Communication between Teammates in Urban Search and Rescue

by

Cade Earl Bartlett

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Masters of Science

Approved April 2015 by the
Graduate Supervisory Committee:

Nancy Cooke, Chair
Subbarao Kambhampati
Bing Wu

ARIZONA STATE UNIVERSITY

May 2015

ABSTRACT

Although current urban search and rescue (USAR) robots are little more than remotely controlled cameras, the end goal is for them to work alongside humans as trusted teammates. Natural language communications and performance data are collected as a team of humans works to carry out a simulated search and rescue task in an uncertain virtual environment. Conditions are tested emulating a remotely controlled robot versus an intelligent one. Differences in performance, situation awareness, trust, workload, and communications are measured. The Intelligent robot condition resulted in higher levels of performance and operator situation awareness (SA).

TABLE OF CONTENTS

Page

APPENDIX Page

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Robots are relatively recent additions to urban search and rescue (USAR) efforts. They have the potential to be useful for tireless searching, positioning sensors, assessing damage, providing survivors with radio transmitters or supplies, guiding tool placement, and determining survivor position and location under rubble (Murphy, 2000). USAR robots are particularly important in exploring "voids", which are openings in rubble piles which are normally very hazardous to human rescuers. However, current robots are essentially manually controlled extensions of their human operators (Murphy, 2004), and are more of an information source than anything else (Murphy & Burke, 2005). The earliest robots used for USAR use were essentially controllable cameras (Casper & Murphy, 2003), and though their added visual information was valuable, about half of the operator's time, both then and now, is spent building situation awareness (SA) from the video feed and other available information and not on the search task itself (Burke, Murphy, Coovert, & Riddle, 2004; Murphy, 2004; Murphy & Burke, 2005). This is obviously not ideal given that every second spent learning the environment is a second longer that someone spends dying in the rubble. Given this fact, current USAR robots are not adequate for first responder use. Developments in usability, automation, and artificial intelligence may change this, but the robot must perform without error or it could damage itself or harm the people it is meant to help (Yanco & Drury, 2004).

One study on USAR robot autonomy by Yanco (2004) showed that providing robots with enhanced automation allowed the operator to pay more attention to the robot feed. However, the operator also paid less attention to location and orientation, and felt lost if they had to resume manual control. This effect can be reduced by increasing the number of operators and pooling the robots under their command (Gao, Cummings, & Bertuccelli, 2012), or by selecting more capable operators. However, more manpower and higher skill is contrary to the desired end goal of USAR robotics, which is to allow a small number of relatively untrained operators to control a large number of USAR robots to effectively complete their tasks. Research has shown that increasing the robot's autonomy can decrease the operator's workload and resource usage in some cases (Sellers, Fincannon, & Jentsch, 2012), making this goal a possibility. It is possible that the increase in autonomy would negate the need for the operator to take manual controls, which in turn would decrease the demand on operator SA.

It is also important that USAR robots be able to work compatibly alongside humans. Improving autonomy alone may not be enough to accomplish this goal. The ideal system would have to be able to work seamlessly with human responders, and communicate with them in a way that is efficient, complete, and effective (Adams, 2005). In other words, the robot must be an effective teammate when placed on a team with humans. Robot agents need to share their SA, goals, as well as commitment to those goals and plans to achieve them. They should also be aware of their own capabilities and be able to self-assess their own and their teammate's progress. They should be able to interact with human language, as well as understand, respond to, and use standard arm and hand gestures (Salas, Fiore, & Letsky, 2013). On top of the explicit communications humans use for such tasks, the interaction would not be complete without implicit cues, which play an important role in time-pressured situations (Shah & Breazeal, 2010). Search and rescue environments are variable and highly stressful, so the robot must be able to quickly and effectively communicate and adapt to changes in plans and the environment

(Cantrell, et al., 2012). They must also be flexible enough to skip goals, deal with open-ended goals such as not knowing how many survivors there are, and allowing only partial goal completion when dealing with time-sensitive situations (Kambhampati & Cooke, 2012). Human communications are also frequently ambiguous, incomplete, or dependent upon building a partial meaning of something as it is being said (Scheutz & Eberhard, 2008). The ability to understand a human might also improve its functionality (Wiltshire, Barber, & Fiore, 2013). Although addressing these issues will effectively integrate the robot into a human team and allow it to perform independent of operator situation awareness, it is important to remember that it will likely still emulate many human team failings that are inherent consequences of how human USAR teams work together.

The human's trust in the robot is an important factor in how the robot is used on the team, and levels of trust can be easily influenced. Perhaps the most obvious factor in the operator's real-time trust is the robot's real-time reliability (Desai M. K., 2013). Another common factor in trust is the level of feedback that the robot provides to the operator, which is essential in helping the operator to correct the robot (Raman, et al., 2013). The optimal level of trust is when the operator's trust in the robot matches its performance, and one study by Kaniarasu (Kaniarasu, Steinfeld, Desai, & Yanco, 2013) found that when the robot provided feedback on its performance the operator's trust more closely matched its abilities. However, another of their studies (Kaniarasu, Steinfeld, Desai, & Yanco, 2012) shows that too much feedback can result in a strange state of lower operator trust in the robot, while still allowing the robot to have uncorrected control. From this one can infer that it is not the quantity of the feedback, but the quality that properly increases trust in automation. Another study shows that providing reasons for errors or failures in automation increases trust (Desai, et al., 2012). For the development of these USAR systems the exchanges of communications during these errors of why it could not carry out the plan—also termed “excuses” (Göbelbecker, Keller, Eyerich, Brenner, & Nebel, 2010) for the purposes of this experiment—will have to be accounted for in order to build and maintain appropriate levels of trust between the robot and the other human teammates.

In this study it is examined whether an intelligent and independent robot acting as a teammate results in better team performance than a remotely controlled robot that follows an operator’s explicit instructions. This question is explored in the context of an uncertain environment, such as is common in USAR situations. The roles of the human operator and robot searcher are played by a team of two human participants. For the purposes of the experiment, the participant in the operator role is referred to as the “External” participant, because he or she works outside of the virtual environment. Similarly, the participant in the role of the search and rescue robot is referred to as the “Internal” participant, because he or she works inside of the virtual environment. Other important measures are also taken into account such as operator stress and workload, the levels of trust between the operator and the robot, and how SA for the human operator and the robot searcher changes with each condition. The experiment also produced a collection of natural language interactions between two humans carrying out such a search and rescue task. The types of communication used are also compared to performance to see what communication patterns may be most indicative of a high performing team.

Despite the expected reduction in SA in the intelligent condition it was expected that there should be better team performance over the remote condition, as the internal takes more decision making into their own hands, reducing the necessity of the external to know the environment. It is likely that the more pro-active internal participant will result in a lower number of “excuse”-related communications. Due to these things which point to less external participation, the external will probably also report lower workload. It is also expected that, in the intelligent condition, due to the increased flow of communication and feedback from the internal participant that the external participant will show an increase of trust when compared to the remote condition. These effects will likely be small until the conditions can be tested with one operator over multiple intelligent or remotely controlled robots.

METHOD

PARTICIPANTS

Participants were recruited from the Arizona State University introductory psychology participant pool and other students around the Arizona State University Polytechnic campus. The 11 participants from the psychology pool were awarded credit for participation, and 9 other students were given monetary compensation of ten dollars an hour. Differences in performance between paid (M=20.44, SD=8.85) and credited (M=19.73, SD=8.33) participants are non-significant [$F(3, 19) = 0.59$, $p = 0.629$]. All participants were to be familiar with standard PC gaming controls (WASD + mouse). Of all participants, 32 were male and 8 were female. Participant ages were from 18 to 48, with an average age of 23.

TASK

Two human participants interacted to complete a simulated search and rescue task in a virtual environment, acting as a human-robot search team. One participant acted as the external participant who worked outside the search and rescue environment, comparable to the operator, whereas the other participant acted as the internal participant who worked inside the search and rescue environment, comparable to the robot. Both sat in the same room and were able to speak freely to each other, but the external participant was unable to see the computer screen—which displayed the virtual environment—due to a divider which was placed between them. The environment resembled an office structure with interspersed green, blue, and pink blocks that represented potential targets. Blue and green blocks were meant to represent survivors, whereas pink blocks represented hazards. Pressing a button on green blocks counted positively towards the team's performance, whereas pressing a button on pink blocks counted negatively towards the team's performance. Blue blocks were time-sensitive, such that pressing a button on them before eight minutes into the scenario counted positively toward performance, but pressing the button after that time counted against performance. Pressing the button on any block more than once counted negatively towards the team's performance. The environment also contained light switches and power-locked doors. A map of this environment was also made. There were intentionally introduced inconsistencies, such as missing walls, additional walls, and misplaced doorways to simulate a damaged building, none of which were depicted on the map.

MATERIALS

The screen recording software used is Ohsoft's oCam Screen Recorder (ohsoft.net) which recorded the virtual environment as well as the voice communications. The virtual environment (see Appendix O for screenshots of the environment) was created using the popular video game Minecraft from Mojang (mojang.com and minecraft.net), which was chosen for its ease in controllability, flexibility, and variability. A ten-problem box folding test (Appendix B) was also used to measure participant spatial awareness abilities (based on those at http://www.iq-test.com/spatial-ability-test.php). A page of role-specific instructions was made for each participant (appendices C, D, E, and F), in addition to a cheat-sheet summary of those instructions (Appendix G) and an instructional sheet that was given to the external participant after 8 minutes informing them that they were no longer to get the time sensitive targets (Appendix J). The map of the virtual environment (appendices H and I) was made with a few inconsistencies such as missing, blocked, or additional doorways or walls. A questionnaire was used afterward to collect demographics, experience, and self-reported TLX of the workload experienced during the task for each participant (appendices K and L). Consent, debriefing, and payment receipt (where applicable) sheets were also used (appendices A, M and N).

MEASURES

Team performance. The score is measured at the team level, and is calculated from the number of rooms on the map (Appendix I) with the correctly marked numbers and types of boxes plus the number of buttons pressed on blue boxes before eight minutes and green boxes, minus the number of repeated presses on a box or pressing a button on a blue box after eight minutes or a pink box:

(Rooms Marked correctly + Correct Presses) – (Repeated Presses + Incorrect Presses) (1)

There were 18 rooms, and 20 boxes, and no mandatory repeated or incorrect presses, making a maximum possible performance score of 38. The questionnaire (Appendix K) was also used to gather self-reported levels of performance.

Performance Covariates. Different scores were calculated from the spatial ability test (Appendix B) for use as covariates based on how many they got right out of 10, and how many they got wrong out of 10, and while these two account for their number of attempts they may not necessarily add up to 10. Each one conveys valuable information about the team. The team's combined number correct is used as an indicator of overall team spatial ability, ignoring how careful they are being to get the right answer. The team's combined number of incorrect spatial ability answers is their additional, but failed attempts, which is a useful indicator of the team's effort or conscientiousness. The team's accuracy is a percentage of how many they got correct out of their attempts (note: not out of total possible), which is useful for getting a more refined and standardized look at the team's spatial ability, taking conscientiousness into account. The difference in number of incorrect spatial ability answers between teammates conveys how different the team's composition is, relative to spatial ability. The team's employment status, as indicated in the demographics questionnaire (Appendix K), was also used. It is interpreted as it is an indicator of the team experience that is gained from interaction with coworkers. The team's employment score was a 0 if neither were employed, 0.5 if one was employed, or 1 if both were employed.

Situation Awareness - External. The score was determined by the number of rooms on the map (Appendix H) with correctly marked numbers and types of boxes. There were 18 rooms, and therefore a maximum score of 18. The questionnaire (Appendix K) was also used to gather self-reported levels of SA.

Situation Awareness – Internal. The score was determined by the number of repeated presses on boxes. The questionnaire (Appendix K) was also used to gather self-reported levels of SA.

Trust. Eight self-report questions on the questionnaire (Appendix K) are used for an aggregate score.

Workload – TLX. The score was determined with a six-factor NASA TLX (Appendix L) for mental, physical, and temporal load, as well as sentiments of accomplishment, difficulty, and stress.

Communications. Screen and voice recording was used to gather communications data. Voice communications were transcribed to text format and encoded into categories of "question", "acknowledgment", "box location", "direction", "observation", "current action", "ask confirm", "confirm", "SA Description", or "suggestion". Communications could also additionally be marked as "excuse related" (see table 1 for description of coding). Total communications and word counts were extracted from this data.

PROCEDURE

Participant teams perform either the "remote" or "intelligent" conditions of the task. In both conditions, following administration of consent forms, participants began by taking a five minute spatial awareness test. Roles of "Internal" and "External" were determined by randomly selecting the role for the first participant to arrive. The internal

Table 1

Explanation of which communications receive which code

Code	Explanation
Question	Asking a question
Acknowledgment	Acknowledging teammate's communication
Box Location	Communicating the location of a box or where there are no boxes
Direction	Explicitly telling teammate to do something or giving a suggestion with a strong expectation that they act on it
Observation	Make an observation, frequently a verbal communication without much meaning
Current action	Talking about current actions
Ask Confirm	Asking for a confirmation on something
Confirm	Providing a confirmation
SA Description	Describing an aspect of the situation such as the virtual environment, how things appear on the map, or time
Suggestion	Providing a minimally explicit or optional form of direction, typically when helping their teammate do their job
Excuse Related	A flag for whether the communication was related to one of the environment's inconsistencies with the provided map

participant was seated in front of the computer running Minecraft and the screen recording software, whereas the external participant was seated near the internal with a divider placed between them to keep the external from view of the monitor. Both participants were given microphones to record their speech, and both were given instructions and materials for their respective roles according to their specific conditions. In the remote condition only the external participant was given the non-matching map, along with instructions to draw a search plan that they must adhere to, or communicate changes to the plan if they deviate from it. The internal participant was instructed to act only on the external participant's explicit directions as they carry out the task. In the intelligent condition the internal participant independently made a search plan and both the internal and external participants were given a copy of it. Both participants kept the map through the task and were instructed to adhere to it, and similarly communicate any changes to the plan if they deviated from it. The internal participant in the intelligent condition was *not* instructed to only act on the external participant's instructions, allowing him or her to act freely and according to the internal's own judgment.

In both conditions, the experiment started with the instructions to press buttons on any green and blue blocks, and then return to the start and press a button on a red exit block to end the mission. The external participant, in both conditions, was instructed to mark the number of each block color in every room, as well as mark any inconsistencies they noticed. After eight minutes they were instructed to not press buttons on blue blocks. They were instructed to never press buttons on pink boxes. After receiving instructions and creating a plan, the internal participant went through a brief training segment in which they were familiarized with their controls and how to properly interact with the virtual environment. They were then given fifteen minutes to find green blocks, and blue blocks for the first eight minutes, and then exit the building through a red exit block. If they did not use the red block by the end of the fifteen minutes the mission was automatically terminated. The mission was followed by a questionnaire and NASA TLX administration. Participants were then debriefed, compensated accordingly, and dismissed.

RESULTS

Data were analyzed as a 2 x 2 ANOVA design (the two conditions of intelligent and remote were between teams and internal compared to external was within teams). Some data were only useful at the team level and were analyzed with t-tests. It was useful to look at communications data both with those methods, and also with basic correlations.

PERFORMANCE

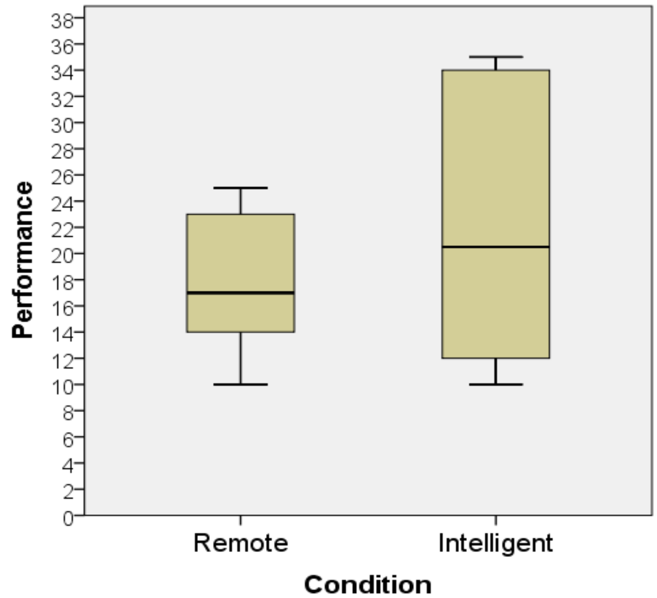


Figure 1 Calculated performance scores between conditions.

A univariate ANOVA was conducted with the five previously mentioned covariates (team correct, team incorrect, team accuracy, and team incorrect difference on the spatial ability test, as well as team employment) to look at the difference in team performance between the remote (M = 17.50, SD = 5.08) and intelligent (M = 22.60, SD = 10.33) conditions. This analysis showed a significant difference at $p < 0.05$ between teams [$F(1, 13) = 7.01$, $p = 0.020$]. The intelligent condition shows a significantly higher performance than the remote condition (figure 1). Performance mean and standard deviation for both combined teams, while not used in this section, should also be noted for future use (M = 20.05, SD = 8.34).

Although the analysis of performance shows a significant result it is important to note that four of the covariates are derived from the same spatial ability data as each other (Appendix B), and are therefore closely related, which may inflate their effect. That being said, the covariates derived from the spatial ability test each convey something different from each other about the team's combined spatial ability.

Also of note, though the overall performance did not appear to have any floor or ceiling effects, a few of its component scores did. These scores were the internal's incorrect and repeated presses. However, for the purposes of this study, the data from the external is the most important.

It is also important to know that a *post-hoc* power analysis revealed that the results for performance were greatly under powered. However, after factoring in the covariates the performance scores still showed a significant difference at an appropriate power at the $p < 0.1$ level (power = 0.805). An alpha of 0.1 was determined appropriate for this task because the Remote condition is actually low-autonomy emulating remote controlled interaction, which would weaken the effect.

SITUATION AWARENESS

A t-test comparing the number of rooms marked correctly on the map was conducted to look at the difference in SA in the external participant between the remote ($M = 6.60$, $SD = 3.66$) and intelligent ($M = 9.00$, $SD = 6.04$) condition. The difference in SA between intelligent and remote conditions was not significant at $p < 0.05$ [$t(14.82) = 1.08$, $p = 0.299$].

A similar t-test analysis was conducted for the internal situation awareness, only using the number of repeated button presses to measure the difference between the remote ($M = 1.5$, $SD = 2.42$) and the intelligent ($M = 1.20$, $SD = 1.62$). This comparison was also non-significant at $p < 0.05$ [$t(18) = -0.33$, $p = 0.748$].

TRUST

2 x 2 ANOVA was used to look at each teammate's trust of each other for intelligent and remote conditions, and within internal and external roles. When comparing how much internal remote ($M = 3.29$, $SD = 0.96$), internal intelligent ($M = 3.04$, $SD = 1.18$), external remote ($M = 3.04$, $SD = 0.65$), and external intelligent ($M = 3.13$, $SD = 0.66$)

trusted their teammate there was no significant difference for either role or condition [$F(1,36) = 0.08$, $p = 0.774$] or for interaction effects [$F(1,36) = 0.36$, $p = 0.552$] at the $p < 0.05$ level.

WORKLOAD

NASA TLX was analyzed both within and between teams with 2 x 2 ANOVA since it is of interest to see how workload changes for both roles between conditions. Workload measures were structured in such a way that a lower number from 0 to 20 was lower workload, while a higher number was a higher workload. A general workload score was calculated by averaging all six individual workload scores together. This general workload score was used for this analysis. When comparing workload for internal remote ($M = 7.03$, $SD = 2.15$), internal intelligent ($M = 6.98$, $SD = 2.84$), external remote ($M = 9.75$, $SD = 2.54$), and external intelligent ($M = 5.88$, $SD = 2.68$) there was no significance between roles [$F(1,36) = 0.99$, $p = 0.325$], but there was a significant effect between conditions [$F(1,36) = 5.83$, $p = 0.021$]. There is also a significant interaction between role and condition on workload [$F(1,36) = 5.54$, $p = 0.024$]. The individual workload differences within roles was dependent on the condition where higher workload was experienced by the external in the remote condition and by

Figure 2 Interaction for workload TLX measures by role as it changes between conditions.

the internal in the intelligent condition. The internal experienced about equal workload between conditions (Figure 2).

COMMUNICATIONS

Communications were transcribed and also analyzed within and between teams with 2 x 2 ANOVA for which conditions and roles used which communications more. The comparison between roles shows the direction of feedback flow for each condition, with more communications from one participant representing a greater flow of feedback from them. Special attention is given to the analysis of "Excuse related" communications measures and how prevalent they are in more or less successful teams.

Words Spoken. 2 x 2 ANOVA shows that even though there was a large difference between the average number of words spoken in the Intelligent condition between the internal (*M* = 656.9, *SD* = 386.4) and external (*M* = 230.5, *SD* = 165.8), combined with the lack of difference in the Remote condition between the internal (*M* = 827.7, *SD* = 286.0) and external (*M* = 937.7, *SD* = 243.1) the number of words spoken over the course of the task was not significantly dependent upon the role [$F(1,36) = 3.151$, $p = 0.084$]. However, the number of words spoken over the course of the task was very significantly dependent upon the condition [$F(1,36) = 24.263$, $p = 0.000$]. The number of words spoken over the course of the task was also dependent on an interaction between condition and role [$F(1,36) = 9.056$, $p = 0.005$].

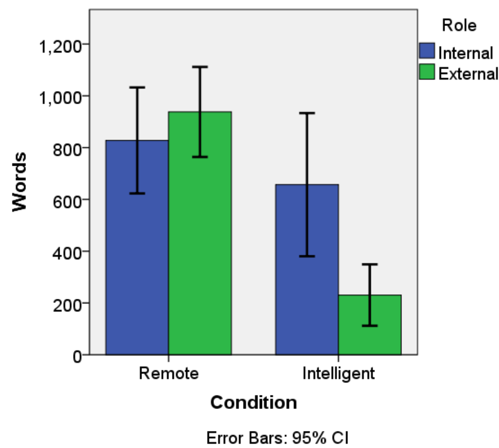


Figure 3 Comparison of word volume between and within conditions.

Communication Difference Between Conditions. Given that there is such a large difference in the number of words between conditions, the difference in types of communications between these conditions is looked at in terms of percentage of each type of communication out of the total number (figure 4). Independent samples t-tests showed the only communications differences that were significant at $p < 0.05$ were for box locations between Remote (M = 0.106, SD = 0.17) and Intelligent (M = 0.225, SD = 0.68) [$t(10.10) = 5.38$, p=0.000], directions between Remote (M = 0.194, SD = 0.090) and

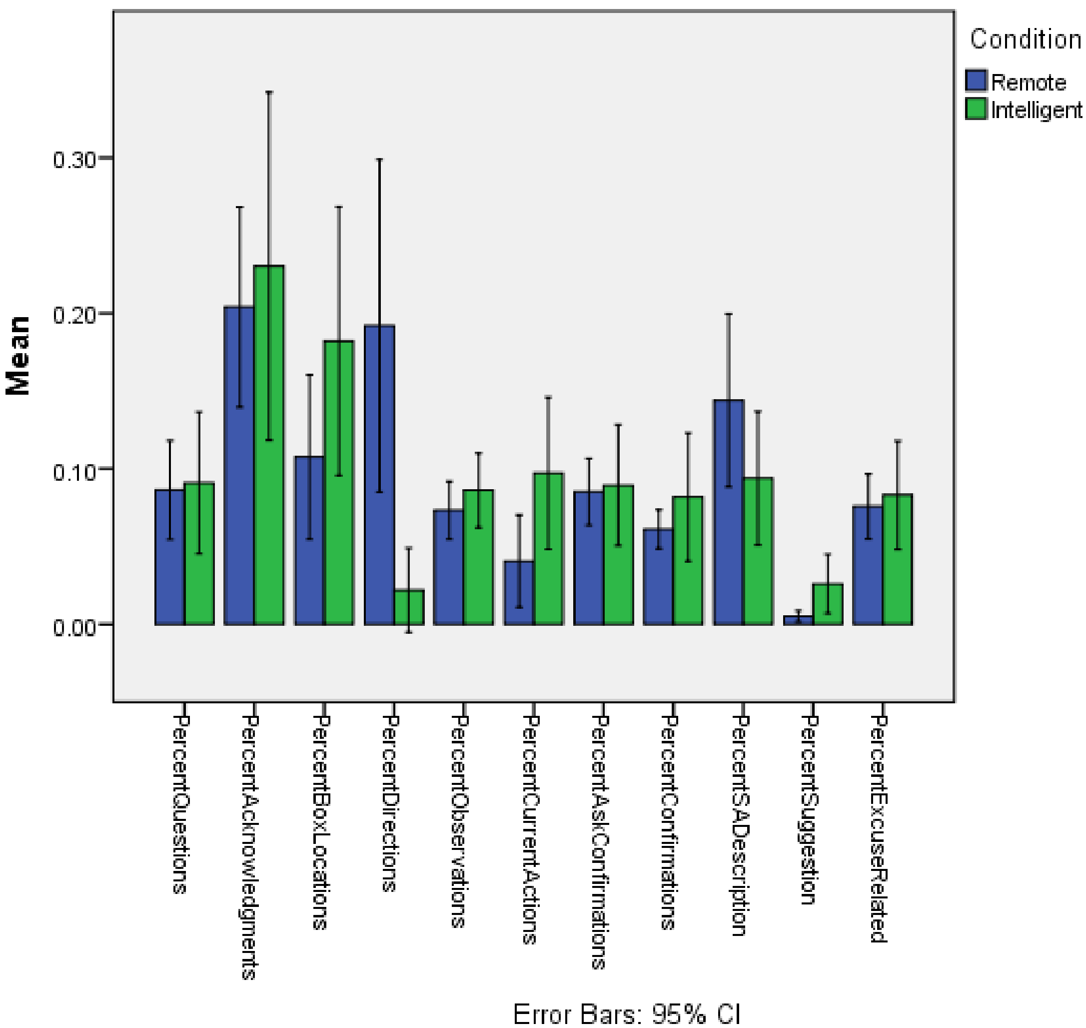
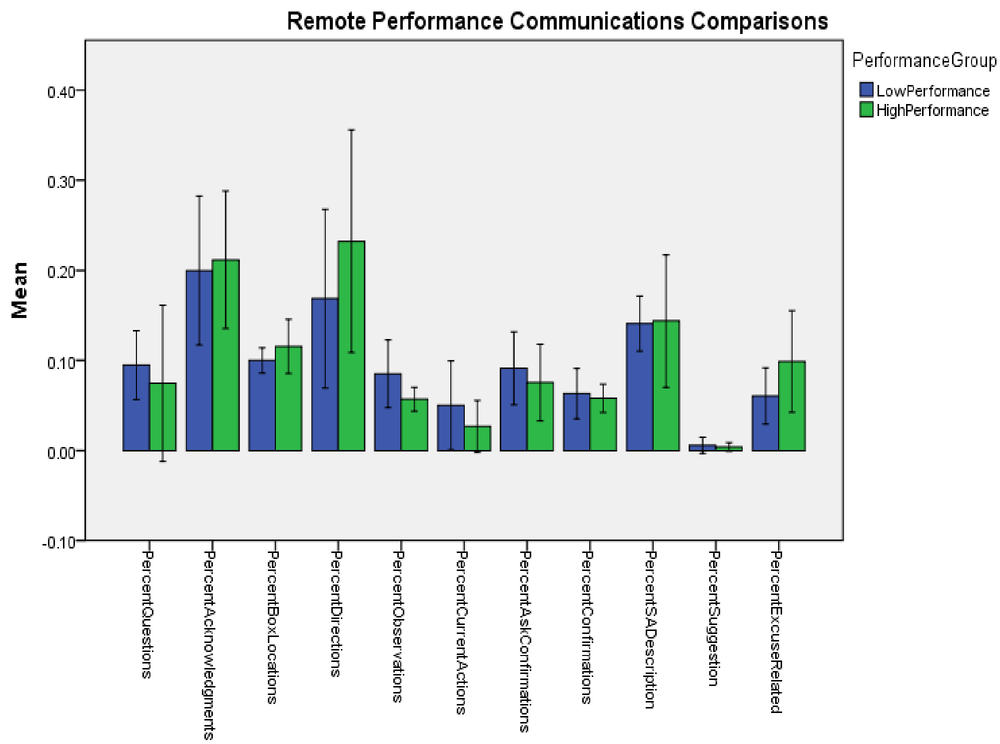


Figure 4 Comparison of coded communications between conditions, by percentage. Each communication was given a 1 if the communication matched its code category. Excuse Related was a separate code that could be flagged in addition to one of the other codes.

Intelligent (*M* = 0.022, *SD* = 0.044) [$t(18) = 5.45$, $p = 0.000$], current actions between Remote (*M* = 0.041, *SD* = 0.039) and Intelligent (*M* = 0.119, *SD* = 0.047) [$t(18) = 4.09$, $p = 0.001$], and suggestions between Remote (*M* = 0.005, *SD* = 0.007) and Intelligent (*M* = 0.020, *SD* = 0.016) [$t(11.91) = 2.59$, $p = 0.024$).

Communications and Performance. A correlation analysis of each communication percentage with performance was performed. At the $p < 0.05$ level, the only significant correlation with the combined performance score was the percentage of excuse related communications (*M* = 0.086, *SD* = 0.051) [$r(18) = 0.639$, $p = 0.002$]. Correlations were broken down further by condition. For the Intelligent condition correlations to performance, percentage of excuse related communications (*M* = 0.097, *SD* = 0.062) remained the only significant correlation [$r(8) = 0.71$, $p = 0.021$]. However, in the Remote condition's correlations to performance the significance of the correlation to percent of excuse related communications (*M* = 0.076, *SD* = 0.036) disappears [$r(8) = 0.297$, $p = 0.404$], and instead has a significant negative correlation to percent of observation communications (*M* = 0.74, *SD* = 0.031) [$r(8) = -0.691$, $p = 0.027$). The twenty teams were divided into two categories of high performance and low performance, divided across the median. When compared in this way with a t-test the only significant difference in communication percentage between low (*M* = 0.060, *SD* = 0.025) and high (M = 0.113, *SD* = 0.057) performing teams at the $p < 0.05$ level was for excuse related communications [$t(18) = 2.74$, $p = 0.014$]. However, given the many varieties of communications people can use, it might also be worth mentioning that the percentage of questions asked by low (*M* = 0.101, *SD* = 0.055) and high (*M* = 0.064, *SD* = 0.037) performing teams almost had a significant difference [$t(18) = 1.83$, $p = 0.085$] at



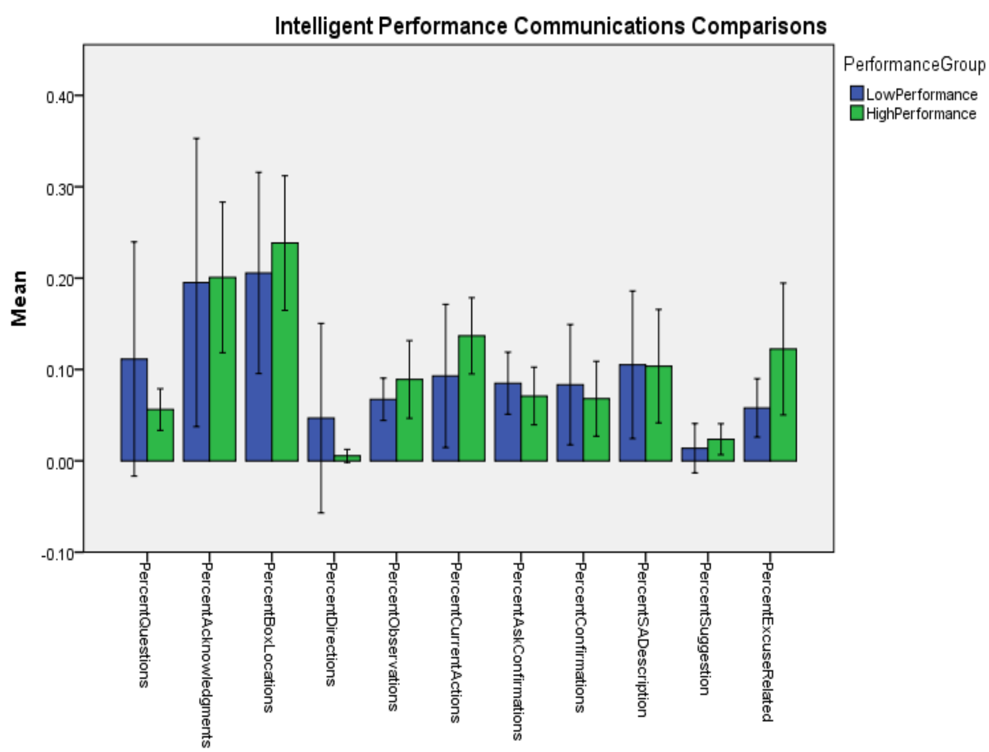


Figure 6 Comparison of low to high performing groups. All teams are first divided into the top and bottom 10 performing times, and then again into their conditions. Remote condition comparing 6 low and 4 high is seen on the top. Intelligent condition comparing 4 low and 6 high on the bottom.

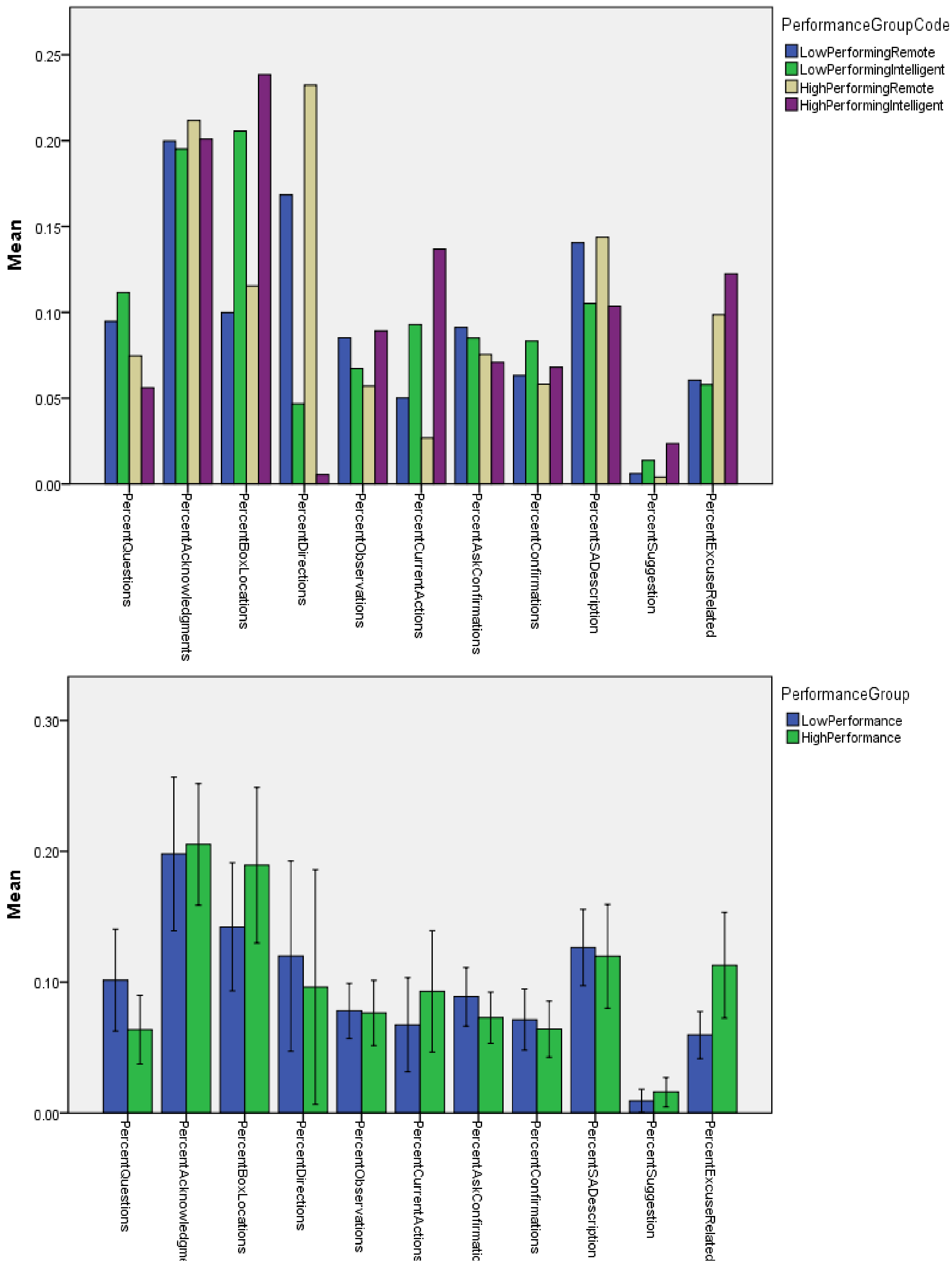


Figure 5 Comparisons of percentages of communication by condition and performance grouping. The top graph splits the 20 teams into the top and bottom 10 performing teams, and further divides them by condition. The bottom graph only compares the top and bottom 10 performing teams, regardless of condition.

the $p < 0.05$ level (figure 5).

These high and low performing groups were further divided into their conditions. The low performing group had six Remote teams and four Intelligent teams, while the high performing group had four Remote teams and Six Intelligent teams. A t-test comparison of percentages of communications did not show any statistical significance between the high and low performing groups of each condition at the $p < 0.05$ level. The closest was the comparison of low ($M = 0.058$, $SD = 0.020$) and high ($M = 0.122$, $SD = 0.069$) performing Intelligent groups for the percentage of excuse related communications [$t(8) = -1.79$, $p = 0.072$] (figure 6).

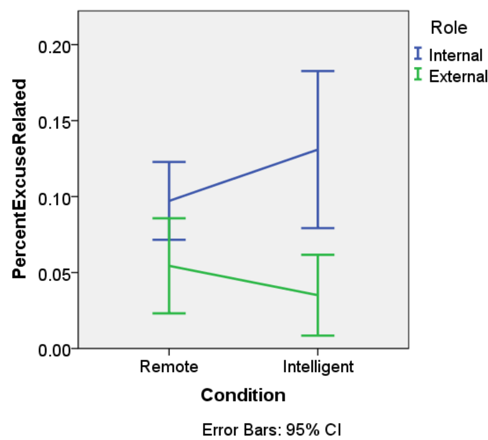


Figure 7 Comparison of percentage of excuse related communications between and within conditions.

Excuse Related Communications Between Conditions. Differences in percentage of excuse related communications between the Remote condition for the internal ($M = 0.097$, $SD = 0.036$) and external ($M = 0.054$, $SD = 0.044$) and the Intelligent condition for the internal ($M = 0.131$, $SD = 0.072$) and external ($M = 0.035$, $SD = 0.037$) using 2 x 2 ANOVA at the $p < 0.05$ level were not found to be significantly dependent upon condition [$F(1, 36) = 0.213$, $p = 0.647$] or the interaction between role and condition [$F(1, 36) = 2.88$, $p = 0.099$] but it was found to be significantly dependent on role alone [$F(1, 36) = 19.60$, $p = 0.000$] (figure 7).

Communications and Trust. Correlation analysis of each individual's self-reported trust score with the number of communications [$r(38) = 0.62$], words [$r(38) = 0.53$], and words per communication [$r(38) = 0.94$] revealed that there was no significant correlation of these self-reported trust measures to volume of communications at the $p < 0.05$ level.

ROOM NUMBERING

Over the course of the experiment, eleven of the twenty external participants were observed to give arbitrary numbers to each of the rooms on the map. Four of these were in the Remote condition, and seven were in the Intelligent condition. Teams that did give room numbers on the map ($M = 9.73$, $SD = 5.50$) were compared to teams that did not give room numbers on the map ($M = 5.44$, $SD = 3.25$) in relation to their external SA with a t-test. While this showed a non-significant difference between the two groups at the $p < 0.05$ level [$t(18) = 2.06$, $p = 0.055$], it did have a significant result for Levene's test of equal variances at the $p = 0.05$ level [$F(1,18) = 9.43$, $p = 0.007$], and without the assumption of equal variances the relationship appears to be significant at $p = 0.05$ [$t(16.57) = 2.16$, $p = 0.045$]. There is a pretty large spread of values for this data (figure 8).

Since the number of rooms correctly marked which is used as the external's SA score is also used to calculate performance it is expected that teams with room numbering behavior ($M = 23.09$, $SD = 9.51$) would have a higher performance than those without ($M = 16.33$, $SD = 4.90$) in a t-test. While their performance is higher

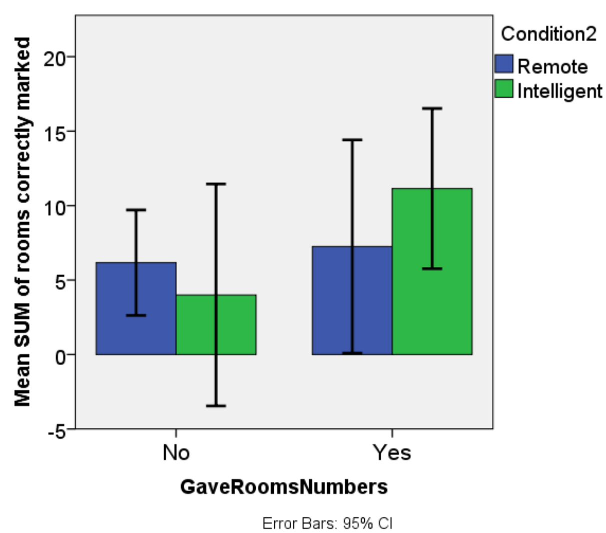


Figure 8 Comparison of the external role's situation awareness measure (SUM of rooms correctly marked) between conditions, based on whether or not the external participant gave numbers to the rooms on the map.

the difference is not significant at $p < 0.05$ [$t(18) = 1.93$, $p = 0.07$], though it is close. Again, Levene's test was significant for this comparison at $p < 0.05$ [$F(1,18) = 6.39$, $p = 0.021$], but the test withiout equal variance assumption was still non-significant at $p < 0.05$ [$t(15.50) = 2.05$, $p = 0.058$], but still very close.

It was also thought that the room numbering behavior might be a driving factor in the communications decrease from the Remote to Intelligent condition, since a room number gives a more concise reference than describing one's surroundings. Number of words spoken in the Remote condition by teams that did not number rooms ($M = 1792.0$, $SD = 476.93$) and teams that did number rooms ($M = 1725.5$, $SD = 287.68$) was compared the number of words spoken in the Intelligent condition by teams that did not number rooms ($M = 1032.0$, $SD = 541.04$) and the teams that did number rooms ($M = 825.43$, $SD = 473.61$) using 2 x 2 ANOVA, looking at significance at the $p < 0.05$ level. As already seen, the number of words spoken was significantly dependent upon condition [$F(1,16) = 14.89$, $p = 0.001$]. However, the number of words spoken was neither significantly dependent upon whether or not they numbered rooms [$F(1,16) = 0.106$, $p = 0.749$] or the interaction between the condition and whether they numbered rooms [$F(1,16) = 0.403$, $p = 0.535$].

DISCUSSION

It was expected that team performance would be superior in the Intelligent conditions, largely because in the Intelligent condition the operator is able to offload SA responsibilities to the "robot," in a sense eliminating the "middle-man". Data show that the Intelligent condition resulted in better performance, as expected, compared to the Remote. Data would also suggest that the "middle-man" of communication was very much eliminated, or reduced in the Intelligent condition, reflected by the great difference in communication volume. As is seen, the Intelligent condition required far less communications than the Remote condition to achieve a higher performance. It was also expected that there would be less excuse-related communication in the intelligent condition due to the independent thinking and acting of the internal participant, and the results showed that, though there was less excuse-related communication over all, by percentage they were about the same.

However, what was not expected was that the external role in the intelligent condition actually had a higher SA than the external in the remote condition. Although the data for this measure at this sample size were statistically non-significant, looking at the wide spread in the SA data may be evidence that the intelligent condition raises the ceiling for the external's situation awareness. However, based on the data from teams that put arbitrary room numbers on the map it would appear that it is quite possible that this increase in SA in the Intelligent condition was largely driven by whether or not the external numbered the rooms when they made the search plan. It would make sense that a team with a shared map with numbered rooms would have an easier way to communicate SA than those without, since they had a simplified way to reference each room. The Intelligent teams that did not use room numbering had a lower average external SA than those that did, and even lower than the Remote condition's external SA. This is more in line with what was hypothesized. However, it should be noted that numbering rooms also matched with higher external SA in the Remote condition, where only the external had the advantage of being able to reference a room by its number on the map. It was thought that this might be reflected in the number of words spoken by teams that did and did not number rooms in the Intelligent condition, since both teammates shared a map and a room number could possibly be used as a more concise reference for SA communications. However, it appears that there was no significant effect on the number of words spoken by these groups. It may just be that whether or not rooms were numbered is a good indicator of individual spatial organization or teamwork. With this in mind, it would be advantageous to the cause of USAR robot autonomy development if increasing autonomy also increased operator situation awareness, or if a method could be derived from this data to negate operator SA losses in situations where they have an independent and autonomous teammate, whether a robot or a human. It might be worth future research consideration, though the effect may level out and perhaps vanish with multiple entities in the robot's role. More participants or greater demands on SA might be sufficient to yield a significant analysis.

Workload is most important for the human operator, and it was expected that the TLX would show a lower external load in the intelligent condition, and a lower internal load in the remote condition, since the external would be doing all of internal's thinking for them. Interestingly, the lowest workloads for both internal and external were in the intelligent condition. It is possible that the actions involved in the internal communicating all the necessary SA information to the external so the external could in turn direct the internal were more stressful than the internal acting independently on retained information. However, for future experiments involving more participants in the internal role with only one operator it may be expected that whether the robots behave autonomously or are all remotely controlled, more robots would likely increase workload.

It was thought that the changes in SA demand and responsibility would drive a change in direction of feedback flow which would, in turn, result in an inverse increase in trust (more internal responsibility = high internal feedback = higher external trust, more external responsibility = high external feedback = higher internal trust). Although there is a difference in means in support of this, the difference is very small, and far from significant. However, trust was only measured through self-report and had a very low power, so either a better trust measure or more participants would be needed to determine how much of a difference these conditions make on trust. Again, it is also possible that these differences will not be overwhelmingly significant until the introduction of more agents.

A number of the communication differences between conditions confirm the effectiveness of the two conditions to emulate a remote controlled robot or an intelligent one. It makes sense that there was a significant difference between percentages of communications between conditions for directions, suggestions, and current actions. All of these can be explained as, and give support to, the effectiveness of the shift of roles with the different conditional instructions given to the participants. As in the Remote condition the internal is instructed to only act on the explicit commands of the external, this behavior would foster the idea that the Remote condition would have more directions, fewer suggestions, and fewer independent reports of current actions, and this is what the statistics show. The difference in box location related communication is likely related to the difference in performance between the two conditions, since the Intelligent condition had a higher performance, which is derived from a greater number of boxes found and reported to the external, which necessitates more box location communications.

What might be more interesting is the difference in communication volume between conditions. Whereas the communications for the internal role are only reduced by about 100 words on average from the Remote to the Intelligent condition, the communications for the external role are reduced by about 700 on average (figure 3). This is combined with the improvement in performance, and the lower external workload scores in the Intelligent condition compared to the Remote. Also of interest is the significant correlation between percentage of excuse-related communications and performance. As with all correlations it is hard to say which one causes the other. Does increasing or ensuring communications in the event of a plan disruption improve performance, or is an increase of these communications merely a side effect of improved performance as the team is able to encounter more roadblocks? It is a question that might be worth future research.

CONCLUSION

Primarily, the study provides a language corpus for teams with different levels of effectiveness that can be modeled in the development of a robot that serves as an effective teammate. The results of the study also provide insight into how much autonomy is necessary for an effective human-robot team in a USAR situation. The Intelligent condition had a higher team performance, a lower workload, and reduced communication volume when compared the Remote condition. Because the Intelligent condition was meant to represent a team with a fully autonomous robot, it would be expected that an increase in USAR robot autonomy would also increase USAR team performance as they conduct real world USAR tasks. For future studies it would be interesting to explore how the two conditions differ with multiple internal participants assigned to one external participant, which would correspond with the end goal of one operator for a team of search and rescue robots.

REFERENCES

Adams, J. A. (2005, September). Human-robot interaction design: Understanding user needs and requirements. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 49, pp. 447-451. SAGE Publications. doi:10.1177/154193120504900349

Burke, J. L., Murphy, R. R., Coovert, M. D., & Riddle, D. L. (2004). Moonlight in Miami: Field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise. *Human–Computer Interaction, 19*(1-2), 85-116.

Cantrell, R., Talamadupula, K., Schermerhorn, P., Benton, J., Kambhampati, S., & Scheutz, M. (2012, March). Tell me when and why to do it! Run-time planner model updates via natural language instruction. *Human-Robot Interaction (HRI). 2012 7th ACM/IEEE International Conference* (pp. 471-478). Boston, MA: IEEE.

Casper, J., & Murphy, R. R. (2003). Human-robot interactions during the robot-assisted urban search and rescue response at the world trade center. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 33*(3), 367-385.

Desai, M. K. (2013). Impact of robot failures and feedback on real-time trust. *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction* (pp. 251-258). IEEE Press.

Desai, M., Medvedev, M., Vázquez, M., McSheehy, S., Gadea-Omelchenko, S., Bruggeman, C., . . . Yanco, H. (2012). Effects of changing reliability on trust of robot systems. *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference* (pp. 73-80). IEEE.

Gao, F., Cummings, M. L., & Bertuccelli, L. F. (2012). Teamwork in controlling multiple robots. *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction* (pp. 81-88). ACM.

Göbelbecker, M., Keller, T., Eyerich, P., Brenner, M., & Nebel, B. (2010, May). Coming Up With Good Excuses: What to do When no Plan Can be Found. *ICAPS*, (pp. 81-88).

Kambhampati, S., & Cooke, N. J. (2012). Planning for Peer-to-Peer Human Robot Teaming in Open Worlds.

Kaniarasu, P., Steinfeld, A., Desai, M., & Yanco, H. (2012). Potential measures for detecting trust changes. *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction* (pp. 241-242). ACM.

Kaniarasu, P., Steinfeld, A., Desai, M., & Yanco, H. (2013). Robot confidence and trust alignment. *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction* (pp. 155-156). IEEE Press.

Murphy, R. R. (2000). Marsupial and shape-shifting robots for urban search and rescue. *Intelligent Systems and their Applications*, *15*(2), 14-19.

Murphy, R. R. (2004). Human-robot interaction in rescue robotics. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 34*(2), 138-153.

Murphy, R. R., & Burke, J. L. (2005). Up from the rubble: Lessons learned about HRI from search and rescue. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 49*, pp. 437-441. SAGE Publications.

Phillips, E., Ososky, S., Grove, J., & Jentsch, F. (2011). From Tools to Teammates Toward the Development of Appropriate Mental Models for Intelligent Robots. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 55*, pp. 1491-1495. SAGE Publications.

Raman, V., Lignos, C., Finucane, C., Lee, K. C., Marcus, M. P., & Kress-Gazit, H. (2013). Sorry Dave, I'm Afraid I Can't Do That: Explaining Unachievable Robot Tasks Using Natural Language. *Robotics: Science and Systems*.

Salas, E., Fiore, S. M., & Letsky, M. P. (Eds.). (2013). *Theories of team cognition: Cross-disciplinary perspectives.* New York, NY: Routledge.

Scheutz, M., & Eberhard, K. M. (2008). Towards a Framework for Integrated Natural Language Processing Architectures for Social Robots. *NLPCS*, (pp. 165-174).

Sellers, B. C., Fincannon, T., & Jentsch, F. (2012). The Effects of Autonomy and Cognitive Abilities on Workload and Supervisory Control of Unmanned Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 56*, pp. 1039-1043. SAGE Publications.

Shah, J., & Breazeal, C. (2010). An empirical analysis of team coordination behaviors and action planning with application to human–robot teaming. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 52*, 234-245.

Wiltshire, T. J., Barber, D., & Fiore, S. M. (2013). Towards Modeling Social-Cognitive Mechanisms in Robots to Facilitate Human-Robot Teaming. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 57*, pp. 1278-1282. SAGE Publications.

Yanco, H. A., & Drury, J. (2004). “Where Am I?” Acquiring Situation Awareness Using a Remote Robot System. *Systems, Man and Cybernetics, 2004 IEEE International Conference on. 3*, pp. 2835-2840. IEEE.

APPENDIX A

CONSENT FORM

CONSENT FORM

INTERACTIONS IN URBAN SEARCH AND RESCUE IN MINECRAFT

INTRODUCTION

The purposes of this form are to provide you (as a prospective research study participant) information that may affect your decision as to whether or not to participate in this research and to record the consent of those who agree to be involved in the study.

RESEARCHERS

Nancy J. Cooke, Professor, Arizona State University, Subbarao Kambhampati, Professor, Arizona State University, Cade Bartlett, MS Student Arizona State University

STUDY PURPOSE

The purpose of the research is to understand effective ways for humans to communicate with each other in a search and rescue planning environment.

DESCRIPTION OF RESEARCH STUDY

If you decide to participate, you will join a study funded by the Office of Naval Research in which you will learn how to interact as a team in the context of a simulated search and rescue task. We will record (via audio and screen recording) your communications and task performance and you will be questioned at the end of the task about your subjective assessment of the task. We anticipate that this study will require roughly 1 hour. Your participation is completely voluntary and you may cease participation at any time. Participants must be between the ages of 18 and 65.

RISKS

There are no known risks from taking part in this study, but as in any research, there is some possibility that you may be subject to risks that have not yet been identified. Your participation in this research study is anonymous, and no identifying information that links your answers to your identity will be collected.

BENEFITS

This research will have implications for the design of interfaces relevant to the field of urban search and rescue. Effective communication and coordination is essential for the success of such missions, and the safety of the team. At the same time you will learn something about search and rescue and human factors by participating in this study.

CONFIDENTIALITY

All information obtained in this study is strictly confidential. The results of this research study

may be used in reports, presentations, and publications, but the researchers will not identify

you. In order to maintain confidentiality of your records, Dr. Nancy J. Cooke will follow these procedures: (1) Each participant will be assigned a number; (2) The researchers will record any data collected during the study by number, not by name; (3) Any original data files will be stored in a secured location accessed only by authorized researchers; (4) consent forms will not link names to ID numbers. Consent forms will also be secured in a separate file.

WITHDRAWAL PRIVILEGE

Participation in this study is completely voluntary. It is ok for you to say no. Even if you say yes now, you are free to say no later, and withdraw from the study at any time. Your participation is voluntary and that nonparticipation or withdrawal from the study will not affect your status in class.

VOLUNTARY CONSENT

Any questions you have concerning the research study or your participation in the study, before

or after your consent, will be answered by Nancy J. Cooke at ASU Polytechnic, 480-727-2418.

If you have questions about your rights as a subject/participant in this research, or if you feel you have been placed at risk; you can contact the Chair of the Human Subjects Institutional

Review Board, through the ASU Office of Research Integrity and Assurance, at 480-965 6788.

This form explains the nature, demands, benefits and any risk of the project. By signing this form you agree knowingly to assume any risks involved. Remember, your participation is voluntary. You may choose not to participate or to withdraw your consent and discontinue participation at any time without penalty or loss of benefit. In signing this consent form, you are not waiving any legal claims, rights, or remedies. A copy of this consent form will be given (offered) to you.

Your signature below indicates that you consent to participate in the above study.

____________________________ __________________________ ____________

Subject's Signature Printed Name Date

INVESTIGATOR'S STATEMENT

"I certify that I have explained to the above individual the nature and purpose, the potential benefits and possible risks associated with participation in this research study, have answered any questions that have been raised, and have witnessed the above signature. These elements of Informed Consent conform to the Assurance given by Arizona State University to the Office for Human Research Protections to protect the rights of human subjects. I have provided (offered) the subject/participant a copy of this signed consent document."

Signature of Investigator_______________________________________
Date_____________

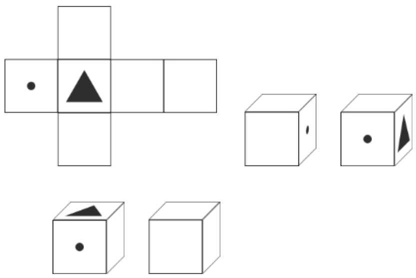
APPENDIX B

SPATIAL ABILITY TEST

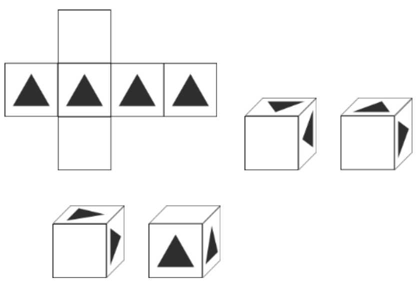
Team#___ P# I / E Date_______

Which cube <u>cannot</u> be made based on the unfolded cube on the left?

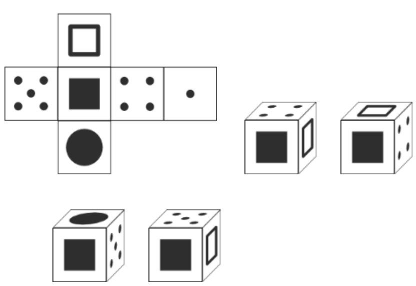
1.



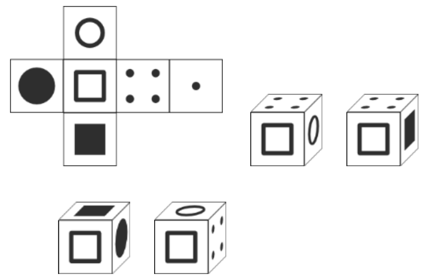
2.



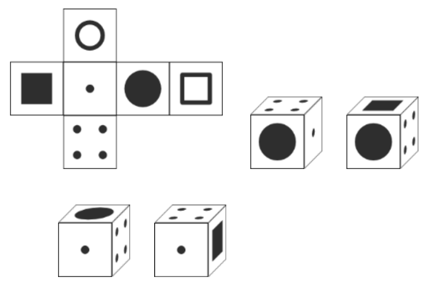
3.



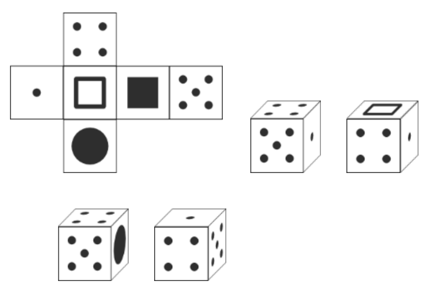
4.



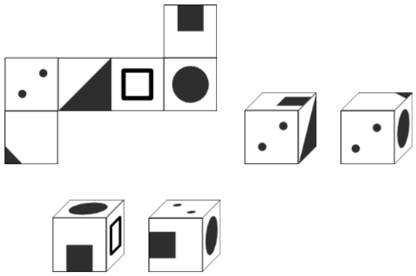
5.

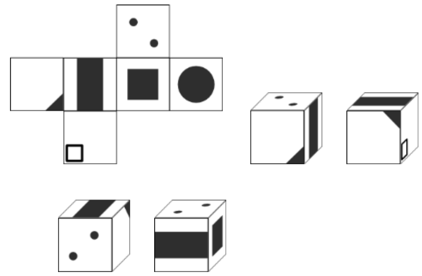


6.



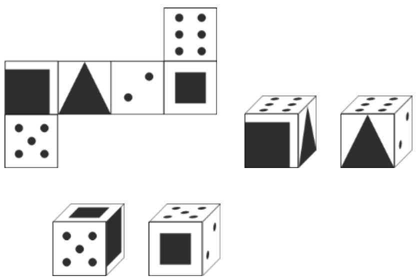
7.

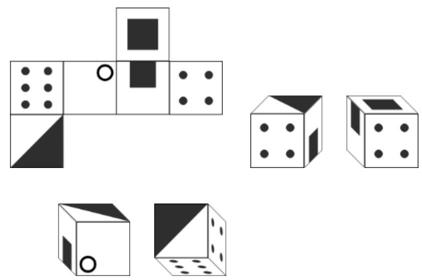




8.

10.





9.

APPENDIX C

INTERNAL REMOTE INSTRUCTIONS

YOUR INSTRUCTIONS

You are about to participate with your teammate in a simulation of a search and rescue scenario. The virtual structure you will explore represents a damaged office building. As a team you are to locate as many victims (represented by green and blue boxes) as possible, mark number of victims and hazards (hazards are represented by pink boxes) on the map, and leave the area safely before your time limit is up.

You are the internal teammate. Follow the directions given by the external teammate who will guide you through the building while you press the buttons on all green and blue boxes. The external teammate will mark the number of pink, green, and blue boxes in each room. Do not perform any action unless instructed by the external teammate. Button presses are tied to your team score – the more the better. There are rules about button pressing.

RULES

1) Buttons should not be pressed more than once, and buttons should NOT be pressed on any pink boxes. Doing so will result in a penalty.
2) Buttons on blue boxes should only be pressed within the first 8 minutes of the mission. Pressing buttons on blue boxes after 8 minutes (**or 420 or fewer seconds remaining** on the timer) will result in a penalty.
3) You will have 15 minutes to complete the task and exit the building. The internal teammate must exit the building before the time is up. This is done by pressing a button on the red box which is located near the starting location. This is mandatory, and failure to do so results in a penalty. This will end the mission immediately, so it is wise to confirm this action with your teammate first. You will receive a warning when there are 60 seconds left and you need to make your way to the exit.

Your teammate is the external teammate. It is the job of the external teammate to work outside of the virtual environment and mark on the map the number of victims (green and blue boxes) and hazards (pink boxes) that you find in each room. Be sure to communicate with your teammate to let him or her know what boxes are located and in what rooms. Your teammate may also need to communicate special instructions to you throughout the mission.

TO GET STARTED WITH THE MISSION

Please ask the experimenter if you have any questions about the instructions. In the next five minutes the external teammate will create a search plan by drawing on the map. Once the plan is complete you are ready to begin the mission and enter the virtual environment. Before you begin the actual mission you will be introduced to the virtual environment through a brief training segment. The real mission begins immediately after completing the training. Remember, you will have 15 minutes to complete the task and exit the building. Good Luck!

APPENDIX D

EXTERNAL REMOTE INSTRUCTIONS

YOUR INSTRUCTIONS

You are about to participate with your teammate in a simulation of a search and rescue scenario. The virtual structure you will explore represents a damaged office building. As a team you are to locate as many victims (represented by green and blue boxes) as possible, mark number of victims and hazards (hazards are represented by pink boxes) on the map, and leave the area safely before your time limit is up.

You are the external teammate. You work outside the virtual environment and your teammate is the internal teammate who works inside the virtual environment. Your task is guide the internal teammate through the virtual environment in accord with a plan that you will make in order to find and press buttons on blue and green boxes. You are to mark on the map the number of victims (green and blue boxes) and hazards (pink boxes) that your teammate finds in each room. Your teammate has been instructed to follow your explicit instructions and so communication is important. Also you need to keep track of where your teammate is during the search. You may also need to communicate special instructions to your teammate throughout the mission.

Your teammate is the internal teammate. It is the job of the internal teammate to follow your instructions for searching the virtual environment. He or she also has to press buttons on the blue or green boxes that they encounter. Button presses are tied to your team score – the more the better. Your teammate has been given the following rules about button pressing.

YOUR TEAMMATE'S INSTRUCTIONS

4) Buttons should not be pressed more than once, and buttons should NOT be pressed on any pink boxes. Doing so will result in a penalty.

5) Buttons on blue boxes should only be pressed within the first 8 minutes of the mission. Pressing buttons on blue boxes after 8 minutes (**or 420 or fewer seconds remaining** on the timer) will result in a penalty.

6) You will have 15 minutes to complete the task and exit the building. The internal teammate must exit the building before the time is up. This is done by pressing a button on the red box which is located near the starting location. This is mandatory, and failure to do so results in a penalty. This will end the mission immediately, so it is wise to confirm this action with your teammate first. You will receive a warning when there are 60 seconds left and you need to make your way to the exit.

TO GET STARTED WITH THE MISSION

Please ask the experimenter if you have any questions about the instructions. You will now take 5 minutes to create a search plan by drawing on the map. During the mission, you will annotate this map with numbers of blue, green, or pink boxes as they are identified during the search. Please keep in mind that the map that accompanies the plan is basically a floor plan of the office building before it was damaged by an earthquake and therefore, the actual building may differ in some ways from the map due to the earthquake damage. Also, some doors may be equipped with an electronic lock and should be in the unlocked position in a state of emergency. If they are not, they can be given power at the breaker box. A dark room may be lit with light switches, which are indicated in the environment as switches with a dim red light above them.

Once you have completed the plan you are ready to begin the mission and your teammate will enter the virtual environment. Before you begin the actual mission, the internal teammate will be introduced to the virtual environment through a brief training segment. The real mission begins immediately after completing the training. Remember, you will have 15 minutes to complete the task and exit the building. Good Luck!

APPENDIX E

INTERNAL INTELLIGENT INSTRUCTIONS

YOUR INSTRUCTIONS

You are about to participate with your teammate in a simulation of a search and rescue scenario. The virtual structure you will explore represents a damaged office building. As a team you are to locate as many victims (represented by green and blue boxes) as possible, mark number of victims and hazards (hazards are represented by pink boxes) on the map, and leave the area safely before your time limit is up.

<u>You are the internal teammate.</u> It is the job of the internal teammate to make and carry out a plan, making changes to the plan as necessary based on changes in the situation. You will work within the virtual environment to locate and press buttons on the blue or green boxes that you encounter. Button presses are tied to your team score – the more the better. There are rules about button pressing and the interior of the building.

RULES

7) Buttons should not be pressed more than once, and buttons should NOT be pressed on any pink boxes. Doing so will result in a penalty.

8) Buttons on blue boxes should only be pressed within the first 8 minutes of the mission. Pressing buttons on blue boxes after 8 minutes (**or 420 or fewer seconds remaining** on the timer) will result in a penalty.

9) You will have 15 minutes to complete the task and exit the building. The internal teammate must exit the building before the time is up. This is done by <u>pressing a button on the red box</u> which is located near the starting location. This is mandatory, and failure to do so results in a penalty. <u>This will end the mission immediately, so it is wise to confirm this action with your teammate first</u>. You will receive a warning when there are 60 seconds left and you need to make your way to the exit.

10) Doors and Lights: Some doors may be equipped with an electronic lock and should be in the unlocked position in a state of emergency. If they are not, they can be given power at the breaker box. A dark room may be lit with light switches, which are indicated in the environment as switches with a dim red light above them.

<u>Your teammate is the external teammate.</u> It is the job of the external teammate to work outside of the virtual environment and mark on the map the number of victims (green and blue boxes) and hazards (pink boxes) that you find in each room. Be sure to communicate with your teammate to let him or her know what boxes are located and in what rooms and any deviations from your plan so that they can keep track of your location during the search. Your teammate may also need to communicate special instructions to you throughout the mission.

TO GET STARTED WITH THE MISSION

Please ask the experimenter if you have any questions about the instructions. In the next five minutes please create a search plan by drawing on the map and when it is complete, then you will have 3 minutes to discuss the plan/map with your teammate. Please keep in mind that the map on which you note your plan is basically a floor plan of the office building before it was damaged by an earthquake and therefore, the actual building may differ in some ways from the map due to the earthquake damage. Once the plan has been discussed you are ready to begin the mission and enter the virtual environment. Before you begin the actual mission you will be introduced to the virtual environment through a brief training segment. The real mission begins immediately after completing the training. Remember, you will have 15 minutes to complete the task and exit the building. Good Luck!

APPENDIX F

EXTERNAL INTELLIGENT INSTRUCTIONS

YOUR INSTRUCTIONS

You are about to participate with your teammate in a simulation of a search and rescue scenario. The virtual structure you will explore represents a damaged office building. As a team you are to locate as many victims (represented by green and blue boxes) as possible, mark number of victims and hazards (hazards are represented by pink boxes) on the map, and leave the area safely before your time limit is up.

You are the external teammate. You work outside the virtual environment and your teammate is the internal teammate who works inside the virtual environment. Your task is to mark on the map the number of victims (green and blue boxes) and hazards (pink boxes) that your teammate finds in each room. Be sure to communicate with your teammate to understand his/her progress in keeping with the plan or any deviations from the plan so that you can keep track of where he or she is during the search. Be sure to develop an understanding of your environment. You may also need to communicate special instructions to your teammate throughout the mission.

Your teammate is the internal teammate. It is the job of the internal teammate to make and carry out a plan, making changes to the plan as necessary based on changes in the situation. Allow your teammate to work freely. He or she also has to press buttons on the blue or green boxes that they encounter. Button presses are tied to your team score – the more the better. Your teammate has been given the following rules about button pressing and the interior of the building.

YOUR TEAMMATE'S INSTRUCTIONS

11) Buttons should not be pressed more than once, and buttons should NOT be pressed on any pink boxes. Doing so will result in a penalty.
12) Buttons on blue boxes should only be pressed within the first 8 minutes of the mission. Pressing buttons on blue boxes after 8 minutes (**or 420 or fewer seconds remaining** on the timer) will result in a penalty.
13) You will have 15 minutes to complete the task and exit the building. The internal teammate must exit the building before the time is up. This is done by pressing a button on the red box which is located near the starting location. This is mandatory, and failure to do so results in a penalty. This will end the mission immediately, so it is wise to confirm this action with your teammate first. You will receive a warning when there are 60 seconds left and you need to make your way to the exit.
14) Doors and Lights: Some doors may be equipped with an electronic lock and should be in the unlocked position in a state of emergency. If they are not, they can be given power at the breaker box. A dark room may be lit with light switches, which are indicated in the environment as switches with a dim red light above them.

TO GET STARTED WITH THE MISSION

Please ask the experimenter if you have any questions about the instructions. Your teammate will now create a search plan on the map which he or she will take 3 minutes to share with you when it is complete. You will then annotate this map with numbers of blue, green, or pink boxes as they are identified during the search. Please keep in mind that the map that accompanies the plan is basically a floor plan of the office building before it was damaged by an earthquake and therefore, the actual building may differ in some ways from the map due to the earthquake damage. Once the plan has been discussed you are ready to begin the mission and your teammate will enter the virtual environment. Before you begin the actual mission, the internal teammate will be introduced to the virtual environment through a brief training segment. The real mission begins immediately after completing the training. Remember, you will have 15 minutes to complete the task and exit the building. Good Luck!

APPENDIX G

CHEAT SHEET

Things to remember:

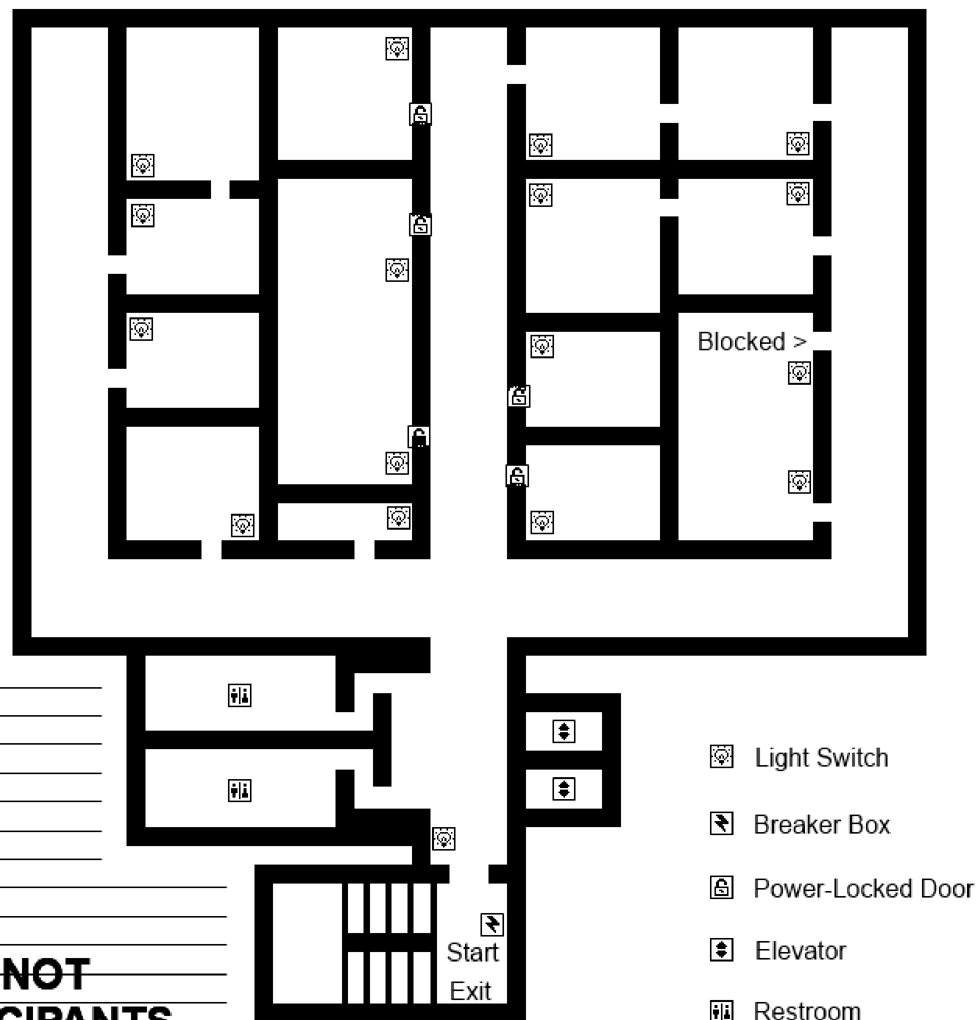
- ☐ Note changes to the plan
- ☐ Always get green boxes
- ☐ Get blue boxes only until 420 seconds remaining
- ☐ Never get pink boxes
- ☐ Exit with the red box before the time is up
- ☐ The map may not match the environment

APPENDIX H

REAL ENVIRONMENT MAP (EXPERIMENTER USE ONLY)

Notes

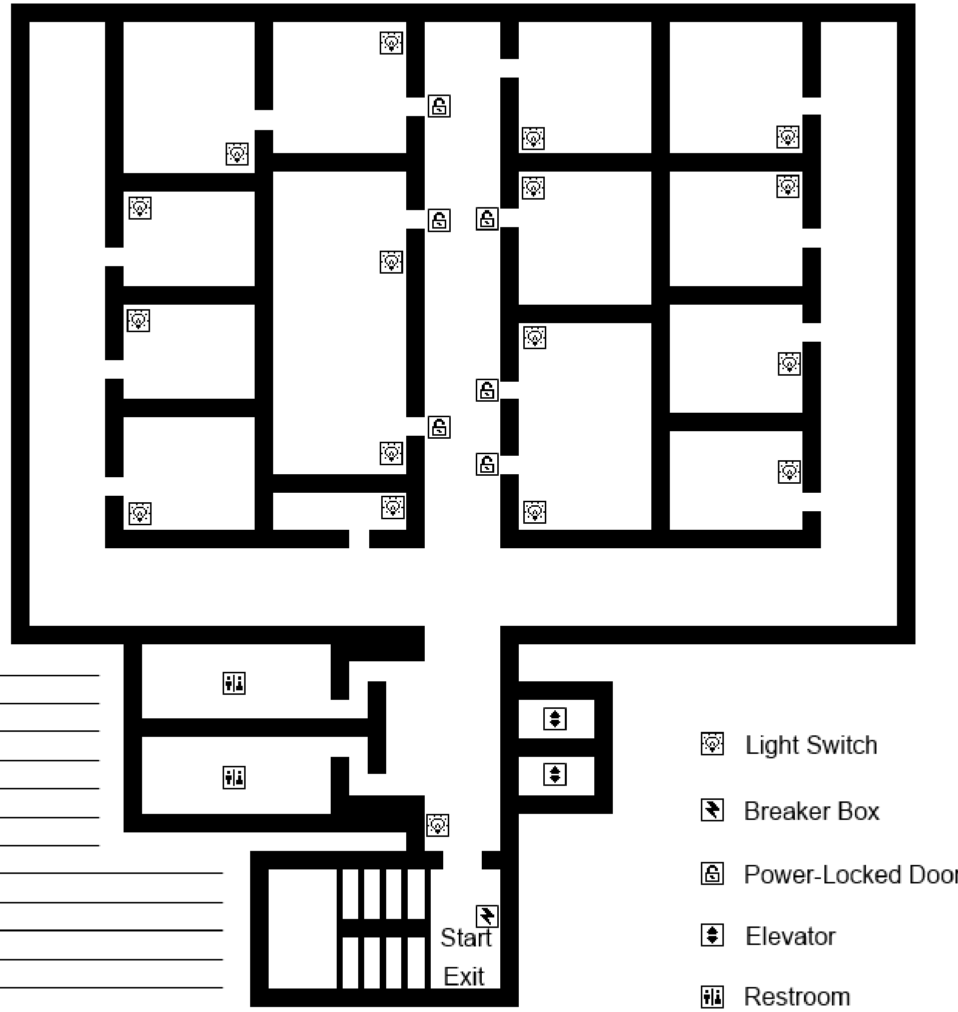
REAL MAP. DO NOT GIVE TO PARTICIPANTS



APPENDIX I

INCONSISTENT ENVIRONMENT MAP (PARTICIPANT USE)

Notes


Start
Exit
Light Switch
Breaker Box
Power-Locked Door
Elevator
Restroom

APPENDIX J

420 SECOND TIME-SENSITIVE TARGET DISCONTINUATION SHEET

Continue to press buttons on green boxes, but now pressing buttons on blue boxes will result in a penalty. All colors of boxes should still be marked on the map. Share this information with the internal teammate.

APPENDIX K

POST-TASK QUESTIONNAIRE

Team#___ P: I / E Date: ___________

Post-Task Questionnaire

Please answer the following to the best of your ability. All answers will be kept confidential and will only be reported statistically (grouped with others' responses). Please feel free to leave a question blank if you feel uncomfortable answering it.

1. What is your age? _________

2. What is your gender? (circle):
    a. Male
    b. Female
3. What is your current level of education?
    a. Less than High School
    b. High School/GED
    c. Some College
    d. 2 year degree
    e. 4 year degree
    f. Master's
    g. Doctoral
    h. Professional (MD, JD, etc.)

4. If you have been or are enrolled in a post high school institution, what is your major? _______________________

5. Are you currently employed?
    a. Yes
    b. No

6. If yes to #5, what is your job title?

____

7. Do you have military experience?
    a. Yes
    b. No

8. Do you have an experience with emergency response (fire department, police department, EMS, etc.)?
    a. Yes
    b. No

9. If yes to # 8, what service?
    ____

10. Do you have an experience with Search and Rescue?
    a. Yes
    b. No
11. If yes to #10, please briefly explain.
    ____
    ____
    ____

12. Do you have any experience with Logistics tasks?
    a. Yes
    b. No

13. If yes to #12, please briefly explain.
    ____
    ____
    ____

14. Have you participated in an experiment with robots before?
    a. Yes
    b. No

15. If yes to #14, please elaborate:
    ____

16. Have you ever taken a course or worked with robotics or remotely controlled vehicles?
    a. Yes

b. No

17. If yes to #16, please elaborate:

____

18. Have you participated in an experiment with remotely controlled vehicles (cars, airplanes, boats, etc) before?
    a. Yes
    b. No

19. If yes to #18, please elaborate:

____

20. How often do you use a computer?
    a. Daily
    b. Every couple days
    c. Once a week
    d. Every couple weeks
    e. Less than once a month
    f. I do not use computers

21. I am proficient with computers.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

22. I was familiar with, and proficient with "WASD" and mouse controls in electronic games prior to participation.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

23. In what way do you use computers? (Circle all that apply)
    a. I do not use computers
    b. Internet
    c. Email
    d. Word processing
    e. Spreadsheets
    f. Computer Games
    g. Other

24. Do you use any of the following technologies? (circle all that apply)
    a. Cell phone
    b. PDA/Blackberry
    c. iPod/MP3 player
    d. Navigation System

25. Do you have any experience in playing electronic games?
    a. Yes
    b. No

26. If yes to #25, how often do you play electronic games?
    a. I have played in the past, but do not play currently
    b. Less than 1 hr/week
    c. 1-5 hrs/week
    d. 6-19 hrs/week
    e. 20-35 hrs/week
    f. More than 36 hrs/week

27. If yes to #25, do you prefer playing single-player or multiplayer games?
    a. Single-player
    b. Multiplayer
    c. Both

28. If yes to #25 and #27, (you prefer playing multiplayer games), do you play as part of a team or as an individual?
    a. Individual
    b. Team
    c. Both

29. Do you work with a team on a regular basis?
    a. Yes
    b. No

30. If yes to #29, in what context do you work with a team and how many individuals make up this team? (Circle all that apply)
    a. Work-related *If circled, provide number of individuals* _________
    b. Sports *If circled, provide number of individuals* _________
    c. Recreation *If circled, provide number of individuals* _________
    d. Other *If circled, provide number of individuals* _________
       Please specify other: ______________________ ______________________ ________

31. How familiar were you with Minecraft prior to participation?
    a. Never seen/played
    b. I have seen it before
    c. I have little experience with it
    d. I am experienced in it
    e. I am proficient in it

32. Which platforms have you played Minecraft on? (circle all that apply)
    a. I have not played Minecraft
    b. Xbox 360 or Xbox One
    c. Touchscreen or mobile device
    d. Computer

Please rate the degree to which you agree with the following statements.

33. Regardless of my score, I feel like I performed well overall.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

34. I was able to use my materials effectively (map, computer, virtual environment, etc.)
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

35. If I were asked to participate in another project like this one, I would like to.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

36. The procedures I employed were the most effective way to complete the missions.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

37. The way I made decisions was the best way to make decisions.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

38. I did **<u>not</u>** like the way I made decisions.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

39. This task was complicated.

a. Strongly Agree
b. Slightly Agree
c. Neutral
d. Slightly Disagree
e. Strongly Disagree

40. This task was easy.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

41. This task was boring.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

42. The user-computer interface was easy to use.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

43. I enjoyed participating in this study.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

44. My teammate did their job well.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

45. I had confidence in my teammate's abilities.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

46. My teammate and I were well coordinated.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

47. My teammate communicated effectively with me.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

48. I communicated effectively with my teammate.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

49. My teammate gave me appropriate feedback on my own performance.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

50. I understood my teammate's communications.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

51. I feel I know the virtual environment well.
    a. Strongly Agree
    b. Slightly Agree

c. Neutral
d. Slightly Disagree
e. Strongly Disagree

52. I trusted my teammate’s abilities to fulfil their role.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

53. I trusted the information and suggestions given by my teammate.
    a. Strongly Agree
    b. Slightly Agree
    c. Neutral
    d. Slightly Disagree
    e. Strongly Disagree

APPENDIX L

WORKLOAD TLX

Task-Load Index

Please make an "X" in the spaces between the lines on the scales in order to answer the questions. Please do not place an "X" on the lines themselves.

1. How mentally demanding was the task?

Very Low Very high

2. How physically demanding was the task?

Very Low Very high

3. How hurried or rushed was the pace of the task?

Very Low Very high

4. How successful were you in accomplishing what you were asked to do?

Very Low Very high

5. How hard did you have to work to accomplish your level of performance?

Very Low Very high

6. How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low Very high

APPENDIX M

DEBRIEFING

Debriefing

Thank you for your participation in our study.

This study you have completed was to help us to understand effective ways for humans to communicate with robots in a search and rescue planning environment. Robots are seeing more use in modern search and rescue situations and are beneficial in searching in places that also put the responders at risk. These robots are currently remotely controlled, but the long-term goal is to have automated robots that can effectively function as individuals on the team.

This study provides us with natural language communication samples that must be accounted for if a robot is to be completely integrated into the team, with specific interest in communications when the conditions of the task change.

In addition, the study also compares the performance differences between remotely controlled and intelligent robots. In one condition the participant in the robot's role is instructed directly follow the operator's explicit directions, while in the other condition they are instructed to act independently. Intelligent and independent robots are expected to perform better, especially as more are operated by a single operator.

As this is ongoing research, please do not discuss this information with anyone that will be participating. You are free to discuss this study with anyone that will not be participating.

Thank you again. Please ask if you have any additional questions.

Sincerely,

Nancy J. Cooke

Professor, ASU

APPENDIX N

PAYMENT RECEIPT

Interactions in Urban Search and Rescue in Minecraft Project Participant Acknowledgement of Payment

---

I, ____________________________________ , participated in a collaborative team study that was conducted in Dr. Nancy Cooke's Cognitive Engineering Research on Team Tasks (CERTT) laboratory located at Arizona State University's Polytechnic Campus.

I was paid [$10/hr] a total of $________

(Circle one of the following that applies):

For completing the study session for which I was scheduled
OR
For completing part of the study session for which I was scheduled
OR
Because other participants did not show up and I was sent home

Print name: ____

Sign name: ____

Phone number: ______________________________________

Date: _____________

APPENDIX O

SCREENSHOTS OF THE VIRTUAL ENVIRONMENT