Developing Human-Robot Team Interdependence In a Synthetic Task Environment

7 authors, including:

Glenn Lematta
Arizona State University

Shawaiz Bhatti
Arizona State University

Erin Chiou
Arizona State University

Nathan J. McNeese
Clemson University

DEVELOPING HUMAN-ROBOT TEAM INTERDEPENDENCE IN A SYNTHETIC TASK ENVIRONMENT

Glenn J. Lematta[1], Pamela B. Coleman[1], Shawaiz A. Bhatti[1], Erin K. Chiou[1],
Nathan J. McNeese[2], Mustafa Demir[1], Nancy J. Cooke[1]
[1]Arizona State University, [2]Clemson University

In future urban search and rescue teams, robots may be expected to conduct cognitive tasks. As the capabilities of robots change, so too will their interdependence with human teammates. Human factors and cognitive engineering are well-positioned to guide the design of autonomy for effective teaming. Previous work in the urban search and rescue synthetic task environment (USAR-STE) used Minecraft, a customizable gaming platform. In this effort, we advanced the USAR-STE by increasing interdependence in dyadic human-robot teaming through the Coactive Design framework. In this framework, we defined required capacities of victim identification in USAR from literature, and used them as inputs for modeling interdependence, and determined recommendations that would enhance interdependence in the task environment. Although Coactive Design is typically used to design interdependence for robots or jobs, we demonstrated how it can also be used to design an experimental team task environment.

INTRODUCTION

In future urban search and rescue (USAR) teams, robots may be expected to fulfill increasingly cognitive and physical capacities (Liu & Nejat, 2013). As USAR robots gain increasing autonomy, the interdependence between the robot and its human teammates may change in unanticipated ways. The demands of the human teammates in victim identification tasks include extended search periods with little sleep and limited visual information, and accountability to teammates, victims, and larger USAR efforts. Balancing multiple priorities, a human operator may not notice every relevant environmental feature in the camera feed as the robot autonomously searches a collapsed structure. This lack of awareness may come at high costs to USAR efforts as coordination breaks down between the human and robot teammates.

Human teams can support cognitive activity advantageously through team interactions (Cooke et al., 2013). However, the ways that autonomous robots and people can support teamwork interdependently has yet to be fully worked out. For instance, there is evidence that humans mirror their autonomous teammate's interactions (Chiou et al., 2016; McNeese, 2018) and consequently, these teams may develop less flexible interactions (Demir et al., 2017). Flexibility in teams is critical for adaptation in dynamic and complex environments (Gorman et al., 2010). Thus, there is a need to design interdependence in human-robot relationships that are more team-like, to best utilize the strengths of each, and their ability to support one another in normal and unexpected situations.

Realizing effective human-robot teaming in USAR is an multidisciplinary challenge. With advances in autonomous control (Liu & Nejat, 2013) and maneuvering capabilities (Whitman et al., 2018) of USAR robots, the technological aspects of the problem space are well-defined. Additionally, findings from USAR field research (Burke et al., 2004; Krujiff et al., 2012; Whitman et al., 2018) described the real-world demands on people and their robot teammates. Yet, there are several questions remaining regarding how future USAR robots may act as team players.

Critical in the implementation of autonomy is understanding how system changes shift interdependency in the task. *Interdependence* was defined as "the set of complementary relationships that two or more parties rely on to manage required (hard) or opportunistic (soft) dependencies in joint activity" (Johnson et al., 2014, p. 47). Johnson et al.'s Coactive Design framework posits that observability, predictability, and directability (OPD) describe eight of 10 challenges for teamwork with automation (Klein et al., 2004). *Observability* means that aspects of one's status, as well as one's knowledge of the team, task, and environment, may be accessed. *Predictability* means that expectations of the consequences of team member actions are accurate. Finally, *directability* is the extent that each partner may influence each other's actions. These three constructs represent general requirements for effective interdependent relationships. A general challenge is to gain insight into the actual capacity requirements of work, which can entail various forms of research.

Field research in USAR may provide ecologically valid data on work-as-done, though it may be difficult for several reasons. Collecting data unobtrusively during field work may be extremely difficult or impossible. Furthermore, since technological changes also affect interdependency (particularly with an autonomous teammate), the ability to model the future landscape of interactions using work-as-done today is limited. Contrasted with field research, tightly controlled experiments may isolate specific phenomena well for increased explanatory power, but they also have limited ecological validity and thus do not often translate well to real-world settings.

One approach to studying complex work systems, that captures the benefits of controlled experiments while increasing ecological validity, is to use a synthetic task environment (STE; Cooke & Shope, 2004). An STE is an experimental testbed that use features abstracted from knowledge about a specific work context to guide the design and development of task artifacts, stimuli, and constraints.

Supporting interdependence means that specific constraints in taskwork drive changes to the team task interface with the intention of improving teamwork. In our work, we describe modifications to an STE. These modifications included changes to the participant's tools, the addition of a robot role performed by a confederate (i.e. a Wizard of Oz study), and updates to the Minecraft virtual environment. These changes were in service of our research objective, which was to test different robot explanation strategies. The following sections describe the method used to develop the original STE and the application of an interdependence framework, Coactive Design (Johnson et al., 2014). We present our results as design changes made to increase and support interdependence for our human-robot dyad.

REDESIGNING A SYNTHETIC TASK ENVIRONMENT

This work uses Cooke and Shope's (2004) methodology for designing an STE. The design process begins with gathering knowledge on the context of work, current research objectives, and other constraints such as logistics and software flexibility. Then, the relevant features of the knowledge are abstracted into task environment architectures, which are implemented in a prototype. This prototype was developed iteratively through feedback and discussion to become a finished product.

Design Inputs

We considered three types of inputs for determining which design architectures to include in the STE: 1) knowledge of victim identification tasks, 2) our research objectives, and 3) other constraints related to the research team's capabilities and resources. A literature search was conducted to address the first two types of inputs. The selection criteria for the first input was any published study in the *USAR* domain that specifically considers the use of *ground vehicles* or *robots*. For the research objectives, we included relevant articles related to *explanations*, *human-robot teamwork*, *team cognition*, and *autonomy capabilities*. These papers either needed to be published in the past 10-years or have been seminal to the research area.

Knowledge on the work domain. Developing the USAR-STE required understanding robot-assisted victim identification tasks, tools, and work environments. Various models of uninhabited ground vehicles have been used in USAR with different morphologies, operating systems, and control structures (Burke et al., 2004; Krujiff et al., 2012; Whitman et al., 2018). In an observational study of a 16-hour USAR training drill, limited communication was found between robots and their operators (Burke et al., 2004). Additionally, multiple priorities and tradeoffs in USAR, and the integration of several data streams (e.g., visual information and communication) were required to carry out victim identification. In addition to these teleoperated robots in the USAR study, different autonomous control structures for low-level autonomy, semi-autonomous control, and adjustable autonomy within the scope of single robots as well as multi-robot teams have emerged (Liu & Nejat, 2013). Research indicated that human-robot teams in USAR are vulnerable to loss of situation awareness and lack a rich communication structure between teammates (Bartlett & Cooke, 2015; Demir et al., 2018).

Research objectives. Design goals in this work included : 1) identifying in the team cognition literature and in studies of all-human teams, individual teammate behaviors that are associated with achieving high levels of individual and team situation awareness, 2) facilitating the study of team cognition, 3) providing a task environment realistic to victim identification and mapping, and 4) addressing design considerations for robot explanations in future USAR human-robot teams. The study of team cognition involves considering interactions that develop over time (Cooke et al., 2013). This means that the testbed should facilitate various measurement of team interactions, such as communication flow and interaction patterns (Cooke & Gorman, 2009). Additionally, because team interactions are largely driven by interdependence between teammates, task interdependence should be a major design consideration for measuring team cognition.

An STE is intended to mirror the same cognitive aspects of the modeled work domain (Cooke & Shope, 2004). Therefore, in the USAR context, the STE should consist of a human-robot team in which a robot searches for victims inside a collapsed structure. The humans should have multiple priorities to consider, such as time, accountability, and environmental dynamics.

Future autonomous USAR robots have potential to fulfill cognitive task functions in victim identification that were previously carried out by people, including reasoning with mental models (Sreedharen et al., 2018), communicating in natural language (Feng et al., 2018), and providing explanations (Chakraborti et al., 2018). When autonomous systems gain degrees of freedom (i.e., the range of possible actions increases), boundaries with respect to a robot's behavior are more complex and uncertain. This means that understanding future human-robot teaming cannot rely completely on prior models of human-robot interaction or human teamwork. One method to quickly test future robot capabilities is through a WoZ experiment method, which uses a confederate or experimenter to act as a robot. This method has been used for robot roles including sensing, communication, and mapping/localization (Riek, 2012). All these cognitive elements are important for victim identification.

Other constraints. Several other constraints to the STE included the flexibility and accessibility of Minecraft. The game features a 'redstone' circuitry system and special blocks called 'command blocks', allowing programmed events to be triggered within the virtual environment. The use of simple 'hopper' timers (Clock circuit, n.d.) enabled synchronization of dynamic events over time. Additionally, a chat and data management system was developed simultaneously to facilitate text-chat communication and data post-processing. Collaboration between researchers involved in the STE design and hardware developers of the chat and data management system yielded a highly customized product, with features such as event timers and predefined key messages. Finally, the laboratory space used included a dividing wall with holes that could be used to connect the workstations.

Coactive Design

In the early prototyping phase of our design, we found that pilot participants would not communicate much with the robot. The little communication that occurred also reflected a more supervisory control relationship. For instance, the robot would only search rooms and identify victims with permission, meaning the team was spending a considerable amount of time and effort communicating. Most exchanges were directions to the robot to go to specific rooms or to stop moving. The breadth of the robot's vocabulary was also unclear in pilot testing. In interviews with pilot participants, they stated that communication via text-chat increased workload and that they would only communicate with the robot when necessary. To address these issues, we used the Coactive Design framework (Johnson et al., 2014) to organize the inputs from our review into required capacities in the team task. These capacity requirements include sensing, movement, interpretation, and decision-making necessary for completing a task. Then, each subtask in the STE is modeled as performed by each team member, including how other team members may support the activity. For each role, a codified color scheme represented a team member's capability in the task (see Table 1). Following the interpretation of these constraints, recommendations for supporting interdependence were defined (e.g., "Share status information at this time"). Table 2 provides a sample of the interdependence framework in our task.

Table 1. Coding scheme for team member role alternatives

Performer	Supporting Team Member
I can do it all. **[4]**	My assistance could improve efficiency.
I can do it all but my reliability is < 100%. **[3]**	My assistance could improve reliability.
I can contribute but assistance is needed. **[2]**	My assistance is required.
I cannot do it. **[1]**	I cannot provide assistance

Table 2. Sample of interdependence modeling with Coactive Design

Capacity	Human Performs		Human Performs	
	H	R	R	H
Sense the environment	2	2	3	3
Localize robot position	3	3	4	4
Map the building	3	3	1	2
Interact with virtual objects	1	2	4	4

**Note that capacities such as interacting with virtual objects and mapping the building are 'hard constraints.' That is, only one agent may fulfill this part of a task.*

DESIGN CHANGES

Increasing Interdependence

Several design changes were made to the environment, task, and roles to constrain individuals and encourage team interactions. Environmental changes included separating the workstations into different rooms and implementing dynamic events in the Minecraft virtual environment. Dynamic events included building collapses to maneuver around (Figure 1) and critical victims with limited time. These require a team to coordinate in novel ways and to re-plan activities as they progress through the victim identification task.

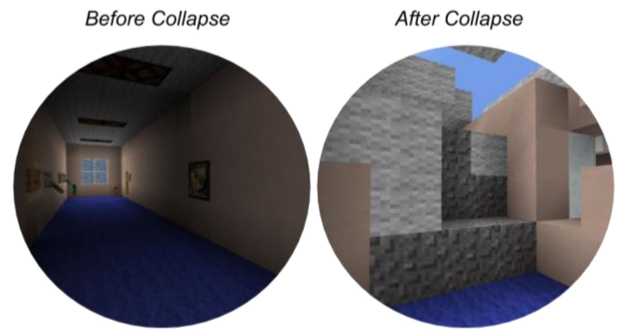

Figure 1. The building before and after a collapse

Changes to the team task involved designing around multiple priorities including time pressure, accuracy, and accountability. We added time pressure by adjusting the overall time limit in the task and introducing a time constraint to tagging a victim for collection based on its triage status. We emphasized accuracy by rewarding participants only for correctly marked map elements, rather than framing the objective around tagging all the victims. Accountability to the USAR team and larger efforts was introduced through the STE narrative, which guided training. This included a need to monitor the robot's behavior and understand the cause of events. For instance, we stated in the training that the USAR team leader may ask the participant why the robot deviated from the plan. These changes should altogether encourage teammates to depend on each other for status information and planning.

The robot's role was to search the building while acting as a team player (see the following section). Effort was placed on defining the boundaries of the robot's performance, given that our research objective is to test advanced autonomy. We concluded that the robot should have a broad understanding of the task context, with rich vocabulary and communicative capability. Additionally, for our research objective of studying robot explanations, the robot needed to behave in ways that were apparently plan deviations (i.e., a script of timed off-plan behaviors), as well as be capable of providing explanations for those behaviors. We defined the participant's role as the human teammate, whose objective was to produce an accurate map of the building and support the robot. Given that they were responsible for map accuracy, this role encouraged careful monitoring of events in the building, and interactions that support team situation awareness, such as sharing status information, plan progress, or challenges.

The USAR-STE

The updated USAR-STE features two team member workstations and one experimenter workstation (Figure 2). The human operator workstation includes a map of the structure before the collapse, camera view, and chat system as tools. Additionally, a quick reference guide is provided which reviews training material. The camera view is a reduced view of the robot workstation's computer that runs the virtual environment (i.e., Minecraft). The STE features text-chat systems hosted on computers in each workstation. These are linked through a chat server application that is highly customizable for specific display features, mission times, and predefined key-messages. The teammate's workstations are in separate rooms. These design elements were evaluated and iterated through pilot testing. However, designing a team task includes understanding task effectiveness, which includes support for interdependence. The use of Coactive Design yielded a concrete model of interdependence and recommendations for supporting the human-robot dyad.

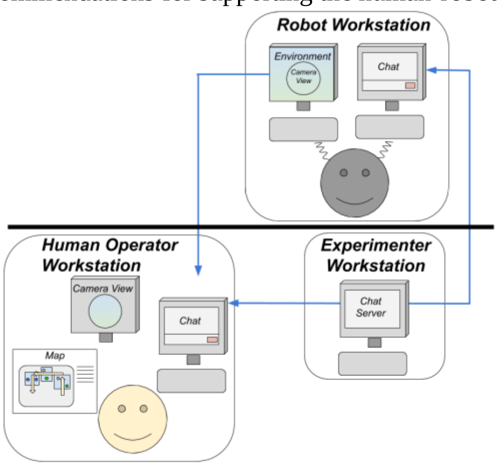

Figure 2. USAR-STE schematic view

Supporting Interdependence

We found opportunities to provide support in OPD through the elaboration of design changes described in the prior section.

Observability. In the USAR-STE, relevant status information included aspects of the environment such as room accessibility, obstacles, and victims. This information was particularly important for integration with the team's plan. Changes were made to the human teammate's map to include a printed route plan. Additionally, space was added to allow the human teammate to track the progress along the plan. Lastly, we included a quick reference guide that allowed participants to review training content during a mission.

Direct observation of robot teammate's view provided real-time feedback of the environmental features. However, our review highlighted that human teammates in USAR may miss details while balancing multiple priorities. A system of communication between the human and the robot could enable greater observability by exchanging status information between agents. Thus, we designed communication in this task to enable status sharing of relevant information, such as room numbers. We also supported observability of plan information by adding the route plan directly to the map. This allowed participants to track the robot's location relative to the plan. *Predictability.* We identified in pilot testing that participants were not aware of their options to communicate with the robot. We revised our training materials to explicitly define conversation topics in the task. These topics included discussing what is happening, the plan, and progress or challenges in achieving goals. Additionally, we defined the robot's capabilities of autonomous planning, situation awareness, and natural language communication. This knowledge-based training was provided in the interactive PowerPoint training. We also added a hands-on training mission, which was a simplified version of the victim identification task guided by an experimenter. The robot would explain how to interact with it throughout this training while the experimenter ensured the participant were adequately trained in their tools. Overall, our knowledge-based and interaction-based training was intended to help participants predict interactions with the robot and the environment.

One concern in pilot testing was inconsistent response times to participant messages (i.e., to type a response). To address this, our chat system was designed with the ability to add up to 30 pre-defined chat messages. This meant that a highly trained experimenter could use the same number of keystrokes to respond to messages predictably. These pre-defined messages covered a broad array of communications in the task. The robot role is designed to understand most possible text-messages a participant may send, including abbreviations (e.g., "u" for "you"). If the human teammate shares unclear information, the robot will state that they do not understand. This communication script enabled information to be exchanged in a more predictable manner.

Finally, the robot had a coordination script for moving and interacting with objects in the virtual environment. These consistent movements, such as following the route plan, waiting before entering a room, and scanning rooms in similar patterns, increase the predictability of the robot behavior. However, we purposefully did not make the robot's behavior completely predictable for our research objectives of testing robot explanations. At times, the robot performed off-plan behaviors for which they could provide explanations. Explanations should improve the human teammate's predictability of the robot over time. Increasing the predictability of normal robot behaviors may also increase participants' sensitivity to off-plan behaviors.

Directability. Our approach to improving directability was to define the robot's capabilities to take any valid action or provide any task-relevant information to the human teammate. Directable actions included interacting with objects, moving to specific rooms, looking in specific locations in the building, and providing information.

The robot provides limited directions to their human teammate during missions. In fact, the only explicit instances of directing the human teammate involve pulling information about plan progress and readiness throughout the mission. Other indirect ways of influencing the human teammate's behavior include making particular aspects of the environment observable for participant's task-relevant activities (e.g., victims, blockages, and openings).

DISCUSSION

In this study, we described our process in redesigning the USAR-STE using the Coactive Design framework. In doing so, we learned several lessons. We had identified the importance of task interdependence for our research objectives early on but lacked a solid framework to make concrete adjustments. Consequently, we had several pilot tests in which the human-robot team acted independently, or the human acted as a supervisor to the robot. Coactive Design was used to identify agent dependencies in the task as well as find ways to modify roles, tools, and the environment. The resulting recommendations for observability, predictability, and directability as well as their interrelation rapidly advanced the design process. For instance, while the robot communicated in natural language, the scope of communication in the task was ill-defined. This reflected in interviews with pilot participants, who generally stated they were not sure about how to interact with the robot. This lack of predictability affected how participants directed the robot (i.e., through suggestions and information pulling) to observe relevant status information. Additionally, Coactive Design helped identify relevant status information that each role needed, which scoped the communication space and helped inform the robot script.

It was critical to our research objective to strike a balance between support for interdependency and task demands. If the robot supported too much of the human teammate's task or completed the task entirely on its own, the task would have resembled supervisory control. Given that we developed the robot teammate role to be capable of planning, situation awareness, and natural language communication, it was possible to design the robot role to complete any aspect of the task besides mapping. Thus, we limited the robot's support to help mostly when asked, and to seldom share information proactively (e.g., sharing plan progress).

Limitations

This work reflects recent trends in autonomy advancement and what program managers and engineers envision future robots to be capable of. Yet, the true capabilities of future robots may be different. The results from data collected in the USAR-STE are to inform future design considerations toward developing effective human-robot teams.

Typically, the design of an STE involves obtaining feedback from subject matter experts (Cooke & Shope, 2004). While three field studies helped depict the environmental conditions of USAR work for our design, a more ecologically valid testbed may have been possible if we had access to information from subject matter experts. Additionally, though we abstracted several features of the cognitive work and tried to reflect the demands of a remote human teammate, it is difficult to argue that an experimental environment would provide the same degree of pressure and risk that would be experienced during the actual work of urban search and rescue.

CONCLUSION

The USAR-STE is presented as a testbed for studying questions related to future autonomous robots for human-robot teamwork in victim identification tasks. Overall, the design of any STE for complex work is challenging and may benefit from systematic design frameworks. In our case, The Coactive Design framework supported developing interdependence in our STE. Our next step is to investigate the role of robot explanations in team situation awareness within this task environment.

ACKNOWLEDGMENTS

This work was partially supported by AFOSR FA9550-18-1-0067 and ONR N000141612892. We acknowledge the assistance of Steven M. Shope and Paul Jorgenson, Sandia Research Corporation who developed and provided the chat system.

REFERENCES

Bartlett, C. E., & Cooke, N. J. (2015). Human-robot teaming in urban search and rescue. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 59, No. 1, pp. 250-254).

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

Chakraborti, T., Sreedharan, S., Zhang, Y., & Kambhampati, S. (2017). Plan explanations as model reconciliation: Moving beyond explanation as soliloquy. *arXiv preprint arXiv:1701.08317*.

Clock circuit/Hopper clock (n.d.).Retrieved from https://minecraft.gamepedia.com/Clock_circuit/hopper_clock

Chiou, E. K., & Lee, J. D. (2016). Cooperation in human-agent systems to support resilience: A microworld experiment. *Human factors*, *58*(6), 846-863.

Cooke, N. J., & Gorman, J. C. (2009). Interaction-based measures of cognitive systems. *J. of cognitive engineering and decision making*, *3*(1), 27-46.

Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013). Interactive team cognition. *Cognitive science*, *37*(2), 255-285.

Demir, M., Amazeen, P. G., McNeese, N. J., Likens, A., & Cooke, N. J. (2017). Team coordination dynamics in human-autonomy teaming. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 236-236).

Demir, M., McNeese, N. J., & Cooke, N. J. (2018). Dyadic Team Interaction and Shared Cognition to Inform Human-Robot Teaming. *Proceedings of the Human Factors and Ergonomics Society*, *62*(1), 124–124.

Feng, W., Zhuo, H. H., & Kambhampati, S. (2018). Extracting Action Sequences from Texts Based on Deep Reinforcement Learning. *arXiv preprint arXiv:1803.02632*.

Gorman, J. C., Cooke, N. J., & Amazeen, P. G. (2010). Training adaptive teams. *Human Factors*, *52*(2), 295-307.

Johnson, M., Bradshaw, J. M., Feltovich, P. J., Jonker, C. M., Van Riemsdijk, M. B., & Sierhuis, M. (2014). Coactive design: Designing support for interdependence in joint activity. *Journal of Human-Robot Interaction*, *3*(1), 43-69.

Klien, 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 Intelligent Systems*, *19*(6), 91-95.

Kruijff, G. J. M., Pirri, F., Gianni, M., Papadakis, P., Pizzoli, M., Sinha, A., ... & Priori, F. (2012, November). Rescue robots at earthquake-hit Mirandola, Italy: A field report. In *Safety, security, and rescue robotics (SSRR), 2012 IEEE international symposium on* (pp. 1-8). IEEE.

Liu, Y., & Nejat, G. (2013). Robotic urban search and rescue: A survey from the control perspective. *Journal of Intelligent & Robotic Systems*, *72*(2), 147-165.

McNeese, N. J., Demir, M., Cooke, N. J., & Myers, C. (2018). Teaming with a synthetic teammate: insights into human-autonomy teaming. *Human factors*, *60*(2), 262-273.

Miller, T. (2018). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*.

Riek, L. D. (2012). Wizard of oz studies in hri: a systematic review and new reporting guidelines. *J. of Human-Robot Interaction*, *1*(1), 119-136.

Sreedharan, S., Chakraborti, T., & Kambhampati, S. (2018, January). Handling Model Uncertainty and Multiplicity in Explanations via Model Reconciliation. In *ICAPS* (pp. 518-526).

Whitman, J., Zevallos, N., Travers, M., & Choset, H. (2018, August). Snake robot urban search after the 2017 mexico city earthquake. In *2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)* (pp. 1-6). IEEE.