Chapter XII
Communications for Agent-Based Human Team Support

Gita Sukthankar
University of Central Florida, USA

Katia Sycara
Carnegie Mellon University, USA

Joseph A. Giampapa
Carnegie Mellon University, USA

Christopher Burnett
University of Aberdeen, Scotland

ABSTRACT

This chapter discusses the problem of agent aiding of ad-hoc, decentralized human teams so as to improve team performance on time-stressed group tasks. To see how human teams rise to the challenge, we analyze the communication patterns of teams performing a collaborative search task that recreates some of the cognitive difficulties faced by teams during search and rescue operations. Our experiments show that the communication patterns of successful decentralized ad-hoc teams performing a version of the task that requires tight coordination differ both from the teams that are less successful at task completion and from teams performing a loosely coupled version of the same task. We conclude by discussing: (1) what lessons can be derived, from observing humans, to facilitate the development of agents to support ad-hoc, decentralized teams, and (2) where can intelligent agents be inserted into human teams to improve the humans' performance.

INTRODUCTION

Teams are a form of organizational structure where the decision-making is a bundle of interdependent activities that involve gathering, interpreting and exchanging information; creating and identifying alternative courses of action; choosing among alternatives by integrating the often different perspectives of team members; implementing a choice and monitoring its consequences. It is well recognized that proficient teams achieve goals and accomplish tasks that otherwise would not be achievable by groups of uncoordinated individuals. While previous work in teamwork theory (Salas and Fiore, 2004) has focused on describing ways in which humans coordinate their activities, there has been little previous work on which of those specific activities, information flows and team performance can be enhanced by being aided by software agents. This chapter reports work on, (a) characteristics and challenges of human teamwork, related to decentralization and self-organization in time stressed situations, (b) study of human teamwork performance that incorporate these challenges in order to establish a baseline, and (c) identification of fruitful ways for agents to aid human teams with these characteristics.

In this chapter, we focus on examining the coordination and self-organization problems faced by decentralized ad hoc human teams. Ad hoc teams are groups that are brought together for the duration of a task and who lack prior experience training together as a team. An ad hoc team can be as simple as a group playing a pick-up soccer game in the park or as complicated as a multinational peacekeeping forces working alongside personnel that lack previous operational experience working together. Much of the previous work on human teamwork for time-stressed situations, most of it in commercial aviation and the military, has focused on (a) teams where the team members already have an assigned role (e.g. a pilot, co-pilot and navigator in cockpit teams), (b) where the team already has a given authority structure, and (c) where the team members were collocated. Recent interest in supporting emergency response teams, military interest in operations other than war, and coalition operations, motivates the need for teams that engage in time stressed tasks, are distributed in space and time, and are ad hoc in their organization.

Some important issues arise in ad hoc teams: when faced with a new task, how do team members that come together as a team for the first time create roles and allocate them to team members, when no organizational structure is exogenously provided? To design and build software agents that can assist ad hoc and self-organizing human teams tackling unfamiliar tasks, we need to address this question. If the supporting agents are insensitive to shifts in the team's organization, they cannot effectively monitor the team's activities. (Please see Chapter 19 for more discussion on the problem of run-time organizational shifts).

Although our research is focused towards the ultimate goal of developing agent assistants for human teams, we believe that this work is also relevant to researchers studying purely agent-based teamwork, especially as agents become increasingly sophisticated and capable of human-like teamwork.

Work in the team literature (Fiore et al 2003) has found that establishing effective communication patterns are the key to creating effective ad hoc agent systems and human teams; this is especially true for distributed ad hoc teams in which communication is cited as a key problem area (Pascual et al., 1999). Our research is an initial step towards the problem of identifying communication patterns of teamwork of ad hoc and distributed teams in time critical situations so that suitable agent aiding strategies could be developed. The identification of the patterns is through communication logs collected from human teams. The results of prior research on team communication have typically been used for developing guidelines for team training. In contrast, we are interested in using team communication results for monitoring team performance by software agents that would be used for team aiding. The focus of our initial human team experimentation is to (a) establish a baseline of human-only teamwork for a given task domain and (b) ascertain the relative importance of different information flows for the team task in order to derive "insertion points" for agent assistance of human teams. These insertion points are not merely limited to coordination and information flows, but include teamwork self-organization, maintenance and task completion.

BACKGROUND

Research in human team performance suggests that experienced teams develop a shared understanding or *shared mental model* to coordinate behaviors by anticipating each other's needs and adapting to task demands (Fiore and Schooler, 2004). Furthermore, for such teams, both tacit and explicit coordination strategies are important in facilitating teamwork processes. Explicit coordination occurs through external verbal and non-verbal communications, whereas tacit coordination is thought to occur through the meta-cognitive activities of team members who have shared mental models of what should be done, when, and by whom (Entin and Serfaty, 1999; Fiore et al., 2001; Hoeft et al., 2006). A team's shared mental model thus allows the team members to coordinate their behavior and better communicate depending on situational demands. Initial theorizing on training shared mental models suggests that for teams to successfully co-ordinate their actions, they must possess commonly held knowledge structures, such as knowledge of teammates' roles and responsibilities along with team tasks and procedures. Due to lack of previous co-training, ad hoc teams face the additional challenge of having to establish shared mental models as they perform the task; not only does this potentially increase the communication demands of ad hoc teams but forming incompatible mental models can hinder progress towards team goals (Cannon-Bowers et al., 1993).

Having agents aid human teams similarly requires the establishment of shared cognition, common ground between humans and agents. Creating this shared cognition between human and agent teammates is the biggest challenge facing developers of mixed-initiative human/agent organizations. The limiting factor in most human-agent interactions is the human's ability and willingness to spend time communicating with agents in a manner that both humans and agents understand (Sycara and Lewis, 2004). Horvitz (1999) formulates this problem of mixed-initiative interaction as a process of managing uncertainties: (1) managing uncertainties that agents may have about the human's goals and focus of attention, and (2) uncertainty that humans have about agent plans and status. Creating agent understanding of human intent and making agents' results intelligible to a human are problems that must be addressed by any mixed-initiative system, whether the agents reduce uncertainty through communication, inference, or a mixture of the two.

Agent Roles in Human Teams

Sycara and Lewis (2004) identify three primary roles played by agents interacting with human teams.

- **Agents support individual team members in completion of their own tasks.** These agents often function as personal assistant agents and are assigned to specific team members (Chalupsky et al., 2001). (Please see Chapter 7, "Software Personal Agents for Human Organizations" for more information about personal assistance agents.) Task-specific agents utilized by multiple team members (e.g., (Chen and Sycara, 1998)) also belong in this category.

- **Agents support the team as a whole.** Rather than focusing on task- completion activities, the agents directly facilitate teamwork by aiding communication and coordination among humans and agents, as well as focus of attention. The experimental results summarized in (Sycara and Lewis, 2004) indicate that this can be a very effective aiding strategy for agents in hybrid teams.
- **Agents assume the role of an equal team member.** These agents are expected to function as "virtual humans" within the organization, capable of the same reasoning and tasks as their human teammates (Traum et al., 2003). This is the hardest role for a software agent to assume, since it is difficult to create a software agent that is as effective as a human at both task performance and teamwork skills.

There are additional research challenges, specific to the team role assumed by the agent. Agents that support individual human team members face the following challenges: (1) modeling user preferences; (2) determining optimal transfer-of-control policies (Scerri et al., 2003); (3) considering the status of user's attention in timing services (Horvitz, 1999). Agents aiding teams (Lenox et al., 1997, 1998, 2000; Lenox, 2000), face a different set of problems: (1) identifying information that needs to be passed to other team members before being asked; (2) automatically prioritizing tasks for the human team members; (3) maintaining shared task information in a way that is useful for the human users. Agents assuming the role of equal team members (Traum et al., 2003; Fan et al., 2005, 2006) must additionally be able to: (1) competently execute their role in the team; (2) critique team errors; (3) independently suggest alternate courses of action. Perhaps because of these challenges, there are very few prior results on human-agent team aiding and teamwork. Examples of tasks that were investigated include target identification (Lenox et al., 1997, 1998), achievement of a military rendezvous plan (Lenox et al., 2000; Lenox, 2000) and delivery of supplies to troops (Fan et al., 2005, 2006). All of this prior work has uniformly found that human-agent teams exhibited superior performance over human-only teams not only in achievement of task objectives but also in performance stability. This finding was typically attributed to the fact that the agents in these works were successful in reducing the cognitive load of the human subjects by performing part of the task.

Team Coordination

Decentralized teams that are connected by some sort of network, called network-centric teams, face exacerbated challenges in teamwork, especially for time critical tasks. Coordination is challenging in network-centric environments because entities are often geographically dispersed and may be unfamiliar with other entities as well as the specific task or mission. This situation leads to what has been called "team opacity" (Fiore et al., 2003) and has been frequently associated with differences in process behaviors, poorer shared understanding, and lean communication, relative to co-located teams (Cooke et al., 2007). In fact, teams often adapt to these situations through spontaneous self-organization of their coordination structure (Cooke and Gorman, 2007).

It is important to note that we do not consider coordination in information theoretic terms (Shannon and Weaver, 1949) in which information is encoded, decoded and passively moved among different team members with some degree of uncertainty based on channel capacity. Rather, coordination involves active communication or mediation among team members in a social network (Friedkin, 1998).

Consequently, our coordination metrics do not measure amount of information passed or uncertainty, but instead extend social network theory or coordination theory by quantifying the effectiveness of coordination patterns.

Team coordination in network-centric teams may be predictive of the performance of the team, and to some degree, the social system in which the team is embedded. However, team coordination is not identical to team performance. Sometimes poor coordination can result in fortuitously positive outcomes and even the best coordination can sometimes fail to prevent a negative outcome.

Based on prior research, coordination improves with team experience and training, but decays over long retention intervals (Cooke and Gorman, 2007). The development of coordination skill is a large part of the development of collective competence of the social group. Coordination, therefore, is a team skill that can be trained. It is also a skill that can be quantified and modeled. The measurement and modeling of the development of coordination in network-centric teams is challenging due to the nonlinearities associated with interactions in complex distributed systems (Cooke et al., 2007), due for example to positive feedback loops among different team members.

AIDING HUMAN TEAMS

One of the goals of agent technology is to create agents that can perform part of the tasks that face the humans so as to improve team performance. Galbraith observed that "the more uncertainty in a task, the more information processing necessary to achieve a given level of performance" (Galbraith, 1977). Hence, having the agents assist either in information processing or decreasing uncertainty should improve the team's performance. Based on experiments of student project teams, Kraut suggests that a human team's resultant state of coordination, defined as the degree to which interdependencies are managed well, is an important predictor of team performance (Kraut et al., 2005). This state of coordination can be created by mechanisms such as communication, shared cognition, and team history. If agents can improve the state of coordination between team members or reduce the cost of achieving a good state of coordination, the team performance should improve.

Ad hoc teams face unique set of barriers to effective team performance, especially if they are also distributed. Pascual et al. (1999) surveyed forty military personnel with extensive experience working in both ad hoc and distributed teams. Communications, achieving situational awareness, engaging in standard teamwork behaviors, and demonstrating leadership were listed as major problems by the majority oft he subjects. Leaders of ad hoc organizations face difficulties in task allocation, anticipating team members' actions, and anticipating team problems. Ad hoc team members commonly experienced communication problems (1) in knowing when to communicate updates to their team members (2) in knowing whom to ask for information (3) in providing and accepting feedback.

Agent Support Systems

The utility of agent support systems has been demonstrated in a number of operational domains. Fan et al. (2005, 2006) have evaluated the use of cognitive agents within a collaborative Recognition-Primed Decision model (RPD) for supporting human teams performing military decision tasks. In one task (Fan et al., 2005), the human teams had to maximize the amount of supplies delivered to the troops by successfully protecting an airport. The agents were able to participate in human teams and increase the amount of supplies delivered when the humans were under time-stress. In a second task, Command and Control teams had to react to incoming threats menacing a metropolis from crowds, insurgents, and improvised explosive devices. The task is difficult because it is 1) real-time and 2) involves context switching. The human-agent C2 teams performed much better than the human-only teams at the same level of task complexity; moreover the human-agent team performances are significantly more stable than the human-only performances.

Prior research for agent aiding in human teams (Lenox et al., 1997, 1998, 2000; Lenox, 2000) has involved two different types of cognitive tasks. In the MokSAF experiments (Lenox et al., 2000; Lenox, 2000), the team mission involves three commanders that, starting from different geographical points, must rendezvous at a particular point and time with a particular force configuration. The planning task is deliberative, iterative and flexible. The commanders must coordinate the number and types of vehicles they plan to move from the individual start points to the rendezvous point. The mission briefing supplied to the commanders provides them with a list of vehicles that should arrive at the rendezvous point. In addition, the commanders are instructed to avoid generating routes that lie on the same path as any other commander, and that they should coordinate their routes through the communication center to avoid this. Each commander selects units for his/her platoon from a list of available units. Commanders have 15 minutes to determine the composition of their platoon, and plan a route from a starting point to the rendezvous point for that platoon. Once a commander is satisfied with the individual plan, s/he can share it with the other commanders and resolve any conflicts. Conflicts can arise due to shared routes, shared resources, or the inability of a commander to reach the rendezvous point at the specified time. The experiments were performed to investigate a number of hypotheses. Can agent-based assistance assist in the completion of team tasks? If assistance is provided in achieving the individual goal, then does this improve the achievement of the team goal? Does agent-based aiding become more effective as the complexity of the intangible aspects of a planning problem increase? Experimental results showed that the agent aiding provided better decision support both for individual route planning and team-based planning as compared to the Baseline (unaided) condition (Lenox et al., 2000; Lenox, 2000).

Work by (Lenox et al., 1997,1998) examined agent aiding in another domain, a target identification task where a moderate fidelity simulation (TANDEM) was used. The TANDEM simulation was developed under the TADMUS (tactical decision making under stress) program of the US Office of Naval Research and simulates cognitive characteristics of tasks performed in the command information center (CIC) of an Aegis missile cruiser. The cognitive aspects of the Aegis command and control tasks which are captured include time stress, memory loading, data aggregation for decision making and the need to rely on and cooperate with other team members (team mode) to successfully perform the task. Instead of interpreting displayed radar signals to acquire diagnostic information about targets, TANDEM participants access this information manually from menus. In the TANDEM task subjects must communicate to exchange parameter values in order to identify and take action on a large number of targets (high workload) and are awarded points for correctly identifying the targets (type, intent, and threat), and taking the correct action (clear or shoot). Extensive experimentation with different types of agent aiding on a large number of human teams (60 teams of three subjects each) concluded that (a) agent aiding significantly increased human team performance, and (b) agent aiding of the team as a whole improved performance more than aiding individual team members.

These evaluations of human-agent teams are encouraging because they demonstrate that agents can produce a measurable difference in human team performance.

Search and Rescue Domain

To create a baseline of decentralized ad hoc team performance in a time-stressed domain, we monitored teams of human subjects performing a collaborative search task in simulation. Search and rescue is a challenging, team task with a potentially high payoff since inadequate team performance can result in fatalities. The collaborative search task that we designed for our experiments, the team scavenger hunt, recreates some of the challenges faced by expert human teams during search and rescue operations. To implement the task, we reconfigured a scenario in the multi-player game and battlefield simulator, Operation Flashpoint (OFP version 1.96) (Flashpoint, 2001), by customizing the pre-game briefing, map, object triggers, and scoring mechanism.

In the team scavenger hunt, human subjects have to read a map, navigate a 3D simulated environment and recover a collection of objects (bottles) within a bounded amount of time. We modeled a bottle search task since, unlike modeling rescuing of victims, it was possible to insert static objects in the game (bottles) and log the interactions of the human subjects. The bottle search task has many common elements with the search and rescue of victims task. The task is designed to evaluate the team's ability to develop and execute a search plan under time-stress. As an experimental task, the team scavenger hunt offers several advantages: (1) it can be learned and executed within a short period of time by novice subjects; (2) it can be simulated within a variety of test beds; (3) it offers a simple team performance metric: number of objects collected.

We first provide a *task analysis* of how civilian human teams perform wilderness search and rescue operations summarized from (Goodrich et al., 2007; Setnicka, 1980). We assume that many aspects of the task analysis are also applicable to different types of search and rescue teams, such as military teams or emergency response teams. A goal-directed task analysis of wilderness search and rescue operations identified the following list of operational goals and sub-goals (Goodrich et al., 2007). The italicized task elements are also applicable to our simulated collaborative search task.

1. Stage preparation
   (a) Reporting party call
   (b) Activation call
   (c) Assemble (prepare for search)
2. Acquire missing person description
   (a) Gather missing person information
   (b) Determine missing person's intent
3. *Develop search plan*
   (a) Create a perimeter
   (b) *Assign priority to clues*
   (c) *Update map information*
   (d) *Create a priority pattern*
   (e) *Organize resources for search execution*
   (f) *Communicate search plan*
4. *Execute search plan*
   (a) *Follow plan*
   (b) *Find signs (or absence of)*

(c) Keep searchers safe
(d) *Communicate acquired information*

5. Recover victims
(a) First aid for victims
(b) Rescue, extract, or recover the missing person
6. *Debrief search team*
(a) *Determine what happened*
(b) *Evaluate how the team can improve*

Wilderness search and rescue operations pose the following challenges to expert human teams (Goodrich et al. 2007): (1) information overload of the incident commander while assimilating information collected by the field teams; (2) the creation of accidental holes in the search pattern due to poor execution of the search plan by the field teams; (3) poor priority assignments in the search plan due to false clues and hunches. In the next section, we describe our experimental version of the collaborative search task, the team scavenger hunt, which tests the ability of human subjects to collaborate to develop and execute a team search plan in a simulated environment.

Experimental Testbed

The experiments focused on the activity of three human players acting through virtual characters in the Operation Flashpoint (OFP version 1.96) simulated physical environment to find and crush liquor bottles in twenty minute test periods in different experimental conditions.

In OFP, terrain around and including the village of *Flers* on the island of *Normandie* was chosen as the focal point for the one practice and three experimental scenarios. The area is a tract of land that is 512 meters long in a north–south direction (N/S), and 768 meters long in an east–west direction (E/W); in all, 393,216 square meters. On the 2-dimensional (2D) Operation Flashpoint map, this area corresponds to 4 map squares N/S, 6 map squares E/W, where each map square corresponds to 128 meters by 128 meters. Exploratory benchmarks determined that, depending on search technique and ability, it could take a single OFP civilian virtual character from sixty to ninety minutes to explore all 24-map squares of this scenario. In twenty minutes, a civilian character can thoroughly explore roughly ten map squares of the surrounding countryside. The village of Flers occupies four map squares; part of the village is organized in a radial street plan and another part has a N/S, E/W grid of streets and buildings. Given the area and layout, we have observed that it requires from ten to twenty minutes for the virtual civilian character to search the area.

Experimental Task

In this section, we analyze the baseline performance and the communication patterns of human ad hoc teams performing the collaborative search task.

Seventeen teams of three paid subjects, each, were recruited to participate in the pilot study. Human subjects self-assessed and reported their abilities to play first person video games in terms of the following classification: novice, medium expertise, or expert. Combined expertise of the teams varied from “two novices and a medium expert” to a team of “three experts.” (See Figure 2.) Note that expertise in playing video games does not imply expertise in search and rescue.

Each team member played the game through an assigned and dedicated laptop. The seating of the three-team members was such that they could not look at each other's screen. The human subjects were forbidden to share computer screens, note sheets or other such aids — they could only describe their locations, intentions and actions in the game by using verbal communications and the 2D OFP map of Flers. All verbal communications were recorded using TeamSpeak (TeamSpeak, 2001).

Time was taken during a practice session to instruct the players on the key and mouse commands for the game. Players were instructed on how to move their characters, find and crush bottles, query bottle counts, and how to use additional aids that are available to their avatars. After sighting a bottle, a player must move their avatar to within a couple of meters of it in order to crush it and get credit for the crush. When they are close enough to crush the bottle, the command to crush that type of bottle, e.g. **Crush Martini Bottle**, will appear in the player's command menu at the bottom right corner of their screen. Feedback to the player is given in multiple ways: (1) the sound of a vehicle crashing into a wall, (2) puffs of oily black smoke emanating from the morphing bottle, (3) the morphing of the bottle into a crumpled form. If the player queries their bottle count, they will see that it has increased by one.

Once a player has crushed a bottle, the command to crush it is removed from their menu, never to appear again for that bottle, even if they happen upon its crushed remains at a later time. If a player encounters the remains of a bottle that was crushed by a teammate, they can choose to invoke the command to crush it in order to avoid false detection of that crushed bottle at a later time. No penalty was assessed for attempting to crush an already crushed bottle.

The five ways of detecting a bottle are:

1. ***Visual detection***, in which the human player "sees" a bottle via the unmagnified vision of their avatar;
2. ***Magnified visual detection***, in which the human player slightly magnifies (roughly, 3X) their avatar's field of vision;
3. ***Visual detection via binoculars***, in which the avatar uses binoculars for a narrower but more distant field of view;
4. ***Non-visual proximity sensing***, in which the player is notified of a bottle's presence whenever their avatar comes within "sensing range" of the bottle. A bottle is sensed based on the expertise of the OFP avatar and if the player is proximate to it. This game effect is useful if the bottle is on the other side of a hedge or if the player accidentally passes the bottle. It does not work if the bottle is in a terrain depression, or more than a few meters away from the player. The notification consists of the player's command menu appearing in the bottom right corner of their screen, with the added command, "Crush *X* Bottle", where *X* indicates the type of bottle.
5. ***Tool tip sensing*** of the bottles from the 2D map view of the world. OFP avatars can navigate the environment in a 2-dimensional map view. When in 2D map view, the player's avatar is represented as two concentric red circles with a radial line indicating the avatar's bearing. If the human user moves the mouse cursor over the area of the map in the vicinity of the avatar, they can detect any objects that they could normally see in the visual detect mode. When an object is detected, a "tool tip" label appears next to it, indicating the object's type.

We evaluated the three experimental conditions:

- # **Bottles Known**, in which the subjects knew how many total bottles they were trying to recover; the performance measure was a team score, where each found bottle counted for one point.
- # **Bottles Unknown**, in which the subjects did not know how many bottles were hidden in the search area; the performance measure was a team score, where each found bottle counted for one point.
- **Bottle Portfolio** condition, where subjects were given a bonus of 100 points, if s/he had collected a "portfolio" of 7 bottles, each of a different kind. The different available types of bottles were told to the subjects by the experimenter: a portfolio was comprised of a set of Martini, Barbera, Jack Daniels, Seagram's, Napoleon, Baileys, and Whisky bottles. The bottle type was visible on the interface, when a subject was close enough to see the bottle. Subjects could also determine the bottle type through their command interface (e.g., "Crush Barbera"), and through the tool tips feature of the 2-D map view. Duplicate bottles counted simply one point.

Figure 1. The coding scheme that was used to label team communications

High-Level Categories	Utterance ID	Description	
Information Sharing: (1)_Info about tasks, logistics (2)_Info about location & movement of self & others	1	Requesting or Communicating Team Members' Location	
	2	Any Reference to Terrain or Map Features	
	3	Question / Indication of Bottle Locations	
	4	Personal Tips	
	20	Situation Assessment of Equipment	
	26	Question about Means of Doing Something	
Self-Organization: (1)_Information About Organizational Structure (2)_Establishing Authority Structure	5	Communication for Role Allocation	Team Planning Before Execution
	6	Communication for Division of Execution Space	
	7	Communication for Role Allocation	Team Planning During Execution
	8	Communication for Division of Execution Space	
	32	Discussion of Capability	
Problem Solving: (1)_Setting goals (2)_Forming hypotheses, predictions (3)_Making plans	12	Hypothesis about Bottle Locations	
	18	Plan or Activity Critiquing / Plan Suggestions	
	23	Considering / Evaluating Alternatives	
	24	Suggesting Individual or Team Strategies	
Meta-Cognition: (1)_Monitoring team progress (2)_Assessing team performance (3)_Reporting on own progress, past actions	9	Bottle Count	
	10	Object Count	
	11	Coverage Progress	
	29	Reporting Status / Activity	
	31	Request for Status Check	
	33	Monitoring Remaining Time in Session	
Team Coordination: (1)_Coordinating with others on on-going task (2)_Directing actions of others (3)_Stating one's intentions (4)_Backup Behaviors	15	Communicating Intent with Respect to Action	
	16	Communicating Intent with Respect to Changing Location	
	21	Commands	
	28	Permission / Request for Agreement to Do Something	
	30	Offer of Help	
	34	Request for Help	
Non-task Related:	13	System Problems: Keyboard Lockups, Game Window Loses Focus, etc.	
	14	Personal Discussion	
Interpersonal Affect:	19	Social Encouragement	
	22	Justifying Behavior, Action, or Plan	
Closed Loop Communications: (1)_Acknowledgment (2)_Disagreement (3)_Elaborations	17	Acknowledgment / Agreement / Answers to Questions	
	27	Disagreement	
	25	Clarification Question	

The three experimental conditions were *counterbalanced*, namely the conditions were executed in varying order among the teams to mitigate order effects. The subjects had 20 minutes to perform the task in each experimental condition.

We created these different experimental conditions in order to observe the communication patterns, their similarities and differences in the different conditions, in particular in the non-portfolio vs. the portfolio conditions. We hypothesized that there would be significant differences since the portfolio condition engenders additional dependencies among the tasks of team members; hence it would require tighter coordination. We also wanted to see whether there were any differences in the high performing teams vs. the low performing teams in the portfolio vs. non-portfolio conditions. Although there is consensus among teamwork researchers that communication patterns alone cannot predict performance, we remain interested in the potential for agent intervention to improve performance by improving communication for low performing teams. Further, we are interested in identifying communications and communication patterns that could be used by agents to understand team goals, individual roles, personal difficulties, strategies and progress to the team goal, so as to provide better assistance. Moreover, we hypothesized that there would be differences in the self-organization of the teams in the portfolio vs. the non-portfolio conditions, again due to the need for tighter coordination in the portfolio condition.

To analyze the coordination demands of the collaborative search task, we recorded all audio communications between team members. The audio files were manually transcribed to produce the logs of utterances that were segmented into conversational moves (Hirokawa, 1983). According to this scheme, a conversational move unit is an uninterrupted utterance of a team member that has a discrete problem solving function. Figure 1 shows the main categories and subcategories of the codes we used plus examples for each. The codes represent task related problems solving and coordination categories that are consistent with current teamwork literature (Fischer et al., 2007). In addition to the communication categories used by other researchers, we included categories of utterances relevant to team self-organization. These aspects were not needed in prior work where, unlike for ad hoc teams, team structure and role allocation were already in place.

Figure 2. The performance of the teams in all three test conditions

Team		# Bottles Recovered / Total # Bottles in Scenario			Portfolio Score	
Number	**Subject Expertise**	**# Bottles Known**	**# Bottles Unknown**	**Portfolio Condition**	**Raw**	**%age**
1	novice, medium, expert	84.62%	78.57%	64.29%	204 / 307 =	66.45%
2	novice, expert, expert	73.33%	54.76%	28.57%	8 / 307 =	2.61%
3	novice, novice, expert	60.53%	66.67%	42.86%	12 / 307 =	3.91%
4	novice, novice, medium	77.27%	35.71%	39.29%	11 / 307 =	3.58%
5	medium, expert, expert	86.36%	73.81%	78.57%	208 / 307 =	67.75%
6	novice, medium, medium	59.52%	52.38%	46.43%	12 / 307 =	3.91%
7	expert, expert, expert	81.82%	76.19%	60.71%	17 / 307 =	5.54%
8	medium, expert, expert	94.00%	80.95%	57.14%	109 / 307 =	35.50%
9	medium, medium, expert	76.60%	80.95%	71.43%	113 / 307 =	36.81%
10	expert, expert, expert	97.78%	64.29%	50.00%	107 / 307 =	34.85%
11	medium, expert, expert	80.85%	90.48%	42.86%	105 / 307 =	34.20%
12	novice, novice, expert	82.93%	83.33%	57.14%	109 / 307 =	35.50%
13	medium, expert, expert	88.24%	69.05%	53.57%	201 / 307 =	65.47%
14	medium, expert, expert	67.44%	73.81%	42.86%	12 / 307 =	3.91%
15	expert, expert, expert	95.35%	90.48%	71.43%	113 / 307 =	36.81%
16	medium, medium, medium	84.62%	76.19%	28.57%	8 / 307 =	2.61%
17	novice, medium, medium	74.42%	71.43%	42.86%	12 / 307 =	3.91%
mean	—	80.33 ± 11.07%	71.71 ± 14.03%	51.68±14.51%	26.08±24.32%	

RESULTS

Data were collected in a three condition repeated measures experimental design and were analyzed using the SPSS software.

Team Performance

Figure 2 reports the performance of all the teams in our initial set of experiments, measured by percentage of bottles crushed by each team. We had each subject self assess their expertise at computer games; this information is reported in the second column. The subjects participated in an initial practice session during which they were learning the user interface (results not shown). We evaluated the performance of the teams on three search tasks presented in counterbalanced order: (1) a session in which the subjects knew the total number of bottles hidden on the map (labeled in the table as # **Bottles Known**), (2) a session in which the subjects did not know how many bottles they were trying to recover (# **Bottles Unknown**), and (3) a session where the subjects were supposed to construct portfolios. For that condition, there were 28 hidden bottles, of which there were 7 unique types, thus enabling the formation of 4 portfolios in the ideal case.

A repeated measures ANOVA shows that team performance in the **#Bottles Unknown** condition was poorer than in the **#Bottles Known** condition ($F_{1,15}= 6.659$, $p=.02$) indicating that knowing the goal (total number of bottles to be found) improved performance. This effect is likely based on the commonly observed distinction between self-terminating and non-terminating searches. Where the bottle count is known the task can be stopped when all bottles have been found. Conversely, if bottles remain to be found, the subjects are aware of this and can intensify their search. Subjects in the # **Bottles Unknown** condition have no such standard and as a consequence they have greater difficulty in regulating their search behavior. Of more interest to us is the difference in coordination demand between the two conditions. In the # **Bottles Unknown** condition, a good team strategy would require team members to search different areas but otherwise allow them to act independently. Taking advantage of the additional

Figure 3. Bottles found in the three conditions

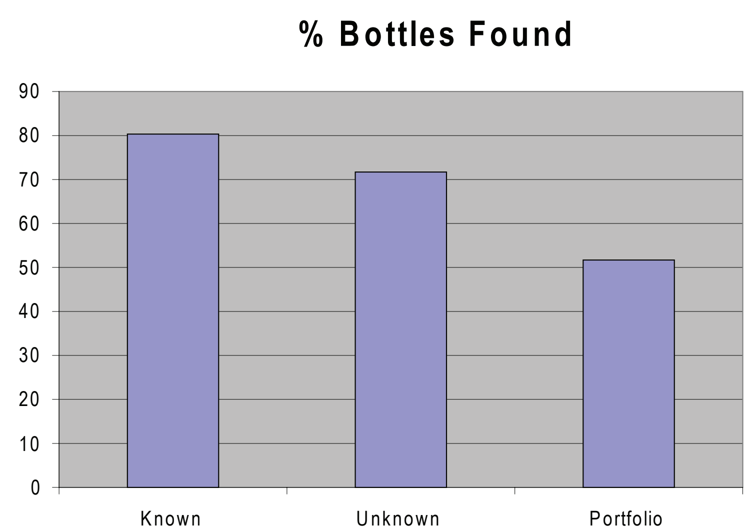


Table 1. Differences in communications across conditions

Communication code	$F_{2,14}$	Sig	Unknown vs. known	Unknown vs. Portfolio	Known vs. Portfolio
(3) reference bottle locations	19.258	p=.0001	N.S.	p=.001	p=.0001
(8) division of search space	6.437	p=.004	N.S.	p=.007	N.S.
(12) hypotheses, bottle locations	4.109	p=.026	N.S.	N.S.	p=.05
(15) action intent	15.021	p=.0001	N.S.	p=.003	p=.0001
(21) commands	7.979	p=.002	N.S.	p=.043	p=.004
(22) justification for action/plan	3.516	p=.042	N.S.	N.S.	N.S.
(25) clarification of question	12.988	p=.001	N.S.	p=.001	p=.014
(29) reporting status	6.202	p=.005	N.S.	p=.01	N.S.
(30) offer of help	10.464	p=.0001	N.S.	p=.001	p=.041
(32) discuss capability	3.630	p=.038	N.S.	p=.05	N.S.
(34) request help	33.549	p=.0001	N.S.	p=.0001	p=.0001

information in the # **Bottles Known** condition, however, requires exchange of bottle counts and other coordinating communications. If we see better performance in the # **Bottles Known** condition, and that superiority depends upon communication and coordination, then we should observe differences in communications patterns as well. The third, **Portfolio** condition, places greatly increased coordination demands on the team. In this condition, the reward associated with collecting individual bottles is dwarfed by the potential portfolio construction reward. Assembling a portfolio, however, may require players to pass up duplicate bottles, locate and report bottles needed for another's portfolio, and other more complex interdependent behavior. We should therefore expect to see substantial differences in communication patterns between the Portfolio and the other two conditions. Because collecting bottles is secondary to assembling portfolios for this group we should also expect lower bottle scores. Figure 3 shows the mean percent bottles found for each of these groups. Differences in performance were found between the three conditions ($F_{2,14}$= 36.617, p<.0001) with paired comparisons showing teams in the portfolio condition to find fewer ($p < .0001$) bottles than in either of the other conditions.

Differences in communications patterns across the three conditions were found for eleven of the thirty-four coded categories using a repeated measures ANOVA with paired comparisons.

As Table 1 shows all significant differences in communications were found between the Portfolio and the two bottle search conditions (N.S means non-significant). The differences involving communications such as bottle locations and hypotheses about locations, intended actions and their justifications, and communications characterizing teamwork such as offering and requesting help are of the sort that might be expected to distinguish the relatively independent bottle collection tasks from the high coordination demand required for assembling portfolios.

Analysis of Team Communication

To assess the communication demands of the collaborative search task, we compiled frequency counts of the different types of team communication.

Five of the 34 utterance categories: (2) reference to terrain or map features, (12) hypotheses about bottle locations, (14) unrelated personal discussion, (15) communication of intended action, and (23) evaluating alternatives were found to be related to team performance in the non portfolio conditions through a series of step-wise regressions. Teams in the # **Bottles Unknown** condition that exerted the least coordination demand produced the simplest model containing only two independent variables, (14) and (15). The frequency of unrelated discussions, β=.433, t_{14}=2.205, p=.045, and communicated intent, β=-.433, t_{14}=-2.305, p=.037, significantly predicted the bottles found with a regression explaining about half of the variance, R^2= .504, $F_{2.14}$=7.116, p=.007, in the number of bottles. Our finding that extraneous communication helped and conveying intent hurt in this condition reinforces our contention that the # **Bottles Unknown** task requires independent search by teammates. Communications such as (15) that encourage unnecessary coordination hurt performance while those that supplant potentially disrupting alternative communications (14) actually helped.

The # **Bottles Known** condition, that stands to benefit from coordination, yields more complex communication patterns. Unlike the # **Bottles Unknown** regression, (15) communication of intent, now contributes positively, β=.323, t_{12}=2.68, p=.02, to predicting bottles found. Related communications (23) involving sharing hypotheses also contributes positively, β=.502, t_{12}=4.162, p=.001, as does (14) unrelated discussion, β=.323, t_{12}=2.68, p=.02, that we have argued may benefit performance by supplanting inappropriate task related communications. Communication code (2) reference to terrain or maps, enters the model negatively, β=-.392, t_{12}=-3.32, p=.006. We speculate that this may be due to the fact that players that talked a lot about terrain features may have been disoriented and lost; such players retrieved few bottles. This four variable regression significantly predicted the bottles found and explained almost all of the variance, R^2= .842, $F_{4,12}$=16.003, p=.0001, in the number of bottles.

Teams in the **Portfolio** condition had a substantially different task since their scores were dominated by the portfolio bonus. Nevertheless, we examined the relation between their communication patterns and *number of bottles found* before looking at the more salient relation between communications and *scores*. (where the score includes the bonus for constructing portfolios). Our **Portfolio** data were fit by a five variable regression on bottles found. (29) reports of status, β=.687, t_{11}=6.17, p=.0001, (17) acknowledgment/agreement, β=.504, t_{11}=4.777, p=.001, and (34) request for help, β=.265, t_{11}=2.522, p=.028, were all positively related to bottles found. (1) communicating location, β=-.518, t_{11}=-4.851, p=.001, and (31) request for status, β=-.353, t_{11}=-3.353, p=.006, were negatively related. As for the earlier conditions we presume that communications (1) and (31) place cooperative demands on team members that detract from the independent task of accumulating bottles. This five variable regression significantly predicted the bottles found and explained almost all of the variance, R^2= .916, $F_{5,11}$=24.029, p=.0001, in the number of bottles.

Regression on score, the more salient performance measure in the **Portfolio** condition, revealed a different set of predictors. Data were fit by a four variable regression on team score with all variables providing positive contributions. Communication types (10), Object account, β=.85, t_{12}=6.827, p=.0001, (19) social encouragement, β=.450, t_{12}=3.615, p=.003, (11) reported coverage, β=.315, t_{12}=2.708, p=.018, and (15) communication of intent, β=.278, t_{12}=2.293, p=.041 contribute positively to the model. This four variable regression significantly predicted team scores and explained a substantial amount of the variance, R^2= .906, $F_{4,12}$=29.075, p=.0001, in scores. These differences in predictors for bottle count and scores suggest that different team processes may be involved. To explore the possibility that teams were using different processes to attain these goals we examined the correlation between number of bottles found and team scores. The positive correlation between bottles found and score, r_{15}=.73, p < .001 suggests that teams with good searchers were better at both finding more bottles and assembling portfolios. This debunks the notion that the negative relation observed between (1) communicating

Figure 4.

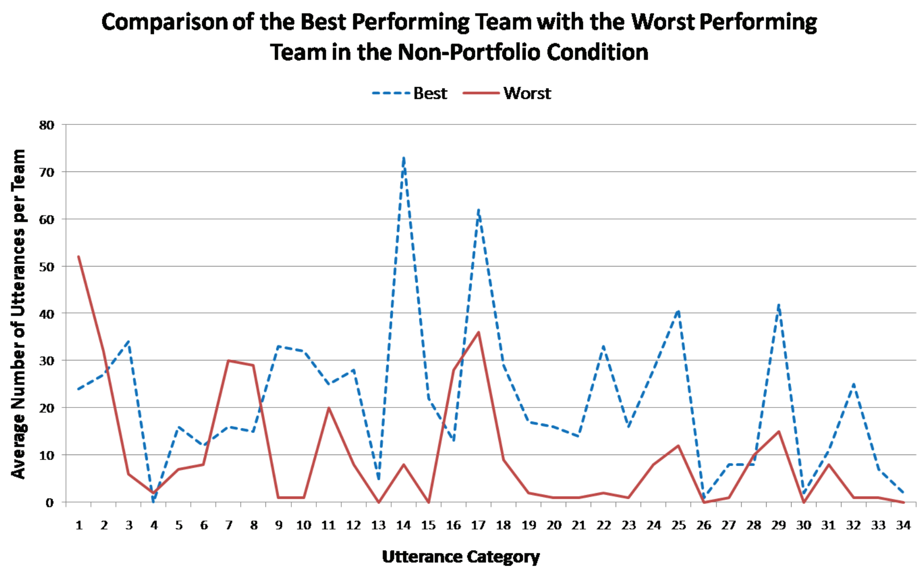
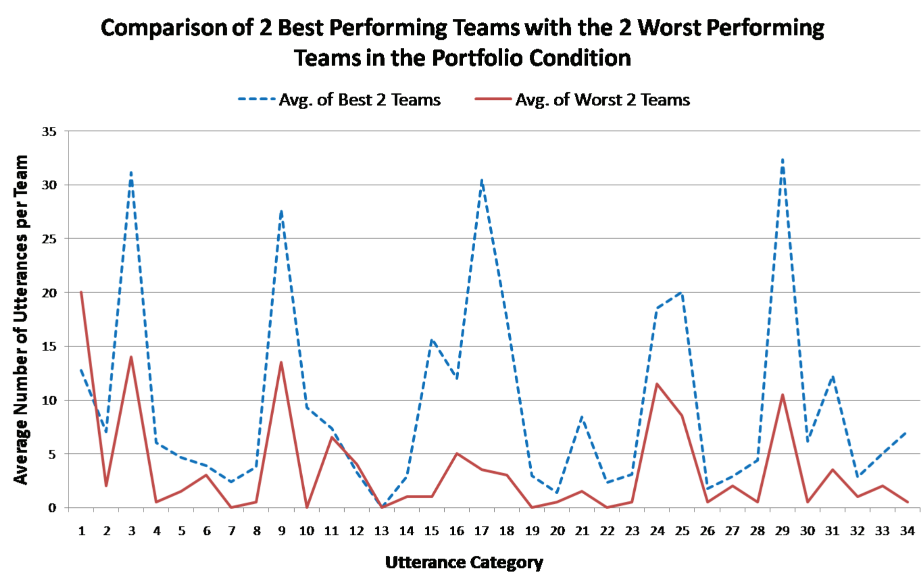


Figure 5.



location and (31) request for status in the regression on bottles found resulted from allowing team members to search independently accumulating more bottles by freeing them from coordination demands. A more likely explanation is that agents that talk a lot about position may have been lost and thus do not retrieve many bottles.

To further explore communication patterns in teams we report the results of the highest and lowest performing teams in the non-portfolio condition, **#Bottles Known** (called in the graphs the "Non-Portfolio Condition") and in the **Portfolio** condition. We do not report the communications pattern of the **#Bottles Unknown** condition since, as Table 1 shows, the differences in the communication among the two non-portfolio conditions were not significant. Moreover, we report here the highest and lowest performing team so that their differences may give us some clue as to the types of agent assistance that would be beneficial (See Discussion section.)

From the graphs of Figures 4 and 5 we make a variety of observations. First, we robustly find that the lower performing teams in both portfolio and non-portfolio conditions communicate less in all communication categories. This is consistent with social theories of teamwork, where communication is considered one of the most important teamwork dimensions. In addition we see that, consistent with theories of teamwork (Salas and Fiore, 2004) the communication frequency peaks of high performing teams in both conditions were for teamwork behaviors such as inventory monitoring (code 9), object counts (code 10), offering suggestions (code18), describing intentions (code 15), acknowledgments (code 17), and monitoring the passing of time (code 33). We also note some differences. In the non-portfolio condition, offering help (code 30) is not utilized much due possibly to the more loosely coupled nature of the task, whereas in the portfolio condition, there is a frequency peak there.

Team Self-Organization

In our experiments, we were particularly interested in observing differences in self-organization between (a) roughly equally performing teams (high/low) in the non-portfolio vs. portfolio condition, and (b) of high performing vs. low performing teams in the portfolio condition. We were also interested in observing any adaptations of the organizational structure during task execution.

To examine this question we first looked at the relationship between expertise with video games within teams and team performance. To consider the possible organizational dynamics we looked at the min, max, and average experience of groups to examine how team composition had affected performance. The rationale of this comparison is that team performance may be dominated by deficiencies of the weakest member (min), strengths of the most experienced member (max), or reflect their independent performance (avg). Our data show clear variation across the three conditions. In the # **Bottles Unknown** condition which promotes independent search all three estimates of team experience were correlated with bottles found: min(experience), $r_{15}=.44$, $p < .05$, max(experience), $r_{15}=.52$, $p < .05$, and avg(experience), $r_{15}=.50$, $p < .05$. In the # **Bottles Known** condition which relies more directly on coordination min(experience), $r_{15}=.63$, $p < .01$ and avg(experience), $r_{15}=.59$, $p < .05$ were significantly correlated but max(experience) was not. This pattern is reversed for the **Portfolio** condition in which max(experience), $r_{15}=.49$, $p < .05$ and avg(experience) $r_{15}=.44$, $p < .05$ were significantly correlated but min(experience) was not. These patterns suggest that substantially different forms of organization were most successful under the different conditions. In the # **Bottles Unknown** condition the independence of team members searches was reflected through correlations with all three measures. In the # **Bottles Known** condition which required team members to track one another's finds but without more elaborate coordination the lack of skill of the weakest member proved predictive of team performance. Finally, in the more complex **Portfolio** condition that required some structure of authority the skill of the most skilled team member proved to be a determining factor in team performance.

In further analyses we looked for differences across two organizational structures: (a) an authority structure and (b) a functional structure. In terms of authority structure, one of the striking observations is that none of the teams elected a leader, although in a few teams a team member volunteered to, for

Figure 6. Transcript of communication between subjects at the beginning of the search. This group of utterances was categorized as an example of team planning before execution. The "64" refers to a row on the map. There are significant pauses between the utterances; during one such pause, one of the subjects changes their mind and decides to cover a different area.

	A - Ann, B - Jon, C – Tom
A	Okay do we want someone to stay in the courtyard and do those bottles?
B	I'll do that and then head east
C	I'll do the same area that I did before.
B	I'll clear the courtyard and then clear the road to the south.
A	I'll work on the northern part and 64.
C	I think I'm going to stay closer to the town and circle around.

example, keep track of time or do some other team supporting behavior. In terms of functional structure, the nature of the search task encourages functional division of the search space (e.g. search the village, search the northern part of the countryside), and allocation of roles according to that spatial division. Figure 6 presents a transcript pertaining to the planning for spatial role allocation in one of the teams. It is a typical example of team planning communication at the beginning of a search session. The three subjects quickly develop a search strategy in which each subject assumes responsibility for covering a certain region. Another way that roles could be allocated would be in terms of ascertaining capabilities of the different team members (e.g. fast runners, adept at using the binoculars). A third way of role allocation would be in terms of whether a team member would be a bottle "marker" or a bottle "picker."

We observe from Figures 6 and 7 that in both non-portfolio and portfolio conditions the frequency of utterances pertaining to self-organization, namely utterances for role allocation (code 5), and division of execution space (code 6) during the planning phase is lower for the low performing teams than for the high performing teams.

In the non-portfolio condition, most of the team planning discussions was related to the division of the execution space: how to allocate the efforts of the team members to cover the entire map within the 20-minute test period. Although some teams agreed on a division of labor at the beginning of the task period, many teams modified their search strategies during execution based on their perceived task progress with respect to area coverage or their assessment of which areas contained a higher bottle density.

Besides the role allocation based on spatial dimensions, in the portfolio condition, we observe additional organizational structure and role specialization. An examination of the transcript of the highest performing team in the portfolio condition shows that the team had a sophisticated plan for role allocation. Below, we present a summary of the utterances of the highest performing team in the portfolio condition that suggest individual specialty roles:

1. Elect a player to be the first to achieve a portfolio. (role = 1st to achieve a portfolio)
2. After the first player achieves a portfolio, a second teammate should be chosen to form a portfolio. (role = 2nd to achieve a portfolio)
3. All teammates mark bottles to help the portfolio achiever make a portfolio. (role = bottle marker)

4. In the second half of the game, revert to self-interested roles ("Crush indiscriminately.") (role = self-interested search and retrieval)
5. Divide the space (intention of covering west, middle, east). (role = forage specific geographic area)
6. Decide whether a teammate should be responsible for keeping track of teammates' inventories in addition to their own inventory. (role = inventory keeper for situation awareness)

The low frequency of utterances concerning role allocation (Figures 6 and 7) of the worst performing teams shows that they do not spend enough time to plan about role allocation. We also observed that in the non-portfolio condition, the worse performing team have a higher frequency of utterances about role allocation *during execution*, meaning that they probably are trying to adapt for the lack of planning. In the portfolio condition, there is no appreciable difference between the role allocation utterances of the best and worst performing teams during execution.

Another interesting observation is that in both the non-portfolio and portfolio conditions, the worst performing teams have much lower frequency of utterances than the best performing concerning discov-

Figure 7. Excerpt of Team 3 transcript for the Portfolio condition. This is an example of a player experiencing a localization failure. Player B is confused about his/her absolute position on the map; Player C is attempting to help B localize. The first column gives timestamps.

MMSS		A - Ann, B - Jon, C - Tom
1357	C	I think I'm half way to clearing my area. I see Ann
1409	B	you see me? I don't see you
1417	C	you are way over by the edge
1425	B	so now I'm lost again
1434	B	I'm south and wanted to go more east

Figure 8. Excerpt of Team 1 transcript for the Portfolio condition. Player C is confused about the correct path to take to a marked bottle. Player C is leading A to the bottle location, costing the team valuable search time.

	A - Ann, B - Jon, C – Tom
A	I am but I don't see it, like I just went past that area
C	Um, where are you? Yeah you're in the wrong place, come back south
A	come back south... oh yes I did overshoot it again, where is that triangular thing
C	Can you see me, in the 3d mode?
A	ummm, no
C	Ok you see me?
A	Yes I see you, ok we're doing that, the lead
C	I think
A	Ok I think I saw something, nope that was something else
C	Just follow me
A	Yeah
C	Right in here, see it? RIght by the fire
A	Ah, no I totally missed this area thanks

ery of capabilities (code 32), for example good runner, good user of binoculars etc. Clearly, discovering the capabilities of different teammates is a precursor to good role allocation. Moreover, discovering of teammates' capabilities is a very crucial need in ad hoc teams that have not trained together.

DISCUSSION

There are too many possible agent support activities that could be implemented and are potentially relevant to the collaborative search task to exhaustively test all possibilities. Here we concentrate on the ones that could be inferred from our experiments. The four basic categories of possible interventions include: 1) agents that help the humans with basic task skills like navigation or providing user interface assistance; 2) agents that monitor task progress such as timekeeping or coverage monitoring.3) agents that check whether all team members are fulfilling agreements on plan and roles related to the team search pattern; 4) agents that help humans with their teamwork skills such as regulating communications with teammates or alerting players to possible assistance opportunities. Looking at the analysis of communication gives us some valuable clues regarding the merits of these possible assistance strategies.

Navigation Failure

Two types of utterances (1: player location) and (2: terrain features) were negatively correlated with performance across multiple conditions and were symptomatic of players experiencing navigation or movement failures. Seeing groups of contiguous utterances of (1) (player location) often indicated that the subjects were having trouble localizing and were straying into each other's territory. In the transcript in Figure 7, player B is clearly off course and has strayed into player C's territory. In the Transcript

Figure 9. Excerpt from the Team 10 transcript for the Portfolio condition (executed 2nd). The team members devise a test to determine who has the best sensing capability, and then discuss a strategy that uses that capability and a possible assist role based on that strategy. There is a high communication frequency. Team plan and role formulations are often fragmentary and interspersed with actions and observations that are unrelated to the plan formulation.

MMSS		A - Ann, B - Jon, C - Tom
0000	A	Who has the sensing thing? It's beginning.
0003	A	Like run off, don't get them all first
0007	B	Leave them there?
0008	A	Then go to the next grid over, see who can see them then we'll know
0015	A	Alright so let's say right at the beginning we all run off to what was it, F-63?
0023	B	Well that little courtyard is where they all are
0024	A	Maybe if we run to the edge right between F and G
0033	A	Then we can see who can see, then we go from there
0040	A	What do we want to do from there?
0043	C	Once we know who it is, the person who can't see well should stay close and get the easy stuff
0050	A	Maybe they should go with the sensing person
0055	C	How much time are we really saving by having someone work with them?
0101	C	Do you (inaudible) stick together so they can just use their eyeballs together?
0110	C	Well I don't have the good sensing and I can still see about half a square
0116	A	So that other part's open but you can't see anything because of the elevation.
0119	A	I'm going, started heading towards the west

of Figure 8, we see that player C helps guide the disoriented player A to the bottles. The failure of one team member to localize penalizes team performance in multiple ways: (a) distracting the lost player from the primary task of searching for bottles; (b) reducing the total coverage of the team when the lost player fails to cover his/her planned area; (c) costing other players search time as they attempt to assist the lost player. Utterances of type (2) were used by players to discuss terrain features when they were (a) impeded by obstacles such as hedges; (b) crossing the line of flags that marked the boundary.

Capability-Based Role Allocation

One of the highest performing teams, Team 10, spent a significant amount of time assessing their character's physical capabilities and allocating roles based on their assessment. Figures 9 and 10 are excerpts from voice log transcripts that were recorded during Team 10's experimental session. Team 10 was unique in a variety of ways:

1. It was the best performing team in the **#Bottles Known** condition.
2. The team members experimented with the capabilities of their virtual characters, such as the use of binoculars, the relative speeds of their respective characters, and the sensing capabilities of the characters in each scenario.
3. The team members experimented with the contexts in which the capabilities could be best used to their advantage, such as in which parts of the terrain binoculars provided the best advantage.

Figure 10. Excerpt from the Team 10 transcript for the Known condition (executed 3rd). The team members evaluate and critique their past performance, as well as revise their team strategy and respective roles. They decide to drop the specific test for determining a character's sensing capabilities, and agree that the best performer should provide a supportive role after finishing his section of eight 2-D map squares. This team improved its performance in successive sessions, and demonstrated the best performance of all teams in this condition.

MMSS		A - Ann, B - Jon, C - Tom
0341	B	Well, OK, strategy, what do you want to do?
0347	C	To me, it doesn't seem like there hasn't been very good coverage from far away.
0350	C	Do you want to break into groups? We just keep running over each other. Should we break into sections?
0356	C	Like we did last time
0400	C	Well we ended up in the same square, at one point.
0401	C	But that's because we never really throw ourselves down and stuck to it. Do you want to do what we talked about in the first place, and just eight, eight, eight? The person who does something well, should quickly, should just trust the map ...
0414	B	Clear, at one point.
0416	C	And then come back and help the others. Don't take the easy ones in the middle and just get out there.
0421	B	Yeah, yeah, yeah, definitely. You just get out there.
0423	C	If you find one, take it. It only takes two seconds to click. Don't, don't wait, don't waste, don't take the chance.
0426	A	Well, this is not one where we have to limit ourselves.
0431	B	No, we just take as many as we can.
0436	C	After the first few seconds, it's pretty obvious if you're the guy or not, so I don't think we need to take the time for that test.
0442	B	Yeah.
0443	C	You know what I mean? At least in the past it's been pretty obvious.

4. The team experimented with team strategies and roles in which capabilities could be leveraged to their performance advantage. For example, in the practice session, the team experimented with the strategy of having a person run from hilltop to hilltop and scan for bottles with their binoculars and depending on their findings, thus direct the search efforts of their teammates.
5. Team 10 displayed a tendency to investigate role specializations where the more capable players provided more support to the team and less capable players performed easier tasks, such as clearing areas where it was easier to find bottles.

The text in Figure 9 provides an example of the team dynamic just before the start of the session. In the previous session, they discovered that each character has a different sensing distance, and they spent some time trying to understand which character could sense the farthest. At the start of this session, player "Ann" suggests that everyone leave the bottles uncrushed in the courtyard where they start, and to run out of the courtyard. If they monitor their character's progress in the 2-D map, at a certain point they will not sense the bottles that they left behind in the courtyard; hence their "sensing distance", expressed in terms of 2-D map squares. Utterances 0003 – 0019 are where Ann explains this testing technique. Later utterances (not shown) represent a discussion about how the team should organize once they determine which character has the best sensing ability.

In the excerpted transcript shown in Figure 10, Team 10 plans their strategy just before beginning the Known condition. Player "Tom" complains (utterance 0347) about the lack of adequate coverage of the extreme areas of the terrain surrounding the village. He re-proposes dividing up the search space (utt. 0350), like last time (utt. 0356), but quickly follows with criticisms about how the team poorly executed their plan, previously. Namely, everyone should concentrate on executing the search in their area of terrain (utt. 0401) without deviating from their sector, and to not just search in the village where bottles are easier to find (utt. 0416). Tom also does not think it is necessary to spend time on testing a character's sensing ability (utt. 0436 and 0443), claiming that it is "pretty obvious if you're the guy or not," and suggests that the role for that person should be to clear their area as quickly as possible using the 2-D map (utt. 0401), and at completion, help the other teammates (utt. 0416).

Figure 11. Excerpt from the Team 16 transcript for the Portfolio condition (executed 1st). There is a disagreement between players Jon and Tom about whether to split up or stay together.

MMSS		A - Jon, B - Ann, C - Tom
0345		[GAME START]
0350	C	Alright so I say we go around the buildings and crush as many bottles as we can
0358	A	Yeah and try to maximize the time, I was going to say we should probably split up, each of us take two vertical columns
0410	C	But it will be harder to look for the bottles
0412	B	Oh the bottles aren't displayed on the map anymore
0415		(Personal Discussion)
0509	C	Yeah so we should all run together
0515	A	OK, I think maybe we would cover more ground if we split up
0526	A	I'm going, started heading towards the west
0531	C	Ok, let's go to the west first

Conflict Management

Although most of the teams were able to decide on a course of action without disagreement, some of the teams had problems agreeing on a search plan. Figure 11 is an excerpt from the transcribed voice log of Team 16's Portfolio condition, which they executed first of the three conditions. This transcription is remarkable in the way that it illustrates an incompatibility of mental models of the teammates. Player Ann suggests that the team should split up, each member responsible for searching the terrain that corresponds to two adjacent columns on the 2-D terrain map (utterance 0358). Player Tom, on the other hand, believes that it would be easier to find bottles if the players searched the same space, together. Ann offered the reason that splitting up would enable the team to cover more ground (utt. 0515). The two players do not resolve their differences, and for the first couple of minutes, Ann would propose an area of terrain for her to search, and Tom would commit to searching the same terrain. After the players each find bottles, and a couple of minutes have passed, Tom finally announces that he will search in the opposite direction opposite to Ann. Precious time was lost, not all terrain was covered, and indeed, Team 16 tied as one of the worst performing teams in the Portfolio condition. (See Figure 2.) In a post-experiment interview with the human subject who played Ann, he confessed that he lost confidence in his ability to propose further strategies or roles to Tom, fearing that they would not be well received. Indeed, Ann and Tom did not begin to discuss the strategy for forming a portfolio of different bottle types until less than half the time remained in the session. Player "Jon," meanwhile, did not participate in the planning or strategizing, and crushed four of the eight bottles that the team collectively found.

Modeling Team Intention

Developers of agent support systems for ad-hoc human teams face substantial challenges, as illustrated by the transcripts above. The main challenge remains: how can team-supporting agents acquire a model of what the human team members intend to do and thereby be enabled to monitor their task execution and coordination as a team? This is related to the problem of agents translating organizational norms into specific reasoning rules, described in Chapter 22.

First, there is information in human team communications that agents can use to understand human intentions, strategies, plans and roles. Although the human communications may be noisy and fragmented, and may not contain explicit relevance to strategy, plan or role as the above transcripts attest to, statistical language understanding techniques could be utilized by the agents to make correct inferences. Second, a significant portion of the vocabulary used to express such team-oriented communications is provided by the context of the execution environment. Therefore, it would alleviate the problem that agents have to infer what the humans are doing, if the agent has knowledge of the execution environment. Additionally, we observed that there is frequently a disconnect between what humans say and what they do. A proclamation that *a person will do X* can be interpreted as an *announcement of intention* as well as an *announcement of commitment*. This gives rise to a requirement that agents be multi-modal, namely they integrate information about the proclaimed intention along with environmental information, such as actual team member location and activity.

Creating a single agent that would have a high level of awareness of the mission and environmental context is very challenging, if not impossible. Our strategy would be to have different agents aiding different aspects of the task so that the agent's scope of representation and inference would be limited, thus facilitating agent implementation. In addition, the above observations motivate considerations of an agent interface design that does not rely exclusively on voice recognition and natural language understanding techniques, but which also incorporates multi-modal and on-demand human input features.

Agent Assistance Strategies

The data analysis, transcripts, observations of human subject behavior, and anecdotal post-experiment interviews with subjects provide suggestions and insights into specific intervention actions that an agent might execute to assist a human team. Here we discuss three general types of intervention actions: 1) aiding individual task work skills 2) regulating team communication 3) assisting with team planning.

Many of the subjects experienced problems with the navigation aspect of the task. A posttest questionnaire revealed that novice subjects were frequently disoriented in the environment, even when switching from 2-D map to virtual reality. When human subjects became lost or disoriented, they spent more time discussing terrain features, their locations, where they wanted to go, and how to get there, than actually dedicated to their task of finding bottles. This is illustrated by transcripts 9 and 10 and also by the regression analysis results where utterance on player location (1) and terrain features (2) was negatively correlated with performance. Agents could aid disoriented team members in a variety of ways: (a) displaying guide arrows to prevent players from wandering in loops; (b) providing directions to players trying to reach a specific point on the map; (c) reporting individual deviations from the teams' stated search plan. Moreover, automatic detection by the agent of these problems could be done through interpretation of utterances that refer to location.

More generally, agents can support human teams in the realm of task execution coaching. The type of coaching could range from reminding teammates to "stick with the plan" or ``their roles," "stay on pace," and to offer tips on improving search strategy, technique, and reminders when subjects overlook features and critical notices.

Meta-level team communication and planning is another area amenable to agent assistance. Although Team 16 eventually converged on a team strategy for the Portfolio condition, they did so when it was too late to have a positive impact on their performance. One of the ways an agent could support decentralized ad hoc teams is to support the team process, such as stimulating conversations about techniques, evaluations of team and individual progress, and requiring teammates to provide rationale for certain suggestions and decisions.

Regulating communication has been demonstrated to be useful in other agent support systems in which agents were used as mechanism for reducing the information overload created by indiscriminate player broadcasts (Fan and Yen, 2007). Indiscriminate communication of intent (15) and unnecessary requests for status (31) are negatively correlated with performance in some of the conditions, although communication of intent can be positively correlated with performance (e.g., the **Bottles Known** condition). The problem is communicating the right information to the right person in a timely fashion. Information routing is a task that agents are well suited for. There are many possibilities for agent aiding in this area, ranging from the simplest ones of having the agent forward status messages to the relevant player or having the agent use activity recognition to infer player goal intention directly from their movement. Interestingly, social type communications (e.g., unrelated discussion and social encouragement) were positively correlated with higher performance in some of the conditions and seemed to occur with players who felt relaxed enough with the interface and the cognitive demands of the task to chat with the other players. The key point is the reduction of cognitive overload, rather than indiscriminate elimination of all communications.

Based on this analysis, adding navigational aids and reducing information overload are likely to be productive and easily implemental agent support strategies for assisting players in our simulated collaborative search task.

We believe that these principles can also be extended to assist team members in real search and rescue operations. For instance, unnecessary radio chatter could be a significant problem for real search and rescue operations since broadcast style communications will reach dozens of people and yet most of the information communicated is likely to be relevant to only a small number of searchers. Becoming disoriented is even more hazardous in real world search domains and can happen even to survival experts. Although the details of the user interface will be different in the real world (e.g., interacting with the user through handhelds or headsets), the basic principles of assistance remain the same.

CONCLUSION

Decentralized ad hoc teams face difficult teamwork challenges, in particular, concerning communication and self-organization. This set of experiments was designed to (1) create a baseline of decentralized, ad hoc team performance in time stressed situations and (2) determine where agent aiding is likely to have the greatest impact.

When faced with the unfamiliar task of portfolio collection, human teams responded by developing various organization and communication strategies. The encoding of the utterances and their analysis allowed us to conclude that agent aiding for situation awareness and self-organization would help low performing teams increase their performance.

FUTURE RESEARCH DIRECTIONS

Our research study was an initial attempt to characterize the communication patterns of ad hoc teams operating in a particular domain. As we move towards a theory of agent-support for ad hoc teams, there are still many unanswered research questions and experiments that remain to be performed. One of the important areas of research could be to study the effect of organizational structure for these very dynamic and time stressed tasks. How do teams with different organizational structures perform on the same task? How do task-domain and organizational-structure interact to affect team performance? Is self-organization viable for highly complex team tasks?

Another interesting open research area is to study the effects of agent-support strategies: how do ad hoc teams respond to proven agent-support strategies (e.g., task monitoring) that have been used successfully with non ad hoc teams? Can agent-support be used to compensate for a lack of organizational structure in highly complex team tasks?

A third research area is identifying linguistic markers so that agents could best infer the intent of the teammates, with respect to the various subtasks and organizational roles of team members. Some research questions include: do teams communicate as they self-organize? Can we identify linguistic markers that are highly correlated with shifts in organizational structure?

We believe that ad hoc teams will become increasingly common in the future as organizations strive to become agile, adaptive, and responsive to rapidly changing global demands. Identifying answers to these questions will substantially benefit the future developers of agent-support systems.

ACKNOWLEDGMENT

Research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

We would like to thank Michael Lewis for his invaluable help with the experimental design and analysis.

We would like to thank Lori R. Price for her assistance with editing and formatting this chapter.

REFERENCES

Cannon-Bowers, J., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In Castellan, J., (eds), *Individual and group decision making*. Lawrence Erlbaum Associates.

Chalupsky, H. et al. (2001). Electric Elves: Applying agent technology to support human organizations. In *Proceedings of the Innovative Applications of Artificial Intelligence Conference*.

Chen, L., & Sycara, K. (1998). Webmate: A personal agent for browsing and searching. In *Proceedings of the Second International Conference on Autonomous Agents*.

Cooke, N., & Gorman, J. (2007). Assessment of team cognition. In Karwowski, P., (ed), *International Encyclopedia of Ergonomics and Human Factors*. Taylor and Francis Ltd.

Cooke, N., Gorman, J., Pedersen, M., & Bell, B. (2007). Distributed mission environments: Effects of geographic distribution on team cognition, process, and performance. In Fiore, S. and Salas, E., (eds.), *Toward a science of distributed learning and training*. American Psychological Association.

Entin, E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors*, *41*.

Fan, X., Sun, B., Sun, S., McNeese, M., & Yen, J. (2006). RPD-Enabled agents teaming with humans for multi-context decision making. In *Proceedings of International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

Fan, X., Sun, S., McNeese, M., & Yen, J. (2005). Extending the recognition-primed decision model to support human-agent collaboration. In *Proceedings of International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

Fan, X., & Yen, J. (2007). Realistic cognitive load modeling for enhancing shared mental models in human-agent collaboration. In *Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-07)*.

Fiore, S., Salas, E., & Cannon-Bowers, J. (2001). Group dynamics and shared mental model development. In London, M., (ed.), *How people evaluate others in organizations: Person perception and interpersonal judgment in industrial/organizational psychology*. Lawrence Erlbaum Associates.

Fiore, S., Salas, E., Cuevas, H., & Bowers, C. (2003). Distributed coordination space: towards a theory of distributed team performance. *Theoretical Issues in Ergonomic Science*, *4*.

Fiore, S., & Schooler, J. (2004). Process mapping and shared cognition: Teamwork and the development of shared problem models. In Salas, E. and Fiore, S., (eds.), *Team Cognition: Understanding the Factors that Drive Process and Performance*. American Psychological Association.

Fischer, U., McDonnel, L., & Orasanu, J. (2007). Linguistic correlates of team performance: Toward a tool for monitoring team functioning during space missions. *Aviation, Space, and Environmental Medicine*, *78*(5), B87 – B95. Section II.

Flashpoint (2001). Operation flashpoint: Cold war crisis. http://en.wikipedia.org/wiki/Operation_Flashpoint. Accessed May 2007.

Friedkin, N. (1998). *A structural theory of social influence*. Cambridge University Press.

Galbraith, J. (1977). *Organization Design*. Addison–Wesley Publication Company.

Giampapa, J., & Sycara, K. (2001). Conversational case-based planning for agent team coordination. In *Case-Based Reasoning Research and Development: Proceedings of the Fourth International Conference on Case-Based Reasoning*, ICCBR 2001, volume 2080, pages 189–203.

Giampapa, J. Esch, S., John, B., Singh, R., & Sycara, K. (2003). Evaluation criteria for the MORSE simulation environment. Technical Report CMU-RI-TR-03-38, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.

Goodrich, M., Quigley, M., Humphrey, C., Adams, J., Gerhardt, D., Cooper, J., Buss, B., & Morse, B. (2007). *Mini-UAVs for visual search in wilderness search: Tasks, autonomy, and interfaces.* Technical Report BYUHCMI 2007-1, Brigham Young University.

Hirokawa, R. Y. (1983). Group communication and problem-solving effectiveness: an investigation of group phases. *Human Communication Research*, *9*, 291–305.

Hoeft, R., Kochan, J., & Jentsch, F. (2006). Automated team members in the cockpit: Myth or reality. In Schulz, A. and Parker, L., (eds.), *Series: Advances in Human Performance and Cognitive Engineering Research*. Elsevier Science.

Horvitz, E. (1999). Principles of mixed-initiative user interfaces. In *Proceedings of SIGCHI.*

Kraut, B., Fussell, S., Lerch, F., & Espinosa, A. (2005). Coordination in teams: Evidence from a simulated management game. *Journal of Organizational Behavior.*

Lenox, T. (2000). *Supporting teamwork using software agents in human-agent teams*. PhD thesis, Westminster College.

Lenox, T., Hahn, S., Lewis, M., Payne, T., & Sycara, K. 2000). Agent-based aiding for individual and team planning tasks. In *Proceedings of IEA 2000/HFES 2000 Congress.*

Lenox, T., Lewis, M., Roth, E., Shern, R., Roberts, L., Rafalski, T., & Jacobson, J. (1998). Support of teamwork in human-agent teams. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics.*

Lenox, T., Roberts, L., & Lewis, M. (1997). Human-agent interaction in a target identification task. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*.

Pascual, R. G., Mills, M. C., & Blendell, C. (1999). Supporting distributed and ad-hoc team interaction. In *People in Control: An International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres*, number 463 in Conference Publication, pages 64 – 71. IEE.

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

Scerri, P., Pynadath, D., Johnson, L., Rosenbloom, P., Schurr, N., & Tambe, M. (2003). A prototype infrastructure for distributed robot-agent-person teams. In *Proceedings of International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

Setnicka, T. (1980). *Wilderness Search and Rescue*. Appalachian Mountain Club.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press.

Sycara, K., & Lewis, M. (2004). Integrating agents into human teams. In Salas, Fiore, S., & Fiore, E. (eds.), *Team Cognition: Understanding the Factors that Drive Process and Performance*. American Psychological Association.

TeamSpeak (2001). Teamspeak communication system. http://www.goteamspeak.com. Accessed May 2007.

Traum, D., Rickel, J., Gratch, J., & Marsella, S. (2003). Negotiation over tasks in hybrid human-agent teams for simulation-based training. In *Proceedings of International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

RECOMMENDED READING LIST

Burstein, M., & Diller, D. (2004). A framework for dynamic information flow in mixed-initiative human/agent organizations. *Applied Intelligence*, *20*(3).

Cannon-Bowers, J., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In Castellan, J., (ed.), *Individual and group decision making*. Lawrence Erlbaum Associates.

Cannon-Bowers, J., & Salas, E., (ed.) (1998), *Making Decisions Under Stress: Implications for Individual and Team Training*. American Psychological Association.

Fan, X., Yen, J., & Volz, R. (2005b). A theoretical framework on proactive information exchange in agent teamwork. *Artificial Intelligence*, *169*, 23–97.

Grosz, B., & Kraus, S. (1996). Collaborative plans for complex group action. *Artificial Intelligence*, *86*.

Horvitz, E. (1999). Principles of mixed-initiative user interfaces. In *Proceedings of SIGCHI*.

Lenox, T., Hahn, S., Lewis, M., Payne, T., & Sycara, K. (2000). Task characteristics and intelligent aiding. In *Proceedings of International Conference on Systems, Man, and Cybernetics*.

Levesque, H., & Cohen, P. (1990). On acting together. In *Proceedings of National Conference on Artificial Intelligence (AAAI).*

Lewis, M. (1998). Designing for Human-Agent Interaction. *AI Magazine*, *19*(2).

Orasanu, J. (1990). Shared mental models and crew performance. In *Proceedings of the 34th Annual Meeting of the Human Factors Society.*

Orasanu, J., & Salas, E. (1993). Team decision making in complex environments. In Klein, G., Orasanu, J., Calederwood, R., and Zsanbok, C., editors, *Decision making in action: Models and methods*. Ablex.

Salas, E., Bowers, C., & Edens, E. (2001). *Improving teamwork in organizations:Applications of resource management training.* Lawrence Erlbaum Associates.

Salas, E., & Fiore, S., (ed.) (2004). *Team Cognition: Understanding the Factors that Drive Process and Performance.* American Psychological Association.

Sierhuis, M., Bradshaw, J., Acquisti, A., van Hoof, R., Jeffers, R., & Uszok, A. (2003). Human-agent teamwork and adjustable autonomy in practice. In *Proceedings of 7th International Symposium on Artificial Intelligence (I-SAIRAS).*

Sycara, K., & Lewis, M. (2004). Integrating agents into human teams. In Salas, E. and Fiore, S., (ed.), *Team Cognition: Understanding the Factors that Drive Process and Performance*. American Psychological Association.

Sycara, K., Paolucci, M., Giampapa, J., & van Velsen, M. (2003). The RETSINA multiagent infrastructure. *Autonomous Agents and Multi-agent Systems*, *7*(1).

Tambe, M. (1997). Towards flexible teamwork. *Journal of AI Research*, *7.*

Warner, N., Letsky, M., & Cowen, M. (2005). Cognitive model of team collaboration: macrocognitive focus. In *Human Factors and Ergonomics Society (HFES) 49th Annual Meeting.*

Warner, N., & Wroblewski, E. (2004). The cognitive processes used in team collaboration during asynchronous, distributed decision making. In *Command and Control Research and Technology Symposium*.

Yen, J., Fan, X., Sun, S., Hanratty, T., & Dumer, J. (2004a). Agents with shared mental models for enhancing team decision-making. *Decision Support Systems.*

KEY TERMS

Ad Hoc Teams: A group of people that are brought together to achieve common goals for the duration of the task but who lack the experience of training together as a team.

Agent: An agent is a system that maintains a computational model of goal-directed and adaptive behavior. For the purposes of this chapter, we use agent to specifically refer to software systems.

Agent Support Systems: An agent-based software system that renders assistance to a human or group of humans that are trying to collaboratively accomplish a particular task. In this chapter, we focus on agent support systems that assist with team tasks rather than individual tasks.

Communication: The process of transferring information between humans or agents. In these experiments, communication between team members was measured by counting conversational moves, an uninterrupted utterance of a team member that has a discrete problem solving function.

Search and Rescue: An operation mounted by emergency services to find a lost person or recover victims from a disaster. Response teams usually have a doctrine or set of guidelines that they follow when constructing a search plan and allocating searchers to regions.

Shared Mental Models: A set of beliefs that a group of people holds in common about how the world works. The process of working together and training together creates shared mental models in human teams and is an important aspect of team cognition.

Teamwork: The state that occurs when groups of humans (or agents) commit to a shared set of goals and roles. Agents can achieve an understanding of teamwork through various formalisms that allow them to reason about establishing commitment through communication.