19th ICCRTS

"C2 Agility: Lessons Learned from Research and Operations"

Decision Science Challenges for C2 Agility

Topic 1 (First Choice)

Topic 3

Topic 8

Elizabeth Bowman

Jeffrey M. Bradshaw

Robert R. Hoffman

Niranjan Suri

Corresponding Author: Elizabeth Bowman

Army Research Lab

B. 321, Aberdeen Proving Ground, MD 21005

410-278-5924

Elizabeth.k.bowman.civ@mail.mil

Decision Science Challenges for C2 Agility

Abstract

In recent decades, decision-makers in many areas, ranging from defense to weather forecasting, have argued that problems could be solved if more data were available. But merely increasing the volume of data gathered has not led to the hoped-for success. On the contrary, it has done just the opposite by overloading systems, networks, and, most importantly, human operators. Such factors not only jeopardize decision-making effectiveness, but also the adaptive capacities needed to assure the resilience of the decision-making process itself. New methods are needed to help decision-makers deal with the overwhelming amount of data being made available to them. In military Command and Control (C2), this challenge is intensified due to the diversity of data produced by multiple sources, the variety of information needing attention, and the time pressures that are a natural consequence of operations. In this paper, we consider strategic interdisciplinary research needed to produce transformational decision science capabilities for Warfighters at all levels across the Department of Defense (DOD). Our focus is on the needs of the human operator(s) who must rely on a variety of technologies to collect, interpret, and assimilate meaning from a variety of inputs in a complex and time-constrained environment. Toward that end we explore fundamentals of decision making and identify research challenges in the areas of human-machine collaborative sensemaking, task-relevant valuation and selection of information, and performance metrics for decision making at a genuine systems level. We conclude with a discussion of emerging military technologies that are designed to improve decision making in military domains and consider future trajectories that harness the power and opportunities provided by extremely large and disparate data.

Introduction

The rapid pace of development in sensor, data processing, and display capabilities, combined with the dynamic and changing nature of the threats and challenges faced by decision-makers, mean that it is more urgent than ever to develop robust and resilient methods to support decision-making. A number of factors, including technology improvements, procedural changes, and methodological changes have vastly increased the volume of data to be analyzed. From a technology perspective, increases in the number of sensors and their quality (e.g., resolution) have contributed to increases in the volume of data collected. Increased digitization of traditional media (e.g., digitization of news and other open source information) and social media (which is already digital) have also resulted in significant increases in the volume of data that can be collected. Procedural and methodological changes, partly triggered by the Global War on Terror (GWOT), have increased the scope and extent of surveillance, thereby also increasing the volume of data that is gathered and available to be analyzed.

In recent decades, decision-makers in many areas, ranging from defense to weather forecasting, have argued that problems could be solved if more data were available. But merely increasing the volume of data gathered has not led to the hoped-for success. On the contrary, it has done just the opposite by overloading systems, networks, and, most importantly, human operators. Such factors not only jeopardize decision-making effectiveness, but also the adaptive capacities needed to assure the resilience of the decision-making process itself. New methods are needed to help decision-makers deal with the overwhelming amount of data being made available to them. In military Command and Control (C2), this challenge is intensified due to the diversity of data produced by multiple sources, the variety of information needing attention, and the time pressures that are a natural consequence of operations.

Apart from the volume of data facing decision makers, the variety of data types available in open source and social media domains provides unprecedented opportunities for contextual understanding of underlying human attitudes, motivations, and behaviors (in both a retrospective and predictive sense). The recent explosion of social networking sites available via mobile communication devices has drastically altered the landscape for text and image analysis. The proclivity of individuals around the globe to use social networking provides a window by which outsiders can rapidly identify, follow, understand (and potentially forecast) previously unknown activities and patterns. These socio-cultural attributes provide much needed context to the traditional military information standards.

In a recent interview appearing in Defense News, Acting Secretary of Defense for Research and Engineering Al Shaffer made the following comment:

> [One] area of technology surprise is in the area of human systems. We have two vectors there. The first vector would be in things like man-machine interface. The second... is in the whole area of cognition — how people take in information and react. That has applications for training. If everybody learns differently, if I can understand how you learn, I can tailor training systems and create a combat-ready person much more rapidly. The other part of cognition that comes into play is if we can reduce the amount of time it takes for a person to recognize a situation and react. [One] area of technology surprise is in the area of human systems. We have two vectors there. The first vector would be in things like man-machine interface. The second... is in the whole area of cognition — how people take in information and react. That has applications for training. If everybody learns differently, if I can understand how you learn, I can tailor training systems and create a combat-ready person much more rapidly. The

> other part of cognition that comes into play is if we can reduce the amount of time it takes for a person to recognize a situation and react.

In this paper, we consider strategic interdisciplinary research needed to produce transformational decision science capabilities for Warfighters at all levels across the Department of Defense (DOD). Our focus is on the needs of the human operator(s) who must rely on a variety of technologies to collect, interpret, and assimilate meaning from a variety of inputs in a complex and time-constrained environment. Toward that end we explore fundamentals of decision making and identify research challenges in the areas of human-machine collaborative sensemaking, task-relevant valuation and selection of information, and performance metrics for decision making at a genuine systems level. We conclude with a discussion of emerging military technologies that are designed to improve decision making in military domains and consider future trajectories that harness the power and opportunities provided by extremely large and disparate data. We turn first to the center of our exposition, human decision making.

Fundamentals of Decision-Making

It is widely held that decisions are made by acquiring information, processing that information, making a commitment to action, and then taking an action (see Hoffman & Yates, 2005). In much of the literature on decision-making, "the decision" is regarded as a point-like thing, a singular commitment that marks the end of a sequence of three or four clear-cut mental operations or stages. While this may sometimes be true, it is generally not true for the sorts of decisions that have to be made in complex sociotechnical work systems. Decisions are rarely simple input-output chains or fixed steps.

Recent research in the fields of "Expertise Studies" and "Judgment and Decision-Making" has explored complex decision-making in complex domains (Hirokawa & Poole, 1996; Yates, 1990). Judgment processes can be formalistic or substantive. Formalistic procedures are exemplified by the application of rules such as those of probability calculus or utility analysis. Such rules are largely indifferent to the content and context of judgment problems. Quite the opposite is true of substantive procedures, which entail the attempt to envision how the world may (or may not) create some event of interest. Research shows that people resort to formalistic procedures only when they cannot use substantive ones, which seem to be more "natural" but more typical of decision-making in complex sociotechnical systems, where context is crucial (Klein, 1989).

Historically, decision research has been dominated by questions about deviations of people's actual decision behavior from what is predicted or prescribed by logical rules such as the expected utility, additive utility, and discounting models. While attempting to form some basis for "rational" decision-making, this perspective reflects a narrow and overly idealistic conception of how people deal effectively with complex trade-offs in real life. Effective human decision processes involve significant cognitive, evaluative, and affective activities that are parallel and interactive. While deciding involves acquiring information, the acquisition of information might itself involve other decisions and deliberations such as sensemaking, noticing problems, generating possible solutions, choosing goals, developing implementation plans, and establishing methods for evaluating outcomes. While most "simple causal chain" theories regard decisions as culminations, decisions are often expressions of contingencies and anticipations of events yet to unfold and might be surprising (Hoffman & Yates, 2005). Thus, there is often no single clear-cut "end" point in decision-making. And even after a decision has been "made," there are contingencies. New decision issues emerge, either in the process of implementing a commitment or because a previous decision led to new threats and opportunities.

In making decisions in complex cognitive work, a number of issues have to be resolved, either tacitly or deliberatively:

- How do people come to recognize that existing or developing circumstances constitute threats or opportunities, and that a decision has to be made?
- Who will decide, and how will they approach that task?
- What kinds and amounts of resources will we invest in the process of deciding?
- What are the different actions we could potentially take to deal with this problem?
- What are the various things that could happen if we took that action, and which ones do we care about?
- Which of the things that we care about actually would happen if we took that action? How should we make the trade-offs that are required to settle on the action we will actually pursue?
- How do we contend with many stakeholders' sentiments concerning what is decided, how it is decided, and how it is implemented?

While a decision aid might seek to get people to a point of commitment, in complex cognitive work situations, events always follow the commitment. A commitment to act does not necessarily have action as its primary functionality. Rather, it is a resolution to accept a particular understanding in the hope that the understanding will serve to help the decider know when to be surprised after the action has commenced and the anticipated contingencies play themselves out, or not.

The modeling component of many decision aids basically involves taking input data; creating tabular representations of entities, attributes, and weightings; and performing mathematical operations predicated on notions of decision analysis. Decision analysis offers comfortable means to describe decision-making in terms of choice among probability distributions. It offers techniques to mathematically specify preferences, derive and evaluate probabilities, and work on equations that balance gain and risk. It provides mathematical methods to achieve consistency by rules of logic. This approach prescribes a decision process that involves identifying promising prospective courses of action and their potential significant consequences (step 1), assessing the utility of those consequences and evaluating the likelihoods of all the recognized potential outcomes (step 2), and then selecting the alternative that is indicated to be best according to a "rational" decision rule (step 3).

Assuming this is what deciding is, then surely people must need help with these things. Over the years, this view has had a substantial influence on the character of the literature on human biases and limitations, and hence it has been formative of entire programs of resarch and development on decision-aiding. But the promise has not caught up to the reality. "Behavior-focused decision aids have had little documented success ... Deciders therefore often ignore such aids because they appear irrelevant to significant decider concerns. And when deciders do try the aids, the results disappoint them because the aids leave untouched the quality dimensions that matter to them" (Fischhoff, 1986, p. 13).

Research Challenge #1: Human-Machine Collaborative Sensemaking

New methods are needed to help decision-makers deal with the overwhelming amount of information being made available to them to conduct their missions (Endsley & Hoffman, 2002; Klein et. al., 2004). They must consider all data as potentially relevant and must integrate cyber, geospatial and non-geospatial data (e.g., computer networks; representation of hypotheses) in a cohesive manner. Shortcomings of current visualization methods include the following:

- *Technology-centeredness*. Current technology focuses on what can easily be shown, rather than presenting what needs to be known — or helping people discover it themselves. For example, depictions of network topology abound, but they do not generally provide the kind of insight into the relevant hypotheses desired by cyber analysts. What analysts ultimately want to know is not just the status of the network, but rather the potential impact of current events and trends on the overall mission.
- *Insensitivity to issues of human perception and cognition*. Often people insist on the need to "see everything." However, the scale of modern systems makes the exclusive reliance on human interpretation impractical.
- *Insufficient interactivity.* Effective decision-making requires more than just "seeing" the data but also requires being able to explore and interact with it. In addition, it requires the capability for the decision-maker to take action when necessary without having to move to a different display or software application.
- *Retrospective emphasis*, showing something that had happened, rather than helping decision-makers anticipate what might happen next through the extrapolation of current trends.
- *Overreliance on visual information.* Only the visual channel is exploited, leaving other sensory modalities underutilized.

From a human-centered computing perspective, the value of a given visualization can only be determined with respect to its usefulness in addressing the questions and issues brought to it by the decision-maker. Though sensitivity to visual aesthetics is a desirable quality for visualization designers, there is no necessary correlation between the attractiveness of a given display and its usefulness. This has been shown to be true in a number of applications. Just because a display is aesthetically pleasing doesn't mean that it's a "better" display. The key insight here is that visualizations do not stand alone, but are part of a larger cognitive and social process of sensemaking.

Sensemaking is a motivated, continuous effort to understand connections (which can be among people, places, and events) in order to anticipate their trajectories and act effectively (Klein, *et al.*, 2006a, b).

Sensemaking is how cognition adapts to complexity (Moore, 2011; Moore and Hoffman, 2011). Figure 1 illustrates what Klein and his colleagues call the "data/frame theory of sensemaking" (2006b, p. 89). At the most basic level, the theory acknowledges that the understanding of situations always occurs with respect to a framing perspective. The frame constitutes a set of more or less coherent hypotheses about the data to be understood, and serves both to determine what counts as data of interest and to shape the interpretation of the data. Note the absence of input and output arrows in the diagram. The sensemaking process can start, or recommence at any point, even though it is often triggered by surprise.

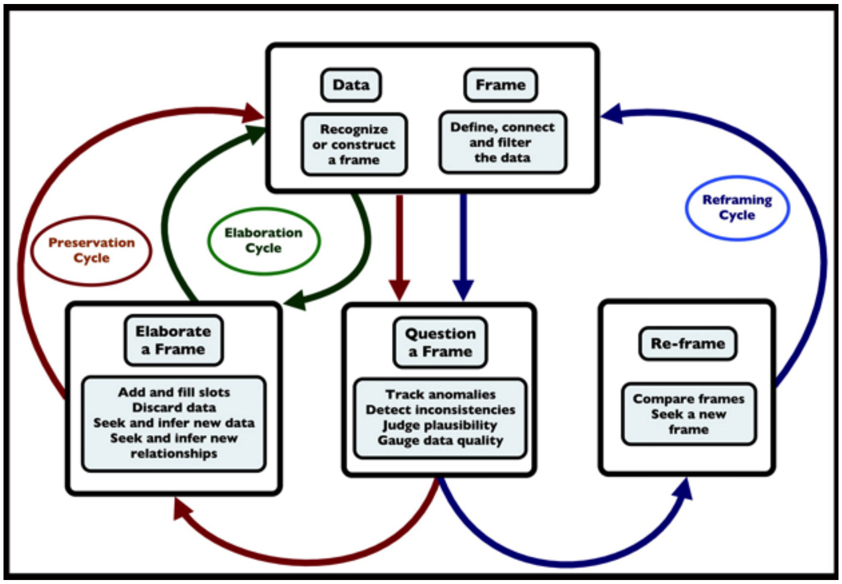

Figure 1. The Data/Frame Model of Sensemaking.

As data accumulate, the sensemaker may be confronted with the question of whether to elaborate a current frame by incorporating new details, or to seek a new frame that better accounts for current findings. The process involved in the ongoing evaluation of a given frame includes the possibility of a closed-loop alternation between backward-looking mental model formation—which seeks to explain past events—and forward-looking mental simulation—which anticipates future events.

Significant research already has been performed in order to discover ways in which the decision-maker's sensemaking process might be shaped in order to help them counteract lines of reasoning that might lead to misconceptions (Weick, 1995).

A basic foundation for sensemaking having recently been laid already in the research literature, a next step is implementation of sensemaking support systems that can harness the joint power of humans and machines. In particular, an understanding is needed of the potential impact (positive and negative) of new forms of visualization and automation as part of the sensemaking process, and how such tools ought to be designed in light of what we already know. We are not asserting here that sensemaking can be automated, only that technology is formative of the sensemaking process. There is a need to examine questions about the role and benefits of computer interaction with people in center stage.

In their discussion of the Data/Frame theory, Klein, *et al.* conjecture that the role of machines in assisting people with sensemaking may not be merely to confirm or disconfirm the accuracy of a particular interpretation with respect to a given frame, but also as an aid in the reasoning process that leads to the possibility of reframing: "The implication is that people might benefit more from intelligent systems that guide the improvement of frames than from systems that generate alternative understandings and hypotheses and foist them on the human" (2006a, p. 89). This conjecture is consistent with the view of Woods (Woods and Hollnagel, 2006), who have adopted a stance to resilience engineering that takes as its basic assumption that "human systems [are] able to examine, reflect, anticipate, and learn [i.e., engage in sensemaking] about [their] own adaptive capacity" (Klein et. al., 2006b, p. 128).

A variety of emerging technological capabilities relate to machine-assisted sensemaking merit thorough investigation. Here are three examples of such topics:

- *Prescriptive guidelines for the design of visualizations that are informed by principles of human perception and cognition.* For example, displays that rely on the ambient visual channel have proven their effectiveness more than a decade ago (Still, *et al.*, 2001, 2004). The ambient channel is used primarily for tasks involving both focus and movement, such as locomotion that can be accomplished without conscious effort or even awareness. For example, ambient vision is used by people to quickly and successfully navigate crowded hallways without conscious thought or to catch a football on the run. Because displays relying on ambient vision occupy a middle ground between displays designed for use by the peripheral and foveal vision channels, they can excel when there exists a large amount of information requiring continual monitoring and response. Other underutilized principles of display design include proportionately scaled symbology, holistic foreground against contextual background, structure from motion, pop-out, and chunking. Performance models that describe what is normal in a given context provide the data necessary to help people or systems recognize what is anomalous in the displays, helping them know when reframing may be advisable.

- *The use of software agents as an adjunct to human sensemaking*. By their ability to operate independently in complex situations without constant human supervision, collaborating teams of software agents can help people perform sensemaking on a scale that would be impossible for other approaches (Bunch, *et al*., 2012), especially when it can be tuned to the idiosyncracies of teams and individuals (e.g., mental ability, cognitive style, experience). Working coactively with people, agents can assist with taskwork of identifying complex or high-tempo patterns of interest in data and tagging them so they can be made visually salient in the display (Bradshaw, *et al*, 2012). In this way, agents can be used to elaborate the current sensemaking frame. "Devil's advocate" agents can be used to seek disconfirming evidence of a hypothesis under consideration, thus assisting in the reframing aspect of sensemaking. In addition, agents can help with process of coordination of teamwork; helping people become aware of pertinent information coming from others, synchronize handoffs, and realize when progress is running ahead or behind expectations (Feltovich, *et al.*, 2007).

- *Incorporating displays and analytics that assist with some of the neglected aspects of deciding*. These include methods to help people determine whether there's a significant decision problem to solve in the first place, develop promising alternatives, envision non-obvious but critical potential side effects of alternatives, and discern how key parties truly feel about possible outcomes of selected options as well as the decision process itself (Hoffman and Yates, 2005). It is precisely these other tough and crucial aspects of deciding that often spell the difference between effective and ineffective deciding and thus are ones where help is required.

- *Enhancing consequential elements of the entire decision process.* When we trace the history of a decision process, it's always possible to identify one or more moments of choice. We can then describe history in terms of causal steps leading up to that moment, creating a simple causal model that might then be amenable to specification in terms of rules. But when we look at deciding as it occurs, a different picture emerges. People can reach moments of commitment that signal their occurrence clearly but are never achieved by following precisely the same path. People are not engaging a cause-effect chain or a rule-based process. They're navigating a space of constraints and issues, involving contingencies and contextual dependencies. We expect that visualizations and analytics to address this ongoing process of decision-making might enhance decision quality (Savikhin *et al.*, 2008).

Research Challenge #2: Task Relevant Valuation and Selection of Information

Given an unbounded amount of available information, even the best decision-maker might make poor decisions. Thus, the problem is one of managing the attentional resources of the decision-maker, and ideally having computational mechanisms filter the available information down to just a limited subset of the available information, a subset that has been verified and validated and determined to be task or goal-relevant. We will discuss this problem using the three following definitions:

- *Perceptual Salience*: The ways in which displays, auditory alerts, and graphical objects can be designed so as to capture attention (e.g., by color coding alerts, etc.). Most of the research on information salience, including the computational models, is aimed at this graphical-display level of understanding, that is, visual salience (Itti & Koch, 2000; Wolfe, 1998), and it commonly implicates or leverages low-level biological mechanisms.
- *Information Salience*: The ability of information to capture the attention of the decision-maker, through either bottom-up (feature- or pattern-driven) aspects of the information, or through top-down (knowledge-driven) aspects of the information. The bottom-up aspect of Information Salience overlaps with Perceptual Salience. The top-down aspect, on the other hand, is commonly associated with the deliberate control of attention.
- *Information Relevance*: The pertinence and utility of the information in actually making a decision; that is, information that should actually shape or determine the course of decision-making. Such information is usually task- or goal-specific, and relates to the context or situation as much as the aims of the consumer of the informational analysis.

Perceptual salience, information salience, and information relevance can work together or against each other. Perceptual salience can be co-incident with information salience, but they are not necessarily the same thing, as extensive practice can also influence information salience (Hoffman and Fiore, 2007). Information salience and Information relevance produce the best decision-making when they are aligned and concordant. In this case, the display elements that are visually salient convey exactly that information that is also highly relevant and will lead to good decision-making. On the other hand, the classic demonstration of misaligned Information salience and information relevance is the Stroop Effect, where naming color words presented in colored fonts causes a conflict between top-down relevance and bottom-up salience, resulting in measurably worse performance (Melara & Algom, 2003). Thus, some approaches aim to aid the decision-maker by applying context awareness to enhance the salience of critical information (Fischer, 2012).

Salience itself is typically considered in the context of visual or auditory tasks in which attention is easily measured (for example, through eye tracking), and the task is simple and fixed and can be carefully controlled. In these contexts, Salience and Attention are often considered to be one and the same, though it is often the bottom-up aspect of involuntary processing that the term Salience is focused on. In these laboratory tasks, millions of trials of experiments have been conducted (Wolfe, 1998), resulting in a fairly clear definition of display factors that capture attention (e.g., color contrasts; see Trafton, *et al*, 2000.). However, eye tracking is not terribly informative. Where a person is looking does not tell you what they are thinking. A better measure is percentage of time eyes are closed during each blink, as a measure of task engagement — unfortunately that is also just as heavily influenced by room illumination level (Halverfson et. al., 2012).

The decision-maker's knowledge and level of expertise plays a key role in determining Information salience, and in shaping top-down attention management (Hoffman and Fiore, 2007). The top-down aspect of information salience allows an expert to quickly sift through reams of information to locate key pieces of information, without being distracted by irrelevant information. They develop what they perceive to be effective strategies for handling excess information processing demands. For example, Woods and Sarter (2010) described the ability of experienced workers in nuclear power plant control rooms to ignore (actually, to filter and hence remain largely unaware of) unimportant auditory cues, while interrupting their processing to attend to those that do matter. More generally, this ability of experts to automatically orient to important information is described as the acquisition of automaticity They also have to know when to go beyond automatic processing, when to know that something important has changed, and need to think critically and problem solve. Schneider and Shiffrin, 1977; Shiffrin and Schneider, 1977) or as perceptual learning (Hoffman and Fiore, 2007), and models of this process have implicated pattern recognition (Best, *et al*., 2007).

Experts thus combine their knowledge and understanding in order to focus on appropriate context-specific goals. They then leverage a top-down driven automatic filtering ability acquired through extensive practice. For individuals having less experience, however, the perceptual salience of information can drive attention to items that are less relevant, or are irrelevant, thus having a negative impact on decision-making. The primary challenge (and opportunity) in aiding non-experts is to identify the appropriate goal and then harness the human innate bottom-up filtering ability by maximizing the congruence between information salience and information relevance.

Findings from studies relating to this research challenge #2 could be directly applied also to challenge #1, improving human-machine collaborative sensemaking, by using the results of research on salience and task relevance to the design of software agents and visualizations that can be better tuned to what the decision-maker needs to know.

From an information processing perspective, being able to help determine relevance computationally when possible would be an important component in the arsenal of techniques that would need to be applied to handle the vast amount of information available (e.g., Matheson, 1988). The value estimation is then used to prioritize information that is provided to the operators. Such prioritization is essential to reduce operator workload. In tactical environments, prioritization is also essential given the bandwidth constraints in place, which makes it impossible to transmit the vast volumes of information available. Prioritization is also essential given that displays in tactical environments tend to be small, thereby limiting the amount of information that can be displayed and navigated. Finally, tactical users may be operating in dangerous environments, where minimizing distractions caused by unnecessary information is important.

A pre-requisite to determining the value of information to a user is to be able to model the user's mission or task at hand. Specifically, once a model is formed of task goals, and the ways in which the task goals may change as progress is made toward a broader mission goal, then information can be matched against individual user contexts and ranked or displayed appropriately. Some initial work along these lines for tactical networking environments is described in Rota et. al. (2010).

Research Challenge #3: Performance Metrics at a Genuine Systems Level

In the context of complex sociotechnical work systems, significant challenges arise with regard to measurement. Some of the challenges that emerge at the system level call into question the extensibility of comfortable notions of human performance measurement when what one really needs to evaluate is the goodness of cognitive work in complex sociotechnical systems.

Discussions of measurement present useful approaches to evaluating validity and reliability, but depend uncritically on distinctions such as that between subjective and objective measurement. Standard treatments (O'Neill, 2007; Stanton *et al.,* 2005; Wilson & Corlett, 2005) focus on measurement of the performance of individuals (for example, response correctness, errors, reaction time, frequencies of behaviors or behavior types, and so on). Most measures of human performance are measures of hits, errors, accuracy, and time. These enter into calculations of efficiency, effort, or other concatenations (for example, relations of speed and accuracy). The design and analysis of complex cognitive work systems needs a measurement approach that goes beyond these things (Hoffman, 2009). Measures of the individual worker's raw performance have important and necessary uses, but are not adequate to the understanding of complex sociotechnical work systems, which is where the payoff for investment really comes.

It is often the things that are easy to measure that are the things that get measured. And once measures (or metrics) are entrenched, then the measures become important programmatically, and things that are genuinely important go unmeasured. Often it can be relatively easy to measure performance in terms of the primary task goals, for example by counting the number of documents analyzed, the number of targets identified, and so forth. It can be easy to evaluate software usability by measuring response time on well-defined and regularly-occurring action sequences, or by collecting practitioner self-ratings of mental workload. As National Institute of Science and Technology (NIST) researcher Jean Scholtz (2005) and others have argued, what is critical and more difficult is to develop methods for measuring work at a meaningful, system level. We need to measure such things as the goodness of the technology, the learnability of work methods, and the resilience and adaptability of the work system.

The importance of systems-level measurement and analysis has been recognized for many years by engineers and others who have studied dynamic systems (Bar-Yam, 2003; Jagacinski & Flach, 2003), and by industrial and human factors psychologists who have witnessed the emergence of complex cognitive systems (e.g., Goguen, 1994; Neville *et al.*, 2007). This trend is exemplified by discussions of such topics as how to integrate human-systems analysis into systems engineering (Deal, 2007) and how to determine the costs of integrating (or costs of not integrating) cognitive systems engineering methodologies into systems engineering (Zachary *et al.*, 2007).

The process of using multiple measures to form meaningful measurement scales is referred to as conjoint or derived measurement. There are many kinds of conjoint measurement structures, defined by such factors as the independence or non-independence of the individual measures (Krantz, 1972). For instance, an evaluation of a new interface might involve measuring: usability based on a rating administered after an initial practice period, plus learnability measured by the number of practice trials it takes for participants to reach a level of 85 percent correct (say) across the trials in the practice period. With regard to cognitive work systems, one might want to evaluate them for resilience by looking at performance on tough cases, or performance when a mission (or activity) gets derailed. One might want to evaluate a cognitive work system with regard to teamwork functions, and look at whether team members can describe other team members' goals, anticipate other team members' needs, or cope with goal conflicts.

But in all such cases, precisely what would one measure? How would one forge a meaningful set of measures and then create a measurement structure that allows one to conjoin the measurements into meaningful scales that map onto the policies that might be used to set metrics? How can a cognitive work system be designed such that effects on intrinsic motivation can be considered (Hoffman *et al.*, 2008)? How can we evaluate the extent to which new technology accelerates achievement of proficiency, or the ability of workers to cope with rare or tough cases (Hoffman et al., 2009; Hoffman *et al.*, 2004)? How might it be possible to track changes in work such as 'the discovery of toolness'? What is it about some new software that workers find valuable? Does software move workers toward new ways of working, even ways not anticipated by the designer? To ask such questions is to take a first step in creating and refining a scheme for systems-level measurement.

But the step from conceptual measurable to an operational definition of a measure is neither direct nor easily come-by in the case of complex cognitive systems. *A critical outstanding need for the study of cognitive work is to extend conceptual definitions of cognitive concepts (functions and supporting functions) to operational definitions.*

Each high-level cognitive function—sensemaking, replanning, problem detection, etc. —will be the basis for multiple measures, and not just one measure. In developing the measures, one searches for domain-specific or appropriate aspects of cognitive functions that can be used to evaluate hypotheses. For example, consider the function of coordinating (Klein, *et al.*, 2004). Members of a team must share some knowledge in order to collaborate, replan, and so forth. Achieving common ground requires establishing a set of shared goals, and assigning roles and responsibilities. Maintaining common ground is a continual process of communicating and coordinating, of updating knowledge and beliefs, of anticipating needs and activities. This requires shared beliefs about what each team member believes and knows, about each team member's intent, capabilities, and so forth. Such a roster of features can be taken as the conceptual definition of sensemaking. The challenge is then to link the conceptual definitions to operational definitions and translate those definitions into design, training, and operational practices. As a case in point regarding systems-level thinking, coordination cannot be measured simply in terms of the frequency and extent of communication. Well-functioning teams may explicitly communicate *less* precisely because they share knowledge and beliefs.

Measurement for complex cognitive systems must be sensitive to trade-offs, that is, comparisons of increases or decreases of one sort against increases or decreases of another sort. For example, there is a trade-off between expanding the expert's range of 'the familiar' with an increased likelihood that when something is a surprise, it will be particularly dangerous. This is not just a matter of using measures that can be placed in ratios, but using measures in the context of experimental designs that allow for the discovery that there are tradeoffs, and what their magnitude and importance are. For example, Wulf, *et al.* (2002) found that training programs that provided rapid and accurate feedback significantly improved the learning curve of the trainees. Metrical guidance based on this finding would lead to a decision to provide rapid and accurate feedback. However, it turned out that the feedback reduced performance when trainees moved into the actual work context. The reason is that the group getting rapid and accurate feedback never had the opportunity to mull over their errors and re-think the problems. Thus they never developed skills in generating their own feedback, and so on the job they were handicapped.

Science and Technology Applications for Research Challenges

Thus far we have examined three critical and intertwined research challenges that characterize the interaction between human and machine analysis of complex factors that are compounded by the proliferation of a variety of information sources. In this data-rich environment automated tools are critical to the efficient workflow of collection, processing, exploitation, and dissemination. Integrating human decision makers into this workflow at critical junctions is a non-trivial task. We now turn our attention to consideration of how specific research activities can address the three research challenges.

Decades of science and technology research in the DOD have produced intelligence capabilities and weapons platforms that far surpass those possessed by other nations, but these are grounded primarily in mathematical, physical, material and engineering sciences (Picucci & Numrich, 2010). As noted in Flynn, Pottinger and Batchelor (2010), US forces must focus on people and their native environment to achieve success in irregular warfare (and some might argue in major combat operations as well). Social network analysis is one accepted method by which analysts can develop and maintain tactical awareness of social and cultural relationships and behaviors. Another is traditional information processing and visualization displays that incorporate nontraditional data such as that contained in social media applications (both text and imagery). In the remaining section of this paper we propose several recommendations for a long-term research effort that leverages data science, decision science, and novel information processing technologies.

Fusing Socio-Cultural Data for Discovery

Irregular warfare, non-state terrorism movements, and unstable environmental patterns that trigger major weather disasters make it difficult for military and government leaders to rely on traditional physics-based sensors alone to plan current and future actions. Such sensors do not provide awareness of dynamic context in the area of operations, and this context is critical to monitor goals, functions, and data needs [11]. Strategies for achieving contextual understanding can include observational data, a priori knowledge models, and inductive knowledge [12]. Contextual understanding is generally achieved through a combination of human and computer processing techniques that take advantage of a person's cognitive ability to fuse and assimilate multiple sources and types of information for new insights [13]. In irregular warfare environments, it is critical to incorporate both hard and soft data to gain an understanding of the delicate balance between individuals and groups in society and the environments (geopolitical, social, agricultural, etc.) upon which they depend. Test and evaluation of methods that fuse hard and soft data are challenging due to the nature of the data, the test environment, and the metrics for determining outcomes [14]. Unlike much physical sensor information, however, data sources for this new type of problem are not classified or difficult to obtain; open source data is available and plentiful. Because it is collected by a diverse group of researchers, it is scattered in a multitude of organizational domains but frequently available with little effort. The challenge becomes correlating data of many different types that represent various aspects of a region of interest. This approach is similar to the signal processing approach of weak signal detection, which is used to extract received signals [15], identify images in noisy backgrounds [16], and conduct remote sensing of land and water resources for sustainable development of natural resources [17]. A novel approach to correlating a variety of data sources to understand problems in an area and forecast conflict is described by [18]. In this example, the potential conflict is the weak signal that is detected through the correlation of diverse datasets describing many features of the region, to include demographic, political, social, economic, educational, agricultural, weather, etc.

The challenges here are several in number. First, a method for collecting various datasets for a region of interest and correlating these for overall understanding and meaning is a nontrivial task. Many of the datasets are based on different scales and involve different referents to the population or the environment. A key challenge in integrating these disparate data is the semantic meaning implicit in the components of the overall structure of the region. Additional research on the use of semantic technologies such as OWL for this purpose may have great potential. Second, a weighting scale must be developed sufficient to provide representational meaning and inferential capabilities to the reasoning tool. Third, a visual representation must be developed sufficient for a human to reason about the correlations; a display that provides all of the facts but does not suggest inferences is insufficient and meaningless. Finally, a performance measurement capability is needed to compare the reasoning analytics to reasonable expectations for use of such a tool. Quantitative and qualitative metrics will be needed for such an application.

Information Processing for Tactical Intelligence

Social networks have become common adjuncts to military operations. However, today's analysts are faced with manual methods for compiling these networks from a variety of text documents. With the rapid accumulation of documents available for analysis, a capability for automated generation of the social network is much in demand. To complicate this matter, such a capability must be able to function over large and poorly formatted databases, which will include multi-sensor data feeds to include text and biometric data. Technology is needed to provide software capable of ingesting multi-source data (text, biometrics) to produce social network relationships and predict high value targets (persons, objects, or locations). Social network algorithms that are adapted to large data sets and highly uncertain data will serve to structure knowledge for discovery and provide intuitive and user-adaptive visualizations that allow references to original data sources. A visualization capability for these large and poorly structured networks will improve layered analysis and interpretation, and improve prediction for decision support.

Tactical Warfighters demand rapid information fusion capabilities to develop and maintain accurate situational awareness and understanding of dynamic enemy threats in asymmetric military operations. The ability to extract meaning in relationships between people, objects, and locations from a variety of text and multi-source datasets is critical to proactive decision making. Because of the changeable nature of these human-centric relationships, these data always represent some measure of uncertainty and reflect reality in an abstract, uncertain, and time-limited way. New algorithms that support visual interpretation of associations in the data are needed due to the complexity of interactions among component parts, the dynamic nature of the relationships, and the short time periods available for analysis and decision making.

Fusion of uncertain and multi-source information is attracting a growing interest of practitioners and researchers [19, 20, 21, 22]. The poor quality of the data stems from the quantity (far in excess of what a human analyst can process), the uncertainty or noisy features of the data (contain errors, inconsistencies, and potential deceptions), and the incomplete and ambiguous (can be interpreted in many different ways). Social network analysis techniques are frequently one output of this fusion process and are an aid to understanding complex relationships. However, as the corpus of information input increases and the quality of the data decreases, the resulting social network graphs will also become more complicated and difficult to analyze. To that end, a final challenge is to develop new algorithms to produce meaningful social networks for predictive analysis.

Representing Cultural Features in a Social Network

Knowledge of sociocultural influences on attitudes and behaviors in a population is critical to effective interactions on the part of military personnel with local populations. As General (Retired) Anthony Zinni, U.S. Marine Corps, former U.S. Central Command Commander stated, "You have to understand the culture you're getting involved in. We never do a good job of culture intelligence, of understanding what makes people tick, what their structure is, where authority lies. Culture bias limits our ability to understand what is going on around us." [32] Sociocultural modeling techniques have progressed in recent years, as evidenced by the plethora of games built on a variety of sociocultural domains [33], yet these models have not been effectively transferred to decision support capabilities.

Modern warfare and conflict environments are drastically different from what was once considered to be the norm [41]. Where once the norm consisted of countryside "battlefield combat with distinct front lines," modern conflict increasingly occurs in urban areas lacking distinct boundaries [41], within foreign cultures, where the focus is centered on the civilian population instead of the battlefield [34]. The discerning combatants from civilians has become increasingly complex in that combatants are frequently dispersed throughout the civilian population and without any clear uniform, it can be extremely difficult to discern friend from foe [41]. To further complicate combatant and civilian distinctions, a civilian encountered as such one day may present as a combatant another. The conventional goal of overcoming an armed enemy is expanding, incorporating goodwill missions where the goal is to win over local civilians. Operating in foreign environments, warfighters often find themselves in unfamiliar situations where they need to know how to resolve a situation appropriately in the context of an alien culture [37]. Due to these expansions in military conflict norms, sociocultural knowledge has become a critical factor for success in modern warfare environments [8]. Soldiers must understand a society's values, motivations, culture, and subcultures within. For mission success, it is critical for all Soldiers and commanders to maintain cultural situational awareness when in a foreign environment [34]. Effective situational awareness depends on the ability to collect data from many distributed, heterogeneous information-sources and through visual analytics, display the data so that it facilitates understanding of evolving events occurring within complex and dynamic environments [36]. The military has applied visualization techniques to enhance decision making [35]. Incorporating culturally significant information into a military database accessible by Soldiers and commanders will enhance decision-making.

The challenges in this area are threefold. First, building and maintaining foreign cultural awareness and understanding societies where military operations are conducted has not always been a priority of the U.S. military [40]. It has been suggested that the lack of preliminary efforts to understand the local populace and culture that our forces operate in resulted in many of the early challenges encountered during Operations Iraqi Freedom and Enduring Freedom (OEF and OEF) [40]. Therefore, we are having to back-track in order to develop and share cultural awareness while still immersed in these foreign cultures. Secondly, there are many challenges regarding data collection, data entry, data management, and quality. For instance, in regards to data entry, a large issue occurs with entity resolution and relationship awareness. How will a user discern two individuals with the same name? Answering these questions as well as exploring how to best visually represent this information to a user will be some challenges developers will confront. Regarding data quality, given that information may be coming from many different sources, data may overlap, be incomplete, or incorrect [38]. Determining how to overcome and compensate for these issues is an ongoing challenge for developers. Developers will have to determine how to fuse information collected from various types of data sources into meaningful information that will enhance human understanding without increasing user stress [36] and while compensating for any data errors [38].

Military analysts are often focused on social groups with an interest in hiding behavior, such as criminal or terrorist networks. In these conditions, innovative methods are needed to identify proxy features of a network that may aid discovery goals. Temporal trends are one such category. Examples may include factors such as frequency of contacts between nodes or clusters, inter-contact time, recurrent contacts, time order of contacts along a path, and delay path of information diffusion. Methods to extract, characterize, and monitor social networks dynamically over time is a research challenge of interest.

Scalability and predictability have been perennial problems in social network analysis. As networks increase exponentially in size and complexity, it is harder to use graphical methods to represent, monitor, and understand network behavior. The representational graphs grow to unmanageable size, contain complex relationships among nodes, and often contain several varieties of nodes. Two promising approaches are being explored by the SNA community, visual analytics and semantic analysis. Visual analytic methods supported by ontology have been shown to reduce the visual complexity of these graphs to enable users to identify important structural and semantic aspects of networks. Research is needed to identify key actors and supported relationships, detect the presence of bridging nodes that can uncover hidden sub-networks, and determine the flow of resources (information, money, influence) within the social network.

Semantic analysis of social network data can be used to identify sub-topic discussions that can lead to effective course of action planning. Social media products allow an analyst to access the insider perspective of the group as represented in the shared information. The extent to which such information can be trusted or accepted as normative behavior for a larger societal group remains a challenging research need. The value of semantic analysis is the ability to identify topics of interest, cluster these in terms of high-frequency text occurrences, assign positive and negative ratings to words and phrases, and trend these in time. For example, one might wish to monitor messages in a social media application to monitor the public's attitudes toward a leader. Analysis of these data should be mined in such a way to show spikes in sentiment, polarization of sentiment, and rates of posted text.

A further component of social network analysis closely related to sentiment analysis is topic or concept extraction. This approach can be used to cluster actors around shared interests or to discover relationships among individuals and topics. As an example of how topic extraction and co-referencing can be useful in the military domain, we might consider the case of an increasingly unstable nation. One might expect fluctuations in political, social, religious, and economic leaders, influencers, and protagonists. Without a cadre of experts on the ground, such intelligence would be difficult to gather. Topic extraction and relationship linking tools can potentially provide a monitoring capability to detect changes in established trends based on newspaper reports and text blogs. Such a capability could be used to monitor the population's expressed support for one leader or another, or for proposed military or diplomatic actions. Irregular warfare, non-state terrorism movements, and uncertain environmental patterns that trigger major weather disasters are examples of events that require military response. In responding, decision makers will use text and imagery analytics to develop the necessary contextual understanding of the region and key elements. Strategies for achieving contextual understanding can include observational data, a priori knowledge models, and inductive knowledge. Contextual understanding is generally achieved through a combination of human and computer processing techniques that take advantage of a person's cognitive ability to fuse and assimilate multiple sources and types of information for new insights. Correlation and aggregation of open source data, such as agriculture, weather, terrain, demographics, economics, social patterns etc. is nontrivial but vital to effective military response.

Many areas of the world where future military action may be required are rich in language or dialect diversity. To fully engage local populations and respond to humanitarian needs, language translation will become critical to text analytics efforts. Natural Language Processing (NLP) and computational linguistics methods and algorithms are needed to develop and improve technologies for machine translation, information extraction, and automated summarization. Also relevant are methods and algorithms to develop and improve technologies for optical character recognition (OCR) and speech recognition as input to machine translation, information extraction, and automated summarization. Development of language data in support of building these technologies and development of metrics to evaluate underlying software algorithms are also needed.

Social media includes web-based and mobile technologies used to turn communication into interactive dialogue across the global internet. Social interaction is now possible on demand through mobile devices with seemingly unlimited applications (apps) to suit a variety of interests. The practical implications of this technology include the choice of how people share information (e.g., through text, images, and video), with whom they share, and the purposes for which they share (e.g., to inform or influence). The ability to access social media sites on mobile apps allows immediate sharing of information, often to the global community. Even if information is shared with a selected friendship group, re-transmission of information to other groups can lead to global sharing within very short timeframes. In these ways, social media communication is used to form identity, share information, and establish relationships. While analysis of these networks can inform trend analysis of various social factors, such study must be somewhat opportunistic and adaptive because these social networks are dynamic and undergoing constant change as the persons, organizations, and activities within them.

Online social media sites continue to be used for political and economic purposes with most major news agencies and businesses advertising their site and encouraging viewers or customers to 'join them online.' In research circles, data from many of these sites is available and used for a variety of purposes, primarily to develop and test extraction and referencing tools. Online sites serve to reduce the cost of collection and provide first-person attribution of the information placed online. A major source of error in collecting human-provided data is the processing that takes place when an observer views and records an action from a cultural outsider's point of view. While we acknowledge that not every contributor to online media sites is completely trustworthy and may exaggerate or minimize statements, in general, statements from first person reporters are preferred over second-hand reports. In certain circumstances, however, deception is clearly the intended purpose (as in scams, for instance). Analysis methods capable of detecting deception would be particularly useful for online data mining. Authorship affiliation technologies applied to various types of text are considered an unmet research need.

Discussion

This White Paper outlined some fundamental research areas that must be pursued to address the larger problem of effective decision-making in the face of vast amounts of data on complex and dynamic situations. We considered the research and development challenges from the standpoint of human-centered computing and suggested a way forward for meeting the challenges, in terms of emerging and promising designs for complex cognitive work systems, including their computational capabilities.

We note that it is relatively easy to build complex systems and then put people into them. Some of our complex, knowledge-intensive systems have now reached the point where that strategy obviously is no longer viable. These are a tendency for technology centeredness, a retrospective emphasis on information, and an overreliance on the visual sensory channel. We noted that a solid foundation has been laid by decision scientists, and the next step of implementing a sensemaking support system capable of harnessing the joint power of humans and machines must now be sought. In that endeavor, we argued that emerging visualization and decision support capabilities should be designed based on the principles of human perception and cognition. Using software agents as sensemaking adjuncts, designing display components that assist with neglected aspects of deciding, and enhancing the consequential elements of the entire decision process are critical features of necessary enhancements. We further considered the complicated nature determining the relevance and value of information. As computational algorithms increasingly segment, parse, aggregate and filter information, they must do so in a way that preserves the value of the information to the human user's tasks and goals. Finally, we examined the challenges associated with measurement in complex socio-technical work systems. We note that a critical outstanding need is to extend conceptual definitions of cognitive concepts (both primary and supporting functions) to operational definitions and to be sensitive to trade-offs that exist in these systems.

Recommendations

In light of the foregoing discussion, we suggest that research activities be undertaken with respect to the following focus areas:

- Promote research on developing decision aids for dealing with massive data that leverage human-machine collaborative sensemaking theories and approaches. Topics of interest would include prescriptive guidelines for and examples of interactive visualizations that are informed by principles of human perception and cognition; the use of technologies such as software agents as an adjunct to human sensemaking; displays and analytics that assist with problem framing, alternative generation, and potential outcome assessment; methodologies and tools that support decision-making as an ongoing process of continuous negotiation and monitoring rather than merely a specific point of commitment.
- Support research that explores computational approaches to determining the value of information, taking into account the context of the users who will consume the information as well as their current task and their historical biases. An equally important adjunct to this topic is determining when users have sufficient information (and not too much) to make good decisions.
- Leverage software agent technologies to develop decision aids that bridge the gap between display salience and task relevance.

- Promote research on developing useful measurement techniques based on a systems-level analysis of the key trade-off functions that determine the effectiveness and resilience of decision-making within the context of complex cognitive work systems.

References

Bar-Yam, Y. (2003). The dynamics of complex systems. Boulder, CO: Westview Press.

Best, B. J., Gonzalez, C., Young, M. D., Healy, A. F., & Bourne, L. E., Jr. (2007). Modeling Automaticity and Strategy Selection in Dynamic Visual Detection. *Proceedings of the Sixteenth Conference on Behavior Representation in Modeling and Simulation* (pp. 3-11). Orlando, FL, USA: Simulation Interoperability Standards Organization. Retrieved from http://repository.cmu.edu/sds/77/

Bradshaw, Jeffrey M., Marco Carvalho, Larry Bunch, Tom Eskridge, Paul J. Feltovich, Matthew Johnson, and Dan Kidwell. "Sol: An Agent-Based Framework for Cyber Situation Awareness." *Künstliche Intelligenz* 26, no. 2 (2012): 127-40.

Bunch, L., Bradshaw, J.M., Carvalho, M., Eskridge, T., Feltovich, P., Lott J., and Uszok A. (2012). Human-Agent Teamwork in Cyber Operations: Supporting Co-Evolution of Tasks and Artifacts with Luna. Invited Paper in Ingo J. Timm and Christian Guttmann (eds.), Multiagent System Technologies: Proceedings of the Tenth German Conference on Multiagent System Technologies (MATES 2012), Trier, Germany, 10-12 October 2012. Berlin, Germany: Springer, LNAI 7598, pp. 53-67.

Deal. S. (Organizer) (2007). Human Systems Integration (HSI) Interest Meeting. Held at the Human Factors and Ergonomics Society 51st Annual Meeting, Baltimore, MD.

Endsley, A. and Hoffman, R. (November/December 2002). The Sacagawea Principle. *IEEE Intelligent Systems*, pp. 80-85.

Feltovich, Paul, Jeffrey M. Bradshaw, William J. Clancey, Matthew Johnson, and Larry Bunch. "Progress appraisal as a challenging element of coordination in human and machine joint activity." Presented at the Engineering Societies for the Agents World VIII, Athens Greece, October, 2007.

Fischer, G. (2012). *Context-aware systems*. New York: ACM Press.
doi:10.1145/2254556.2254611

Fischhoff, B. (1986). Decision-making in Complex Systems. In E. Hollgagel, G. Mancini, and D. Woods (Eds.), Intelligent Decision Support in Process Environments (pp. pp. 61–85). Springer-Verlag.

Goguen, J. (1994). Requirements engineering as the reconciliation of social and technical issues. In M. Jirotka & J. Goguen (Eds.), Requirements Engineering: Social and Technical Issues (pp. 165-200). Burlington, MA: Elsevier.

Halverfson, T., et al. (2012). Classifying Workload with Eye Movements in a Complex Task. *Proceedings Of The Human Factors And Ergonomics Society 56th Annual Meeting* (pp. 168-173). Santa Monica, CA: human Factors and Ergonomics Society.

Hirokawa, R.Y and Poole, M.S. (Eds.) (1996). Communication and Group Decision-making. City: Sage Publishing.

Hoffman, R. R. (2009). Some challenges for macrocognitive measurement. In E. Patterson and J. Miler (Eds.), Macrocognition metrics and scenarios: Design and evaluation for real-world teams (pp. 11-28). London: Ashgate.

Hoffman, R. R. and Fiore, S. M. (2007, May/June). Perceptual (Re)learning: A leverage point for Human-Centered Computing. *IEEE Intelligent Systems,* pp. 79-83.

Hoffman, R. R., Fiore, S. M., Klein, G., Feltovich, P. & Ziebell, D. (2009). Accelerated learning IEEE: Intelligent Systems, 18-22.

Hoffman, R. R., Lintern, G., & Eitelman, S. (2004). The Janus Principle. IEEE Intelligent Systems, 2(March/April), 78-80.

Hoffman, R. R. and Yates, J. F. (July/August, 2005). Decision(?)Making(?). IEEE: Intelligent Systems, pp. 22-29.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40* (10-12), 1489-1506. doi:10.1016/S0042-6989(99)00163-7

Jagacinski, R. J., & Flach, J. M. (2003). Control theory for humans: Quantitative approaches to modeling performance. Mahwah, NJ: Erlbaum.

Kaempf, G. L., Wolf, S., Thordsen, M. L., & Klein, G. (1992). Decision-making in the AEGIS combat information Center. Report under contract N66001-90- C-6023, Control and Ocean Surveillance Center, Naval Command, San Diego, CA.

Kaplan, A. (1964). Power in perspective. In R. L. Kahn & E. Boulding (Eds.), Power and conflict in organizations. London: Tavistock .

Klein, G. (1989). Recognition-primed decisions. In W. B. Rouse (Ed.), *Advances in Man-Machine Research*, Volume 5. Greenwich, CT: JAI Press.

Klein, G., Woods, D. D., Bradshaw, J. D., Hoffman, R. R., & Feltovich, P. J. (2004). Ten challenges for making automation a “team player” in joint human-agent activity. IEEE Intelligent Systems, 6 (November/December), 91-95.

Klein, G., Feltovich, P. J., Bradshaw, J. M., & Woods, D. D. (2006). Common ground and coordination in joint activity. In W. R. Rouse & K. B. Boff (Eds.), *Organizational simulation* (pp. 139–184). New York: Wiley.

Klein, G., Moon, B., & Hoffman, R. R. (2006, July–August). Making sense of sensemaking: 1. Alternative perspectives. *IEEE Intelligent Systems,* pp. 22–25.

Koopman, P., & Hoffman, R. R., (2003). Work-arounds, make-work, and kludges. IEEE Intelligent Systems, 6(November/December), 70-75.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). Foundations of measurement, Volume 1: Additive and polynomial representations. New York, NY: Dover.

Matheson, J.E. "Using influence diagrams to value information and control." Presented at the *Proceedings of the Conference on Influence Diagrams for Decision Analysis, Inference, and Prediction*, Berkeley, California, University of California at Berkeley, 9-11 May, 1988.

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110* (3), 422-471. doi:10.1037/0033-295X.110.3.422

Moore, D.T., 2011. *Sensemaking: a structure for an intelligence revolution.* Washington, DC: National Defense Intelligence College.

Moore, D.T., and Hoffman, R.R. (2011). "A practice of understanding." Chapter 5 In D. T. Moore (Au.), *Sensemaking: A structure for an intelligence revolution*. Washington DC: National Defense Intelligence College Press.

Neville, K., Hoffman, R. R., Linde, C., Elm, W. C., & Fowlkes, J. (2008). The procurement woes revisited. IEEE Intelligent Systems, 1(January/February), 72-75.

O'Neill, M. J. (2007). Measuring workplace performance. Boca Raton, FL: Taylor and Francis.

Patterson, E. S., Roth, E. M., & Woods, D. D. (2001). Predicting vulnerabilities in computer-supported inferential analysis under data overload. Cognition, Technology, and Work, 3(4), 224-237.

Picucci, P. M. & Numrich, S. K. (2010). Mission-driven needs: Understanding the military relevance of socio-cultural capabilities. In B. Johansson, B., S. Jain, J. Montoya-Torres, and E. Yucesan, edg., Proceedings of the 2010 Winter Simulation Conference.

Rota, S., Benincasa, G., Interlandi, M., Suri, N., Bonnlander, B., Bradshaw, J., Tortonesi, M., Watson, S., and Boner, K. (2010). Supporting Information on Demand with the DisServicePro Proactive Peer-to-peer Information Dissemination System. In Proceedings of the IEEE Military Communications Conference, pp. 984-991.

Savikhin, A., Maciejewski, R., & Ebert, D. S. (2008). Applied Visual Analytics for Economic Decision-Making. IEEE Symposium on Visual Analytics Science and Technology, (pp. 107-114).

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84* (1), 1-66. doi:10.1037/0033-295X.84.1.1

Shiffrin, R. M., & Schneider, W. (1977a). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84* (2), 127-190. doi:10.1037/0033-295X.84.2.127

Scholtz, J. (2005). Metrics for the evaluation of technology to support intelligence. In Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting (pp. 918-921). Santa Monica, CA: Human Factors and Ergonomics Society.

Stanton, N. A., Salmon, P. M., Walker, G. H., Baber, C., & Jenkins, D. P. (2005). Human factors methods: A practical guide for engineering and design. Burlington, VT: Ashgate Publishing.

Still, D. L., and L. A. Temme. (2001). "OZ: A human-centered cockpit display." Presented at the Proceedings of the Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL, November 26-29.

Still, David L., Tom Eskridge, and L.A. Temme. (2004). "Interface for non-pilot UAV control." Presented at the Human Factors of UAVs Workshop, Mesa, AZ.

Trafton, J. G., Kirschenbaum, S. S., Tsui, T. L., Miyamoto, R. T., Ballas, J. A., & Raymond, P. D. (2000). Turning pictures into numbers: Extracting and generating information from complex visualizations. *International Journal of Human Computer Studies, 53* (5), 827-850.

Weick, K. (1995). *Sensemaking in Organizations*, New York: Sage Publications.

Wilson, J. R., & Corlett, N. (Eds.) (2005). Evaluation of human work (3rd edn). Boca Raton, FL: Taylor and Francis.

Wolfe, J. M. (1998). What Can 1 Million Trials Tell Us About Visual Search? *Psychological Science*, *9* (1), 33-39. doi:10.1111/1467-9280.00006

Woods, D. D., & Hollnagel, E. (2006). *Joint cognitive systems: Patterns in cognitive systems engineering*. Boca Raton, FL: CRC Press.

Woods, David D., & Sarter, N. B. (2010). Capturing the dynamics of attention control from individual to distributed systems: the shape of models to come. *Theoretical Issues in Ergonomics Science*, *11* (1-2), 7-28. doi:10.1080/14639220903009896

Wulf G., McConnel N., Gärtner M., & Schwarz A. (2002) Enhancing the learning of sport skills through external-focus feedback. Journal of Motor Behavior. 34(2): 171-82.

Yates, J.F. (1990). Judgment and Decision-making, Prentice Hall.

Zachary, W., Neville, K, Fowlkes, J., & Hoffman, R. R. (2007). Human total cost of ownership: The Penny Foolish Principle at Work. IEEE Intelligent Systems, 2(March/April), 22-26.