Using Field Experimentation to Influence Tactical Command and Control System Design

Lawrence G. Shattuck
Department of Behavioral Sciences and Leadership
United States Military Academy
West Point, NY 10996
USA

Email: larry-shattuck@usma.army.mil

ABSTRACT

In spite of the efforts of members of the human factors profession, military command and control systems continue to be developed that are not user-centered. An important reason for this problem is the relationship between the human factors practitioners and the other participants in the system development process. The nature of this relationship is discussed and four suggestions to improve it are made. One of those suggestions is to consider the use of field experimentation as a means to develop user requirements and to test early system prototypes. An example of a field experiment is described in which 84 Army officers participated in a 2 x 2 between subjects design. Novice and experienced officers were given an operation order that depicted either a defensive or an offensive scenario. Their task was to solicit data from the researcher, construct and report their understanding of the tactical situations and describe the decisions they intended to make. Differences in performance were detected between novices and experienced officers and across scenario types. Each of these differences had design implications for the development of information displays and decision aids. The implications for the design of command and control systems are explained.

1.0 INTRODUCTION

The human factors profession is concerned with the discovery and application of information "about human behavior, abilities, limitations, and other characteristics to the design of tools, machines, systems, tasks, jobs, and environments for productive, safe, comfortable, and effective human use" [Sanders 1993]. Practitioners acquire the necessary education and training then dedicate themselves to improving the living and working conditions of humans. Although the profession is less than sixty years old, it has experienced significant growth. Advanced degrees in human factors and related fields are offered in universities all over the world. There are professional societies for practitioners to join, human factors conferences and symposia to attend, and hundreds of journals and books to read which are dedicated to improving living and working conditions.

In spite of this rapid growth, human factors practitioners are far outnumbered by the designers and engineers that actually build human-machine systems. Although there are human factors positions in many organizations, much of the system development happens without the involvement of these practitioners. Organizations often bring human factors practitioners into the design process when it is too late to make substantive contributions to system development. Their involvement is relegated to making very minor adjustments to systems that are nearly ready for production when, in fact, these systems may be seriously flawed (from a human factors perspective) at the theoretical or conceptual level. When this happens,

Paper presented at the RTO SCI Symposium on "Critical Design Issues for the Human-Machine Interface", held in Prague, Czech Republic, 19-21 May 2003, and published in RTO-MP-112.

it is almost certain that the systems, when fielded, will create unexpected and unnecessary challenges for the organizations that acquired them and the operators that use them.

To enhance its influence on system design and development the human factors profession has advocated the development and use of guidelines. These guidelines are developed by human factors practitioners and intended to be used by designers and engineers with little or no human factors expertise. On the surface, this might appear to be a viable solution. However, guidelines are often inadequate, context free, nd underspecified. There is no doubt that the researchers who developed the guidelines had the best of intentions and based their work on empirical studies. A close examination of these studies reveals that they are often conducted in sterile laboratory settings using college or university students with questionable motivation who perform tasks at which they are novices. The findings of these studies are generalized and then published for use by designers and engineers. Guidelines may specify a font size, weight limit, or physical dimension but they often do not include information about the contexts on which the recommendations are based. Designers and engineers invoke these guidelines and insist they have applied human factors principles to their systems. Generic guidelines often will do more harm than good.

Those human factors practitioners who are embedded in organizations that design and build systems may find themselves marginalized. Designers and engineers turn to human factors practitioners for suggestions or recommendations only to be informed that an answer can only be given after an extensive task analysis or comprehensive usability testing. These human factors practices certainly are important methods to define user requirements and analyze system prototypes. However, the production schedule imposed by management will not allow for such empirical investigation. The result is that the designers and engineers take their best guess and continue on with system development. By the time the human factors practitioners return with the answers it is too late. Development has progressed to the point that it would be cost prohibitive to go back and make changes, regardless of the benefit that would be derived from such changes.

The challenge that human factors practitioners face in having an impact on system design is particularly acute in software systems. A recent military wargame was heavily dependent on simulation and modeling. A few days before the wargame was to begin programmers were working feverishly to link various software models and complete the interface designs. Among the many problems they were trying to resolve was access to a particular class of weapons by commanders at various echelons. The solution that was supposed to be built into the model was one in which access to a particular system was granted by a central authority for a single mission or a discrete time period. At the end of that period, control was to be returned to the central authority. Further use of that weapon system would have to be renegotiated.

Researchers expected that such an allocation scheme would have an impact on coordination activities, workload, and resource management. However, the solution that was implemented was quite different. In their haste to fix this problem and without consulting with human factors practitioners they implemented an expedient solution in which access codes were issued to subordinate commanders by the central authority. These codes were not time dependent. Once a code was issued it remained in effect for the duration of the simulation event. The effect of this decision by the programmer was that all coordination activities between the subordinate commanders and the central authority after the initial request for access were unnecessary. The central authority lost control of the resources and there was almost no coordination within the C2 system with respect to the allocation of the weapons.

A well-intentioned programmer, under severe time constraints developed what he thought was a reasonable approach to resource allocation and tailored the human – machine interactions to support his scheme. His approach was inconsistent with the doctrine being tested and observers who were unaware of his modifications drew erroneous conclusions about both the doctrine and the need for collaboration. Unfortunately, while some of the programmer's work may be modified quickly to more accurately reflect doctrine, some artifacts such as the display screens may remain unchanged until much later when operators complain about how hard it is to use the system. At this stage only minor modifications are possible. Woods [Woods 1999] referred to this as sweeping up after the parade.

In his presidential address to the 1999 annual meeting of the Human Factors and Ergonomics Society, Woods [Woods 1999] identified several indications that the human factors profession is sweeping up after the parade. Human factors practitioners tend to:

- React after the fact. Human factors practitioners often react to the designs and prototypes developed by others.
- Bounce among hot buttons. Human factors practitioners are often directed to work on one hot issue after another without achieving theoretical grounding in any of these areas.
- Called in only when others reach impasses. Rather than being part of the development team, human factors practitioners are called on to fix problems; they leave once the problems are fixed, and are not viewed as being part of the team.
- Respond to call for help with "I can test that…" Human factors practitioners are prone to return to their laboratories to derive answers empirically when what the developers really need is a quick answer that is 'in the ballpark.'
- Miss windows of opportunity. Because they are not an integral part of the development team, human factors practitioners miss opportunities to contribute and to observe interesting cognitive and behavioral phenomena in the designing, building, testing, and implementation stages.
- Do their best work in the aftermath of surprises. While human factors practitioners are skilled at performing retrospective analyses, it would be more advantageous for all parties if they identify system inadequacies at the earliest opportunity in the development process.

In spite of these challenges and shortcomings, no other profession is better suited to serve as an advocate for the user than the human factors profession. During the 1980s the U.S. Army developed the MANPRINT (Manpower and Resource Integration) program to ensure the needs of the soldiers were considered in design, development, and acquisition of new systems. However, each year systems that are less than user-centered are designed, built, and implemented. Some of these systems even have had human factors practitioners involved in the process. How, then, can the human factors community have a greater impact on system development? The debate on this topic has raged on for as long as human factors has been a profession. In his keynote address to the 1998 annual meeting of the Human Factors and Ergonomics Society, Donald Norman offered several solutions. Among his ideas were that human factors practitioners needed to be better integrated into the science, engineering, design, and business communities [Norman 1998]. As a discipline, the human factors community would do well to heed Norman's advice. As someone who is both a military officer and a human factors practitioner involved in the study and design of command and control (C2) decision support systems, I offer four additional specific recommendations to my military and human factors colleagues and to those design and engineering organizations that develop C2 systems.

- Improve communication between the scientific community and the domain of interest.
- Integrate human factors practitioners as full partners in the design process from the very beginning.
- Leverage findings and developments from other domains by working at both the theoretical level as well as the applied level.

- Encourage field experimentation whenever possible to determine user and system requirements and to evaluate the system.

While the main topic of this paper is the use of field experimentation in designing command and control systems, it is appropriate to consider the other three recommendations first.

2.0 BETTER COMMUNICATION

A group of military and civilian analysts met recently to discuss their data collection plan for an exercise. One of the areas of interest was situation awareness. After a short time it became obvious that the definition used by the scientific community and the definition used by the military and civilian analysts were quite different. An often cited scientific definition of situation awareness (SA) is "the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and their status in the near future" [Endsley 1988]. The military and civilian analysts (and indeed many military personnel) refer to situation awareness and situation understanding (SU) as two different cognitive states. Figure 1 describes the relationship between Endsley's scientific definition and the military practitioners' definition.

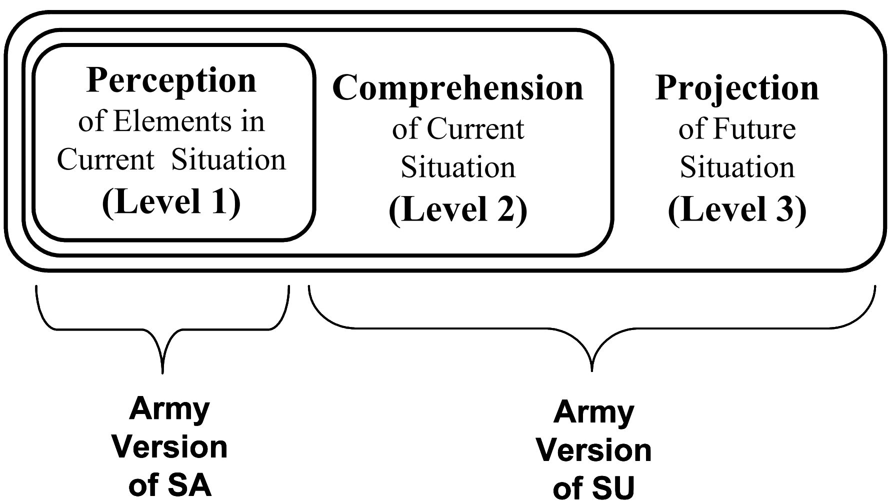

Figure 1: Comparison of Endsley's Scientific Definition of SA (Perception, Comprehension, Projection) to the Military Practitioner's Definition of SA and SU.

This difference is not trivial. Definitions that are written into military doctrine or operational concept documents become the basis for design, development, testing, and evaluation. Bad doctrine can lead to poorly designed systems. Doctrine writers may be unfamiliar with the scientific literature. But this is not necessarily their fault. Sound empirical work on the part of human factors practitioners often is published in relatively obscure journals or in technical reports with limited distribution. The human factors community must find better methods to communicate with practitioners in the various domains that could benefit from their research. Military practitioners, for example, should not be expected to struggle through scientific articles to extract a single nugget of useful information. Military practitioners must communicate their needs to the scientific community and the scientific community must listen and provide the information in a form that is clear and concise. Further, the context and limits of the scientific findings should be evident to the military practitioners.

3.0 LEVERAGING WORK FROM OTHER DOMAINS

There is a broad generalization that captures a difference between human factors practitioners who work in academic settings and those who work in industry. Those in an academic setting have the luxury of working at the theoretical level within a domain and are less constrained by deadlines. Those who work in industry must be more pragmatic and are often focused on solving a particular problem. There is tremendous benefit to be derived from combining these efforts.

The field of human factors has existed for nearly sixty years and includes thousands of practitioners who have worked in many different fields. On the surface, many of these fields may seem to be very different. After all, what do space, tactical military command and control, nuclear power plants, and air traffic control have in common? Actually, quite a bit. Each of them have been studied extensively by human factors practitioners. And each of them is a type of distributed supervisory control system. The problems and challenges that exist in one domain are likely to exist in one form or another in the other domains. The task for the human factors practitioner is to step back from the details of the problem and to see the emergent pattern which suggests a more general class of problems that exists in other domains. In doing so, the human factors practitioner is able to leverage the methodologies, design principles, and best practices that have proven to be successful in those other domains.

4.0 FULL PARTNERS

System development starts long before metal is bent or the first line of software code is written. It begins with theories, concepts and visions. All too often human factors practitioners are left out of these discussions. Committing to a theory, concept, or vision sets the design process in motion. The wrong theory, concept, or vision will result in a poorly designed system. For example, a vision for the future battlefield might be to introduce autonomous robotic agents in order to reduce the danger to which humans are exposed. With this vision in mind, designers set to the task of developing agents that can act independently. However, autonomous agents on the battlefield is an erroneous vision. It is no more practical to have autonomous robotic agents on the battlefield than it is to have autonomous human agents on the battlefield. A more appropriate vision might be a battlefield in which human and machine agents collaborate with one another similar to the way human agents collaborate in accomplishing the mission. Human agents communicate intent, build and repair each other's knowledge, coordinate activities, and engage in mutually supportive goal-directed behavior. The latter vision leads to an entirely different, and it can be argued, more productive approach.

If human factors practitioners are brought into discussions early they can assist in developing or shaping the theories, concepts, and visions. Just as important, they can employ empirical methods to examine the envisioned approach prior to the organization committing to that approach, when the cost of modifying the vision is minimal. Involving human factors practitioners as full partners throughout the development of a system can help ensure that each decision in the lifecycle is made on an empirically-based, user-centered perspective.

If human factors practitioners are given full partner status in the system development process they must make substantive contributions or they will not be invited back. Human factors practitioners must be both proactive and creative about how they test throughout the life cycle of the system. Early in the design process they must ensure that the theoretical approach is correct and they must focus on user requirements. Later, when the earliest prototypes and mock-ups are available they must solicit user feedback through laboratory and field experimentation. At this point changes to the prototypes can be made with minimal cost. As prototypes mature, user testing focuses on refinement, documentation, and training issues. As stated above, throughout the lifecycle, human factors practitioners must be able to communicate effectively to their system development partners. In addition to presenting their findings in a clear and concise manner, they must also demonstrate they are cognizant of issues related to management, business, design, engineering, and marketing.

Critical to effective testing and evaluation is getting access to domain experts. Human factors practitioners working on C2 system development often find it difficult to gain access to tactical commanders and staff officers, may retreat to the laboratory, and settle for college students or ROTC faculty. Whenever possible, military organizations should provide access to researchers and consider it an investment in the next generation of military leaders. It is more likely that military commanders will grant access if human factors practitioners can offer some incentive in return.

Field experimentation through the use of staged worlds or simulations can provide military commanders and staffs opportunities to sharpen their tactical skills, decision making processes, and knowledge of doctrine. A study by Shattuck [Shattuck 1995] investigated the process of by which tactical commanders formulate, communicate, interpret, and implement intent. The study required tactical commanders and staffs to develop operations orders and respond to unanticipated events on the battlefield. Senior commanders described how they expected their subordinate commanders to respond to the events, the subordinate commanders responded to those events, and the senior commanders then had an opportunity to observe and evaluate those responses. Senior commanders found the exercise to be invaluable in gaining insight into the thought processes of their subordinate commanders and in identifying sources of misunderstanding and miscommunication. The researcher collected valuable data and the tactical commanders sharpened their warfighting skills.

5.0 USE OF FIELD EXPERIMENTATION

Military command and control, like most distributed supervisory control systems, is difficult to study in its natural setting. The temporal and physical separation would necessitate a large team of researchers positioned at multiple locations to observe the activities of the remote supervisors and local agents as they implement plans and respond to feedback from the environment. One approach to overcoming the difficulties associated with direct observation of actual command and control systems is to create and study ‘scaled’ worlds that lend themselves to field research in more naturalistic settings.

Traditional laboratory research procedures such as randomization of participants, control, and manipulation of variables are not feasible in natural laboratories. Nevertheless, there are other legitimate methods for conducting meaningful investigation in field settings. If one considers research methods as procedures employed to shape the conditions of observation, then three classes of methods emerge (see Figure 2).

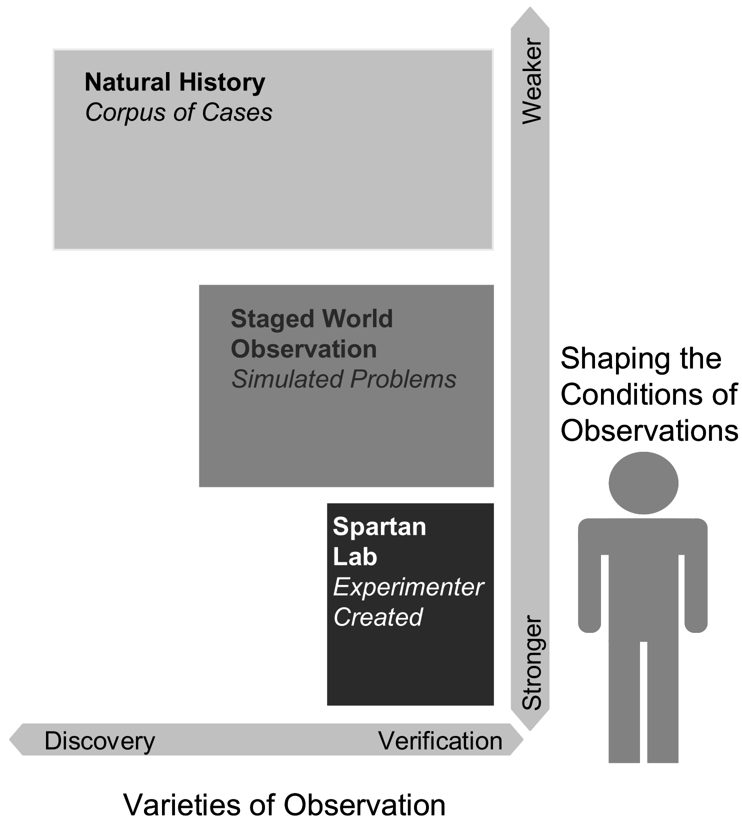

Figure 2: Three Classes of Studies Used by Researchers to Make Observations. Each class has its own strengths and limitations.

Natural history methods are based on a diverse collection of observations in situ. Experiments in the field begin with *scaled world simulations* that capture or 'stage' what is believed to be the critical, deeper aspects of the situations of interest [De Keyser 1998]. *Spartan laboratory techniques* focus on only a few variables of interest and their interactions in experimenter created situations. These three classes of methods have some common features but also are distinguishable in many ways.

All three classes of methods begin with observation. These methods pursue basic experimental values which are modified based on the uncertainties and purposes of each class [Salomon, 1991]. In addition, each method establishes warrant – standards of proof – for the observations that are made. Each method establishes standards of quality for observations, making distinctions between competent and incompetent investigation practices. Each method establishes a means to facilitate generalizability. Finally, each method must reconcile the target – test mapping issue. The researcher must ensure that the mapping between the test situation and the target situation is appropriate for the conclusions to be made from the observations.

As previously stated, natural history observations of command and control systems are problematic. Yet, a spartan laboratory experiment would create a target – test mapping in which the complexity of a command and control system would be trivialized. While both natural history methods and spartan laboratory experiments can make important contributions to our understanding of command and control systems, a field experiment or a scaled world simulation provides the most viable opportunity to learn about the dynamics of these systems in context while retaining the researcher's ability to focus on situations of interest.

Field experiments or scaled world simulations are occasionally characterized as weak or limited versions of 'real' experiments. More accurately, in this class of studies, experienced investigators effectively capture or 'stage' what is believed to be the critical aspects of the situation of interest. The mapping is explicit because it represents a kind of hypothesis about what is important in the target domain. Rather than employing traditional laboratory research design practices such as randomization and subject sampling, these studies focus on problem sampling. Investigators design scenarios that represent the challenges faced by domain practitioners. Further, simulated worlds afford repetitive observations. Ironically, however, repeated observations of a given scenario, do not result necessarily in the same outcomes. Simulation participants often will employ the robust set of problem solving strategies representative of actual practitioners in the target domain.

A typical practice employed by staged world investigators is to introduce a disturbance into the system. The purpose of introducing such disturbances is to afford investigators the opportunity to trace the process as the system moves towards a resolution or a new equilibrium. The investigator can infuse a new technology or may choose to stage an event to which the participants must respond. Events may be either *anomalous* or *unanticipated*. An anomalous event is one which causes the system to deviate from normal operation but from which the system should be able to recover or adapt successfully. System designers foresee anomalous events and develop procedures that can be employed by operators in response to such events. An unanticipated event is one which pushes the system to its limits (or beyond) because the designers did not expect it. Operators may not be equipped to respond to unanticipated events if system designers did not provide operators with sufficient flexibility.

Military command and control is somewhat unique among distributed supervisory control systems in that this domain includes an intelligent, uncooperative, unpredictable enemy that attempts to exploit weaknesses in the plans of the friendly forces. Detailed battle plans often do not remain viable much past the onset of hostilities. When confronted with unanticipated events, subordinate commanders must initiate an event-driven decision-making process. They construct their understanding based on their analysis of battlefield data. The remainder of this paper describes a field experiment in which military officers reasoned about battlefield situations and made decisions based on the data they selected and integrated. The experimental results provide valuable insights into how to design decision aids for commanders.

5.1 Field Experiment Investigating Cognitive Integration

Shattuck, et al., [Shattuck 2000] proposed that cognitive integration is an important part of situation awareness. The researchers defined cognitive integration as the process of extracting data from disparate sources and combining them in meaningful ways to create a veridical, holistic view of the environment. Although cognitive integration has not been formally investigated by other researchers, there are several members of the cognitive engineering community that recognize it as a construct vital to building situation awareness. Cognitive integration, by its very nature, does not lend itself to studies in sterile laboratory settings with naive participants. Military command and control (C2) environments exhibit many of the factors described by Orasanu and Connolly [Orasanu 1993] as characteristic of naturalistic settings.

- Ill-structured problems
- Uncertain, dynamic environments
- Shifting, ill-defined, or competing goals
- Action/feedback loops
- Time stress

- High stakes
- Multiple players
- Organizational goals and norms

To study cognitive integration effectively in naturalistic environments it is important to consider the domain's context. Woods, et al., [Woods 1994] states, "understanding cognition then depends as much on studying the context in which the cognition is embedded and the larger distributed system of artifacts and multiple agents, as on studying what goes on between the ears." Woods, et al. also recognized the role that integration plays in situation awareness. "Situation awareness is about the timely perception of critical elements of the situation, about information integration and management, and about anticipating future situations." Endsley's [Endsley 1988] definition of situation awareness was mentioned earlier in this paper and is similar to the definition put forth by Woods and his colleagues.

More recent work by Jones and Endsley [Jones 2000] includes a detailed discussion of the role of mental models and schema. In their study of representational errors they state that these errors "reflect problems with the assimilation of information into a person's current mental model." They suggest that mental models rely on long term memory stores referred to as schema. "Schema are organized representations of a body of knowledge constructed from past experience with objects, scenes, or events." [Jones 2000]. They acknowledge the importance of integration when they state, "currently observed information is frequently integrated with the schema prototype." They also include an error taxonomy for each level of situation awareness. They describe Level 2 error as "improper integration or comprehension of the situation."

While the preceding discussion is evidence that there is increasing interest in cognitive integration, much of this interest is at an abstract level. Little empirical research exists on how decision makers employ cognitive integration to help them construct understanding in evolving contexts. Even more important is how the results of the empirical research can contribute to the design of future displays for command and control. Figure 3 (below) served as a general hypothesis for the field experiment. Decision makers either seek or are provided with data. They combine these data into informational elements in order to understand what is happening on the battlefield and to make decisions.

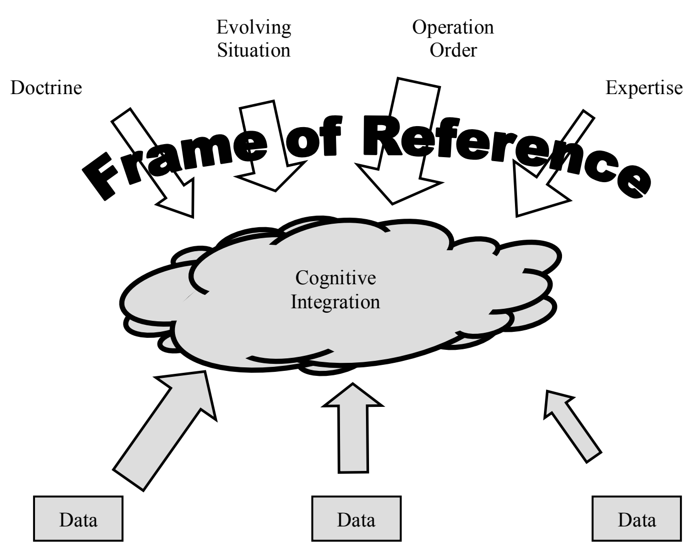

Figure 3: Model Used to Conduct Research on Cognitive Integration (Shattuck, et al., 2000).

The cognitive integration is not performed devoid of context. There are four elements that serve as a filter or frame of reference which influences the selection of data elements and the way in which those elements are integrated. The element with the most immediate (temporal) influence is the evolving situation. Decision makers are influenced by the events that are unfolding before them. These events are meaningful within the context of a particular plan they are attempting to execute or a specific goal they are trying to achieve. In tactical military operations, the operation order provides detailed information on the mission to be accomplished, including descriptions of who, what, when, where, how, and, in some cases, why. Operation orders are governed by doctrine. Military doctrine is broad guidance that influences the conduct of tactical operations. Doctrine, by its very nature, is usually context-free, having been written long before it is used by writers who are far removed from the battlefield. The fourth influence is the expertise of the decision maker. Each decision maker brings to a particular event a unique set of experiences.

These experiences, coupled with the evolving situation, the operation order, and the decision maker's knowledge of doctrine, form a frame of reference. The data that are processed – whether because they have been pushed to the decision maker or pulled from the staff and subordinate commanders - are filtered by this frame of reference. The different sizes of arrows represent the varying degrees to which the elements in the figure might contribute to the integration. For example, a novice decision maker may have little knowledge of friendly or enemy doctrine and may have fewer prior experiences to call upon. In addition, it is not likely that all data will be seen as having equal importance. Decision makers, based on their frame of reference, may seek out or weigh certain types of data over others.

5.1.1 Method

The cognitive integration field experiment was conducted as a 2 (novice vs. experienced officers) x 2 (defensive vs. offensive scenarios) between subjects design. Data were collected over a two year period from 1999 – 2001.

5.1.1.1 Participants

Experienced officers were students at the U.S. Army War College. All of these students were male Army officers in the grade of colonel or lieutenant colonel. Their military specialties were infantry, armor, or field artillery. All had at least 19 years of military service. The colonels and lieutenant colonels were considered to be experienced participants because all of them had successfully completed two years of battalion command. In total, they had been to a combat training center nearly 300 times (mean = 6.95, median = 4.0). Half of the officers participated in the defensive scenario while the other half participated in the offensive scenario. Forty-two officers at the United States Military Academy also participated in the study. These men were infantry, armor, or field artillery captains or majors who had been company commanders and had not yet attended the Command and General Staff College. The captains and majors were considered to be novice participants because none of them had battalion command experience. In total, they had been to a combat training center (CTC) nearly 400 times (mean = 8.90, median = 4.0). While the novices actually had spent more time at the CTCs, the experienced group had been to CTCs as battalion commanders 44 times (mean = 2.19, median = 1.0).

5.1.1.2 Equipment

National Training Center (NTC) scenarios were developed for defensive and offensive operations as drivers for the data collection. Equipment included maps of the National Training Center, acetate overlays, an operation order for each scenario, index cards containing data elements, and situation reports.

5.1.1.3 Procedure

The researcher explained the general procedures of the study and then gave the officer (novice or experienced) an operation order (defensive or offensive), a military map, and overlays. Once the officer read and understood the operation order (OPORD) he was given a situation report for Time Period 0. The situation report provided the officer with some general information about the battlefield at the commencement of the military operations. The researcher told the officer he could ask for up to 15 data elements (of a possible 138 for that time period) to help fill in gaps or clarify his knowledge of the scenario. The data elements were recorded on index cards. Data elements available to the participants consisted of friendly, environmental, and enemy information. There were 120 friendly cards available for each time period that contained data on units two levels below battalion (e.g., platoon level). There were three environmental cards that contained data on weather, visibility, and trafficability. In the defensive scenarios there were 15 enemy cards that provided data on the location, equipment and strength of five different enemy echelons. In the offensive scenarios, the enemy cards contained data on activities at specific locations on the battlefield. These locations were referred to as Named Areas of Interest (NAIs). After the officer received the cards, the researcher asked him to describe his understanding of the tactical situation. In addition, the researcher asked the participant what decisions he would make, if any, at that point in the scenario, given his current understanding of the battlefield situation. Upon completing his description, the researcher gave the officer another situation report that moved the scenario forward to Time Period 1. The researcher invited the officer to ask for up to 15 more cards from a second set of 138 cards. After the cards were drawn, the officer again described the tactical situation and again discussed the decisions he would make, if any, at that time. This process was repeated for Time Periods 2 and 3.

5.1.2 Results

This research effort yielded both quantitative and qualitative data, including more than 100 hours of videotaped verbal protocol data across all four conditions. These data provide opportunities to compare performance across different levels of expertise and scenario types. Results discussed here are those that are relevant to the design of displays in support of tactical command and control.

- Novices selected significantly more cards than experts in the defensive scenario, $F(1, 40) = 13.667$, $p<.05$. However, although novices also selected more cards in the offensive scenario, the statistical analysis failed to reveal significant findings, $F(1, 40) = 1.806$, $p>.05$.
- Novices and experts differed (but not significantly) in the types of cards they selected (friendly versus enemy).
- Novices and experts were consistent in their card selection based on the evolving context (i.e., across time periods).
- Novices and experts were consistent in their card selection across scenarios (offense versus defense).

Quantitative analysis of the data cards drawn revealed three significant trends.

- Participants favored certain categories of cards over others. For example, in the defensive scenario, when units were stationary, participants were more likely to draw cards containing data on equipment status rather than fuel status.
- The data cards drawn were a function of the evolving context. Cards drawn were more likely to provide information on areas of the battlefield at which fighting was ongoing or on units engaged in the fighting.

- There were differences in the number and type of cards drawn by experienced and novice participants. Trends indicated that novices drew more cards and those cards were more likely to provide them data on friendly units.

5.1.3 Discussion

Given present and future technological capabilities, there are several ways system developers might choose to design and build decision aids for soldiers engaged in command and control activities. Robust communications networks and the ubiquitous nature of data processing systems would suggest that all available data should be pushed to tactical decision makers. After all, more data ought to lead to better, faster, more informed decisions. The results of this field experiment clearly demonstrate that this simply is not true. A theoretical approach based on the principle of providing maximal data to users would lead to developing a system that would overwhelm and frustrate users. Performance would likely decrease. The findings of the field experiment do suggest several important principles that should be carefully considered by system developers. A few of these principles are discussed below.

Military doctrine encourages commanders to track battlefield activities two levels down. It is not unlikely that battalion commanders would have to monitor 24 platoon-size elements. Yet, the field experiment revealed that most participants wanted data only one level down – approximately seven company-size elements. Excessive detail was not helpful in constructing understanding of the battlefield. There will be times when decision makers will need to have access to detailed battlefield data to answer specific questions but it should not be the norm to have displays show such data. Instead, this field experiment suggests decision makers would be better served by less detail but with the capability to 'drill down' for more detail when needed.

The field experiment also made it apparent that participants were most interested in those parts of the battlefield where activity was the greatest. Figure 4 (below) illustrates this point clearly. The figure depicts the card selection of 21 experienced officers in the offensive scenario. During time periods H+0 and H+1 the battalion was traveling in a box formation with Team A and Company B in the lead, Team C following Team A, and Team D following Company B. Participants sought more data on Team A and Company B than on Teams C or D. At H+4, Company B hit unexpected enemy resistance. Card selection shifted dramatically as participants sought to determine what happened to Company B. By H+6, Company B and Team D had switched positions in the box formation so that the battalion was then maneuvering with Team A and Team D in the lead, Team C following Team A, and Company B following Team D. Card selection shifted accordingly so that participants were more likely to seek data on Teams A and D than on Company B or Team C. This finding also has some important implications for design of information displays and decision aids. Not all data is of equal importance to the decision maker. At certain times data from one unit may be much more important than data from other units. Displaying data from all units as if they were of equal weight might be an easier solution from a designer's or programmer's perspective but it could result in screens that are cluttered with data that are a distraction to the decision maker for a given context.

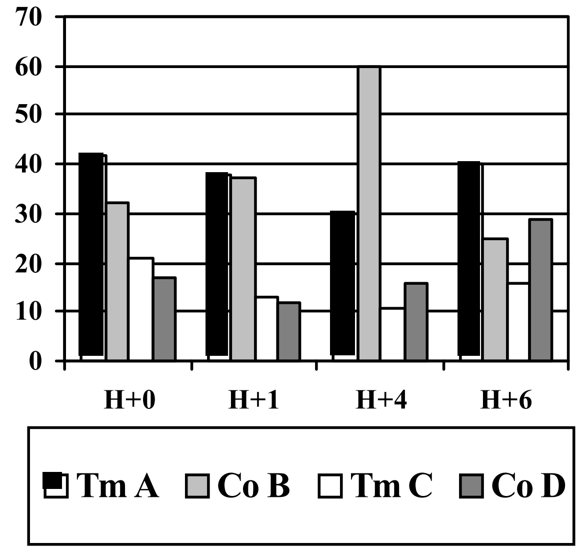

Figure 4: Total Number of Cards Drawn by Experienced Officers in the Offensive Scenario for Friendly Major Combat Units by Time Period.

Another important finding with respect to the design of decision aids and information displays was that the friendly data sought in the defensive scenario and the offensive scenario were different. The same five friendly data elements (unit location, personnel strength, equipment availability, fuel status, and ammunition status) were available in both scenarios. In the defensive scenario, the equipment cards were selected far more than any others. In post-experiment interviews, when some participants were asked why they were more likely to select equipment cards, they said they inferred if a tank were available, there were enough soldiers, fuel, and ammunition to operate it. This simplifying assumption was a way for them to reduce the amount of data they had to process. Participants were thinking in terms of weapon systems rather than in terms of the components that comprised the system. In the offensive scenario, participants tended to ask for cards that specified unit locations. Again, the design implications for these empirical findings are important. Various types of military operations place different demands on organizational resources. Command and control aids that are developed using a 'one size fits all' approach will not prove useful to decision makers. Instead, information displays and decision aids must be tailored to the type of operations in which a unit is engaged. At the very least, the displays and aids should be designed so that they can be tailored quickly and easily along the appropriate dimensions.

The field experiment also revealed some differences between novices and experienced participants. (Novice participants had no battalion command experience while all the experienced participants had successfully completed two years of battalion command.) The novice participants tended to select more total number of cards than experienced participants, suggesting perhaps that they were employing bottom-up processing techniques while the more experienced participants were engaging in pattern recognition. Another difference between the two groups was that novices tended to concentrate on understanding the status of friendly forces while the more experienced participants appeared to have a more balanced approach, selecting approximately the same number of friendly and enemy data cards. The implications of these findings for design of information displays and decision aids are less clear. Do system developers build displays and aids to facilitate decision making of experienced commanders or do they develop systems that critique and compensate for novice commanders? This is a vital but difficult question to answer. Part of the answer may be found in the following discussion of the nature of expertise in the military domain.

Although there were some differences between experienced and novice decision makers, the differences were not as stark as one might have thought. For this reason, the former battalion commanders in this paper have been referred to as *experienced* rather than *expert* commanders. Klein [Klein 1999] stated the following concerning expertise. "Because of their experience, experts have learned to see all kinds of things that are invisible to others. That is why they can move freely in their domains while novices must pick their way carefully through the same terrain. He then listed eight things that "experts can see that are invisible to everyone else."

- Patterns that novices do not notice.
- Anomalies – events that did not happen and other violations of expectancies.
- The big picture (situation awareness).
- The way things work.
- Opportunities and improvisations.
- Events that either already happened (the past) or that are going to happen (the future).
- Differences that are too small for novices to detect.
- Their own limitations.

The specific type of expertise examined by the field experiment described herein involved the decision making processes used by the military officers. Naturalistic Decision Making (NDM) is described by Zsambok [Zsambok 1997] as "the way people use their expertise to make decisions in field settings." She described the study of NDM as examining "how experienced people, working as individuals or groups in dynamic, uncertain, and often fast-paced environments, identify and assess their situation, make decisions, and take actions whose consequences are meaningful to them and to the larger organization in which they operate."

Drillings and Serfaty [Drillings 1997] wrote that the NDM theories, especially the Recognition-Primed Decision Making (RPD) developed by Klein, recognizes the valuable role that experience plays in the decision making process. They also acknowledge RPD's relevance to the military domain. RPD, as described by Klein [Klein 1993] argues that when experts are forced to make decisions rapidly, they generate a workable option and then determine whether or not that option will succeed by engaging in mental simulation.

Serfaty, MacMillan, Entin, and Entin [Serfaty 1997] also examined decision making in the military domain. They stated "the memory consists of an array of 'patterns,' with information items grouped and indexed by their relevance for problem solving in the domain of expertise." Further, they indicated "an expert commander has a mental model of the tactical situation that differs in measurable ways from that of the novice" and "the model allows the expert to detect what essential information is missing." However, rather than expert decision makers generating a single course of action initially, Serfaty, et. al, stated that experts use their experiences to "constrain" the number of plausible courses of action.

If we consider the nature of the military domain versus other domains in which expertise is traditionally studied, we gain insight into the basis for the less than expected differences that existed in this experiment between experienced and novice participants. Expert chess players play thousands of games of chess. Senior commercial pilots land their planes hundreds of times. Expert firefighters have been to countless fires. In each of these domains, the practitioners have done the same task (albeit under varying conditions) over and over again. In the military domain, however, even the very best battalion commanders only stay in that position for two years. And during those two years, they will have only one opportunity to take their unit to a combat training center such as NTC. Even though the biographical data collected from the participants indicated that some officers had been through more than 40 'rotations' at the NTC, they were performing tasks other than commanding a battalion in virtually all but one of those 'rotations.'

The results of this study seem to suggest that domain expertise is specific to particular tasks within that domain. Expertise is not developed solely by virtue of having been in the military for 20 years. Developing expertise requires practicing those tactical decision making skills at a given level of complexity (e.g., battalion, company, etc.). There are probably many other domains in which the distinction is not between expert and novice practitioners but between experienced and novice practitioners. In these domains, practitioners have been members for years, even decades, but have not had the opportunity to experience a large corpus of cases which, seemingly, is necessary for decision makers to apply naturalistic decision making methods. The challenge for human factors practitioners, then, is to consider how to design information displays and decision aids that improve the performance of both experienced commanders and staffs and those who are relatively new to their responsibilities.

6.0 CONCLUSIONS

Field experimentation is but one research method human factors practitioners can employ to gain insight into the true nature of user requirements in command and control systems. However, keen insights into user requirements are of no value if they are provided to system designers late in the development process. The field experiment described herein was not dependent on any particular hardware system. It was based on military domain doctrinal concepts and fundamental human factor principles. Such experimentation could be performed very easily in the earliest stages of system development and provide valuable information to designers, programmers, and engineers.

As previously stated, human factors practitioners must be full partners in the development process. They should be able to speak the language of designers, scientists, engineers, and business people [Norman 1998]. And rather than using hackneyed test and evaluation procedures that are conducted too late in the process to provide any useful information, they must employ a variety of creative procedures that anticipate the needs of the designers, programmers, and engineers. Such procedures will deliver the right information to system developers at the right time and result in information displays and decision aids that lead to superior performance in future command and control systems.

7.0 REFERENCES

[De Keyser 1998] De Keyser, V. & Samercay, R. (1998). Activity Theory, Situated Action and Simulators. *Le Travail Humain 61*(4), 305-12.

[Drillings 1997] Drillings, M., & Serfaty, D. (1997). Naturalistic Decision Making in Command and Control. In C.E. Zsambok & G.A. Klein (Eds.), *Naturalistic Decision Making* (pp. 71 – 80). Mahwah, NJ: Erlbaum.

[Endsley 1988] Endsley, M. (1988). Design and Evaluation for Situation Awareness Enhancement. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97-101). Santa Monica, CA: Human Factors and Ergonomics Society.

[Jones 2000] Jones, D. & Endsley, M. (2000). Overcoming Representational Errors in Complex Environments. *Human Factors*. *42*:3, 367-378.

[Klein 1993] Klein, G.A. (1993). A Recognition-Primed Decision (RPD) Model of Rapid Decision Making. In G.A. Klein, J. Orasanu, R. Calderwood, & C.E. Zsambok (Eds.), *Decision Making in Actrion: Models and Methods* (pp. 138 – 147). Norwood, NJ: Ablex.

[Klein 1999] Klein, Gary A. (1999). *Sources of Power*. Cambridge, MA: The MIT Press.

[Norman 1998] Norman, Donald A. (1998). The Post Disciplinary Revolution: Design and Human Factors – Heal Yourselves. An invited Keynote Address at the 1998 Annual Meeting of the Human Factors and Ergonomics Society. Retrieved April 14, 2003, from http://www.jnd.org/dn.mss/hfes-id-talk.htm.

[Orasanu 1993] Orasanu, J. & Connolly, T. (1993). The Reinvention of Decision Making. In G. Klein, J. Orasanu, R. Calderwood, & C. Zsambok (Eds.) *Decision Making in Action: Models and Methods*. Norwood, NJ: Ablex Publishing Corporation.

[Salomon, 1991] Salomon, G. (1991). Transcending the Qualitative/Quantitative Debate: The Analytic and Systemic Approaches to Educational Research. *Educational Researcher*. *20*(6) 10-18.

[Sanders 1993] Sanders, Mark S., & McCormick, Earnest J. (1993). *Human Factors in Engineering and Design*. New York: McGraw Hill, Inc.

[Serfaty 1997] Serfaty, D., MacMillan, J., Entin, E.E., & Entin, E.B. (1997). The Decision Making Expertise of Battle Commanders. In C.E. Zsambok & G.A. Klein (Eds.), *Naturalistic Decision Making* (pp. 233 – 246). Mahwah, NJ: Erlbaum.

[Shattuck 1995] Shattuck, L.G. (1995). The Communication of Intent in Distributed Supervisory Control Systems. Unpublished Doctoral Dissertation, The Ohio State University, Columbus.

[Shattuck 2000] Shattuck, L.G., Graham, J.M., Merlo, J.L., and Hah, S. (2000). Cognitive Integration: A Study of How Decision Makers Construct Understanding in Evolving Contexts. In *Proceedings of the Human Factors Society 44th Annual Meeting* (pp. 116-119). Santa Monica, CA: Human Factors and Ergonomics Society.

[Woods 1999] Woods, D.D. (1999). Presidential Address to the Annual Meeting of the Human Factors and Ergonomics Society. Retrieved on April 14, 2003, from http://csel.eng.ohio-state.edu/hf99/.

[Woods 1994] Woods, D.D., Johannesen, L., Cook, R.I., and Sarter, N. (1994). *Behind Human Error: Cognitive Systems, Computers and Hindsight*. (CSERIAC State of the Art Report 94-01). Dayton, OH: Crew Systems Ergonomic Information and Analysis Center.

[Zsambok 1997] Zsambok, C.E. (1997). Naturalistic Decision Making: Where Are We Now? In C.E. Zsambok & G.A. Klein (Eds.), *Naturalistic Decision Making* (pp. 3 – 16). Mahwah, NJ: Erlbaum.