Abstract

Based upon the integration of constructs from organizational and cognitive science we present a theoretical framework for understanding memory function in the context of human-agent teams. To support the development of true Human Systems Integration, we use this approach to meld robust concepts in human cognition with human agent team research. Our goal is to illustrate the theoretical and practical importance of these concepts to team cognition in general and augmented cognition in particular. We discuss this through theory in human memory and memory failures and integrate approaches to illustrate their value to developing research plans for augmenting cognition.

1 Integrating Systems and Humans

From the organizational sciences the field of team research has matured substantially over the latter part of the 20th Century. Similarly, the cognitive sciences have grown tremendously upon a strong theoretical and empirical foundation. Only in the last decade have these two fields begun to more formally interact to produce what is now being called “team cognition” (see Salas & Fiore, 2004). Based upon the integration of constructs from organizational and cognitive science we present a theoretical framework for considering human-agent team functioning. In order to support the development of true Human Systems Integration, we use this approach to meld robust concepts from the cognitive sciences with human agent team research. Our goal with this is to illustrate their importance to team cognition in general and augmented cognition in particular. We first discuss this through the lens of HSI and then narrow our focus to memory and memory failures. Finally we integrate these approaches to illustrate their value to developing research plans for augmenting cognition.

1.1 Human Systems Integration

Over the last decade, the Department of Defense has made increasing use of findings from the cognitive and computational sciences within its “human-systems integration” (HSI) program. This is a broad based concept for systems acquisitions programs requiring a level of analysis able to model how tools can support the human in his/her tasks. This includes not only single operators engaged with a given system (Salas & Klein 2001) but also encompasses how teams interact over time and space with distributed technologies (cf. Fiore et al., 2003) and similarly encompasses how intelligent agents are being integrated with modern systems (e.g., McNeese, Salas, & Endsley, 2001; Sycara & Lewis, 2004). HSI doctrine has developed to ensure that both the design of systems and their eventual development are able to fully support the human operator (Clark & Goulder, 2002; Freeman & Paley, 2001; Freeman, Pharmer, Lorenzen, Santoro, & Kieras, 2002; Pharmer, Dunn, & Santarelli, 2001).
Within this context an important development is that of “human-centered work system design,” an effort that emerged out of expert systems research in the 1980’s and which has since evolved into research in a variety of complex domains (Clancey, 2002) including semiautonomous missions to the moon for NASA (e.g., Clancey, 2004; Sierhuis & Clancey, 2002). This approach applies research and theory to better understand how the human interacts with, and is impacted by, their systems. More specifically, “rather than abstracting human behavior as work processes or tasks... [this models] people’s activities comprehensively and chronologically throughout the day (p. 32, Sierhuis & Clancey, 2002). By focusing on how interaction is actually organized and the associated details of such work, this approach takes a broader perspective by considering not only the technologies involved in a task, but also the human operators of these technologies and how they actually use them rather than are believed to use them. As such, this approach is as much anthropological as it is cognitive and engineering – effectively integrating disciplines to create models that appropriately simulate how work really occurs in complex socio-technical systems.

1.2 Overview of Paper

It is this form of human-centered theorizing and design that is foundational to understanding human-agent teams. Specifically, the technology-based characteristics present in such teams have the potential to attenuate the processes and the products occurring during human-agent interaction. For our initial efforts we consider agents broadly, following Fiore et al. and defining them as “ranging from computer-based intelligent decision-support systems with no or minimal anthropomorphism, to highly anthropomorphic machines, such as android robots, robotic animals, and robotic swarms or packs which display group behaviors” (Fiore, Jentsch, Becerra-Fernandez, Salas, & Finkelstein, 2005, p. 1). Understanding the cognitive and social processes emerging within human-agent teams is critical to developing the appropriate tools and techniques for augmenting cognition. To support our efforts we view human-agent teams as a socio-technical system, akin to the way we have viewed distributed teams (see Fiore et al., 2003). While much of the research in Augmented Cognition does emphasize the human, we suggest the research base can be strengthened by more fully exploring human-systems integration and human-centered design. Following this human-centered approach to melding systems with the human (Clancey, 1997; Hoffman, Hayes, & Ford, 2002; Shafto & Hoffman, 2002), we next discuss a theoretical framework that enables us to elucidate a small set of the factors that support human-agent process and are, therefore, targets for augmenting cognition.

2 Understanding Memory Function in Dynamic Environments

Dynamic interaction and distribution over space and time, the rule in technologically-dependant human-agent team environments, forces members to integrate sensory input across differing modalities as they attempt to coordinate their actions. Although the emergence of human-agent teams has led to substantial flexibility in operations, they may also result in undue cognitive load, that is, a workload over and above that experienced in co-located teams (cf. Fiore et al., 2003). When considering that data and interaction can come from, or result from, agent team members, this adds a layer of complexity to coordinative efforts. Specifically, this places additional limits on team members’ ability to attend to cues pertinent to their tasks because their teammates are either geographically-dispersed or not human. Although this generally impacts a number of individual and team processes (see Fiore et al., 2003), in this paper we discuss its potential for interacting with the human operator and lead to additional workload that may impact the memory processes of the human team members.

To describe the aforementioned effect we use the general term memory failures. These failures are suggested to occur because the cues normally relied upon by co-located, human-human teams to support memory processes are now attenuated in some way by the human-agent distributed work. This, in turn,
may produce faulty coordination leading to poorer performance. In short, we argue that this new interaction environment consisting of humans and agents who are not co-located, may alter and even hinder memory performance for distributed team members. A necessary first step in understanding this phenomenon is the development of a classification of the types and causes of memory failures experienced by team members in these environments. Following earlier work on memory failures (Herrmann, Gruneberg, Fiore, Schooler, & Torres, in press), we suggest that human-agent teamwork would benefit from an investigation of the qualitative nature of the memory failures, as well as the proximal (direct) and distal (indirect) causes of these failures. In Table 1 we describe the broader goals for our approach in understanding memory failures.

Table 1. Goals Associated with Understanding Memory Failures in Human-agent Teams

<table>
<thead>
<tr>
<th>Broad Goals</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differentiating Memory Failures</td>
<td>Classification of the differing memory failures so as to develop categories that can differentiate lapses in memory</td>
</tr>
<tr>
<td>Distinguishing Memory Causes</td>
<td>Development of a taxonomy to parse the differing causes of memory failures in human-agent teamwork</td>
</tr>
<tr>
<td>Proposing Guidelines</td>
<td>Specific guidelines able to inform both system designers and operators so as to identify how cognition may be augmented to avoid situations leading to these memory failures</td>
</tr>
</tbody>
</table>

Despite the importance of understanding how memory failures may hinder performance in complex environments there is relatively little research that has been conducted on this phenomenon. As an example of the importance of memory failures, studies of the Aviation Safety Reporting System documented that failures in memory were related to over 10% of the reported errors (Endsley, 1999). Further, these studies found that failures in memory led to problems with decision making, and subsequently could be related to up to 50% of fatal and 35% of non-fatal accidents (Jones & Endsley, 1996). Finally, failures in memory were related to 11% of situation awareness problems as reported in the Aviation Safety Reporting System (Jones and Endsley, 1996). Unfortunately, because no comprehensive system for understanding memory failures is in place, we do not know the nature of these failures.

Some studies have tried to understand memory failures occurring outside the laboratory. For example, classic research on absented-mindedness used diary studies to ascertain the frequency and nature of these memory failures (e.g., Reason & Lucas, 1984). Others have explored how devices such as techniques for reminding can reduce failures (e.g., Beal, 1988). Some have investigated what particular memory improvement methods work in alleviating failures (e.g., Herrmann, Brubaker, Yoder, Sheets, & Tio, 1999; Herrmann, Buschke, & Gall, 1987). Finally, recent studies have investigated “everyday memory failures” by exploring the relation between failure type and cause (Fiore, Schooler, Whiteside, & Herrmann, 1997; Herrmann et al., in press).

It is this latter set of studies on which we base the remainder of this paper. We follow the work of Herrmann and colleagues to highlight its relevance to understanding cognition (and failures in cognition) in human-agent teams. We use this theoretical framework because we argue that a human-centered approach to augmented cognition is necessary to design the technologies that can effectively augment cognition. Specifically, only when we fully understand the limitations of human cognition when operating in human-agent teams, can we understand the particular scaffolds that can support cognition. Research in the area of everyday memory failures uses classification which includes differentiating between lapses that are failures in prospective memory or retrospective memory. Prospective memory is generally referred to as “memory for the future,” or remembering to engage some action at some future time.
Retrospective memory failures are the more familiar type of failures, that is, failing to recall something that had been previously learned.

Although we do not deny the importance of retrospective memory to cognition and coordination in complex operational environments, for two reasons we focus this paper on prospective memory. First, prospective memory has been studied substantially less than retrospective memory. Recent papers by Kvavilashvili and Ellis (1996) and Ellis and Kvavilashvili (2000) document the changing patterns of interest in this topic within the cognitive sciences (see also Brandimonte, Einstein, & McDaniel, 1996). Second, prospective memory represents an area of cognition where technology may be able to make a significant improvement in functioning through augmented cognition (cf. Herrmann et al., 1999; Herrmann et al., 1987). In sum, although there has been recent interest developing in prospective memory and even some investigations of memory failures in operational environments, we still do not fully understand the nature and causes of these failures. More importantly, in the context of human-agent teams, little if any research has attempted to relate this construct to teams. Fiore et al. (in press) have begun to lay the foundation for understanding this phenomenon in teams and we next discuss their framework.

In sum, memory performance in human-agent teams represents not only an important area of inquiry, but also a rich theoretical area from which to consider how to use agents to augment human cognition. Our approach represents an adaptation of a paradigm developed by Herrmann and colleagues (see Fiore et al., in press; Herrmann et al., in press) for understanding memory failures outside of the laboratory.

3 A Framework for Understanding and Augmenting Memory in Human Agent Teams

Programs in augmented cognition attempt to produce diagnostic methods for understanding cognitive processes for what can be termed a form of dynamic scaffolding of cognition. This generally describes research and development in non-invasive techniques for measuring cortical activation that can be linked to a variety of higher- and lower-level cognitive processes (Schmorrow, 2002; Schmorrow & Kruse, 2002; St. John, Kobus, Morrison, & Schmorrow, 2004). While the short- and medium-term efforts are looking at detection accuracies, in the longer-term, augmented cognition will need to better meld with operationally relevant and time-stressed events. Thus, although these programs are still in their early stages of development, the computational methods and engineering systems will soon be at appropriate levels of sophistication to meet these goals. What we suggest is that, simultaneous to these developments, we must better understand human memory and successful and unsuccessful memory performance in operationally complex environments. These research tracks can develop independently but eventually be integrated when the theories and the technologies are themselves appropriately developed. Towards that end, we present a set of potential research principles for pursuing augmented cognition within human-agent teams. As discussed, for our initial foray into this area we narrowly focus on prospective memory.

We argue that intelligent agent technology, given its increasing ubiquity as human-agent teams become more prevalent, represent a viable means for augmenting team cognition. Specifically, given that intelligent agents are being designed as team members and a large body of research is already in progress with respect to cognitive engineering and decision making, it is only prudent that we additionally consider how these agent team members can be used to better augment team cognition.

3.1 Memory Failures Framework

The classification of memory failures and their causes can enable a fuller understanding of team cognition in human-agent teams. Comprehensive programs aimed at augmenting human cognition at the individual and team level must address two primary criteria associated with memory failures. First, research must
determine the quantity, and the qualitative nature of, the memory failures. Second, the causes of the memory failures must be determined. Towards this end, the memory failures framework based upon Herrmann and colleagues multi-modal approach to memory (Herrmann, 1996; Herrmann & Parente, 1994) can aid in addressing these criteria. The multimodal framework was developed to account for the multitude of factors associated with memory in everyday environments (i.e., outside the laboratory). Building upon this approach we suggest that understanding effective memory functioning in the complex environments in which human-agent teams operate must encompass physiological and psychological factors that contribute to cognition. Further, we suggest that social and technological factors must be considered within this broader conceptualization.

As shown in Figure 1, this framework presents a general classification of memory failure causes. These can arise due to endogenous or exogenous factors and can be either proximally (i.e., directly) related, or distally (i.e., indirectly) related to the memory action (e.g., Herrmann et al., in press). Table 2 shows this distinction and presents a description of each. As shown, causes can emerge due to something internal or external to the human. Exogenous factors influence memory failures due to system or technological problems (e.g., poor understanding of system functioning). Exogenous factors also include the environmental context which could encompass both social (e.g., lack of coordination with agent team member), or natural problems (poor weather). Furthermore, endogenous factors encompass problems arising out of psychological or physiological states, that is, factors internal to the human (e.g., stress). Proximal factors are the causes directly disrupting encoding, retention, or remembering processes while distal factors arise from processes not directly related to memory (e.g., physiological states), but which may hinder memory processing if compromised.
<table>
<thead>
<tr>
<th>Directness of Cause</th>
<th>Location of Cause</th>
<th>Endogenous (Internal)</th>
<th>Exogenous (External)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proximal (Direct)</td>
<td>Cognitive problem arising from task overload</td>
<td></td>
<td>External contextual problem such as cue missing</td>
</tr>
<tr>
<td>Distal (Indirect)</td>
<td>Physiological problems arising due to mission duration</td>
<td></td>
<td>External problem such as social distraction</td>
</tr>
</tbody>
</table>

3.2 Research Guidelines for Augmenting Prospective Memory in Human-agent Teams

As can be see in the Fiore et al. framework presented in Figure 1, what is complicated is that the “agent” team members cut across both the technological and the social components of this framework. As such, research must determine the degree to which there may be additive or even multiplicative effects on coordination that are arising from this socio-technical system. By adopting this framework to the study of memory performance in complex environments we can better understand what failure causes may be amenable to technologically-based interventions. Specifically, it may be feasible to use augmented cognition technology to either mitigate the occurrence, or minimize the effects of, memory failures, both at the individual and team level. Thus, despite gaps in our understanding of prospective memory in complex operational environments utilizing human-agent teams, enough research exists to allow us to introduce preliminary guidelines for consideration in augmented cognition research.

Theories on prospective memory (PM) distinguish between forms of prospective memory that are tied to how it is that the action must be completed. We use these distinctions to suggest how agents can be used to augment these differing forms of prospective memory actions. Specifically, research in prospective memory has noted that the actions associated with prospective memory fall into two major categories (see Ellis, 1988). First are those actions that require precise execution at a given point in time, referred to as “pulses.” Other actions can be completed at any time over a wider time frame, referred to as “steps.” A related distinguishing characteristic has to do with whether time or events are considered to be the driving factor in the memory action. This distinction is important because both event-based and time-based prospective memory tasks are diagnosable and similarly represent pertinent targets for augmented cognition. Finally, recent research suggests that matches between ongoing processing and processing required for the prospective memory tasks can be facilitative (see Meier & Graf, 2000, see also Marsh, Hicks, & Hancock, 2000). Following the tenets of Transfer Appropriate Processing theory (see Roediger, Gallo, & Geraci, 2002), this approach notes how synchrony between processes required for the memory task, and the actual operational task processing requirements, can be crucial for performance. This understanding adds an additional level of diagnosticity through which to view human-agent team cognition as it elucidates when memory processes may be compromised. Using the aforementioned theorizing, in Table 3 we present a representative sample of research and development guidelines for augmenting cognition within human-agent teams.
Table 3. Representative Sample of Research Guidelines for Augmenting Cognition Based upon Prospective Memory Theory

<table>
<thead>
<tr>
<th>Guidelines</th>
<th>Guideline Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guideline 1. PM Pulses</td>
<td>Agents used to augment successful prospective memory <strong>pulses</strong> must be able to diagnose mission chronology to prompt the human team members appropriately.</td>
</tr>
<tr>
<td>Guideline 2. PM Steps</td>
<td>Agents used to augment successful prospective memory <strong>steps</strong>, although requiring less rigidity in programming, must be able to monitor task executions over a broader task space to determine when to prompt.</td>
</tr>
<tr>
<td>Guideline 3. PM Events</td>
<td>Augmenting cognition for prospective memory should map event-based prospective memory tasks onto mission <strong>parameters</strong> so human-agent team members are able to diagnose when <strong>critical events</strong> have occurred and provide reminders appropriately.</td>
</tr>
<tr>
<td>Guideline 4. PM Timing</td>
<td>Augmenting cognition for prospective memory should map time-based prospective memory tasks onto mission <strong>chronology</strong> so human-agent team members are able to determine when the requisite <strong>time period</strong> has passed and provide reminders appropriately.</td>
</tr>
<tr>
<td>Guideline 5. PM Processing Matches</td>
<td>Research should determine how agent members of a team can be made aware of the nature of the processing required for a given prospective memory task to determine when memory aids are, or are not, warranted. For example, aids may be warranted when there is a mismatch in processing between the ongoing task and the prospective memory task.</td>
</tr>
</tbody>
</table>

4 Conclusions

In this paper we have presented a preliminary means with which to understand causes of memory failures by illustrating a broad set of categories of causes potentially occurring in human-agent teams. We additionally presented a finer distinction of types of prospective memory tasks. These were used to derive a set of research guidelines for using agents to augment team cognition. Following the general rubric of human-systems integration, and more specifically, human-centered work design, we illustrated how the research base in augmented cognition can be strengthened. We have described how theories on prospective memory emerging out of the cognitive sciences can be used to explore the unique challenges emerging from this new form of organizational structure. Prospective memory presents a rich theoretical and practical area of inquiry in which to explore how agents may be used effectively augment cognition. These are only preliminary guidelines that may be used for research investigating how agents may be designed to target these memory failures so as to attenuate the negative consequences sometimes emerging from human-agent teamwork.

From the perspective of viewing science **strategically**, this approach can be construed of as more of a short- to medium-term effort. In particular, augmented cognition and the research driving us towards that goal is pursuing more of a medium- to long-term emphasis in that the transition from the laboratory to the field is still somewhat in the future. Our argument was that the human-centered approach we have proposed can be pursued simultaneously to other augmented cognition efforts so that convergence can be reached at an earlier date. We bring up this point in our concluding section because there are related efforts to consider from the perspective of long-term research planning. In particular, a growing body of literature is beginning to document the brain regions activated prior to and during prospective memory tasks. For example, West, Herndon, and Ross-Munroe (2000) find that an initial stage of prospective memory known as the noticing component activates the occipital-parietal region. Further, they find that sections of the frontal cortex may be more responsible for a directed search component of this task.
Burgess, Quayle, and Frith (2001) varied the intentions associated with a prospective memory task and demonstrated differing areas of activation dependent upon whether the intention was being maintained or was actually realized (see also Burgess & Shallice, 1997; McDaniel et al., 1999). As such, from a programmatic perspective, research and development integrating human-agent teams with augmented cognition can similarly consider this developing literature coming out of cognitive neuroscience.

In sum, there is a growing convergence on our understanding of prospective memory, an important cognitive process supporting the operator in dynamic environments. We argue that it be a strategic target for the efficacious use of agent technology to augment cognition. We suggest that others pursue this approach to human-agent team research so as begin to place bounds around the cognitive and social consequences affecting interaction and team development when work is technology-mediated. More specifically, research must first understand where the problems in team cognition actually are occurring prior to attempting to augment that cognition in human-agent teams.

References


Acknowledgements

Writing this paper was partially supported by Grant Number SBE0350345 from the National Science Foundation and by contract number N61339-04-C-0034 from the United States Army Research, Development, and Engineering Command, to the University of Central Florida as part of the Collaboration for Advanced Research on Agents and Teams. The opinions and views of the authors are their own and do not necessarily reflect the opinions of the University of Central Florida, the National Science Foundation, the RDECOM-STTC, the U.S. Army, DOD or the U.S. Government. There is no Government express or implied endorsement of any product discussed herein. All correspondence regarding this paper should be sent to Dr. Stephen Fiore, via email at sfiore@ist.ucf.edu or via regular mail at University of Central Florida, 3100 Technology Parkway, Suite 140, Orlando, FL 32826.