Revealing unexpected effects of rescue robots' team-membership in a virtual environment

Corine H.G. Horsch
Delft University of Technology
C.H.G.Horsch@tudelft.nl

Nanja J. J. M. Smets
TNO/ Delft University of Technology
Nanja.Smets@tno.nl

Mark A. Neerincx
TNO/ Delft University of Technology
Mark.Neerincx@tno.nl

Raymond H. Cuijpers
Eindhoven University of Technology
R.H.Cuijpers@tue.nl

ABSTRACT

In urban search and rescue (USAR) situations resources are limited and workload is high. Robots that act as team players instead of tools could help in these situations. A Virtual Reality (VR) experiment was set up to test if team performance of a human-robot team increases when the robot act as such a team player. Three robot settings were tested ranging from the robot as a tool to the robot as a team player. Unexpectedly, team performance seemed to be the best for the tool condition. Two side-effects of increasing robot's team-membership could explain this result: mental workload increased for the humans who had to work with the team-playing robot, whereas the tendency to share information was reduced between these humans. Future research should, thus, focus on team-memberships that improve communication and reduce cognitive workload.

Keywords

USAR, human-robot interaction, teamwork, team performance, virtual environment, shared situation awareness, team identification.

INTRODUCTION

Urban search and rescue (USAR) is the first response when human-made structures collapse (Murphy, 2004). USAR teams operate in a chaotic and unstructured environment, where access to facilities like hospitals, power supply, etc. is limited. The work is dangerous and the workload of USAR teams is often very high due to little preparation time, unknown territory, physical challenges, and emotional demands (De Greef T.E., Der Kleij, Brons, Brinkman, & Neerincx, 2011). Robots could be helpful by mapping the disaster area and searching for victims. Due to high workload and human resource limitation, there is a need for robots that (semi-) automatically perform their part of the work; more as a team member instead of just as a tool (Fong, Thorpe, & Baur, 2003). A team member is an interdependent agent, robot or human, performing coordinated tasks toward the achievement of specific task goals (Nieva, Fleishman, & Rieck, 1985). Although the level of robot autonomy has been increasing over the years, overall performance enhancement is seldom accomplished (Kaber, Onal, & Endsley, 2000; Scholtz, Antonishek, & Young, 2003). Instead of developing robots as autonomous tools, making them more like team members could possibly increase such performance (Fong et al., 2003). Klein and colleagues (2004) defined ten challenges for making an autonomous agent a team player. By making an autonomous agent or an USAR robot more like a team member, performance related concepts, like shared situation awareness (SSA) and the extent to which members feel a part of the team; team identification (TI), might be influenced.

SSA is the reflection of how similarly team members view a given situation (Bolstad, Cuevas, Gonzalez, & Schneider, 2005). The gut feeling is that SSA and performance correlate. Some studies confirm this relationship between SSA and performance (Bolstad et al., 2005), while other experiments do not find a relationship (De Greef T.E. et al., 2011). Other research has shown that people work hard on collective tasks when the group is valued highly (Barreto & Ellemers, 2000). Team members with a higher team identification (TI) feel more intertwined with the team's performance. We want to know whether team performance (TP) actually increases when a robot meets the challenges of Klein and colleagues. We expect TP to increase when the level of team membership of the robots increases (*H1*). As discussed before, SSA and TI are both related to performance. So these concepts may be able to explain part of the difference in TP that is provoked by the level of team membership. Therefore, we expect that the difference in TP can be partly explained by SSA and TI (*H2*).

METHOD

In order to test the two hypotheses, an experiment was set-up in which a human-robot team performed a search task in a virtual world. The USAR team in this experiment consisted of two persons, a flying robot and a ground robot. The first task for the team was to find six victims in a virtual destroyed office as fast as possible. The second task was to make a map of the environment. This experiment had a between subject design. Pairs of participants were randomly assigned to one of the three conditions. The level of team membership of the robots was varied in line with the ten challenges of Klein and colleagues (2004). In condition one, collaboration between the participants and the robots was not supported. In condition three, full collaboration was supported. In condition two, the robots showed on the map where they were in the building, as this functionality was expected to support SSA, which was expected to be an important explanatory factor for performance.

Participants

In this experiment 76 persons participated (average age: 26 years , 52 male and 24 female) in 38 teams. 20 teams did not know the other participant, in the other 18 teams the participants did know each other. In general, the participants were not or little experienced in playing first person shooter games.

Materials

For this virtual experiment the Unreal Development Kit and UsarSim were used. UsarSim is a virtual simulation environment for USAR robots. Figure 1 depicts some screenshots of the destroyed office environment that was created, and a screenshot of the Organization Awareness Display; the interactive map. On the map participants could add messages, like a victim report, an obstacle etc. Moreover, the maps were linked, so they could see the actual positions of the robots as well as each other's actions. Lastly, walkie-talkies were used for communication between the participants, who were located in a different cubicles.

Figure 1. Screenshots of the virtual environment in Unreal and the Organization Awareness Display. From left to right: an example of a destroyed office from the ground robot's view, an example of an victim lying under a desk, the UAV (flying robot), and the interactive map

Procedure

After an elaborate description of the mission and the robots, the participants decided themselves who operated which robot. In condition two and three the participants also got a short manual which explained the interactive map. In order to get familiar with the systems, the participants were given a few minutes for training, after which the real task started. During the mission the experiment leader (as a Wizard of Oz) took care of the collaboration tasks of the robot. The mission ended after 22 minutes or when a team found all victims. After that, the participants filled out several questionnaires. In total the experiment lasted around 60 minutes.

Measurements

Team performance was measured in two ways. First, team performance was measured by *time performance;* the time the teams needed to find three victims. Second, an *integrated performance* score was calculated which included the difficulty of found victims; the area searched; and the number of obstacles indicated on the map. SA was measured with 12 questions (e.g. How many rooms have you searched together?). The answers of both participants in a team were subtracted, so, the closer the score was to zero the better the team's SSA was. TI was measured by the average of ten 7-point scale items based on questions used by Van Der Vegt & Bunderson (2005). Furthermore, a demographic and manipulation check questionnaire were filled out. Lastly, 27 collaboration questions were asked. These questions covered coordination, communication, trust, and satisfaction, which are all seen as important aspects for effective collaboration. In addition, communication between the team members was recorded. Later on the messages were categorized into the following categories: providing information, asking a question, giving an assignment, positive answer, negative answer, confirm message, ask for repetition of former message, and unknown.

RESULTS

To check our manipulation of team membership between the three condition a one-way independent ANOVA (condition: three levels) on the manipulation check score was conducted. The data was negatively skewed and therefore a log transformation was done[1]. There was an effect of condition on the level of team membership. In condition one the level of team membership was lowest (M = .31, SD = .23), in condition two team membership was highest (M = .15, SD = .19), and condition three fell in between (M = .23, SD = .20, $F(2, 73) = 3.79$, p = .027, r = .31). A planned contrast revealed the manipulation check score differed for having a robotic system compared to a paper map (condition 1 vs. condition 2 and 3), $t(73) = -2.41$, p = .019, r = .27.

Team performance per condition

To investigate the relationship between the level of team membership and team performance (*hypothesis 1*), we analysed the integrated and time performance scores for the three conditions. Results revealed no effect of the condition on the integrated performance. Moreover, when time performance was assessed, there was a trend that the higher the level of team membership of the robots the longer it took to find the victims, $F(2, 29) = 2.45$, p = .104, r = .38. In other words, the direction of the found effect is opposite to the one expected. A planned contrast revealed that having a robot with team member aspects (condition 2 and 3) increased the time needed to find three victims compared to only having a paper map (condition 1), $t(29) = 1.99$, p = .056, r = .35.

The explanatory power of SSA and TI for team performance

In order to investigate if SSA and TI could explain a difference in performance (*hypothesis 2*) a mediation analysis as put forward by Baron and Kenny (Baron & Kenny, 1986) was conducted through the PROCESS script[2]. The conditions were used as independent variable, performance as dependent variable, and SSA and team identification served as mediators. A mediation analysis was conducted for both performance measures. These analyses did not show that SSA and TI could explain a difference in performance.

Differences in teamwork between conditions

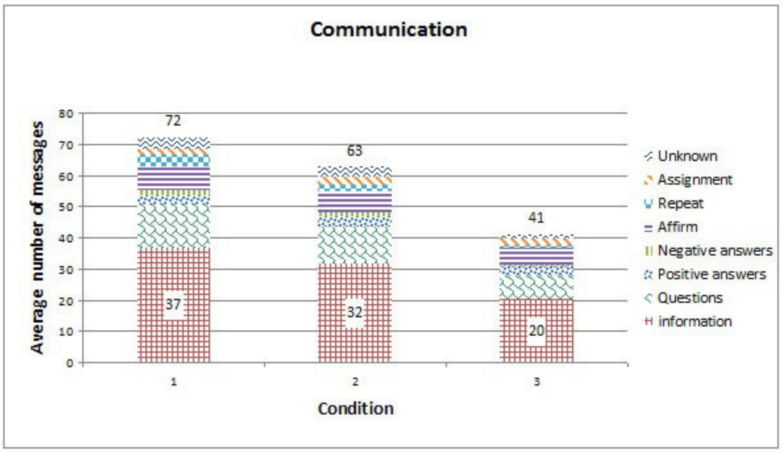

Figure 2. Vocal communication between the participants within a team. The total number of messages sent by participants differed between the conditions. In condition one it were 37 messages, in condition two it were on average 32 messages, and in condition three it were only 20 messages.

To assess teamwork the scores for coordination, communication, trust, and satisfaction were analysed per condition. The teamwork concept scores did not differ between the conditions except for vocal communication.

[1] Note that the data of the manipulation check questions (MC) was negatively skewed (not normally distributed), and therefore the scores were subtracted from the maximum score (7) and after that a log transformation was done: $MC_{log} = log_{10}(7-((\Sigma (MC_1..MC_6))/6)))$

[2] Hayes, A. F. (2012). An analytical primer and computational tool for observed variable moderation, mediation, and conditional process modelling. *Manuscript submitted for publication*.

In figure 2 the vocal communication per condition is presented. There was an effect of condition on the number of total messages, $F(2,35) = 3.42$, $p = .044$, $r = .40$. Furthermore, one of the subcategories of communication was different. The number of information messages decreased over the conditions (condition 1: $M = 37$, $SE = 3.9$, condition 2: $M = 32$, $SE = 5.6$, condition 3: $M = 20$, $SE = 2.7$, $F(2,35) = 4.10$, $p = .025$, $r = .40$).

DISCUSSION

In order to investigate whether team performance increases when a robot is made a team player we performed a study in which the level of team membership of the robots was varied between three conditions. Results showed that time performance indeed differed across conditions, whereas the integrated performance score did not. We expected that performance would increase when team membership of the robots increased (*H1*). However, the opposite was found; time performance decreased when the robots behaved more like team players. Shared situation awareness and team identification were measured in order to investigate if the difference in performance could be explained by these concepts. Results did not provide evidence for this expectation (*H2*). In addition, team work concepts were assessed per condition. Results showed that participants communicated less, when the level of team membership of the robots was higher.

Team performance decreased instead of increased when the robots behaved more like a team member. A possible explanation might be that mental effort for the participants increased due to the robots' 'complexity'. The Cognitive Task Load (CTL) model (Neerincx, 2003) distinguishes three general factors that have an effect on mental effort and performance (Grootjen, Neerincx, & Van Weert, 2006). These factors are time occupied, task switches, and level of information processing. The time the participants were occupied during the mission was equal in all three conditions. Participants in every condition were engaged in the mission all the time. On the other hand, the interactive map probably induced more task switches (more actions had to be performed to add a victim) and required a higher level of information processing. This results in a higher mental workload and a decrease in performance. Furthermore, it could be argued that using the interactive map takes more time than the paper map and therefore a difference in time performance is found between the conditions.

Another possible explanation for the worse team performance in condition three compared to condition one is the difference in vocal communication. The results showed that in condition three only 20 messages were sent on average, whereas 37 messages were sent on average in condition one. Because of the interactive map, there was less need to communicate for the participants. So, the tendency to *not* share information may arise and the decreased performance over the conditions may be explained by this lack of communication.

The second hypothesis stated that SSA and team identification would explain team performance at least to a certain extent. However, team performance did not increase when shared situation awareness or team identification increased or decreased. The mediation analysis showed that the differences in performance could not be explained by SSA or team identification. Maybe other concepts that possibly differed between conditions, for instance participants' confidence in successful completion of the mission, might be able to explain team performance. However, this explanation is speculative. By following the guidelines of Klein et. al (2004) we tried to ensure human-robot collaboration. However, during the experiment it became clear that it is really hard to accomplish actual collaboration. For example, it is not clear if and how the advice given by the robot is experienced as realistic team member behavior. When participants were asked about the robot's positive and negative aspects, only one comment was made about the speech of the robot, other comments only considered robot features. Apparently, the robot features affected the participants more than the collaborative behavior, which suggests that the robots' team behavior was not that obvious to the participants. It is therefore questionable if or to what extent real human-robot collaboration occurred.

Team identification (TI) has not been studied much. Shapiro and colleagues (2002) argue that more studies should include TI, because TI could explain some unexpected performance differences in certain situations (Shapiro, Furst, Spreitzer, & Von Glinow, 2002). TI in this experiment, however, could not explain the difference in team performance. A possible reason could be that some teams did know each other, whereas others did not. The results of this experiment showed a trend in team identification between short-term and long-term teams. Teams that knew each other scored higher on team identification ($M = .28$, $SD = .15$)[3], than teams that did not know each other ($M = .38$, $SD = .21$, $F(1,37) = 2.88$, $p = .100$, $r = .27$). Therefore, it seems better to include only short-term or long-term teams in future research when team identification is the subject of research, because acquaintance seem to affect team identification. Another improvement would be to include a workload measure when different robot configurations are evaluated. As argued before extra workload due to the extra system might explain the decrease in performance.

[3] Note that these statistics concern the transformed team identification score (inverse log transformation, similar to the manipulation check score)

CONCLUSION

Robots are and will be deployed in USAR situations, they can explore dangerous environments where people cannot go. USAR situations give researchers a dynamic and unknown environment to test their robots in which human and robot behavior can reveal underlying teamwork mechanisms. In this experiment three robot settings were tested ranging from the robot as a tool to the robot as a team player. The hypothesis was that team performance would increase when the robot behaved more like a team member. However, the results showed the opposite effect; the teams with the robot as a tool performed better. Possible explanations are an increase in mental workload for the humans. Furthermore, it was hypothesized that SSA and team identification would be able to explain this increase in performance. However, the decrease in performance could not be explained by SSA or team identification. The higher level of robot automation seemed to induce *not* sharing information. Future research should focus on other aspects of teamwork like communication. The results of this experiment suggest that better vocal communication within a team could explain higher performance. In addition, cognitive workload measures should be included in future experiments, since workload could be a confounding factor.

ACKNOWLEDGMENTS

This paper describes research done under the EU-FP7 ICT247870 NIFTi project (http://www.nifti.eu).

REFERENCES

1. Baron, R. M., & Kenny, D. A. (1986) The moderator--mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*, *51*, 6, 1173–1182.
2. Barreto, M., & Ellemers, N. (2000) You cannot always do what you want: Social identity and self-presentational determinants of the choice to work for a low status group. *Pers Soc Psychol B*, *26*, 8, 891–906.
3. Bolstad, C. A., Cuevas, H., Gonzalez, C., & Schneider, M. (2005) Modeling shared situation awareness. *14th Conference on BRIMS*. Los Angles, CA.
4. De Greef T.E., D., Der Kleij, R., Brons, L., Brinkman, W. P., & Neerincx, M. A. (2011) Observability Displays in Multi-Teams. *10th International Conference on Naturalistic Decision Making*, Orlando, FL.
5. Fong, T., Thorpe, C., & Baur, C. (2003) Collaboration, dialogue, human-robot interaction. *10th International Symposium of Robotics Research*. Lorne, Victoria, Australia.
6. Grootjen, M., Neerincx, M. A., & Van Weert, J. C. M. (2006) Task-based interpretation of operator state information for adaptive support. *Foundations of Augmented Cognition,* San Fransisco, CA.
7. Kaber, D. B., Onal, E., & Endsley, M. R. (2000) Design of automation for telerobots and the effect on performance, operator situation awareness, and subjective workload. *Hum Factor Ergon Man, 10,* 4, 409–430.
8. Klein, G., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., & Feltovich, P. J. (2004) Ten challenges for making automation a team player in joint human-agent activity. *IEEE Intell Syst*, *19*, 6, 91–95.
9. Murphy, R. R. (2004) Human-robot interaction in rescue robotics. *IEEE T Syst Man Cy C*, *34*, 2, 138–153.
10. Neerincx, M. A. (2003) Cognitive task load design: model, methods and examples. *Handbook of Cognitive Task Design*, 283–305.
11. *Nieva*, V. F., Fleishman, E.A., Rieck, A. (1985) *Team dimensions: Their identity, their measurement and their relationship*
12. Scholtz, J., Antonishek, B., & Young, J. (2003) Evaluation of operator interventions in autonomous off-road driving, *Proceedings of PerMIS '03*, Gaithersburg, MD.
13. Shapiro, D. L., Furst, S. A., Spreitzer, G. M., & Von Glinow, M. A. (2002) Transnational teams in the electronic age: are team identity and high performance at risk? *J Organ Behav*, *23*, 4, 455–467.
14. Van Der Vegt, G. S., & Bunderson, J. S. (2005) Learning and performance in multidisciplinary teams: The importance of collective team identification. *Acad Manage J*, *48*, 3, 532–547.