A Model of Human Teamwork for Agent-Assisted Search Operations

Dr. Gita Sukthankar
School of EECS
University of Central Florida
4000 Central Florida Blvd.

gitars@eecs.ucf.edu

Dr. Katia Sycara
Robotics Institute
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA 15213

katia@cs.cmu.edu

Mr. Joseph A. Giampapa
Robotics Institute
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA 15213

garof@cs.cmu.edu

Mr. Chris Burnett
Department of Computing Sciences
University of Aberdeen
Aberdeen, Scotland, UK

cburnett@csd.abdn.ac.uk

ABSTRACT

Coalition forces are engaged in distributed collaborative decision making in time-pressured, high-stakes situations. Providing automated decision support for such environments is a very challenging problem, due to shortening decision cycles, the changing nature of threats, opponent tactics, and environmental unpredictability. Intelligent agents have the promise to provide timely assistance in various areas of decentralized, collaborative decision making, such as information gathering, information dissemination, monitoring of team progress and alerting the team to various unexpected events. In order to fulfil the promise of agent technology in providing effective team assistance, better understanding of robust human-agent teamwork is crucial. The goal of our research project is to develop a theoretically grounded and empirically tested framework to allow for effective agent support for human teams that are engaged in adaptive teamwork in dynamic environments.

In order to (a) establish an experimental baseline of the performance of human-only teams and (b) understand where agents can provide the best utility in supporting human teamwork, we designed a scenario and experimentally evaluated team work where human teams performed a time-stressed, collaborative search task in a multi-player gaming environment. The collaborative search task recreates some of the challenges faced by human teams during search and rescue operations in the real world. In our experiments, we analyze (1) verbal communication between team members and (2) team coverage patterns. By ascertaining the information processing and coordination requirements of this team task, we can identify ``insertion points'' for agent assistance to human teams.

The search patterns demonstrated by the experimental subjects exhibited similar problems to the behavior of actual search and rescue teams: (1) the creation of accidental holes in the search pattern due to poor execution of the search plan, and (2) poor priority assignments in the search plan due to false clues and hunches. We have identified that this is a promising area for agent assistance. By having agents monitor and track individual team members' coverage, gaps in the team coverage are exposed earlier in the search process allowing repairs to be made in a more timely fashion. Our model predicts that aiding the state of coordination between team members will result in task performance improvement.

1 Introduction

This work is the initial step in our research plan towards addressing the fundamental question of how software agents can best aid distributed human teams performing collaborative decision making for time stressed critical tasks in uncertain and dynamic environments. Team 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. Software agents can fill a critical need for (1) supporting human team members in accessing, filtering, and synthesizing information from disparate sources; (2) increasing team situation awareness; (3) aiding the formation of shared mental models; (4) supporting team coordination in making decisions related to resources, tactics, and goals to meet the overall planning objectives. Building effective human-agent teams requires overcoming several important scientific challenges that to date have not been addressed: (1) the creation of mutual understandability between humans and agents; (2) the development of coherent team interactions; (3) establishing human trust in agent judgments.

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 (23) has focused on describing ways in which humans coordinate their activities, there has been little focus on which of those specific activities and information flows can be enhanced by being performed by software agents. 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 potentially include teamwork maintenance and task completion. In this paper, we describe our analysis of a collaborative search task, a team scavenger hunt, performed by human subjects in a multiplayer gaming environment. The results of this analysis will inform the future development of software agents to assist human teams performing search tasks.

Proposing and experimentally validating theories and increasing understanding of human-agent teamwork is a scientific problem of longstanding importance to computer science, human-computer interaction, collaboration science and psychology. Additionally, facilitating collaborative team decision making has become crucial in the military due to increased decentralization of the C2 process, the requirement for increased collaboration, decision-action speed, and the rapid restructuring of joint and coalition commands for different types of conflicts. Supporting collaboration and joint decision-making is extremely challenging in the face of shortening decision cycles, the changing nature of the threats and personnel downsizing, thus requiring increased task automation and making the understanding of robust agent aiding of human teamwork a crucial problem. Agent assistance will be particularly critical to military teams, especially coalition operations, as their operations become more agile and situation specific. As unfamiliar forces are brought together for different coalition missions, the infosphere they establish between their networked information systems will become a primary mechanism for coordination. In this uncertain environment agent support of teamwork becomes crucial. Because the domain independence of teamwork agents would allow them to be rapidly deployed across a broad range of tasks and settings, creating technology that can support highly dispersed human teams is a particularly high payoff area for the US and UK military.

This paper is organized as follows. Section 2 gives some background on human teamwork and how agents can be integrated into human teams. Section 3.1 describes the problems faced by expert human teams in performing search and rescue operations. In Section 3.2 we describe our team search task and simulation environment; Section 4 describes our experimental procedure and manipulations. In Section 5 we present our preliminary findings and describe some promising research directions on agent-assisted human teamwork.

2 Supporting Human Teamwork

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 (10). Furthermore, for such teams, both tacit and explicit coordination strategies are important in facilitating teamwork processes. Explicit coordination occurs through external verbal and nonverbal communications, whereas tacit coordination is thought to occur through the metacognitive activities of team members who have shared mental models of what should be done, when, and by whom (5; 8; 15). 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 coordinate their actions, they must possess commonly held knowledge structures, such as knowledge of teammates' roles and responsibilities along with team tasks and procedures.

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 (27). Horvitz (16) 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.

2.1 Agent Roles in Human Teams

Sycara and Lewis (27) 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 (1). Task-specific agents utilized by multiple team members (e.g., (2)) also belong in this category.
- **Agents support the team as a whole.** Rather than focusing on task-completion activities, these agents directly facilitate teamwork by aiding communication and coordination among humans and agents, as well as focus of attention. The experimental results summarized in (27) 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 (29). 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 (24); (3) considering the status of user's attention in timing services (16). Agents aiding teams (21; 20; 19; 18), 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 (29; 7; 6) 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 (21; 20), achievement of a military rendezvous plan (19; 18) and delivery of supplies to troops (7; 6). 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.

2.2 Teams in the Network-Centric Battlefield

The network-centric battlefield demands intense coordination among network effectors (humans and automation) that are part of a larger interconnected social organization. In this context we define coordination as the timely and adaptive distribution of information among network effectors. We think of team coordination as analogous to cognitive processing at the individual level. 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" (9). and has been frequently associated with differences in process behaviors, poorer shared understanding, and lean communication, relative to co-located teams (4). In fact, teams often adapt to these situations through spontaneous self-organization of their coordination structure (3).

It is important to note that we do not consider coordination in information theoretic terms (26) in which information is encoded, decoded and passively moved from effector to effector with some degree of uncertainty based on channel capacity. Rather, coordination involves active communication or mediation among effectors in a social network (12). 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 battlefield settings is 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. Coordination is, however, a precursor of team performance, and in our view, a critical precursor for the network-centric battlefield, in that effector competencies, as well as effectors themselves, are dispersed across the battlefield.

Based on our experimental data coordination improves with team experience and training, but decays over long retention intervals (3). 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 networked command and control is challenging due to the nonlinearities associated with interactions in complex distributed systems (4). For instance, coupled effectors have capabilities for contributing secondhand information to the information available in the local environments of other, reciprocating effectors. This positive feedback mechanism entails nonlinear changes in overall system state as information is adaptively dissipated through the system.

2.3 Improving the Performance of Human Teams

We hypothesize that to improve the performance of human teams, agents must do some combination of the following:

- reduce information processing costs;
- decrease uncertainty in the task;
- improve coordination between team members
- directly accomplish part of the team task

Galbraith observed that "the more uncertainty in a task, the more information processing necessary to achieve a given level of performance" (13). Hence, having the agents assist either in information processing or decreasing uncertainty should improve the team's performance. Moreover, in cases where the task is time-stressed, having the agents simply perform part of the task for the humans has the potential to improve team performance as well, particularly in cases where the task reward is an increasing function rather than a thresholded one. 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 (17). 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.

In addition to our primary hypotheses, we believe that the following dimensions affect the state of coordination between team members (31; 30):

.1. Collaboration system characteristics
.(a) Synchronous versus asynchronous collaboration: Is the collaborative process conducted in a same-time manner or are participants collaborating at different times?
.(b) Proximity of collaborators: Are the participants located proximally or are individuals geographically distributed?
.2. Team characteristics
.(a) Command structure: Are the participants organized in a hierarchical or flat structure?
.(b) Homogeneity of knowledge: Do all participants possess the same knowledge or is there information asymmetry?
.(c) Team size: How many individuals are required to collaborate on a team?
.3. Task dimensions
.(a) Collaborative output: Is the goal of the team to deliberate and process information or to determine a course of action (COA)?
.(b) Time stress: Is the team subject to time pressure?
.(c) Task complexity: How large and complex is the task?
.(d) Task familiarity: Is the task a onetime or a recurring event?
.(e) Nature of constituent subtasks: e.g., whether subtasks involve planning, decision making, cognitive conflict, creative and intellective subtasks etc.

To evaluate our model of human-agent teamwork, we created a team task with the following characteristics: (1) synchronous, (2) geographically distributed, (3) flat command structure, (4) asymmetric information, (5) small team size, and (6) time-stressed. Other than differences in the command structure and team size, the task possesses similar characteristics to the tasks performed by real search and rescue teams, described in the next section.

3 Collaborative Search

For our initial set of experiments, we monitored teams of human subjects performing a collaborative search task in simulation. Search and rescue is a challenging, time-stressed team task with a potentially high payoff since inadequate team performance can result in fatalities. By developing software agents capable of improving human team performance on collaborative search tasks, we can positively impact coalition search and rescue operations.

3.1 Wilderness Search and Rescue Operations

In this section we provide a task analysis of how civilian human teams perform wilderness search and rescue operations summarized from (14; 25). We assume that many aspects of the task analysis are also applicable to military search and rescue teams, although military teams have access to different equipment and also often face the additional problem of rescuing victims from enemy territory. A goal-directed task analysis of wilderness search and rescue operations identified the following list of operational goals and subgoals (14). 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*

Civilian search and rescue operations are directed by an incident commander who develops the search plan and collates information collected by the search teams. The search teams include trained volunteers who search the areas by foot or vehicle, along with technical specialists who search special types of terrain (e.g., divers to search water or climbers to scale cliffs). In civilian search and rescue situations, the search team starts by constructing a profile of the missing person to guide the team's search priorities. Depending on the person's age, physical condition, and wilderness experience, certain areas are marked as being higher or lower priority in the search plan. For instance, victims in poor physical condition are more likely to drift to downhill regions, whereas a victim with wilderness experience in good physical condition will move uphill to get his/her bearings. Victims with special limitations (e.g., autistic children) have unusual inclinations, such as avoiding roads and moving away from noise, that need to be taken into consideration by the search teams.

When executing a wilderness search plan, the teams employ four distinct types of search: hasty, constraining, high probability region, and exhaustive. During hasty search, the searchers rapidly check high probability areas to determine the missing person's location or direction of travel. This type of search is often used in the initial part of the search plan. During constraining search, the searchers attempt to build a perimeter bounding the victim's location; an example of constraining search would be having searchers check a large snowy field for tracks to localize the victim to one side of the field. Hasty search and constraining search are used by the incident commander to find clues and establish search priorities. After search priorities have been established, the incident commander divides the search area into regions and deploys search teams to search high probability regions. Exhaustive search is done by having the searchers form a line and walk abreast through an area; this type of search is used to find clues such as clothing or wrappers after other forms of search have failed.

Wilderness search and rescue operations pose the following challenges to expert human teams: (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. We believe that software agent assistance can potentially reduce the information overload of the incident commander and minimize errors during the execution of the search plan. 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.

3.2 Experimental Task: Team Scavenger Hunt

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 multiplayer game and battlefield simulator, Operation Flashpoint (OFP version 1.96) (11), 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 (Figure 1). 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 testbeds; (3) it offers a simple team performance metric: number of objects collected.

The team scavenger hunt task can be made arbitrarily complicated by adjusting the following parameters: (1) task uncertainty, (2) reward function, (3) adversaries. Task uncertainty is increased if subjects are not provided with maps and have to simultaneously explore the area while searching for objects. Another way to increase task uncertainty is to have subjects locate objects based on clues or probability distributions, rather than precise locations. Varying reward functions can be used to elicit different types of team behavior. Individual players can be awarded incentives for high performance vs. having the rewards split equally among team members. A simple reward function is to have the reward be a linear function of items acquired across all team members; another option is to award points for portfolios of objects. A portfolio of objects is a collection that contains a specified number of unique objects with desired characteristics, e.g., a portfolio consisting of a table, a chair and a telephone, all of the same color. Having a portfolio based reward system makes a subject's optimization problem harder because it penalizes locally greedy acquisition strategies. Adding adversaries to the task forces the players to replan to overcome unexpected obstacles. The game can be made more dynamic by adding mobile objects, automated adversaries to hinder the searchers, or having teams compete against each other.

In our initial version of the experimental task, the subjects have some uncertainty— they are provided with a terrain map, but only objects within a certain visibility range are revealed on the map. The current version of the game requires having the searchers collect static objects; subjects are rewarded based on their total team score, rather than their individual score, at collecting objects within an adversary-free environment.

3.3 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 a twenty minute period. OFP is distributed with a simple but versatile scenario editor that greatly facilitates the creation of multiplayer military and civilian scenarios and missions.

Terrain around and including the village of *Flers* on the island of *Normandie* was chosen as the focal point for the one practice and two experimental scenarios (Figure 2). 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 2dimensional (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

Figure 1: Subject world view during bottle collection. This is a zoomed-in view that has an increased density of bottles for illustrative purposes; the actual 3D environment is much larger and contains a much lower bottle density.

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.

4 Experiments

4.1 Procedure

Seventeen teams of three persons, 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 Table 2).

Each team member played the game through an assigned and dedicated laptop. All three members of a team sat at the same large table arranged in such a way that they could not look at each other’s screen. The human subjects were forbidden to share

Figure 2: The 2D terrain map available to players in the Operation Flashpoint simulation environment. In addition to the terrain map, the subjects are provided with simulated versions of binoculars, compass, and watch.

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, though face-to-face, were logged using TeamSpeak (28).

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 2dimensional 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.

Each experimental session was composed of a twenty minute practice period and three twenty minute search tasks. We evaluated the three experimental conditions: (1) **# Bottles Known** in which the subjects knew how many total bottles they were trying to recover; and (2) **# Bottles Unknown** in which they did not know how many bottles were hidden in the search area. By knowing the total bottle count, we hypothesize that subjects have a better sense of task progress and can assess their individual search performance. Comparing the team performance of subjects with the bottle count information vs. no bottle count information might predict the benefits of introducing agents to teams for search tasks.

4.2 Analysis of Team Communication

To analyze the coordination demands of the collaborative search task, we logged all communication between team members. We looked at the following categories of communication:

- **increasing situation awareness (SA)** This category includes all communications that increase the team members' situation awareness. Examples include communicating one's location, querying teammates for their positions, and discussions about terrain features or object locations.
- **sharing hints (Hints)** Occasionally subjects shared personal search techniques with their teammates, such as scanning large regions in a 2D map view or using binoculars while standing on high terrain features.
- **team planning** This category includes any discussion proposing, accepting or declining team search strategies; for example, team members often took responsibility for covering a certain region or suggested that other team members should redirect their search to a different area. We separated team planning into two categories: (1) planning before execution (**Pre Plan**) and (2) planning during execution (**In Plan**). Within these categories we examined two types of communications: (1) role allocation and (2) division of execution space.
- **monitoring task progress (Monitor)** Often subjects exchanged information about object counts, coverage progress, or time left remaining in the session.
- **sharing world beliefs (Beliefs)** Sometimes the subjects discussed their hypotheses about the relative frequency distributions of the bottles in different regions and speculated about the existence of bottle caches.
- **miscellaneous** Some of the communication between team members was not directly related to the experiment, such as social interaction or complaints over system issues (e.g., unexpected key lockups or display slowdowns).

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.

Figure 3: 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.

Most of the team planning discussions were 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 time 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 or their assessment of which areas contained a higher bottle density. The transcript shown in Figure 3 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.

To assess the communication demands of the collaborative search task, we compiled frequency counts of the different types of team communication (Figure 4). We believe that the categories of increasing team situation awareness and monitoring task progress are amenable to agent assistance. Our model of human-agent teamwork predicts improved team performance if we can reduce the cost of information processing for the team.

4.3 Team Search Patterns

During the experiments, we had the subjects report their search patterns and self-assess their coverage of the area by annotating a hardcopy printout of the map; many subjects used these notes to track their coverage progress. We observed a variety of search strategies among the subjects: (1) scanning the map by quadrants; (2) following terrain features such as roads or hedges; (3) focusing effort in regions with higher bottle counts. Figure 5 shows an example of one subject's search strategy. For each group of subjects, we measured (1) number of quadrants covered per subject; (2) number of quadrants covered per team (the union of each subject's coverage areas). The number of quadrants covered per team is a good measure of team coordination. On average each subject was able to cover 49% (mean of both conditions in Table 1) of the region within 20 minutes; therefore, all three team members were required to perfectly cover the region. Figure 5 shows an example search pattern that was reported by a subject in Team 4. This annotation was used to estimate the number of quadrants that the subject was able to cover; team coverage was determined by examining the union of all team

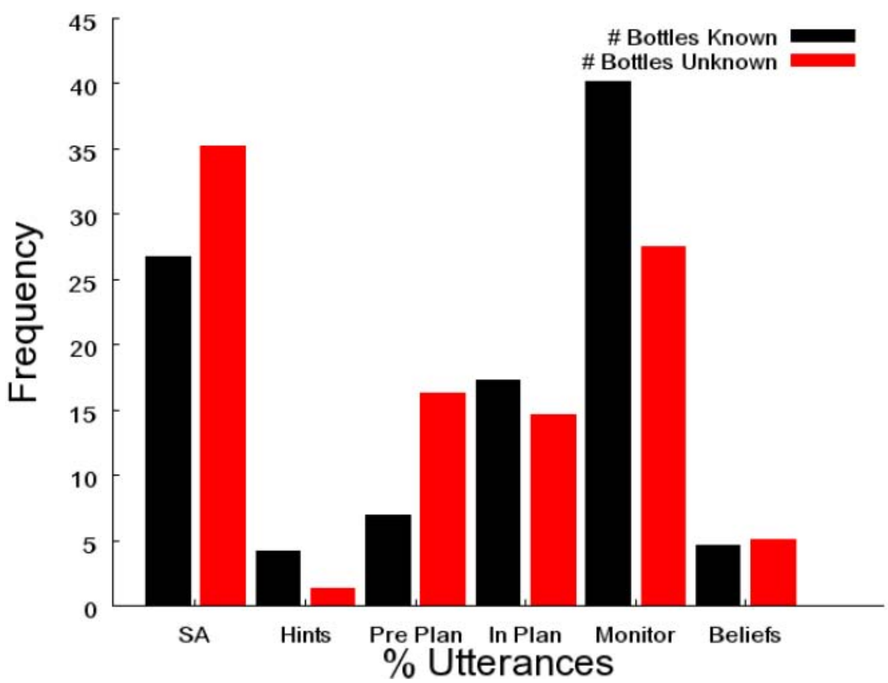

Figure 4: Communication frequency averaged across four teams of subjects. Subjects appear to spend more time monitoring their task progress in **# Bottles Known** condition, whereas in the **# Bottles Unknown** condition more communications are devoted to increasing the team's situation awareness. Also, the subjects

appear to be spending more time planning prior to execution in the **# Bottles Unknown** condition.

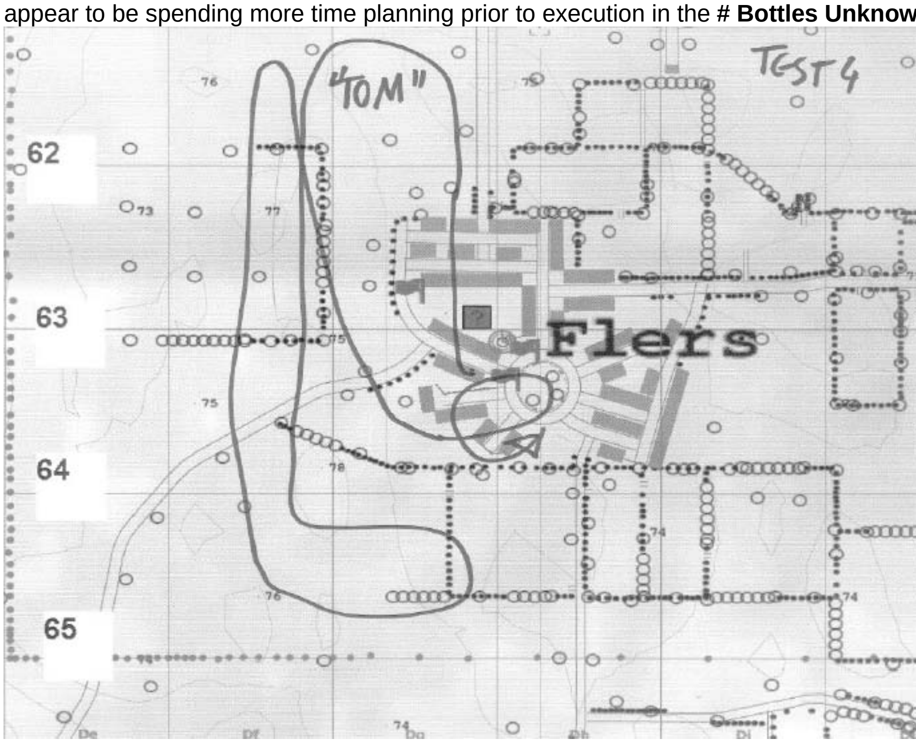

Figure 5: Map of search area annotated by one of the subjects in Team 4. The numbers and letters at the edge of the map are the coordinates for quadrants. After each search, the subjects were asked to report their search pattern by drawing on a printout. These annotations were used to calculate the individual and team coverages (see Table 1).

members' individual coverages at the quadrant level. Table 1 contains the individual (Columns A, B, and C) and team terrain coverage performance for all eight teams. Team scores are lower than the sum of individual scores due to coverage area overlap. On average, team coverage was 60% of the map for both the **# Bottles Known** and **# Bottles Unknown** conditions.

To be successful at covering the entire region, teams had to effectively divide the execution space and be proactive at diagnosing and repairing accidental gaps in the search pattern. Instead of attempting to cover the entire region, some teams hypothesized that certain quadrants had a high bottle density and focused on thoroughly searching those quadrants at the expense of less promising areas. The search patterns demonstrated by the experimental subjects exhibited similar problems to the behavior of actual search and rescue teams: (1) the creation of accidental holes in the search pattern due to poor execution of the search plan, and (2) poor priority assignments in the search plan due to false clues and hunches. This is a promising area for agent assistance; by having agents track individual team members' coverage, gaps in the team coverage are exposed earlier in the search process allowing repairs to be made in a more timely fashion. Our model predicts that aiding the state of coordination between team members will result

Team #	Number of Bottles is ***Known*** Ann	Jon	Tom	Team	% Ovrlap	Number of Bottles is ***Unknown*** Ann	Jon	Tom	Team	% Ovrlap
1	50%	25%	33%	75%	44%	42%	33%	42%	79%	47%
2	58%	50%	29%	100%	38%	50%	46%	25%	96%	26%
3	21%	54%	38%	75%	50%	33%	46%	38%	79%	47%
4	33%	42%	38%	100%	13%	33%	46%	50%	100%	29%
5	54%	38%	46%	88%	57%	42%	46%	67%	96%	61%
6	71%	50%	38%	96%	65%	63%	67%	38%	83%	100%
7	38%	42%	25%	88%	19%	33%	42%	29%	75%	39%
8	83%	33%	54%	100%	71%	29%	21%	79%	96%	35%
9	88%	63%	63%	100%	113%	63%	63%	67%	100%	92%
10	58%	75%	54%	100%	88%	33%	54%	50%	88%	57%
11	46%	54%	79%	100%	79%	42%	58%	79%	96%	87%
12	21%	58%	38%	83%	40%	50%	42%	42%	92%	45%
13	58%	25%	50%	79%	68%	38%	17%	63%	79%	47%
14	17%	71%	25%	88%	29%	38%	38%	42%	75%	56%
15	92%	88%	54%	100%	133%	79%	58%	67%	96%	113%
16	83%	54%	33%	96%	78%	92%	71%	71%	96%	143%
17	54%	38%	42%	96%	39%	54%	33%	38%	96%	30%
Mean		49%		92%	60%		49%		89%	62%

Table 1: Terrain Coverage

Team #	Subject Expertise	# Bottles Known	# Bottles Unknown	Highest % Score Executed 2nd?
1	novice, medium, expert	84.62%	78.57%	0
2	novice, expert, expert	73.33%	54.76%	1
3	novice, novice, expert	60.53%	66.67%	1
4	novice, novice, medium	77.27%	35.71%	1
5	medium, expert, expert	86.36%	73.81%	0
6	novice, medium, medium	59.52%	52.38%	1
7	expert, expert, expert	81.82%	76.19%	0
8	medium, expert, expert	94.00%	80.95%	1
9	medium, medium, expert	76.60%	80.95%	1
10	expert, expert, expert	97.78%	64.29%	1
11	medium, expert, expert	80.85%	90.48%	1
12	novice, novice, expert	82.93%	83.33%	1
13	medium, expert, expert	88.24%	69.05%	1
14	medium, expert, expert	67.44%	73.81%	1
15	expert, expert, expert	95.35%	90.48%	1
16	medium, medium, medium	84.62%	76.19%	0
17	novice, medium, medium	74.42%	71.43%	0
Mean	—	80.33 ± 11.07%	71.71 ± 14.03%	70.59%

Table 2: Bottles Retrieved

in task performance improvement. Another potential assistance strategy would be to have agents help the subjects form better priority assignments by noting the number of bottles found in each quadrant and informing team members about quadrants with higher bottle densities.

4.4 Team Performance

Table 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. During each session, we evaluated the performance of the teams on three search tasks:
(1) an initial practice session during which the subjects were learning the user interface (results not shown), (2) a session in which the subjects knew the total number of bottles hidden on the map (labeled in the table as **# Bottles Known**), (3) a session in which the subjects did not know how many bottles they were trying to recover (**# Bottles Unknown**). Based on these preliminary results, it appears that knowing the total bottle count improved the performance, which indicates that this might be a promising area for agent assistance.

5 Discussion

This initial phase of experiments was designed to (1) create a baseline of expected team performance and (2) determine where agent aiding is likely to have the greatest impact. From our preliminary results we noted a few trends:

- The categories with the highest communication traffic were situational awareness (e.g., communicating one's location to the team) and task monitoring (communicating bottle counts, time, and coverage). In the **# Bottles Known** condition, subjects actually had fewer task monitoring communications, but more communications relating to situational awareness. Pre-planning seemed to increase in the **# Bottles Unknown** condition.
- The team coverage did not differ in the **# Bottles Known** or **# Bottles Unknown** conditions; subjects reported that they were searching about the same amount of the map, although their bottle retrieval performance was lower.
- Team performance, measured by number of bottles retrieved, was poorer in **# Bottles Unknown** condition.
- Gaming expertise was predictive of individual bottle collecting performance, but not of an individual's terrain coverage.

Based on these results, we plan to focus our agent aiding on these areas:

- **reducing the cost of communication between teammates by having agents assist the subjects at increasing situational awareness and monitoring task progress**. We believe that this will free the humans' time to communicate about other aspects of the task, such as sharing search hints and team planning.
- **improving the coordination between team members at dividing the execution space** Subjects were not accurately able to self-assess their team coverage and often left holes in their search patterns. Helping teams accurately monitor team coverage is a very promising future area for agent assistance.

Listening to the recordings of the players was very valuable and gave us some insights. All teams adopted the common sense strategy of forming a team plan and dividing the search space. Most teams also replanned during execution when the following events occurred: (a) subjects finished their assigned coverage areas, (b) when new bottles were discovered, (c) as the deadline approached. Some teams evaluated themselves on terrain coverage whereas others focused on total bottle count. Often subjects formed hypotheses about areas with high bottle counts, similar to following false hunches in search and rescue operations, and speculated about the existence of hidden bottle caches. Although teams were allowed to self-organize, none of them elected a commander. Some players voluntarily assumed roles such as timekeeping or tallying bottle counts.

In the future, we plan to evaluate agent aiding in a version of the task that requires tighter coordination. By examining a task with more interdependencies, we believe that we will observe more planning, especially during execution. In the new version of the task, each subject has to retrieve a portfolio of seven bottles, one of each type (whiskey, martini, etc.). We hypothesize that coordination confers a huge benefit to the subjects if they pool information about bottle types that they have already acquired, the location of bottles that they do not need, and their portfolio requirements. Without team coordination, it is hard for even expert gamers to collect a portfolio of bottles, since it is more likely that they will collect a larger number of bottles without being able to find one of each type.

6 Conclusion

Our ongoing research objectives include understanding how agent-based team support affects team performance in critical, time-stressed situations and the impact of agents on the adaptive decision-action cycle of the team. To understand the effects of agent aiding on team support, we designed a collaborative search task, the team scavenger hunt, that recreates some of the challenges faced by expert human teams during search and rescue operations. 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 testbeds; (3) it offers a simple team performance metric (4) it can be extended in different ways to evaluate concepts like team trust and adversarial reasoning. The team scavenger hunt problem touches on some interesting problems in artificial intelligence such as the multi-agent traveling salesman problem and preference satisfaction over sets of objects. We believe that it is a useful benchmark problem for other groups studying team behavior and agent assistance. This initial set of pilot experiments has allowed us to create a baseline of non-assisted team performance and also has given us some valuable clues as to where agent aiding is likely to have the greatest impact.

Acknowledgment

Research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF0630001. The authors would like to thank Prasanna Velagapudi for his helpful advice on Operation Flashpoint, Mike Lewis for his suggestions on experimental design, Rahul Sukthankar for his proofreading, Michael Goodrich for useful discussions on search and rescue operations, and our subjects for their participation.

References

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

[2] L. Chen and K. Sycara. Webmate: a personal agent for browsing and searching. In *Proceedings of the Second International Conference on Autonomous Agents*, 1998.

[3] N. Cooke and J. Gorman. Assessment of team cognition. In P. Karwowski, editor, *International Encyclopedia of Ergonomics and Human Factors*. Taylor and Francis Ltd., 2007.

[4] N. Cooke, J. Gorman, M. Pedersen, and B. Bell. Distributed mission environments: Effects of geographic distribution on team cognition, process, and performance. In S. Fiore and E. Salas, editors, *Towards a science of distributed learning and training*. American Psychological Association, 2007.

[5] E. Entin and D. Serfaty. Adaptive team coordination. *Human Factors*, 41, 1999.

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

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

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

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

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

[11] Operation Flashpoint: Cold war crisis. `http://en.wikipedia.org/ wiki/Operation Flashpoint`, 2001. Accessed May 2007.

[12] N. Friedkin. *A structural theory of social influence*. Cambridge University Press, 1998.

[13] J. Galbraith. *Organization Design*. Addison–Wesley Publication Company, 1977.

[14] M. Goodrich, M. Quigley, C. Humphrey, J. Adams, D. Gerhardt, J. Cooper, B. Buss, and B. Morse. MiniUAVs for visual search in wilderness search: Tasks, autonomy, and interfaces. Technical Report BYUHCMI 20071, Brigham Young University, 2007.

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

[16] E. Horvitz. Principles of mixed-initiative user interfaces. In *Proceedings of SIGCHI*, 1999.

[17] B. Kraut, S. Fussell, F. Lerch, and A. Espinosa. Coordination in teams: Evidence from a simulated management game. *Journal of Organizational Behavior*, 2005.

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

[19] T. Lenox, S. Hahn, M. Lewis, T. Payne, and K. Sycara. Agentbased aiding for individual and team planning tasks. In *Proceedings of IEA 2000/HFES 2000 Congress*, 2000.

[20] T. Lenox, M. Lewis, E. Roth, R. Shern, L. Roberts, T. Rafalski, and J. Jacobson. Support of teamwork in humanagent teams. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, 1998.

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

[22] D. Roberts, G. Lock, and D. Verma. Holistan: A futuristic scenario for international coalition operations. In *Proceedings of Knowledge Systems for Coalition Operations (KSCO 2007)*, 2007.

[23] E. Salas and S. Fiore, editors. *Team Cognition: Understanding the Factors that Drive Process and Performance*. American Psychological Association, 2004.

[24] P. Scerri, D. Pynadath, L. Johnson, P. Rosenbloom, N. Schurr, and M. Tambe. A prototype infrastructure for distributed robotagentperson teams. In *Proceedings of International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2003.

[25] T. Setnicka. *Wilderness Search and Rescue*. Appalachian Mountian Club, 1980.

[26] C. Shannon and W. Weaver. *The mathematical theory of communication*. University of Illinois Press, 1949.

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

[28] Teamspeak communication system. `http://www.goteamspeak.com`, 2001. Accessed May 2007.

[29] D. Traum, J. Rickel, J. Gratch, and S. Marsella. Negotiation over tasks in hybrid humanagent teams for simulationbased training. In *Proceedings of International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2003.

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

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