Designing for Robust and Effective Teamwork in Human-Agent Teams

Fei Gao, Missy Cummings and Erin Solovey

Abstract We investigated the impact of team structure, task uncertainty, and information-sharing tools on team coordination and team performance in human-agent teams. In applications such as search and rescue, command and control, and air traffic control, operators in the future will likely need to work in teams together with robots. It is critical to understand how these teams could be robust against uncertainty and what influences team performance. We conducted two experiments in which teams of three operators controlled simulated heterogeneous robots on the same testbed. Experiment 1 investigated the impact of team structure and uncertainty of task arrival processes on team coordination and performance. Experiment 2 explored the usage of information-sharing tools under different uncertainty levels. In Experiment 1, it was found that divisional teams were more robust against the uncertainty on task arrival processes. However, this robustness was achieved with an overall worse performance compared to functional teams. Three reasons for the degraded performance were identified, namely duplication on task assignment, under-utilization of vehicles, and infrequent communication. In Experiment 2, it was found that information-sharing tools reduced the duplication on task assignments, improved overall task performance, and reduced workload. These results provide insights for achieving robust and effective teamwork. This goal can be achieved by using a team structure that could adapt to uncertainties together with effective information-sharing tools. These findings could inform the design of robust teams and the development of information-sharing tools to improve teamwork.

Introduction

With the development of automation technology, operators' tasks often shift from manual control of a single task to supervising multiple tasks and agents, which can require monitoring, coordination, and complex decision-making. However, the required cognitive load for working with multiple agents could easily exceed the capacity of a single operator, even with high levels of automation. There is an increasing demand for teams of humans to perform tasks that are less efficiently done or impossible to do by individual human.

Teams have the potential of offering greater adaptability, productivity, and creativity than any one individual can offer and provide more complex,

Fei Gao
77 Massachusetts Ave., E40-206, Cambridge, MA, 02139
e-mail: feigao@mit.edu

Missy Cummings
Duke Institute for Brain Sciences, Hudson Hall, Duke University, Durham, NC 27708
Phone: (919) 660-5306
e-mail: m.cummings@duke.edu

innovative, and comprehensive solutions (Gladstein 1984). However, working as a team imposes extra workload related to coordination and communication, and teams can fail for many reasons (Salas and Fiore 2004). Factors such as a poor combination of individual efforts, a breakdown in internal team processes (e.g., communication), and an improper use of available information have been identified as potential sources of team failure (Salas et al. 2005).

Effective teamwork in highly dynamic environments requires a delicate balance between giving agents the autonomy to act and react on their own and restricting that autonomy so that the agents do not work at cross purposes (Work et al. 2008). To achieve robust and effective teamwork, we must understand the nature of such teamwork, including team structure, team processes and dynamics, and their impact on team performance. In this study, we investigated the teamwork across multiple operators working together with multiple heterogonous autonomous vehicles using two experiments. In Experiment 1, the impact of team structure on team performance under different levels of uncertainty was investigated. Reasons for inefficient coordination were identified. In order to improve the coordination, in Experiment 2, four interface design conditions were compared using the same testbed to see whether facilitating information-sharing within the team could improve team coordination and team performance.

Related Work

Autonomous systems affect teamwork in two primary ways. First, autonomous systems affect the way a task can be completed through task interdependence and work assignment. Second, automation affects how information is presented and shared among team members, which further influence the way team members coordinate and communicate. Previous work has identified many important factors that include team structure, shared mental model and team situation awareness, as well as communication.

Team Structure

Team structure is an important factor hypothesized to affect team effectiveness (Lewis et al. 2011). Team structure affects the manner in which the task components are distributed among team members (Naylor and Dickinson 1969), as well as team communication and coordination. The team structure that is suitable for a specific scenario largely depends on the task characteristics and resources available (Macmillan et al. 2004). For a team of operators working together with multiple heterogeneous autonomous vehicles, there are several ways to organize the vehicles. One common method is functional organization, in which individuals specialize and perform certain roles. For example, one person is responsible for searching and another person is responsible for responding to targets. By specialization on the part of each member, groups are able to tackle problems more efficiently. The clear task responsibility also reduces the need for coordination. One major downside of functional organization is the difficulty in shifting workload flexibly to break up unexpected bottlenecks.

Another way to organize the team is divisional organization, in which each working unit can be responsible for all type of tasks. In divisional organization, each member is allocated with some resources of each type. By creating self-contained tasks, it reduces the amount of information processed within an organization when the level of uncertainty is high (Galbraith 1974). For example, a company can have several divisions each responsible for one product. Each division has its own set of functional units like research, design, marketing etc. Divisional structure was designed in order to have a fast response to the market (Macmillan et al. 2004). In one command and control scenario, it was found that the effectiveness of teams using the divisional and functional structures depends on the nature of the tasks to be accomplished and the uncertainty in the situation. Specifically, functional teams perform better when the environment and tasks are predictable. Divisional teams have a higher level of robustness and perform better when the environment and tasks have more uncertainty (Macmillan et al. 2004).

Shared Mental Model and Team Situation Awareness

Whether working as an individual or a team, developing and maintaining a high level of situation awareness (SA) is critical in autonomous vehicle control. SA includes perception of the elements in the environment, comprehension of the current situation, and projection of future status (Endsley 1995). Team coordination poses extra SA requirements. Team members need to be aware of their teammates' situation in addition to their own. If two or more team members need to know about a piece of information, it is not sufficient if one knows the information perfectly while others know nothing at all. The degree to which each team member possesses the SA required for his or her responsibilities was defined as team SA (Endsley 1995). To develop team SA, each team member needs to understand the impact of other team members' task status on one's own functions and the overall mission, as well as how their own task status and actions impact on other team members. Based on such comprehensions, team members should also be able to project what fellow team members will do to plan their actions effectively (Endsley and Jones 1997).

The quality of team SA affects team communication, coordination and performance directly or indirectly. Blickensderfer, Cannon-Bowers, and Salas (1997) found that teams that shared expectations regarding member roles and task strategies before a radar tracking task communicated more efficiently during the task and achieved higher overall performance outcomes. Previous research identified several ineffective team SA processes that should be avoided, including one member leading others off, insufficient sharing of pertinent information, failure to prioritize the tasks and adhere to the main goal, and relying on unreliable expectations (Bolstad and Cuevas 2010). There are several ways to improve team SA. From system design perspective, team SA can be improved by tools to facilitate team communication, shared displays or shared environments, etc.

Situation awareness has become a critical top of concern when designing a human-machine interface. A system that improved situation awareness should provide a proper amount of information accurately based on the user's situation awareness needs. For teams, one important aspect of the interface design is to facilitate information-sharing among team members. Efforts had been made to improve team situation awareness using team displays for command and control teams, forest fire fighting teams, teams in operating rooms as well as in workspace (Biehl et al. 2007; Bolstad and Endsley 1999, 2000; Parush et al. 2011; Parush and Ma 2012). A team display used in forest fire fighting scenario improved situation awareness and performance, particularly when there was a communication breakdown (Parush and Ma 2012). However, it was also found that the use of an abstracted shared display enhanced team performance, while the use of shared displays that completely duplicated the other team members displays decreased performance and increased workload (Bolstad and Endsley 2000). Despite the potential benefits, a team display aiming to enhance team situation awareness should be carefully designed to avoid an overly complex interface.

Communication

Communication, an important coordination mechanism, influences the share of information among team members. Communication relates to building an accurate understanding of team members' needs, responsibilities, and expected actions (Macmillan et al. 2004), which allows them to anticipate one another's needs so that team members can coordinate effectively (Stout et al. 1999). If the team members don't communicate sufficiently, they may not develop a clear understanding of the situation, which may result in delayed actions, errors, and a suboptimal distribution of team resources.

On the other hand, communication takes time and carries a coordination cost. It can represent a type of process loss, which means team performance could be lower than the combination of individual performance due to the extra work on team coordination (Steiner 1972). Research has investigated the negative effects of communication in terms of increased workload and decreased performance. In a team of six persons performing a joint task force mission of air-based and sea-based operations, it was found that a lower need for coordination and a lower communication rate were associated with better performance (Macmillan et al.

2004). In another study, excessive word usage was found to have a negative association with team performance (McKendrick et al. 2013).

The appropriate amount of communication is impacted by factors such as task characteristics, team structure, level of workload, etc. (Bowers et al. 1996; Oser et al. 1991). In general, an ideal balance is to communicate enough to exchange the required information without too much increase on coordination overhead. In order to reduce coordination overhead, a strategy teams often use under high workload is to switch from explicit communication to implicit coordination (Orasanu 1990; Stout et al. 1996). Instead of communicating explicitly to control teammates, such as proposing actions, prompting or requesting information (Entin and Serfaty 1999), implicit coordination is adopted. Some effective implicit coordination strategies include periodic situation assessment, offering information without explicit request, and providing information to indirectly guide teammates' actions are some effective implicit communicate strategies (Entin and Serfaty 1999; Orasanu 1990; Shah and Breazeal 2010; Stout et al. 1996).

Experiment 1: Team Structure and Robustness

As discussed previously, different team structures have advantages depending on the nature of the task and environment. Human-agent teams often work under uncertainty. One major source of uncertainty is task load. The arrival time and types of tasks are often unpredictable and balancing the tasks and workload among team members can significantly affect outcomes. In Experiment 1, we investigated the communication and coordination process and performance of human-agent teams with different team structures and under different levels of task load uncertainly.

Testbed

The software testbed for our study is Team Research Environment for Supervisory Control of Heterogeneous Unmanned Vehicles (TRESCHU), a video game-like simulation of unmanned vehicle control by a team of three operators. The simulation included three ground control stations, with one operator assigned to each station controlling three vehicles. The three operators were referred to as Alpha (A), Bravo (B), and Charlie (C). The scenario was search and rescue operation in which operators must identify contacts as either friendly or threats, and respond to them appropriately – friendlies must be dropped aid packages, and threats must be neutralized.

Each mission scenario required a team of operators and autonomous vehicles to handle contacts that appeared intermittently over the map. There were three ground control stations, with one operator assigned to each station controlling three vehicles. New contacts appeared on the map as Unknowns. Operators were required to send a scouting vehicle to identify the unknown as either Friendly or Threat, after which they could assign a rescue vehicle or a tactical vehicle to respond. Once assigned, the vehicle would autonomously travel to that particular contact location on the map in a straight line and would continue until either the vehicle reached its assigned destination or the operator re-assigned the vehicle elsewhere.

Once a vehicle arrived at a contact, the operator performed one of three tasks that depended on the vehicle and contact type: scout, rescue, or tactical. All three tasks involved a birds-eye view of the terrain. In the scout task, there were two items of interest presented in the upper left corner of the screen. The operator's task was to select the one item that appears somewhere in the overview map. The rescue task involved controlling the position and movement of crosshairs and dropping aid packages to friendly contacts on the ground. The crosshairs were relatively steady but the projectiles were falling slowly and susceptible to the wind. The tactical task required the operator to center the crosshairs over a stationary threat on the ground and to neutralize it.

The scout task required visual search ability. The rescue and tactical tasks required hand-eye coordination. These three tasks had different levels of difficulty. Rescue tasks were the hardest and took the longest time. Scouting tasks were the easiest and took the shortest time.

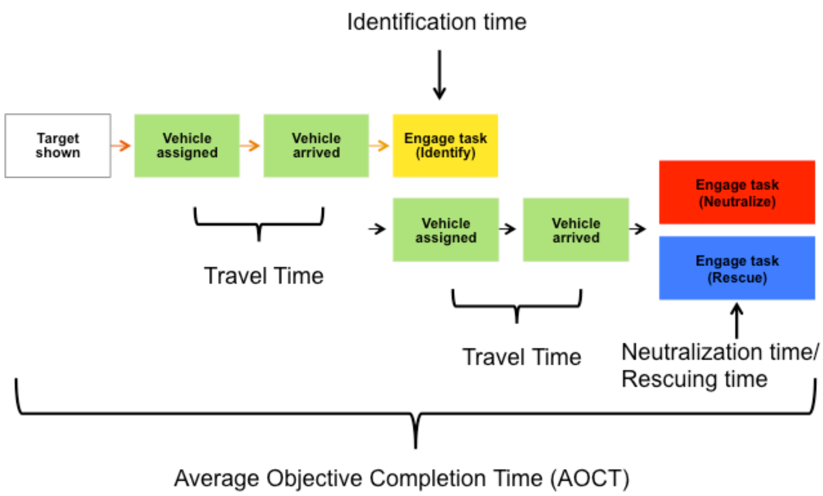

Fig. 9.1. Task workflow

Because of the need to first identify the contact before completing one of the other two tasks, two vehicles were required to complete each scenario, one for scouting and one for the rescue or tactical task. The rescue or tactical vehicle could be assigned before or after the scouting task. The timeline for processing a task is shown in **Fig. 9.1**. The time between the appearance of an unknown contact and the time it was neutralized or aided was called objective completion time. Team performance was measured by averaging the objective completion time (referred to as AOCT later) of all contacts during the mission.

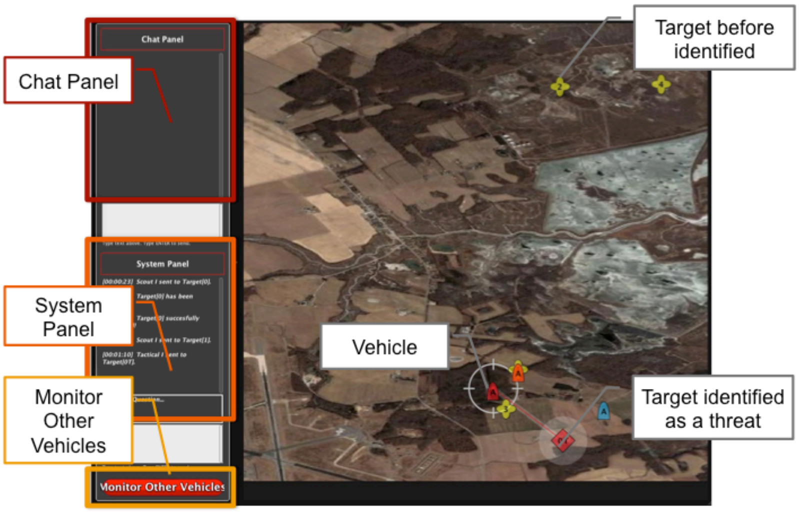

Fig. 9.2. Team Research Environment for Supervisory Control of Heterogeneous Unmanned Vehicles (TRESCHU) Interface

The interface contained four parts: a Map, a Chat Panel, a System Panel and a Monitor Other Vehicles Button (**Fig. 9.2**). The Map represents the geographical area that the operators were responsible for, the three vehicles under their control, and all the contacts that need to be handled by the team. TRESCHU has three kinds of contacts (Unknown, Friendly, Threat) and three corresponding vehicles (Scout, Tactical, Rescue). Types of vehicles were differentiated by color, and types of contacts were differentiated by both color and shape. The operators were able to communicate with each other via instant messaging in a chat interface window.

Operators typed messages into the chat, which would then appear on all the other operators' chat panels instantly. Chat messages were labeled with the operators unique IDs, which corresponded to the labels for each operator's vehicle icons. The System Panel would occasionally send messages to a particular operator, such as a confirmation message that the operator had assigned a particular vehicle to travel to a particular location. It also sent the operator an error message when he or she attempted to claim or engage a vehicle already claimed or engaged by other operators. Operators were unable to see the location of other operators' vehicles unless they explicitly commanded the interface to do so with the Monitor Other Vehicles button.

Experiment Design

A 2x2 mixed design experiment was conducted where the independent variables were team structure (divisional, functional) and the inter-arrival time of unidentified contacts (constant, erratic). The two conditions of inter-arrival time were designed to simulate different uncertainty levels in task load for human-agent teams. The time between successive exogenous events (the inter-arrival time) was 30 seconds for the constant treatment. For the erratic factor level, the inter-arrival times were generated from a bimodal distribution where the means of the modes were set at 75 seconds and 225 seconds from the start of the trial, with a standard deviation of 15 seconds. We use this instead of a more random arrival process (e.g. passion arrival process) to generate two peak times. Ten teams of three participants each completed all four treatments. The experimental trials had a total of 16 exogenous events (unidentified contacts emerging).

The second independent variable was team structure. A functional team was one where the operators have rigidly defined roles and responsibilities. For instance, when all of the vehicles of one type were assigned to one and only one operator, then that operator was given the full responsibility for performing the tasks that only that vehicle can do. This formed sequential dependency in which team members performing tasks in the later steps had to wait until the tasks in the earlier step were completed. If one of each vehicle type was allocated to a single operator instead, then that team structure would be considered divisional since any operator can perform any task that arises, provided that he or she had an appropriate vehicle available. This formed pooled dependency in which independent works of team members were combined to represent team output (Thompson 1967).

Results

Thirty participants participated in the experiment and were tested in groups of three. They went through the four combinations of independent variables in randomized sequence with each session lasted about 15 minutes. The initial experiment results showed that functional teams performed significantly better than divisional teams (($F(1, 24) = 1.484$, $p < 0.01$), as shown in **Fig. 9.3.** The interaction effect was also significant with functional teams performed better with constant arrival, while divisional teams performing better with erratic arrival ($F(1, 24) = 10.47$, $p = 0.04$) (Mekdeci and Cummings 2009). Although divisional teams showed their robustness against the uncertainty of task arrival, their performance was not as good as desired, especially under the constant arrival process. In this effort, we further investigated the teamwork process to identify several reasons for the poor performance: duplicated work, underutilization of vehicles, and infrequent communication.

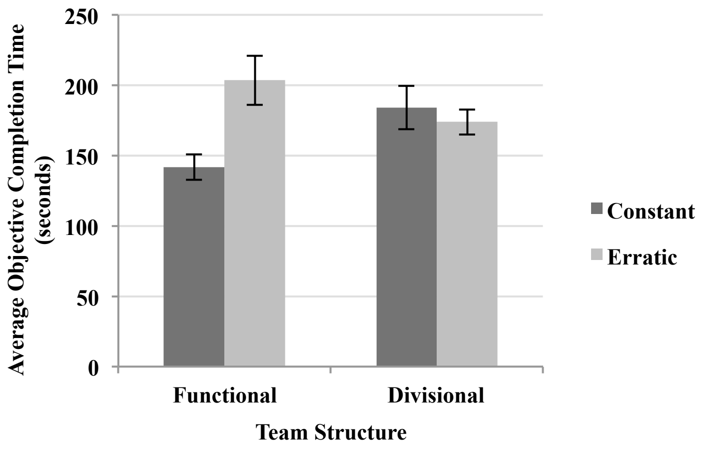

Fig. 9.3. Team performance under different conditions

Duplicated Work

Further analysis on teamwork process shows that teams that had worse performance also tended to have poor team coordination. One example of such poor coordination occurred in divisional teams, where we observed duplicated vehicle assignments. We analyzed who assigned vehicles to each contact after it appeared and later after it was identified. In **Fig. 9.4**, each column represents each contact. Each square in a column represents a vehicle assigned to this contact. Green, blue and red corresponding to operator 1, 2 and 3. We can see that often multiple operators assigned identification vehicles to the same contact before the contact was identified. Similarly, there were several times that multiple operators assigned vehicles to the same contact after it was identified. This resulted in a waste of resources.

This required conflict resolution within the team by explicit communication, which cost time. In the communication transcript, we observed messages such as these:

Bravo: *B 1(Meaning Bravo is taking contact 1).*
Alpha: *A 1(Meaning Alpha is taking contact 1).*
Charlie*: Let B take it - he's closer, A you take 2*
Alpha: *Redirecting to 2*

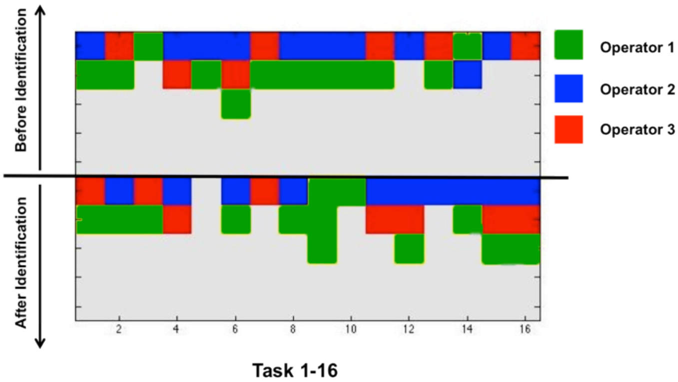

Fig. 9.4. Vehicles Assigned to Each Task by Operators

We can see that both Bravo and Alpha wanted to work on contact 1. This conflict was resolved via communication as Charlie asked Bravo to work on contact 1 and Alpha on contact 2. Note that there was no explicit assignment for a team leader. Leadership emerged organically within the team. Communication was necessary in this case but cost extra time. If the conflict was not resolved, it may have happened that some tasks would have multiple operators working on them while others were ignored. This kind of duplication was quantitatively analyzed based on vehicle assignment conflicts and task engaging conflicts. Vehicle assignment conflicts happened when two vehicles were assigned to the same contact. A Mann-Whitney test shows that team structure had a significant effect on the number of vehicle assignment conflicts (Chi-sq = 4.89, df = 1, p = 0.027) with more conflicts in Divisional teams (Mean = 16.65, SD = 6.38) and less conflicts in Functional teams (Mean = 9.75, SD = 11.36). Task engagement conflicts happened when two operators tried to engage the same contact to perform payload tasks. Similarly, team structure had a significant effect (Chi-sq = 23.92, df = 1, p<0.001) with more conflicts in Divisional teams (Mean = 3.40, SD = 2.80) and zero conflicts in all Functional teams.

Under Utilization of Vehicles

A second reason for the poor performance of divisional teams was the under-utilization of the vehicles. **Fig. 9.5** shows the working process of a divisional team with constant task arrival. Each column is the timeline of one emergent contact from its appearance until it was neutralized or provided aid packets. Green is for vehicle travel time, yellow is for identification time, red is for time to neutralize the contact, and blue is for the time to complete rescue task. Dark grey is for assignment waiting time, during which the contact was waiting to be assigned a vehicle.

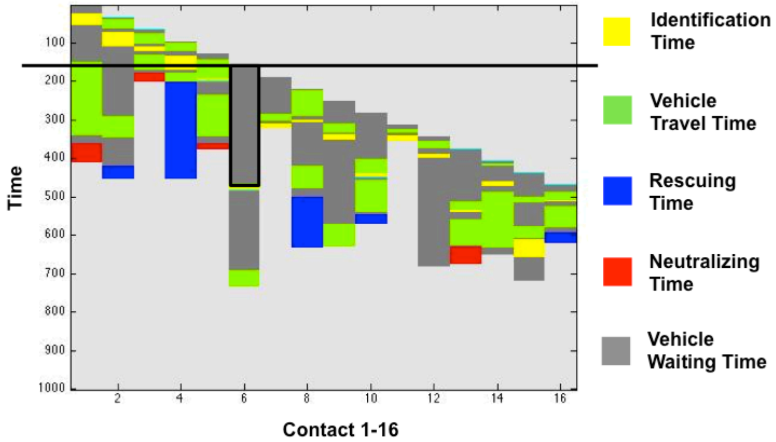

Fig. 9.5. Timeline of Task Completion in a Divisional Team

Table 9.1. Coordination and Vehicle Utilization

Time Flow	Alpha	Bravo	Charlie
Start	Assigned to 4	Arriving at 2T	Identifying 3
Process	"ok, I will get 4 too." "and 3F"	"alpha, you take 0T"	
	Assigned to 0T	Destroy 2T	
		"i got 4T"	
	Assigned to 3F	Assigned to 7	Assigned to 4, 3F, 4T, 0T
			"who got 0F"
			Assigned to 1F
Result: #Idle	0	2	3

The dark grey periods in **Fig. 9.5** are nonproductive time, which happened a lot for this team. Contacts are numbered, with a letter T added after the identification for threats or F for friendly. For example, 1 is an unidentified contact. It is updated as 1T if identified as a threat or 1F if friendly. The longest idle time was highlighted by the black box in Figure 5. We looked at the log of communication and the actions of operators during this time, which are summarized in **Table 9.1**. Operator Alpha assigned a vehicle to an unknown contact 4, and later to a threat contact 0T and a friendly contact 3F. All three vehicles operated by Alpha were busy. He also reported his actions to his teammates via chat messages. Operator Bravo was working on threat contact 2T. After that, he worked on another threat contact 4T, and later assigned a vehicle to an unknown contact 7. Only one vehicle controlled by Bravo was busy at one time. Operator Charlie assigned vehicles to several contacts (4, 3F, 4T, and 0T) after he finished identifying unknown contact

3. However, all of these contacts had already been claimed by the other two team members. Operator Charlie then asked about task allocation and assigned a vehicle to contact 1F. During this time, none of the vehicles operated by him were busy and none of them were assigned to contact 5. From these we can see the vehicles were not used to their full capacity. While there were enough idle vehicles, some tasks had no vehicle assigned to them.

Infrequent Communication

We found that chat density had an influence on the task assignment waiting time, which is the nonproductive time between the appearance of a contact and the time it was assigned a vehicle. We conducted a partial correlation analysis for Average Objective Completion Time (AOCT), average assignment waiting time, and the number of chat messages. Team structure, arrival process, and trial sequence were controlled in order to separate the influence of communication. Although chat density did not have a significant correlation with the overall objective completion time, it negatively correlated with average assignment waiting time with $r = -0.427, p = 0.009$. Average assignment waiting time correlated with AOCT with $r = 0.392, p = 0.016$. In other words, communication indirectly influenced team performance by reducing the nonproductive time. Thus teams that communicated infrequently likely led to poorer performance.

In this study, divisional teams were designed to create working units with a higher level of autonomy. They showed robustness against uncertainly but poorer performance overall. Three reasons related to team coordination were identified for the poor performance, namely duplicated work, underutilization of resources and infrequent communication.

Experiment 2: Information-Sharing

Working as a team on time-constrained tasks in an uncertain environment brings many challenges. To achieve high performance, team situation awareness, communication and coordination are critical. It is important that team members understand what their teammates are doing and get the required information in a timely manner. While explicit communication can be time consuming, supporting implicit information-sharing via the user interface could be more effective and efficient.

These considerations motivated a second experiment. Based on the three reasons identified for the poor performance in divisional teams in Experiment 1, we conducted Experiment 2 to study how teams can be structured and supported by technology to be both flexible and efficient. Specifically, we investigated the effect of enhanced information-sharing tools under different uncertainty levels in divisional teams.

Independent Variables

A 2x4 repeated measures experiment was conducted. The first independent variable was the uncertainty level, which was defined by inter-arrival time of unidentified contacts (constant, erratic), as in Experiment 1.

The second independent variable was information-sharing condition. It was designed to see whether team performance could be improved by enhancing team situation awareness, implicit coordination and communication through information-sharing. In Experiment 1, it was found that *divisional* teams were more robust to uncertainty but had overall performance degradation when compared to *functional teams,* due to duplication of task assignment, under-utilization of resources, and infrequent communication. Information sharing (or lack thereof) was the source of this discrepancy.

There are different ways to share information in teams. The most common way is explicit communication, which is supported by chat panel in the testbed. However, explicit communication is time consuming and poses extra workload on human memory. In Experiment 1, teams with worse performance also communicated less. Without the aid of an information-sharing feature, team members had to rely on explicit communication, which was inhibited when the task load was high, ultimately resulting in degraded performance. To this end, four conditions were compared in the experiment: baseline, icon differentiation, status list, and both:

- In the baseline condition, no additional information-sharing mechanism was provided.
- In the icon differentiation condition, contacts that had been assigned a vehicle would change color to white and reduce in size, as shown in **Fig. 9.6**. We wanted to separate the contacts that had already been claimed from others. We used the color white because it is neutral and has enough contrast with the darker background. We also wanted to minimize them so that team members could devote their resources to contacts that had not yet been claimed.
- In status list condition, the IDs of contacts that had not been assigned any vehicle were listed in a table by three categories, as shown in **Fig. 9.7**. The list could be hidden by clicking the checkbox on top of the list. The status list conveyed the same information as in icon differentiation. However, there were two major differences in terms of visualization. Unlike using white to differentiate the contact, people could still tell the type of contact from the icon color. However, looking at the status list required longer eye movement and extra time in visual search to match an ID in the list to an icon on the map.

- In the last condition, both the icon differentiation and the status list were presented.

Fig. 9.6. Icon Differentiation

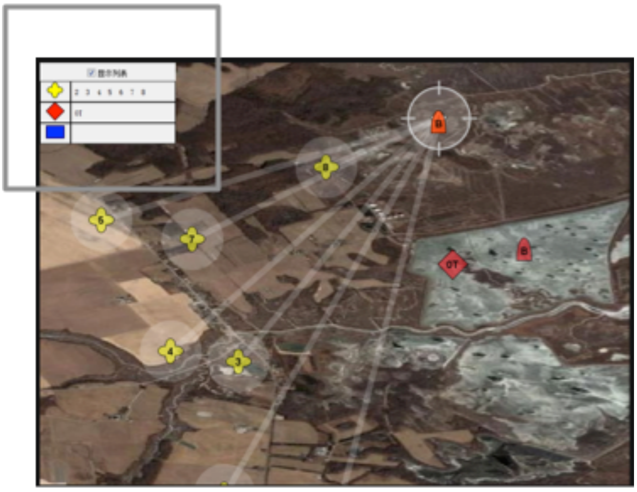

Fig. 9.7. Status List

Dependent Variables

Dependent variables include team performance, measures of team coordination processes, subjective workload and user preference. Team performance was measured using the Average Objective Completion Time (AOCT), as in Experiment 1. Segments of the objective completion time, including identification time, neutralizing task time, rescue task time, vehicle travel time, and assignment waiting time, were also calculated. Measures of team coordination processes included time spent in monitor mode, communication time, vehicle assignment conflicts, and task engagement conflicts. Subjective workload was measured using the NASA-TLX rating (Hart 2006). User preference was user's ranking of the four information-sharing conditions based on their preference.

Participants

A total of 81 participants, participated in the experiment. Participants were tested in groups of three. Data from three groups were removed because the test was not completed due to system errors. The remaining 72 participants were aged 18 to 28 years old, with an average of 22.3 years old and a standard deviation (SD) of 1.64. Among them, 27 were female and 45 were male. All of the participants were undergraduate or graduate students. Participants were asked for the number of hours they played electronic games per month on average. They also rated their experience on visual searching games, first person shooting games, real time strategy games, and team games respectively on a five-point Likert scale, with high values indicating more game experience. Self-report team game experience was found to have a significant correlation with AOCT ($r = -0.417$, $p = 0.042$), and average workload in the team ($r = -0.494$, $p = 0.014$). In other words, teams that had more team game experience tended to finish the tasks faster and with lower workload. Game experience on other categories was not significantly correlated with either performance or the workload.

Procedure

Participants were tested in groups of three under a single uncertainty level: either constant or erratic arrival process. The participants were in the same room, but could not see other team members' displays. The experiment began with a training session introducing the testbed interface, tasks, and the mission goal. Participants then practiced for a complete session under the baseline condition. After that, they were instructed to discuss their team strategy for five minutes before beginning the four test sessions with different information-sharing conditions. Their sequences were randomized and counter-balanced. Before each test session, the information-sharing feature used in this session was explained. Each trial was completed when all 16 contacts were processed. Subjective workload was measured using NASA-TLX rating at the end of each session. Participants ranked the four conditions based on their preference and provided comments after all the sessions were completed.

Results

Data logged during the experiment were post processed to obtain performance and process data. The results were analyzed based on the four experiment sessions from four aspects: task performance, team coordination measures, subjective workload and user preference. Data in the training session was not included in the analysis.

Team Performance

MANOVA was used for the analysis of team performance. No multivariable outlier was found. The assumption of homogeneity covariance matrices was satisfied across the four information-sharing conditions (Box's M = 72.82, df = 63, $p = 0.43$), but not across the two uncertainty levels (Box's M = 61.8910, df = 21, $p < 0.001$). However, since all the cells had equal sample size, MANOVA was still used because of the correlation among dependent variables. Significant differences were found among the two uncertainty levels on the dependent variables (Pillai's criterion = 0.703, $F(6, 22) = 6.72$, $p < 0.001$). The combined dependent variables were also significantly affected by the information-sharing condition (Pillai's criterion = 0.66, $F(18, 66) = 2.99$, $p < 0.001$). The interaction effect between information-sharing condition and uncertainty level was not significant.

Univariate analyses of variance (ANOVA) for each dependent variable were conducted as follow-up tests to the MANOVA. Using the Bonferroni method for controlling Type I error rates for multiple comparisons, an alpha level of 0.008 was used. All the time related variables were measured in seconds. For AOCT (**Fig. 9.8**), information-sharing condition was found to have a significant effect ($F(3, 66) = 4.35$, $p = 0.007$). The condition in which both icon differentiation and status list were presented resulted in the fastest objective completion time (Mean = 242.03, SD = 56.97), followed by status list condition (Mean = 262.20, SD = 62.80), icon differentiation (Mean = 277.38, SD = 48.03) and baseline (Mean = 292.01, SD = 59.02). The main effect of uncertainty level and the interaction effect on AOCT were not significant.

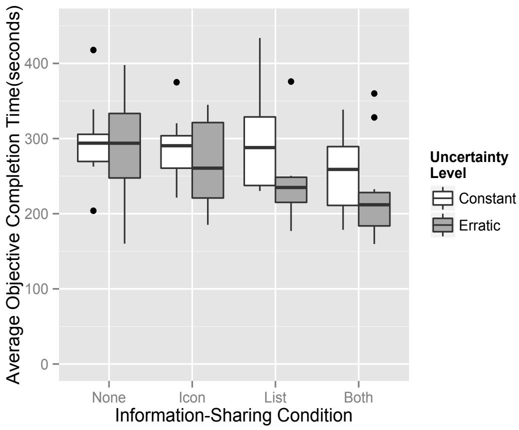

Fig. 9.8. Boxplot of Average Objective Completion Time (AOCT)

Uncertainty level and the information-sharing condition did not have significant impacts on the time to complete payload tasks (identification, rescuing or neutralization), and the assignment waiting time. Total vehicle travel time (**Fig. 9.9**), which was sum of the time between when a vehicle was assigned to a contact to the time this vehicle arrived, was not significantly affected by uncertainly level. The information-sharing condition had a significant effect on travel time (F (3, 66)= 6.10, p < 0.001). The condition with both the status list and icon differentiation had the shortest travel time (Mean = 39.83, SD = 6.63), followed by status list condition (Mean = 42.29, SD = 9.74), icon differentiation (Mean = 47.30, SD = 9.97) and baseline (Mean = 47.75, SD = 7.22). Since the speed of the vehicles were preset by the system, a decrease on travel time means less distance travelled. The reason was the better coordination enabled by information-sharing tools, which either reduced chances that two vehicles travelled to the same contact or matched the contacts with vehicles better based on the distances between them.

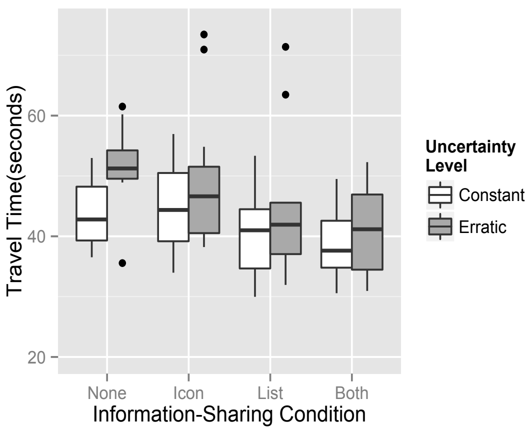

Fig. 9.9. Boxplot of Vehicle Travel Time

Team Coordination

Team coordination was measured from four aspects: time spent in monitor mode, communication time, vehicle assignment conflicts, and task engagement conflicts. A significance level of 0.05 was used.

For total time spent in monitor mode (**Fig. 9.10**), uncertainty level ($F (1, 22) = 4.57$, $p = 0.044$) and information-sharing condition ($F (3, 66) = 3.25$, $p = 0.027$) both had significant effects. Teams with erratic arrival process spent longer time (Mean = 260.71, SD = 97.71) in monitor mode than those with constant arrival process (Mean = 201.17, SD = 74.01). Information-sharing tools reduced the time participants spent in monitor mode. The condition in which both icon differentiation and status list were presented resulted in the shortest time spent in monitor mode (Mean = 206.04, SD = 89.44), followed by status list condition (Mean = 225.25, SD = 102.11), icon differentiation (Mean = 231.58, SD = 89.42) and baseline (Mean = 260.88, SD = 79.79). This was because the information-sharing tools facilitated task assignment and team coordination. Because of the decision-aiding tools, participants could observe the status of contacts more directly from the interface instead of using the monitor mode to figure it out.

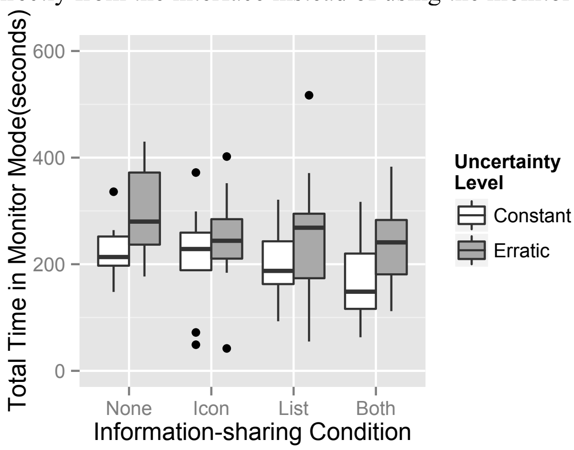

Fig. 9.10. Time Spent in Monitor Mode

Similarly, the information-sharing condition had a significant effect on the amount of communication, as measured by the number of chat messages (Friedman chi-squared = 23.005, df = 3, $p < 0.001$). When there was no information-sharing tool, team members had to communicate more for task assignment and coordination (Mean = 13.75, SD = 12.42). When information-sharing tools were presented, the need for explicit communication was reduced. The average number of chat messages was 5.79 (SD = 6.83) with icon differentiation, 5.21(SD = 5.44) with status list, and 5.00 (SD = 5.18) when both were presented. Arrival process did not impact the amount of communication significantly.

Content of the communication was coded and categorized into five categories: leadership, information prompt, information request, strategy, and other. Leadership contains requests for another team member to work on a certain contact or area, as well as the confirmation and denial of these requests. Information prompt includes reporting the area or contact one is working on, reporting the places one is going to, and negotiation in case of conflicts. Information request includes asking if there is a team member working on a certain contact and who the team member is. Strategy contains discussion on general strategies, such as which area each team member should be responsible for. All the other communications were included in the last category. These are usually not related to the working process, such as open comments and summary about the mission at the end of trials. For each category, number of chat messages was summed across different teams within each information-sharing condition. As shown in **Fig. 9.11**, while the baseline condition had the most communication for all categories, its difference with other conditions was the largest for information prompt. In other words, information-sharing tools reduced communication amount mostly by reducing need to report ones' intentions and actions. The presentation of the information-sharing tools was counterbalanced so the impact from any learning effect was limited.

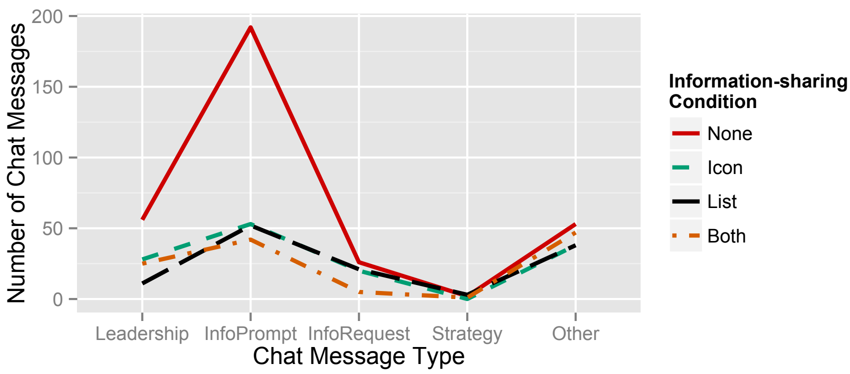

Fig. 9.11. Amount of Communication by Message Type

Vehicle assignment conflicts (**Fig. 9.12**) and task engagement conflicts (**Fig. 9.13**) reflected the result of team coordination more directly. As shown in Study 1, duplicated work was identified as one reason that contributed to the poor performance of divisional teams. In this study, we found that information-sharing tools significantly affected both vehicle assignment conflicts (Friedman chi-squared = 51.859, df = 3, $p < 0.001$) and task engagement conflicts (Friedman chi-squared = 13.268, df = 3, $p = 0.004$). In other words, information-sharing tools reduced duplicated work in teams, making the teamwork more efficient. Among the four information-sharing conditions, icon differentiation and the one with both icon differentiation and status list presented had the least number of conflicts. Status list did not result in much improvement comparing to baseline condition. Uncertainty level also had a significant impact on vehicle assignment conflicts (W = 734.5, p = 0.002) and task engagement conflicts (W = 862, p = 0.032). Erratic arrival process resulted in more conflicts compared to constant arrival process.

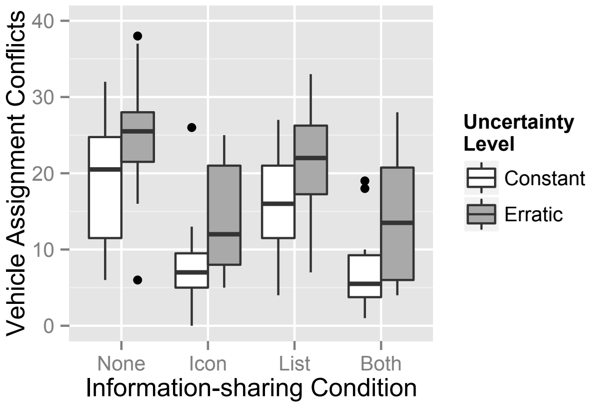

Fig. 9.12. Vehicle Assignment Conflicts Task Engagement Conflicts

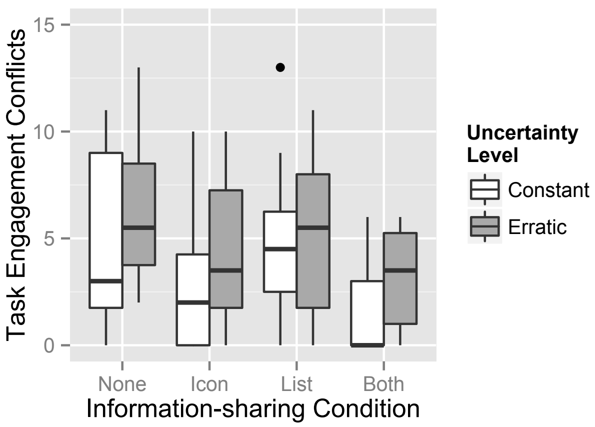

Fig. 9.13. Task Engagement Conflicts

Participants could choose to hide the status list by clicking the check box on its top. We calculated the time that the status lists were hidden for the two conditions when status list was presented. On average, teams chose to hide the status list for 20% of the total mission time (SD = 20.05). The list-hidden time and the total time in monitor mode were positively correlated (r = 0.38, p = 0.007). In other words, participants that chose to hide the status list spent more time in monitor mode as compensation.

Workload

Subjective workload was rated on a scale from zero to one hundred using the NASA-TLX rating. Information-sharing condition was found to have a significant effect on average subjective workload in teams (F (3, 210) = 3.57, p = 0.015). The condition in which both icon differentiation and status list were presented resulted in the lowest workload (Mean = 45.87, SD = 12.74), followed by status list condition (Mean = 47.45, SD = 12.08), icon differentiation (Mean = 47.78, SD = 12.42) and baseline (Mean = 49.76, SD = 12.36). The main effects of uncertainty level and the interaction effect were not significant (**Fig. 9.14**).

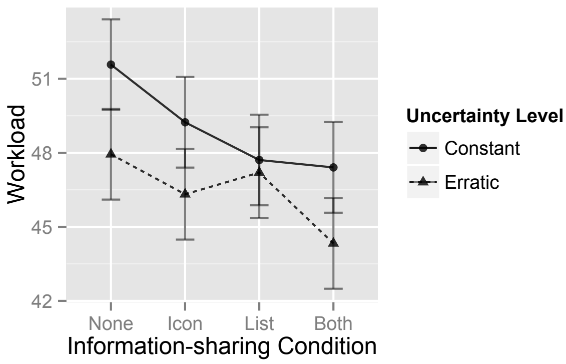

Fig. 9.14. Subjective Workload

User Preference and Comments

Participants' ranking of the four information-sharing conditions were analyzed using Kruskal-Wallis test. There was a significant difference among the four conditions (Chi-sq = 62.05, df = 3, p < 0.001). Icon differentiation and the one with both icon differentiation and status list ranked equally as the top choices, followed by status list, and the baseline condition, which required the participants to use the monitor mode button.

Participants commented that the baseline condition was not convenient, difficult for coordination, easy to have task assignment conflicts, and required extra communication. On the positive aspect, some participants thought this condition was the most interesting because it was challenging. The interface was clearer with no distractions. For icon differentiation, they commented that it was easy to observe which contact has been claimed and assign tasks accordingly.

However, it was difficult to determine the type of contacts with the change of color. For the status list, the information was also useful for team coordination, but was less easy to interpret than the icon differentiation. It was most useful for detecting new contacts or checking whether some contacts had been forgotten. The negative side was that it blocked part of the map, although it could be closed. When both icon differentiation and status list were presented, besides the advantages and disadvantages of each tool, some participants felt the information provided in these two could compensate for each other. On the other hand, the interface was more complex and had more distractions. These comments could be used to further improve the design of information-sharing tools.

Discussion

In Experiments 1 and 2, the ranges of AOCT were different. Overall, participants spent longer time to finish the tasks in Experiment 2 as compared to Experiment 1. This is likely attributed to the difference on screen resolutions used in the two experiments and participants' operating skills. In Experiment 2, the screen resolution was 1024 x 768 compared to 1270 x 960 in Experiment 1. As a result, part of the map could not be shown on the display. Participants had to move the map in the main interface and images for payload tasks to view different parts of them. In Experiment 1, identification time was 12.55s ($SD = 4.29$s) comparing to 16.91s ($SD = 4.88$s) in Experiment 2. In Experiment 1, neutralization time was 13.90s ($SD = 6.11$s) comparing to 23.48s ($SD = 7.62$s) in Experiment 2. Rescuing time was 34.74s ($SD = 14.60$s) in Experiment 1 comparing to 52.88s ($SD = 17.43$s) in Experiment 2. Although a direct comparison between the two experiments was not possible in terms of the visual task, the increase of the overall objective completion time should not affect the assessment on the effectiveness of information-sharing tools.

In Experiment 1, we found that teams communicating infrequently had worse performance. In Experiment 2, teams that performed better with information-sharing tools also had less communication. This is because the information-sharing tools served as an implicit communication channel. With these tools, information on contact status and task assignment could be retrieved directly from the interface, reducing the need for time-consuming explicit communication. When such tools were not available, infrequent communication could not provide sufficient information for team coordination, resulting in duplication on task assignment and suboptimal use of team resources.

Conclusion

In this study, we conducted two experiments to investigate the impact of team structure, uncertainty on task load, and information-sharing tools on team coordination and team performance. In Experiment 1, it was found that divisional teams were more robust against the uncertainty for task arrival processes in terms of team performance. However, this robustness was achieved with an overall worse performance as compared to functional teams. Three reasons for the poor performance were identified, namely duplication on task assignment, under utilization of vehicles, and infrequent communication. In an effort to achieve robust and effective teamwork, the usage of information-sharing tools was explored in Experiment 2. It was found that information-sharing tools reduced the duplication on task assignments, improved overall task performance, and reduced workload in divisional teams.

The conclusions of this study could be useful for the design of human-agent team structure and the development of tools to support teamwork. Consistent with previous research, divisional teams were better able to cope with uncertainty. This reflects on their robustness against different task arrival process. However, divisional teams could have worse performance if the responsibilities of team members are not clear and the communication is not efficient. By providing information-sharing tools for divisional teams, their performance could be improved by reducing the chances of duplicated work and improving coordination, achieving effective and robust teamwork.

The four information-sharing methods resulted in performance improvement at different levels. All reduced the average time required to complete a task and the workload of operators. The best result was achieved when both a status list and icon differentiation were presented. Although we intended to design the two mechanisms to convey the same information, experimental results showed that they actually compensated for each other. People use these two mechanisms in different ways. Icon differentiation was more effective when people wanted to decide whether to work on a specific task. The status list was more effective when people wanted to get an idea of overall progress and strategically allocate tasks among team members. The specific interface design used in this study was not optimized, which could be improved using further usability studies. The key message is that by facilitating information-sharing among team members, the advantage on flexibility and robustness of divisional teams can be maintained while the disadvantages in terms of coordination cost can be limited.

Reference

Biehl, J. T., Czerwinski, M., Smith, G., and Robertson, G. G. (2007), *FASTDash: a visual dashboard for fostering awareness in software teams.* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.

Blickensderfer, E., Cannon-Bowers, J. A., and Salas, E. (1997), *Training teams to self-correct: An empirical investigation.* Paper presented at the 12th annual meeting of the Society for Industrial and Organizational Psychology, St. Louis, MO.

Bolstad, C. A., and Cuevas, H. M. (2010), Team coordination and shared situation awareness in combat identification. *SA Technologies. Marietta, GA*.

Bolstad, C. A., and Endsley, M. R. (1999), *Shared mental models and shared displays: An empirical evaluation of team performance.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Bolstad, C. A., and Endsley, M. R. (2000), *The effect of task load and shared displays on team situation awareness.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Bowers, C. A., Oser, R. L., Salas, E., and Cannon-Bowers, J. A. (1996). Team Performance in Automated Systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human Performance: Theory and Applications* (pp. 243-266). New Jersey: Lawrence Erlbaum Associates.

Endsley, M., and Jones, W. M. (1997). Situation Awareness Information Dominance & Information Warfare: DTIC Document.

Endsley, M. R. (1995), Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*(1), 32-64.

Entin, E. E., and Serfaty, D. (1999), Adaptive team coordination. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 41*(2), 312-325.

Galbraith, J. R. (1974), Organization Design: An Information Processing View. *Interfaces, 4*(3), 28-36. doi: 10.1287/inte.4.3.28

Gladstein, D. L. (1984), Groups in Context: A Model of Task Group Effectiveness. *Administrative Science Quarterly, 29*(4), 499-517.
Hart, S. G. (2006), Nasa-Task Load Index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(9), 904-908. doi: 10.1177/154193120605000909
Lewis, M., Wang, H., and Chien, S. Y. (2011), Process and Performance in Human-Robot Teams. *Journal of Cognitive Engineering and Decision Making, 5*(2), 186-208.
Macmillan, J., Entin, E. E., and Serfaty, D. (2004). Communication Overhead: The Hidden Cost of Team Cognition. In E. Salas & S. M. Fiore (Eds.), *Team Cognition: Understanding the Factors That Drive Process and Performance*: American Psychological Association (APA).
McKendrick, R., Shaw, T., de Visser, E., Saqer, H., Kidwell, B., and Parasuraman, R. (2013), Team Performance in Networked Supervisory Control of Unmanned Air Vehicles Effects of Automation, Working Memory, and Communication Content. *Human Factors: The Journal of the Human Factors and Ergonomics Society*.
Mekdeci, B., and Cummings, M. (2009, September 21-23), *Modeling Multiple Human Operators in the Supervisory Control of Heterogeneous Unmanned Vehicles.* Paper presented at the PerMIS'09, Gaithersburg, MD, USA.
Naylor, J. C., and Dickinson, T. L. (1969), Task Structure, Work Structure, and Team Performance. *Journal of Applied Psychology, 53*, 10. doi: 10.1037/h0027350
Orasanu, J. (1990). Shared mental models and crew decision making: DTIC Document.
Oser, R. L., Prince, C., Morgan Jr, B. B., and Simpson, S. S. (1991). An analysis of aircrew communication patterns and content: DTIC Document.
Parush, A., Kramer, C., Foster-Hunt, T., Momtahan, K., Hunter, A., and Sohmer, B. (2011), Communication and team situation awareness in the OR: Implications for augmentative information display. *Journal of biomedical informatics, 44*(3), 477-485.
Parush, A., and Ma, C. (2012), *Team displays work, particularly with communication breakdown: performance and situation*

awareness in a simulated forest fire. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Salas, E., and Fiore, S. M. (2004), *Team Cognition: Understanding the Factors That Drive Process and Performance*: American Psychological Association (APA).

Salas, E., Sims, D. E., and Burke, C. S. (2005), Is there a Big Five in Teamwork? *Small Group Research, 36*(5), 555-599. doi: 10.1177/1046496405277134

Shah, J., and Breazeal, C. (2010), An empirical analysis of team coordination behaviors and action planning with application to human‚Äìrobot teaming. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 52*(2), 234-245.

Steiner, J. D. (1972), *Group Process and Productivity*. New York: Academic Press.

Stout, R. J., Cannon-Bowers, J. A., and Salas, E. (1996), The role of shared mental models in developing team situational awareness: Implications for training. *Training Research Journal, 2*(85-116), 1997.

Stout, R. J., Cannon-Bowers, J. A., Salas, E., and Milanovich, D. M. (1999), Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 41*(1), 61-71.

Thompson, J. D. (1967), *Organizations in action : Social science bases of administrative theory*. New York: McGraw-Hill.

Work, H., Chown, E., Hermans, T., and Butterfield, J. (2008). *Robust team-play in highly uncertain environments*. Paper presented at the Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 3, Estoril, Portugal.