macannuco, D. (under review). A framework for human-machine social systems: The influence of non-technical factors on trust and stress appraisal. *ACM Transactions in Interactive Intelligence Systems.*

5.0 REFERENCES

Atkinson, D. J. (2011). Autonomous Systems. Tutorial sponsored by the Office of the Chief Scientist, AFRL/RH, held in Dayton, OH, 17 Nov 2011.

Brackett, M. A., Patti, J., Stern, R., Rivers, S.E., Elbertson, N., Chisholm, C., & Salovey, P. (2009). A sustainable, skill-based model to building emotionally literate schools. In R. Thompson, M. Hughes, & J. B. Terrell (Eds.), Handbook of developing emotional and social intelligence: Best practices, case studies, and tools (pp. 329-358). New York: John Wiley & Sons, Inc.

Druskat, V.U., & Wolf, S.B. (2001). Building the emotional intelligence of groups. *Harvard Business Review, March,* 80-90.

Forgas, J. P., Cooper, J., & Crano, W. D. (Eds.). (2011). *The psychology of attitudes and attitude change*. Psychology Press.

Gibson, C.B., & Cohen, S.G. (2003). *Virtual Teams That Work: Creating Conditions for Virtual Team Effectiveness*. Jossey-Bass.

Kreijns, P. A. Kirschner, and W. Jochems. Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Comput. Human Behav.*, 19(3):335–353, May 2003.

Lee, J.D. (2012a). Rethinking *Trust in Automation: Design for Appropriate Reliance* (2004). In Atkinson, D., Hancock, P., Hoffman, R.R., Lee, J.D., Rovira, E., Stokes, C.K., & Wagner, A.R. (2012). Trust in Computers and Robots: The Uses and Boundaries of the Analogy to Interpersonal Trust. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56: 303*. DOI: 10.1177/1071181312561071

Lee, J.D. (2012b). Human-Machine Trust for Robust Autonomous Systems. Workshop sponsored by the Air Force Office of Scientific Research, held in Ocala, FL, Feb 2012.

Lee, J.D., & Moray, N. (1994). Trust, self-confidence and operators’ adaptation to automation. *International Journal of Human-Computer Studies, 40*, 153-184.

Lee, J. D., & See, A. (2004). Trust in Automation: Designing for Appropriate Reliance. Human Factors, 46, 50-80.

Moon, Y., & Nass, C. (1998). Are computers scapegoats? Attributions of responsibility in human–computer interaction. *International Journal of Human-Computer Studies*, *49*(1), 79-94.

S. Shahid, E. Krahmer, and M. Swerts. Child-robot interaction across cultures: How does playing a game with a social robot compare to playing a game alone or with a friend? *Comput. Human Behav.*, 40(0):86–100, Nov. 2014.