NATURAL SPEECH TOWARD HUMANS AND INTELLIGENT AGENTS DURING A SIMULATED SEARCH AND RESCUE MISSION

A.M. Sinatra*, V.K. Sims, M.G. Chin, H.C. Lum, L.U. Ellis, G. Hancock, M. Raymond, A. Baruch, and S. Colombo
University of Central Florida
Orlando, FL, 32816

A. Leonessa
Virginia Polytechnic Institute and University
Blacksburg, VA, 24061

I. Hudson
RDECOM STTC
Orlando, FL, 32826

ABSTRACT

Recently, robots and artificial intelligence have been utilized in situations that are hazardous to humans. This experiment examined differences in performance, communication and perception when humans work with human or non-human "intelligent agent" teammates while engaged in a simulated search and rescue mission. Participants were found to speak more when working with a non-human partner and when more information on their partners' progress was shared. Results suggest that sharing information on the status of a mission may lead participants to interact differently with a non-human "intelligent agent" than a human.

1. INTRODUCTION

As technology continues to develop, robots are becoming increasingly useful assets to humans. Indeed, within the past few decades, researchers have begun to focus more on the complex cognitive demands placed on humans as they interact more with robot teammates (Cuevas, Fiore, Caldwell & Strater, 2007). Robots are not only able to perform a variety of tasks that prove beyond normal human capabilities such as very complex mathematical computations; they are also able to perform tasks in environments that are extremely detrimental to human health and safety (i.e., hazardous environments or terrorist group infiltration). While the use of robots is advantageous in these circumstances, it is important to understand how humans will interact with their robot versus human teammates. This is especially important as the current trend continues of increased human-automation teamwork in a number of influential fields, including the military, agriculture and medicine (Hinds, Roberts & Jones, 2004).

Does the nature of a person's teammate, whether he or she is another human or a non-human "intelligent agent," affect how a human perceives him or her? Does the type of the teammate also affect how a person interacts with, directs or stops the teammate? These questions must be answered in order to ensure the effective acceptance as well as use of robots and artificial intelligence systems in our daily lives.

An essential component in any cooperative team interaction is how a person communicates with their teammate no matter whether that teammate is human or not. Research indicates that "good communication distinguish[es] high performance teams from low performance teams" (Aubert & Kelsey, 2003) and in high consequence situations such as search and rescue missions, good communication is key. The effectiveness of a human/intelligent agent team may also depend on personal attitudes that may differ depending on gender (males being more accepting of robot technology than females) and past experiences (Nomura, Suzuki, Kanda & Kato, 2006).

Research has supported the idea that, while working with a robot partner, participants spend more time trying to understand the state of the robot rather than guiding it (Burke, Murphy, Coovert & Riddle, 2004). In addition, when working on a collaborative task with a robot, studies have found that people are more likely to trust a robot with more responsibility when it resembles a human, rather than when it looks like a conventional machine (Hinds, Roberts & Jones, 2004). Some research has looked at whether or not people are willing to help robots (Huttenrauch & Eklundh, 2006). Research has been done on giving directions to robots, and the point of view that teammates normally attribute to them (Imai, Hiraki, Miyasato, Nakatsu, & Anzai, 2003).

2. CURRENT RESEARCH

In our study, we first set out to determine if there was a difference in communication with a teammate depending on the type of entity that the participant was

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
DEC 2008	**N/A**	**-**

4. TITLE AND SUBTITLE	
Natural Speech Toward Humans And Intelligent Agents During A Simulated Search And Rescue Mission	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Central Florida Orlando, FL, 32816**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002187. Proceedings of the Army Science Conference (26th) Held in Orlando, Florida on 1-4 December 2008

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**7**	

working with; either human or artificial intelligence. To this end, we investigated the natural speech used during communication by assessing fundamental sound characteristics such as pitch and intensity, as well as exactly what participants said to their teammates (e.g., was a command given or a question asked?). We also measured attitudes about the team's performance, the teammate's helpfulness during the mission, etc., because it has been shown that communication to robots may be limited if attitudes and emotions towards these entities are negative (Nomura et al., 2006).

In our study, each participant remotely controlled a rescue vehicle and worked with a teammate who was controlling another robotic vehicle during a simulated search and rescue mission. Participants were either told that their teammate was either another human or "artificial intelligence software" running on a computer in another room. Besides examining the impact of teammate type, we were also interested in examining whether the type of information shared between teammates during the mission affected communication between teammates. During the mission, outlines of the walls of the buildings that the teammate's robotic vehicle "found" were either made visible or not to the participant. The targets "uncovered" during the mission were always shown regardless. It was hypothesized that natural language communication would vary as a function of type of teammate and type of information shared.

3. METHOD

3.1 Setup of Equipment

The setup required the use of three computers in two separate rooms. In the first room, the participants sat at a chair facing a computer monitor. The participants were given a wireless keyboard only when it was necessary, as to ensure they would not move the various windows displayed on the computer monitor (See Figure 1 for a snapshot of the participant's monitor). During the entire experiment, the researcher was able to watch the participants' monitor through the use of a dual-output video card that allowed simultaneous display of identical information on two different monitors. The participants were given a joystick to control their "vehicles" during the mission. An external microphone was provided to the participants to communicate with their teammates. In a second room, a confederate maintained two additional computers used to run the simulation program and provide simulated communication from the "teammate."

3.2 Search and Rescue Mission Software

A simulation program was specifically designed for this study to represent a search and rescue situation in a collapsed building. There was a blank square map area that represented the area that the participants needed to search (Figure 3 is a screenshot of the map of the buildings and where the targets were situated in Mission 1). Up was considered north on the map, with down being south, right being east, and left being west. When the mission began the map was blank except for two "vehicles" in the opposite bottom or "southern" corners. A vehicle was represented by a small colored-triangle, and target people were represented as blue dots. There was a gray circle, or "light halo" that extended around the vehicle that allowed for targets to be illuminated and found. As the vehicle moved, the "halo" illuminated the "walls" of the building that it came in contact with, and made them visible on the map. One vehicle was operated by the participants. The participants were told that their teammate (that was either "artificial intelligence software" or human) was operating the other vehicle, but it was actually run by a pre-recorded motion file. During the mission participants either saw or did not see the walls of the building that their teammate had found.

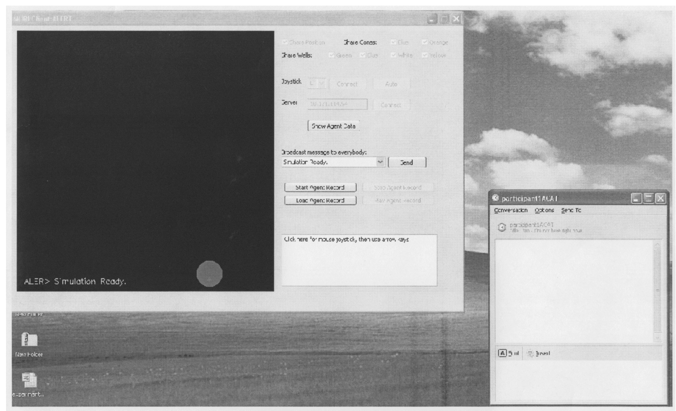

Fig. 1: Screenshot of the participant's monitor.

3.3 Design

The independent variables were teammate type (intelligent agent versus human), information sharing (visible walls versus non-visible walls) and gender (male versus female). The dependent variables were number of targets found, amount of times the participants talked to the teammate and answers on a post questionnaire about the experience with the teammate.

3.4 Procedure

Forty-five undergraduates (33 males, 12 females) at a large state university located in the southeastern U.S. participated. Participants were recruited from two sources: the psychology department's research participant pool (20 participants), and an undergraduate engineering course (25 participants). Participants' ages ranged between 18 and 35 (M =21.13 , SD =3.41).

Participants were randomly assigned to one of four conditions: Intelligent Agent Shared Wall, Intelligent

Agent No Wall, Human Shared Wall, and Human No Wall. See Figure 2 for comparison of views in the No Wall versus Wall condition. Each participant was run through the experiment one at a time although they were told there was another participant present in the experiment in the "human teammate" conditions.

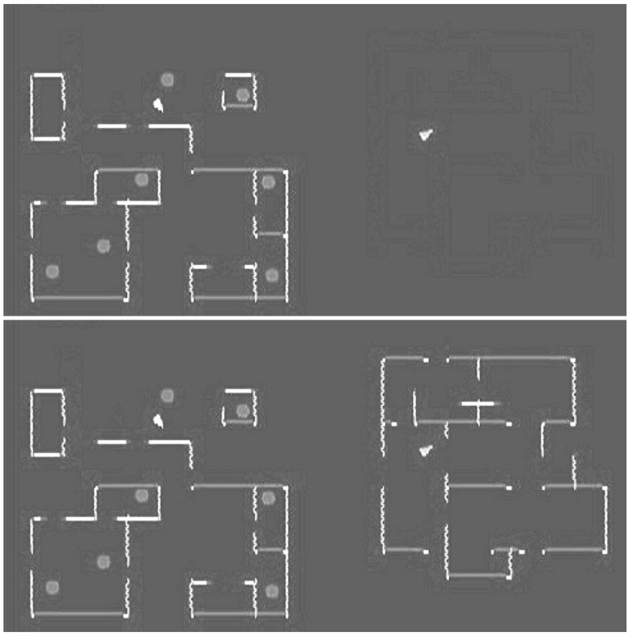

Fig 2.: (Top Left) Participant's teammate in the "No Wall" condition (Bottom Left) Same view of teammate in "Wall" Condition.

Once the participants arrived they were given an informed consent form as well as a demographics questionnaire. An experimenter read the participants detailed instructions regarding the experiment from a script to ensure standardization. The participants were told about the details of the mission they would be performing, and were informed they would be working with either a human or artificial intelligence teammate. Further, participants were also told that they would or would not be able to see the walls of the buildings that their teammate had found.

The participants were told that one of the purposes of the experiment was to investigate how teammates naturally communicate with each other during a search and rescue mission. They were further told that this information would help to develop future communication systems, and that they should feel free to speak naturally into a microphone during the mission. The participants' voices were then calibrated for the voice-recognition program by having them read a short speech into the microphone. Afterwards the voice-recognition software was tested by the participants. The participants stated aloud, "I am now testing the voice software", then waited for it to be transcribed into the instant message window on the screen. The participants then pushed enter on a keyboard to send the text message to their partner.

The participants were instructed that they must work efficiently with their teammate to find people who have been trapped in a collapsed building. These people were designated by blue dots (targets) on the map. The participants were further told that they would receive automatic alerts from the simulation program on the map screen during the mission and that only they would see these alerts. They were informed that it was important to share this information with their partners, as their partners were not aware of this information. The procedure required two experimenters to properly set it up and run it. Experimenter 1 worked with the participants, calibrated the supporting equipment, read instructions and told the participants how to use the joystick. The joystick controlled the participants' vehicle that started in the southwest portion of the screen/map. Experimenter 2 was in another room and set up a pre-recorded path for the "teammate" vehicle to follow. The map that the participants viewed during the mission was in the upper left corner of the display monitor, with the instant message program open in the bottom right corner (see Figure 1).

3.5 Practice Mission

In order to familiarize the participants with the software program, joystick, and speech recognition software, a practice mission was first run. The practice mission was a short mission in which the participants searched for target people in the form of blue dots. The participants were told that there were 11 people to be found, five by themselves, and six by the teammate. The voice recognition and instant message software was active during the entire mission and as soon as the "teammate" found its first target, the instant message, "I have found one person, how many have you found?" was sent. This message was sent to encourage the participant to speak to the teammate and become accustomed to the process of sending messages. After the participant successfully found all of their target people, an instant message was sent saying "I am now ready for the next real mission."

3.6 Mission 1

Participants were told that there were twelve targets (i.e., people) to find in a collapsed building (See Figure 3 for location of targets). They needed to work efficiently with their teammate as there was limited time to find all of the targets. They were shown a rough map of the area that indicated that there were four buildings, one in each corner. The participants' vehicle was in the southwest area whereas the teammate was in southeast. An audio recorder was then turned on. When the participants were ready they said "start", sent the message then used the joystick to get started.

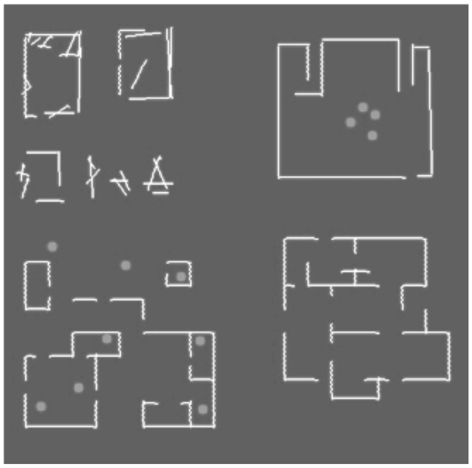

Fig. 3: Fully illuminated map of the buildings and uncovered targets in Mission 1.

The mission was five minutes in length. After one minute had elapsed, an alert message was sent to the participants from the teammate that there were no people in the southeast building. At two minutes an alert message was sent from the teammate telling the participants that data suggested there are four people in the northwest building. If the participants had not spoken to the teammate by this time, another message was sent by the teammate asking how many people had been found by the participants. Participants were then sent an alert message when there were two minutes, one minute and 30 seconds left in the mission. When the participant communicated with the partner, standard responses were sent back by the "teammate" or Experimenter 2 as an instant message. These included such responses as "I'm searching," "my vehicle is functioning properly," and "OK" depending on the specific messages sent by the participants. While the teammate sometimes needed to send different messages in response to the various messages sent by the participants, the teammate's vehicle always followed the same course and did not alter its course to comply with any of the directions given by the participants.

After the mission was over the experimenter gave the participants a post-experimental questionnaire about how they felt about their teammate's performance, their willingness to work with the teammate again and other performance-based questions. Finally, the participants were given a debriefing form and allowed to ask any questions they had about the experiment.

4. RESULTS

The data were analyzed using a series of 2 (Teammate: Human or Intelligent Agent) x 2 (Information Sharing: Walls or No Walls) x 2 (Participant Gender: Male or Female) between-subjects ANOVAs. Two different coders listened to and transcribed the audio recordings of what participants said. Inter-rater reliability was high for both total number of words spoken ($r(43)=.998$, $p<.001$) and for total number of times that the participant spoke ($r(43)=.985$, $p<.001$). Two additional coders watched the video recordings of the missions and assessed how many targets were found. Inter-rater reliability was also high for these two coders ($r(40)=.987$, $p<.001$).

4.1 Targets

The 2(Teammate) x 2(Information Shared) x 2(Gender) ANOVA using number of targets found as the dependent variable revealed a main effect for gender, $F(1,34)=4.77$, $p=.036$. Females found fewer targets ($M=6.65$, $SE=0.91$) than males ($M=8.83$, $SE=0.41$).

4.2 Amount Spoken

The 2(Teammate) x 2(Information Shared) x 2(Gender) ANOVA using amount of times the participants spoke as the dependent variable found a main effect for type of teammate, $F(1,37)=7.09$, $p=.011$. Participants spoke more to non-human intelligent agents ($M=12.75$, $SE=1.56$) than they spoke to humans ($M=7.38$, $SE=1.28$). In addition to the main effect for type of teammate, the interaction between teammate and information shared was also significant, $F(1,37)=7.43$, $p=.01$. This interaction effect was further analyzed using a simple main effects analysis. As can be seen by examining Figure 4, the amount of times the participants spoke was higher for an intelligent agent teammate than for a human teammate only when the information about the walls uncovered by the teammate was available to the participant ($F(1,37)=10.40$, $p=.003$).

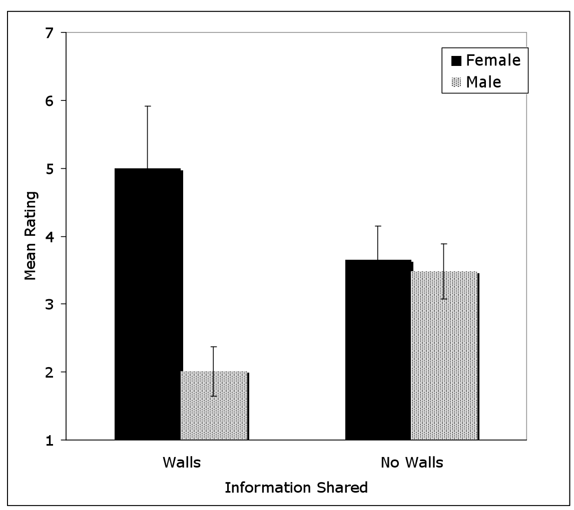

Fig. 4: Number of times the participant spoke as a function of information shared and teammate type

4.3 Pre Questionnaire

There were significant differences for gender regarding questions about video game expertise and experience playing video games. Males (*M*=5.77, *SE*=.26) rated their enjoyment of video games higher than females (*M*=4.50, *SE*=.31), $t(43) = 2.71$, $p = .01$. Males (*M*=5.06, *SE*=.29) also rated their own video game expertise higher than females (*M*=2.75, *SE*=.39), $t(43) = 4.32$, $p < .001$; males (*M*=5.24, *SE*=.30) agreed that they spent more time playing video games as a child than females (*M*=3.42, *SE*=.31), $t(43) = 3.434$, $p = .001$; males (*M*=4.67, *SE*=.33) agreed that they spent more time than females (*M*=2.58, *SE*=.47) playing games now, $t(43) = 3.422$, $p = .001$; males (*M*=5.91, *SE*=1.28) reported playing video games for more hours per week than females (*M*=1.15, *SE*=.43), $t(43) = 2.208$, $p = .033$; males (*M*=12.80, *SE*=.89) reported playing video games for a greater number of years than females (*M*=7.05, *SE*=1.57), $t(41) = 3.137$, $p = .003$; and males (*M*=6.93, *SE*=.55) reported starting to play video games at a much earlier age than females (*M*=10.23, *SE*=1.23), $t(37)=2.83$, $p=.007$.

4.4 Post Questionnaire

When participants were asked to rate their agreement with the statement, “My teammate is smart” on a 7-point Likert scale (1=”Strongly Disagree” and 7=”Strongly Agree”), there was a main effect for gender, $F(1,36)=7.21$, $p=.011$. Females gave higher ratings (*M*=4.32, *SE*=0.52) than males (*M*=2.75, *SE*=0.27). There was also an interaction between information shared and gender, $F(1,36)=5.73$, $p=.022$). This interaction effect was further analyzed using a simple main effects analysis. As can be seen by looking at Figure 5, the mean rating was higher for females than for males only when the wall information was shared ($F(1,36)=9.22$, $p=.004$).

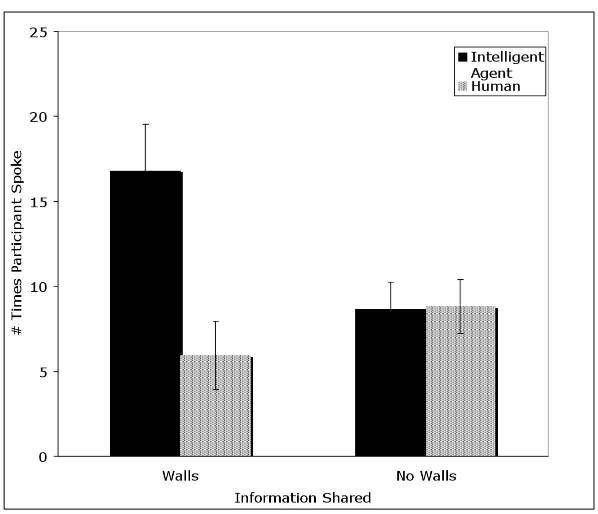

Fig. 5: Mean “smart” rating as a function of information shared and gender

4.5 Vocal Analysis

Pitch was analyzed using a vocal analysis software program called PRAAT (Boersma & Weenink, 2005). PRAAT was used to assess minimum, maximum, mean and standard deviation of pitch (measured in Hz) that were then used as dependent variables. A series of 2 (Teammate: Human or Intelligent Agent) x 2 (Information Sharing: Walls or No Walls) x 2 (Participant Gender: Male or Female) between-subjects ANOVAs revealed only two main effects for Gender when maximum pitch and mean pitch were used as dependent variables. For maximum pitch, females (*M*=318.56, *SE*=22.12) were higher than males (*M*=215.66, *SE*=12.16), $F(1,34)=16.63$, $p<.001$. For mean pitch, females (*M*=213.05, *SE*=13.27) were higher than males (*M*=151.01, *SE*=7.29), $F(1,34)=16.79$, $p<.001$. These main effects were not considered further due to the natural pitch range differences between males and females (pitch range for males is 75-300 Hz and 100-500 Hz for females).

5. DISCUSSION

This research provides support for the idea that working with a human partner may be different than working with a non-human intelligent agent or robot partner. The participant may give different attributions to a non-human intelligent agent than a human, such as intelligence or competence. Indeed, research has found that even when robots display similar “personalities” as their human counterparts, people still feel that the robot is not human-like (Woods, Dautenhahn, Kaouri, te Boekhorst, & Koay, 2005). Further, the amount of information shared with the participant influences the amount of communication that exists.

One of the most interesting findings of this study is the difference in behavior of participants toward their teammates when information was shared. In the No Wall conditions, participants saw their teammates move through a blank area for the duration of the experiment without ever seeing what their teammates were uncovering. In the Wall conditions, participants saw their teammates illuminate the walls of the building, and then could see that it was empty as their teammates continued to navigate back and forth through it for the duration of the experiment. This manipulation influenced how smart the participants felt their teammates were, as well as how much they spoke during the mission. It was found that participants spoke to non-human intelligent agents significantly more than humans. Further investigation indicates that there is an interaction between the wall condition and the amount of speaking. When the participants saw their teammates navigating in blank space and could not see any walls or guidance, they spoke to both human and non-human intelligent agent teammates an approximately equal number of times. In this case they had no way to gauge what their partners were doing and may not have realized that they kept going over the same path numerous times. When the participants could see that their teammates were navigating back and forth through an empty building (with the walls visible) the amount of communication toward non-human intelligent agents was significantly higher than the amount of communication toward humans. A possible reason for this is that the participants attributed thought and reasoning to a human as a teammate, but not to artificial intelligence. They may believe that human teammates have reasons for repeatedly navigating through the same building, and they do not communicate with them or direct them as much as they do non-human teammates. With a non-human intelligent agent they may believe that something is wrong with it, that it does not have an understanding of what they are saying (so they repeat statements) or that it does not understand how to complete the task.

The shared information manipulation also indicated a difference in agreement with the statement "My teammate is smart." When asked if the teammate was smart, men rated their agreement with the statement lower than women did when information was shared. When no information was shared both males and females gave similar ratings. These findings suggest that it is important in a search and rescue mission situation that as much information is shared as possible. Further, people may be more apt to think less of a non-human intelligent agent or put different attributions to its actions. Therefore it is important to give the partner as much information as possible so that they can understand what a non-human partner is doing. This may lead to higher cooperation and more trust in a robot or artificial intelligence teammate.

Males and females may have different experiences in working with teammates to achieve a goal in a computer or video game setting, and may then form an opinion about their partner's abilities based on these experiences. Gender differences also existed in the amount of targets found in the mission. We speculate that greater experience and enjoyment of video games may be the reason that males found more targets. Along the same lines, because males were more familiar with video games, they may have needed to use less "cognitive resources" when using the joystick to control their vehicle, and that may have made it easier for them to find the targets during the simulated search and rescue mission.

This study indicates that there are differences in the amount of communication that is given to a non-human intelligent agent partner, and that the sharing of information can change perception and trust in a teammate. It is important to investigate how and when the communication given to artificial intelligence and robots differs than that given to humans. The next step in this research is to determine the type of things that are said to robots and non-human intelligent agents in comparison to humans. It appears that when given information about what their partner is doing participants treat non-human intelligent agents and humans differently. If we can identify how and why they do this, we can help to design efficient protocols as well as training procedures for working with non-human intelligent agent and robot teammates. This has important implications for any field that depends on humans and robots to work together as a team. Through this and future research we are working toward improving the quality and benefits of working with artificially intelligent teammates in a variety of tasks.

CONCLUSIONS

This study indicates that the amount of communication with an artificially intelligent partner is different than of humans. In cases when more information was given about a non-human intelligent agents performance there was an increased amount of communication. This suggests that the types of information provided to a partner can greatly impact their interaction with the non-human partner. Further, perceptions of artificially intelligent partners were found to be different than that of humans. These differences in communication and perceptions should be taken into account when partnering humans with non-human intelligent agents in real world military situations.

ACKNOWLEDGMENT

This research was supported by the Institute for Simulation and Training and the U.S. Army Research Laboratory.

REFERENCES

Aubert, B. and Kelsey, B., 2003: Further understanding of trust and performance in virtual teams. *Small Group Research*, **34**, 575-618.

Boersma, P. and Weenink, D., 2005: Praat: Doing phonetics by computer (Version 4.3.14) [Computer Program]. Retrieved December 18th, 2007, from http://praat.org/

Burke, J.L., Murphy, R.R., Coovert, and Riddle, D.L. 2004: Moonlight in Miami: A field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise. *Human-Computer Interaction*, **19**, 85-116.

Cuevas, H., Fiore, S., Caldwell, B. and Strater, L., 2007: Augmenting team cognition in human-automation teams performing in complex operational environments. *Aviation, Space, and Environmental Medicine*, **78**, B63-B70.

Hinds, P., Roberts, T. and Jones, H., 2004: Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human-Computer Interaction*, **19**, 151-181.
Huttenrauch, H. and Eklundh, K.S., 2006: To help or not to help a service robot: Bystander intervention as a resource in human-robot collaboration. *Interaction Studies*, **7(3)**, 455-477.
Imai, M., Hiraki, K., Miyasato, T., Nakatsu, R. and Anzai, Y., 2003: Interaction with robots: Physical constraints on the interpretation of demonstrative pronouns. *International Journal of Human-Computer Interaction*, **16(2)**, 367-384.
Nomura, T., Suzuki, T., Kanda, T. and Kato, K., 2006: Measurement of negative attitudes towards robots. *Interaction Studies,* **7**, 437-454.
Woods, S., Dautenhahn, K., Kaouri, C., te Boekhorst, R. and Koay, K.L., 2005: Is this robot like me? Links between human and robot personality traits. *Proceedings of 2005 5th IEEE-RAS International Conference on Humanoid Robots.*