Implications of Decision Making Research for Decision Support and Displays

Jeffrey G. Morrison	Richard T. Kelly & Ronald A. Moore	Susan G. Hutchins
NCCOSC RDT&E Division (NRaD)	Pacific Science & Engineering Group, Inc.	Naval Postgraduate School
San Diego, CA	San Diego, CA	Monterey, CA

Abstract

A prototype decision support system (DSS) was developed to enhance Navy tactical decision making based on naturalistic decision processes. Displays were developed to support critical decision making tasks through recognition-primed and explanation-based reasoning processes, and cognitive analysis was conducted of the decision making problems faced by Navy tactical officers in a shipboard Combat Information Center. Baseline testing in simulations of high intensity, peace keeping, littoral missions indicated that experienced decision makers were not well served by current systems, and their performance revealed periodic loss of situation awareness. A study is described with eight expert Navy tactical decision making teams who used either their current system alone or in conjunction with the prototype DSS. When the teams had the prototype DSS available, we observed significantly fewer communications to clarify the tactical situation, significantly more critical contacts identified early in the scenario, and a significantly greater number of defensive actions taken against imminent threats. These findings suggest that the prototype DSS enhanced the commanders' awareness of the tactical situation, which in turn contributed to greater confidence, lower workload, and more effective performance. Significant work remains to be done in learning how to optimally design and train users of such systems.

INTRODUCTION

It is our premise that for human decision makers to be effective, they must foremost be able to access the data necessary to make a decision when it is needed, where it is needed, and in the form it is needed. The data must be integrated and organized so that they become useful as information to the user. Information must be meaningful, timely, and organized in a way that is consistent with how it is going to be used. The effects the operational environment may have on the human operators of a system complicate understanding how an operator uses information. The U.S. Navy has recognized the criticality of such factors in assuring mission effectiveness and minimizing incidents of blue-on-blue (friendly) or blue-on-white (neutral) engagements. As noted in Chapter 1 of this book, the *Tactical Decision Making Under Stress* (TADMUS) program was initiated in response to one such incident: the accidental shoot-down of an Iranian Airbus aircraft by the USS Vincennes in 1988. The congressional investigation of this incident suggested that emotional stress may have played a role in contributing to this incident, and the TADMUS program was established to assess how stress might affect decision making and what might be done to minimize those effects. In any human-machine system, there are three possible approaches to addressing the human-machine system integration problem. You may select particular people based on their skills and abilities. You may train people to enhance their skills and provide them knowledge to make optimal use of the system and their own capabilities and limitations. Finally, you may design the machine component of the system to better

accommodate and support the people using that system. A fundamental premise of the TADMUS program was that the Navy had evolved a fundamentally effective system for selecting its commanders and tactical decision makers in the Aegis Combat system. It was, however, plausible that we could make some significant contributions in improving the utility of these complex person-machine systems through the application of emerging theories of cognition and models of decision making to the areas of training and system design. Other chapters in this book address many of the training tools and techniques developed as part of the TADMUS program to facilitate effective decision making in real-time tactical systems. This chapter will address the on-going work in developing improved human-system integration (HSI) for tactical decision makers.

Cognitive Analysis of Decision Tasks

Recent theories of decision making emphasize the importance of situation assessment for good decision making in naturalistic, event-driven situations. Moreover, they stress that decisions regarding actions to be taken are a by-product of developing the situation awareness that precedes action selection. Early TADMUS work focused on a family of cognitive theories which have come to be known as "naturalistic" decision making (Klein, 1989; Klein, Orasanu, Calderwood, & Zsambok, 1993). Naturalistic decision making differs from that found in the artificial intelligence / expert system literature in that these models are typically focused on emulating the outcomes of expert decision making by emulating the process a human decision maker might use in reaching the outcome. An example of these models is the analytic approach taken by many computers in playing chess. These programs typically consider all possible moves and counter-moves in a computationally intensive manner, and then make a move based on the best solution from this exhaustive analysis. Human decision makers do not use this exhaustive, analytic approach, particularly not expert decision makers. The human expert will typically look at a situation, and use some general heuristic derived from his/her previous experience to choose an action. With regard to tactical decision making in a single ship, peace-keeping mission, Klein's work found that usually the situation itself either determines or constrains the response options and that experienced decision makers make up to 90% of all decisions without considering alternatives. If the situation appears similar to one that the decision maker has previously experienced, the pattern will be recognized and the course of action is usually immediately obvious. This has come to be known as "recognition-primed decision making" (RPD; Klein, 1989, 1993). On the other hand, if the situation does not seem familiar, a more complex form of decision making will be involved where the decision maker considers general classes of explanations, selects from those that seem plausible to create a working hypothesis, and then rapidly adjusts this hypothesis after evaluating it. This less common form of decision making is referred to as "explanation-based reasoning" (EBR). In effect, this is decision making by telling a story to explain the discrepancies between expectations and what actually happens. As with RPD, the reasoning is not exhaustive but fairly short and concise, and the expert decision maker rapidly develops a reasonable hypothesis to explain the situation.

Additional support for these findings was found in an analysis of how experienced commanders make tactical decisions in realistic situations. Research was conducted to determine the decision requirements for command-level decision makers in the combat information center

(CIC) of an Aegis cruiser. Analysis of 14 incidents from actual problems revealed 183 decisions. Of these, 103 concerned situation assessments. Decision makers arrived at 87% of their situation assessments through feature matching and the remaining 13% through story generation (Kaempf, Wolf, & Miller, 1993). Feature matching strategies involve comparisons of observed data to sets of distinctive features or cues held in memory and based on the decision maker's training and experiences. When a match occurs, as a result of conscious or more automatic processes, the meaning of the observed situation becomes evident along with the appropriate responses. Story generation, on the other hand, involves more active EBR processes in which decision makers attempt to build a coherent story that accommodates the observed data, thereby providing a plausible interpretation of the situation. Story generation typically occurs when the patterns of observed data are unfamiliar or insufficient to allow a match to features of known situations held in memory.

The other eighty decisions that were identified from analysis of the real-world incidents mentioned above, involved course of action selection. Selecting courses of action involves determining what actions need to be undertaken in order to deal effectively with a particular situation. Kaempf, et al. (1993) distinguished between strategies for course of action selection where decision makers recognized the appropriate courses of action for the situation, selected a course of action from among multiple options, or generated a single, custom course of action that fits the details of the situation. These course of action decisions served a variety of functions, although relatively few were intended to end the incident. Twenty were intended as a final course of action decision, 14 were implemented to obtain more information, 22 to manage resources, and 24 to put themselves in a more favorable tactical position. A recognition-based strategy was used by decision makers to develop a final course of action most of the time. This strategy accounted for 95% of the actions taken in the 14 simulated incidents. The decision makers generated and compared multiple options in only 5% of the cases. In line with these findings, the TADMUS program has adopted the position that decision support systems should assist in the decision making *process* and focus on aiding the situation assessment portion of the decision making task.

Baseline tests in representative littoral scenarios further corroborated these analyses of decision making in tactical decision making among expert decision makers (Hutchins & Kowalski, 1993; Hutchins, Morrison, & Kelly, 1996). An analysis of communications patterns indicated a predominance of feature matching strategies in assessing the situation, typically followed by the selection among preplanned response sets (tactics) that were considered to fit the situation. These tests also suggested that experienced decision makers were not particularly well served by current systems in demanding missions. Teams exhibited periodic losses of situation awareness, often linked with limitations in human memory and shared attention capacity. Environmental stressors such as time compression and highly ambiguous information increased decision biases, e.g. confirmation bias, hypervigilance, and task fixation, and were correlated with tactical errors in executing the missions. Hutchins cited several specific problems associated with short term memory limitations, including:

(a) Mixing up track (contact) numbers (track being recalled as 7003 vs. 7033) and forgetting track numbers;

(b) Mixing up track kinematic data (track recalled as descending vs. ascending in altitude, closing vs. opening in range, etc.) and forgetting track kinematic data; and

(c) Associating past track related events/actions with the wrong track and associating completed own-ship actions with the wrong track.

Problems noted by Hutchins that related to decision biases included:

(a) Carrying initial threat assessment throughout the scenario regardless of new information (framing error); and

(b) Assessing a track based on information other than that associated with the track, e.g., old intelligence data, assessments of similar tracks, outcomes of unrelated events, past decision maker experiences, etc. (e.g. confirmation bias).

DECISION SUPPORT SYSTEM DESIGN

In order to address these problems, development of a prototype decision support system (DSS) became the focus of the TADMUS HSI development effort. The objective of this effort was to evaluate and demonstrate display concepts derived from current cognitive theory with expert decision makers in an appropriate test environment. The focus of the DSS was on enhancing the performance of tactical decision makers (viz., the Commanding Officer (CO) and Tactical Action Officer (TAO) working as a team) for single ship, air defense missions in high density, ambiguous littoral warfare situations. The approach taken in designing the DSS was to analyze the cognitive tasks performed by the decision makers in a shipboard CIC, and then to develop a set of display modules to support these tasks based on the underlying decision making processes naturally used by the CO-TAO team.

Given that the CO and TAO decision makers were behaving in a manner consistent with those predicted by "naturalistic" decision making theory (Klein, 1993), this theory became central to the design of a human-computer interface to improve tactical decision making. A prototype DSS was developed with the objectives of: (1) minimizing the mismatches between cognitive processes and the data available in the CIC to facilitate decision making; (2) mitigating the short comings of current CIC displays in imposing high information processing demands and exceeding the limitations of human memory; and (3) transferring the data in the current CIC from numeric to graphical representations wherever appropriate. It was determined that the DSS should not filter or extensively process data; i.e., it should support rather than aid (automate) decision making and leave as much decision making with the human decision makers as possible. The design goal of the DSS was to take the data that are already available in the system and present them as meaningful information where, when, and in the form needed relative to the decision making tasks being performed.

DSS-1 Design

Version 1 of the DSS (DSS-1) was designed expressly for the evaluation of display elements to support feature matching, story generation (viz., RPD and EBR) with the goal of reducing errors, reducing workload, and improving adherence to rules of engagement. The

design was significantly influenced by inputs from subject matter experts to ensure its validity and usefulness for the operational community. It was implemented on a Macintosh computer to operate independent of, synchronized with, or linked to a scenario driver simulation. The design of DSS-1 was constrained by practical and research requirements. It was never intended to represent a display to be used aboard ship. The display was designed to complement an existing Aegis geo-plot display so that performance could be evaluated with the Aegis geo-plot as a baseline condition for comparison purposes.

Figure 1 shows the DSS-1 prototype display. The DSS is a composite of a number of distinct display modules. Modules are arranged in a tiled format so that no significant data are obscured by overlapping windows. The DSS was conceived as a supplementary display to complement the existing geo-plot and text displays in current CICs. DSS modules have been discussed and demonstrated in detail elsewhere (cf. Moore, Quinn, & Morrison, 1996). Nevertheless, three of the modules will be discussed in more detail as an illustration of how the information requirements of tactical decision making tasks were mapped with cognitive processes described in naturalistic decision making theory to generate the DSS.

Track Profile

The track profile module consists of two graphical displays in the upper portion of the DSS that show the current position of a selected track in both horizontal and plan-form displays. Information requirements addressed by this module included the need to: (1) see where the track is relative to own-ship, (2) see what the track has been doing over time, (3) recognize whether the track can shoot you, and (4) recognize whether you could shoot the track. An important aspect of this display is that it shows a historical plot of what the track has done in space and time (the history is redrawn each time the track is selected). This greatly reduces the short term memory requirements on the CO and TAO in interpreting the significance of the selected track. This historical dimension of the display allows the decision maker to see what the track has done and primes his recognition of a likely mission for that track which could account for its actions. In addition, the profiles show own-ship weapon and track threat envelopes displayed in terms of range and altitude so that the decision maker can visualize and compare mental models (templates) as he considers possible track intentions and own ship options.
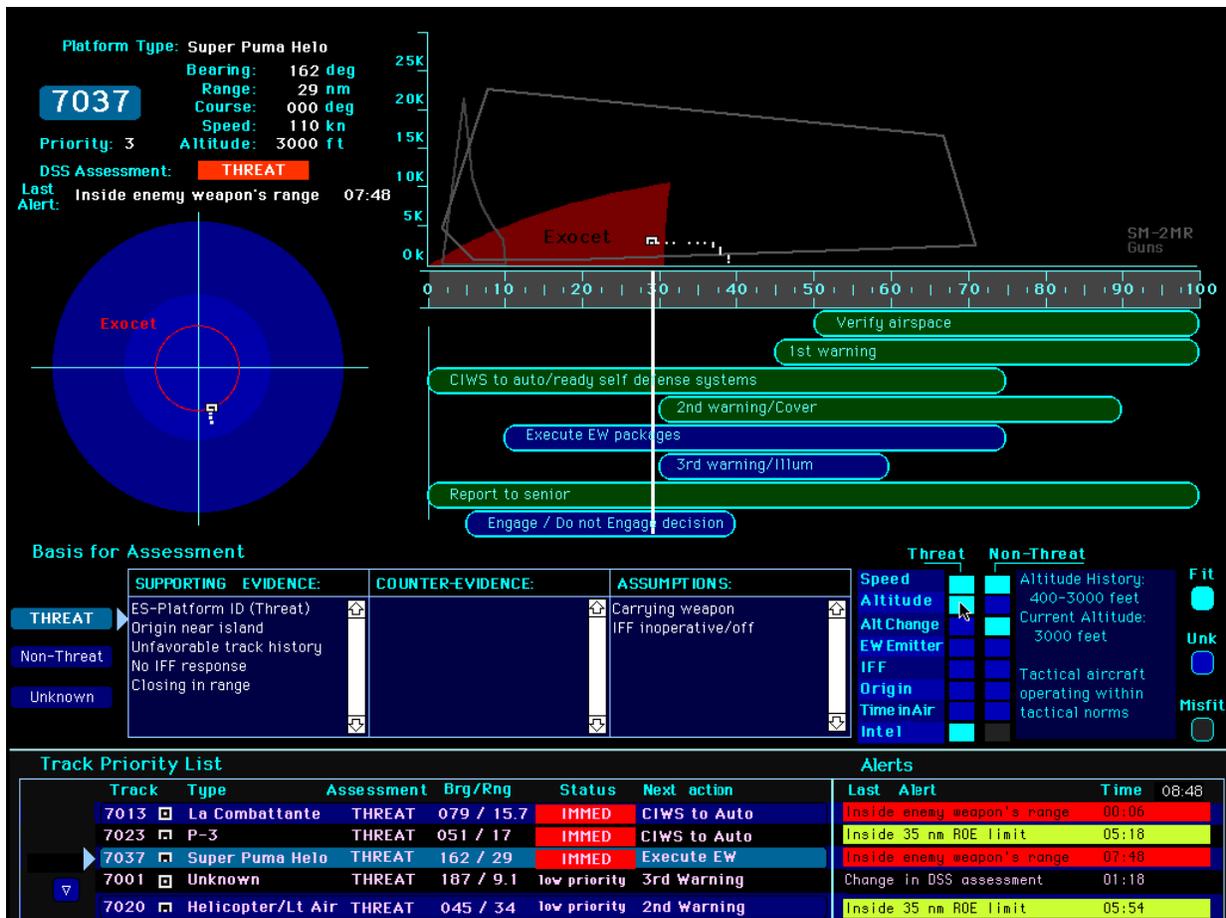
Platform Type: Super Puma Helo
Bearing: 162 deg
Range: 29 nm
7037
Course: 000 deg
Speed: 110 kn
Priority: 3   Altitude: 3000 ft
DSS Assessment: THREAT
Last Alert: Inside enemy weapon's range 07:48

25K 20K 15K 10K 5K 0k

Exocet

SM-2MR Guns

0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100

Exocet

Verify airspace
1st warning
CIWS to auto/ready self defense systems
2nd warning/Cover
Execute EW packages
3rd warning/Illum
Report to senior
Engage / Do not Engage decision

Basis for Assessment

Threat   Non-Threat   Fit

SUPPORTING EVIDENCE:	COUNTER-EVIDENCE:	ASSUMPTIONS:
ES-Platform ID (Threat)		Carrying weapon
Origin near island		IFF inoperative/off
Unfavorable track history		
No IFF response		
Closing in range		

THREAT
Non-Threat
Unknown

Speed
Altitude
Alt Change
EW Emitter
IFF
Origin
Time in Air
Intel

Altitude History: 400-3000 feet
Current Altitude: 3000 feet
Tactical aircraft operating within tactical norms

Unk
Misfit

Track Priority List   **Alerts**

Track	Type	Assessment	Brg/Rng	Status	Next action	Last Alert	Time	08:48
7013	La Combattante	THREAT	079 / 15.7	IMMED	CIWS to Auto	Inside enemy weapon's range	00:06	
7023	P-3	THREAT	051 / 17	IMMED	CIWS to Auto	Inside 35 nm ROE limit	05:18	
7037	Super Puma Helo	THREAT	162 / 29	IMMED	Execute EW	Inside enemy weapon's range	07:48	
7001	Unknown	THREAT	187 / 9.1	low priority	3rd Warning	Change in DSS assessment	01:18	
7020	Helicopter/Lt Air	THREAT	045 / 34	low priority	2nd Warning	Inside 35 nm ROE limit	05:54	

Figure 1. DSS Version 1

Response Manager

The response manager is located immediately below the track profile and is tied to it via a line indicating the track's current distance from own ship. It represents a Gantt chart type display showing a template of pre-planned actions and the optimal windows in which to perform them. The display serves as a graphical embodiment of battle orders and doctrine, and it shows which actions have been taken with regard to the selected track. The display is intended to support RPD and serves the need to: (1) recall the relevant tactics and strategies for the type of track being assessed, (2) recognize which actions need to be taken with the track and when they should be taken, and (3) remember which actions have been taken and have yet to be taken for the selected track.

Basis for Assessment

This module is located in the lower left area of the DSS and is intended to support EBR (story generation). The basis for assessment module presents the underlying data used to generate the DSS's threat assessment for the displayed track. The display shows three categories of assessment on which decision makers focus: potential threat, non-threat, or unknown. The

decision maker selects the hypothesis he/she wishes to explore, and data are presented in a tabular format within three categories: supporting evidence, counter evidence, and assumptions. These categories were found to be at the core of all story generation in which commanders engage while deciding whether a track with the potential to be a threat is, in fact, a real threat. This EBR related to threat assessment is also typically one of the decision making tasks performed when deciding whether to fire on a track or not. The display was designed to present the relevant data necessary for a commander to consider and evaluate all likely explanations for what a target may be, and what it may be doing (i.e., "intents") through the generation of alternative stories to explain the available and missing data regarding the track in question. The display is also intended to highlight data discrepant with a given hypothesis to minimize confirmation and framing biases. Assumptions listed are those assumptions necessary to "buy into" the selected assessment. As a result, the basis for assessment module was expected to be particularly effective in helping sort out and avoid "Blue-on-Blue" (i.e., mistakenly shooting friendly contacts) and "Blue-on-White" (i.e., mistakenly shooting neutral contacts) engagements.

Other DSS-1 Modules

Track Summary. The track summary module is located in the upper left of DSS-1. It summarizes current data related to a selected track. It is intended to support "quick look" RPD processes via data related to track kinematics and overall status (threat, neutral or friendly) of the target. It was designed in response to the mental conversion and readability deficiencies of character read-out (CRO) displays in the current CIC displays. This window also supports a drop down track list to rapidly select the individual track to be displayed, listed in priority order.

Comparison to Norms. The comparison to norms module is located adjacent to the basis for assessment module. It determines how well the observed data for a track fit the established threat / non-threat norms for that type of track. For instance, if the historical data seen for a track classified as hostile are consistent with established normal operating range (tactical norms) for a hostile track, then there is a good fit of the data to the hypothesis with regard to the parameter. If the data are outside tactical norms for what the track would be expected to do given its classification, but within its operational capabilities, then it is a questionable fit. If the data are clearly inconsistent with that expected for that type of track with the current classification there is a poor fit. The relative fit is shown by one of three colors in columns of color-coded "chips". Theoretically, the appearance of any misfit colors would signal the decision maker that there is ambiguous or conflicting data about the track, and cause him to consider alternative explanations / classifications for the track. In addition, this module supports EBR by allowing tactical decision makers to access detailed historical information summarizing the data categories for this track.

Track Priority List & Alerts List. This module is located at the bottom of the DSS display and is intended to support the decision maker in his attempt to maintain general situation awareness, i.e., the "big picture." It is intended to help the decision maker manage his own cognitive resources / attention. The track priority list presents the four highest priority tracks as well as a fifth track of special interest to the decision maker. The alerts list is linked to the track

priority list. It normally displays the last alert issued for each displayed track. Clicking and holding the alert pulls up a list of all alerts in the system that have been issued for that track.

DSS-2 Design

Figure 2 shows the current DSS-2 prototype display. The DSS-2 is a composite of several display modules, which have been adapted from DSS-1 based on user comments and performance tests. The DSS-2 is implemented with two 1024x768 CRT touch screen displays. The display features an integrated geo-plot and a variety of modules designed to solve specific decision making problems encountered by the tactical decision maker. These modules will be described below as an illustration of how the information requirements of tactical decision making tasks were integrated with the cognitive processes described by naturalistic decision making theory.
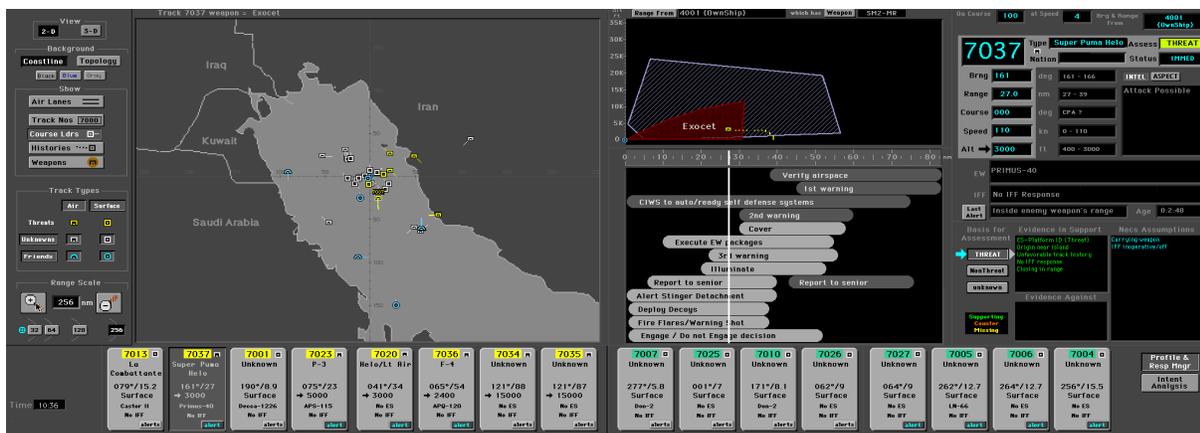
Figure 2. TADMUS DSS Integrated Display

Geo-Plot

The geo-plot occupies the left side of the DSS-2 display, shown in Figure 2. A close-up of the geo-plot is provided in Figure 3. The display uses variable-coded Navy Tactical Data System (NTDS) symbology (Nugent, 1996; Osga & Keating, 1994; Rausche, 1995) to represent the position of air, surface, and sub-surface tracks over a geographic region. The module is intended to be the primary focus of decision makers and is designed for quick decision making associated with situation awareness and RPD processes. It does, however, also contain other geographic detail that could support EBR processes. Symbols are color and shape coded to indicate track identification and threat evaluation, and may be toggled between NTDS symbols and track numbers. Track numbers are the "language" of the CIC and allow rapid location of various tracks. Conceptually, the map consists of several layers, which may be altered to suit particular mission requirements. Shown is a desaturated map (Jacobsen, 1986; Van Orden & Benoit, 1994) which provides sufficient spatial reference for most tasks while minimizing screen clutter and excessive color for most decision making problems. For strike or search and rescue missions, a topographical map may be overlaid to provide a more relevant context in which to frame tactical decisions. Likewise, overlays for infrastructure (highways, population densities,

power grids, etc.) may be added when necessary to further enhance the decision making context. The use of a 2-dimensional representation ensures that tracks can be located quickly with precision. An optional 3-dimensional display could be shown when appropriate for assessing general spatial relations. This display has not as yet been fully implemented, and no data have been collected to validate its potential utility. Controls for altering the geo-plot are arranged along the left edge of the module. In addition to selecting map layers, scaling, and panning controls, the decision maker may supplement the display with velocity leaders to show the relative speed of all the tracks, with course histories which show a track's path over time relative to landmarks, with air corridors, and with other tracks. Weapon threat envelopes may be displayed for potential threat tracks along with own ship to rapidly assess the criticality of a threat.

Note that DSS-1 operated in conjunction with a geo-plot, but that display was independent of the DSS and had much less functionality, particularly with respect to overlays and alternate map types. Based on extensive comments and decision performance data from operational users, the DSS-2 was designed to incorporate the geo-plot as an integral part of the command display. This not only simplifies the human-computer interface dialog by reducing the control actions required but also promotes better situation awareness by making it easier for decision makers to access and use all available tactical data displayed on the DSS.
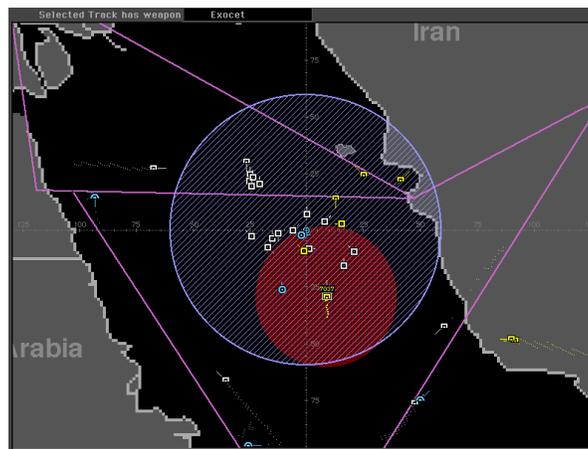
Figure 3. Geo-Plot with Desaturated Map and Variable Coded Symbology

Multi-CRO Access Panel

Across the bottom of the DSS-2 display is a series of buttons for quick access to the highest priority tracks. A close-up view of one of these buttons is shown in Figure 4. The buttons serve as miniature character read outs (Mini-CROs), displaying critical identification and kinematic information about the track and allowing the status of the most critical tracks to be monitored without additional interaction with the system. The buttons are arranged by a fairly simple algorithm in terms of their threat priority, with highest priority on the left and lower priority on the right. As situations evolve, the movement of the buttons quickly draws attention to the changes, which helps preclude attention fixation to a single track or task when the decision maker is under stress. Thus, the collection of Mini-CROs in this panel is intended to support

RPD processes. The relative position of each Mini-CRO, the color coding of threat level, and the summary of critical track information all support rapid feature matching by experienced tactical decision makers.
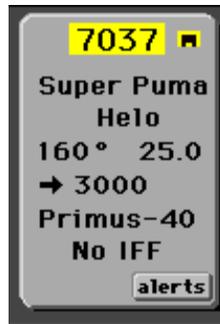

Figure 4. Sample Mini-CRO

In the lower right of each track button is an alert button. This alert button is cyan-colored when a new alert has occurred and gray when there are no new alerts. Pressing and holding the alert button generates a pop-up window to display a chronological list of alerts that have occurred for the selected track. Track age is shown rather than the time at which they occurred (as is the case with current systems) because decision makers are more interested in how old the alert is rather than when it occurred. This feature, which may be activated by users on demand, contributes to their periodic need to engage in EBR processes. This implementation is an adaptation of the Alerts List in DSS-1, based on feedback from operational users.

Track Profile

The track profile, which is shown in Figure 5, is substantially the same as in DSS-1. The track profile module complements the geo-plot by showing a horizontal display of track altitude and range from own ship. Information requirements addressed by this module include the need to: (1) see where the track is now, (2) what the track has been doing over time, (3) recognize whether the track can shoot you, and (4) recognize whether you could shoot the track - all at a glance. The track profile also shows own-ship weapon and track threat envelopes displayed in terms of range and altitude so that the decision maker can visualize and compare mental models (templates) as he considers possible track intentions and own ship options. Thus, it supports RPD processes by experienced tactical decision makers.

To further facilitate these evaluations, the DSS-2 track profile incorporates two pull-down lists. The "perspective picker" allows the decision maker to jump to other friendly (including air, surface, or land) forces in the area so as to assess the possibility that other assets are the target of interest to a potentially hostile track, and to assess whether those friendly units could assist in engaging a prospective threat. When an alternative perspective is chosen, all modules in the display reflect the perspective and capabilities of the chosen asset. The second pull-down list is a "weapons picker" that further elaborates on capabilities and limitations of own ship and other friendly units showing how different weapons could be used against the selected track. In the event that a weapon system goes off-line, such information would automatically be reflected in the weapons displayed in the pull-down list. These capabilities are expected to play a significant role in joint (multi-service) and coalition (multi-national) operations.
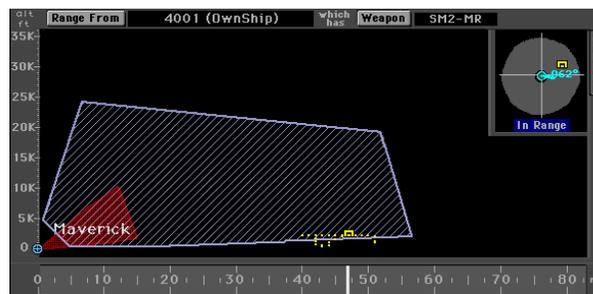

Figure 5. Track Profile with Aspect Inset

Embedded within the track profile is an inset window which shows own ship heading relative to the selected track. The display quickly shows radar cross section and weapons cut outs for assessing whether own ship should be maneuvered to optimize these parameters.

Response Manager

The response manager is located immediately below the track profile and is linked to it via a line indicating the contact's current distance from own ship. As shown in Figure 6, it provides a Gantt chart type display, similar to that used in DSS-1, showing a set of pre-planned actions and the optimal range windows in which to perform them. The display serves as a graphical embodiment of battle orders and doctrine, and shows which actions have been taken

with regard to the selected track. The display is intended to support RPD and serves the need to: (1) recall the relevant tactics and strategies for the type of target being assessed, (2) recognize which actions need to be taken with the target and when they should be taken, and (3) remember which actions have been taken and have yet to be taken for the selected target.
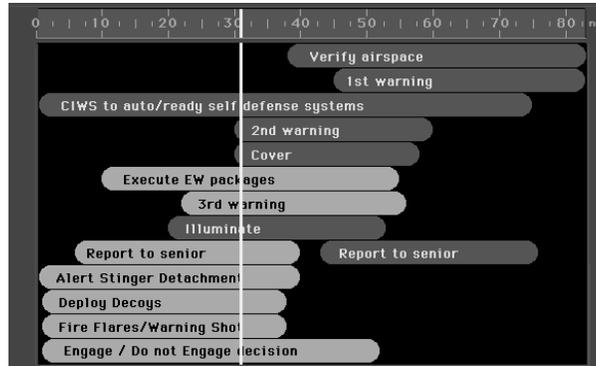

Figure 6. Response Manager

Track Summary

The track summary module, shown in Figure 7, provides a more detailed summary of current and historical data for the currently selected track than was provided in DSS-1. Revisions were motivated by feedback from operational users about the DSS-1. The revised display provides a quick-look at the track's kinematics as well as ancillary data, such as available intelligence, electronic warfare (EW) and Identification-Friend-or-Foe (IFF). Current information is shown as a cyan color, while historical or supplementary data is grayed-out. Kinematics (quantitative data) are read down from the track number while track ID (verbal data) are read across. Embedded within all CROs is a unique feature of the DSS - an altitude trend arrow that shows increasing, constant or decreasing altitude. This is a critical feature of assessing threat intent that must be inferred by the tactical decision maker using conventional systems. There is a large pop-up window which may be accessed to provide a larger view of (possibly) more detailed alerts. This module is expected to be used when more detailed information is required for EBR.
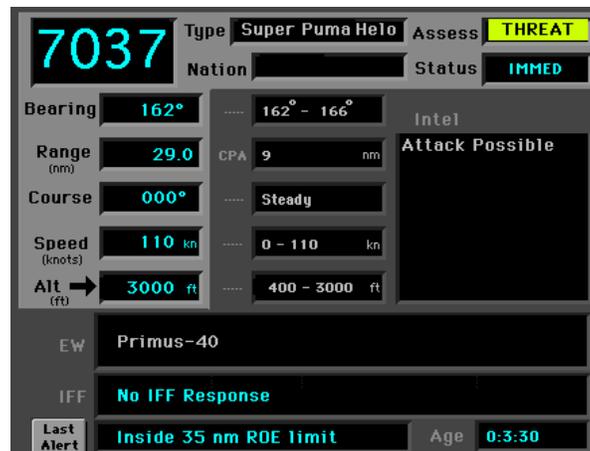
Figure 7. Track Summary

Basis for Assessment

The basis for assessment module is shown in Figure 8. It is similar in format to DSS-1, although substantial revisions are still under development. The basis for assessment module was explicitly designed to support EBR by providing a detailed list of evidence for and against the current assessment of the selected track. It also presents unknown information and implicit assumptions being made in accepting the assessment of potential threat, non-threat, or unknown. This module supports EBR (story generation) by allowing the decision maker to explore alternative hypotheses and to see how the available data do or do not support them. The basis for assessment module presents the underlying data used to generate the DSS's threat assessment for the displayed track. The display was designed to present the relevant data necessary for a commander to consider and evaluate all likely explanations for what a target may be, and what it may be doing (i.e., assess "intents") through the generation of alternative stories to explain the available and missing data regarding the target in question. The display is also intended to highlight data discrepant with a given hypothesis in order to minimize confirmation and framing biases. Assumptions listed are those necessary to "buy into" the selected assessment. Further, the assumptions are intended to prompt the decision maker to consider ways to resolve ambiguity. For instance, in order to assess a track as a threat, it may be necessary to assume that the track is carrying a weapon.   The decision maker could then use organic assets such as friendly aircraft in the vicinity to fly out to the track to assess whether this is actually the case. As a result, the basis for assessment module is expected to be particularly effective in helping sort out and avoid "Blue-on-Blue" and "Blue-on-White" engagements.
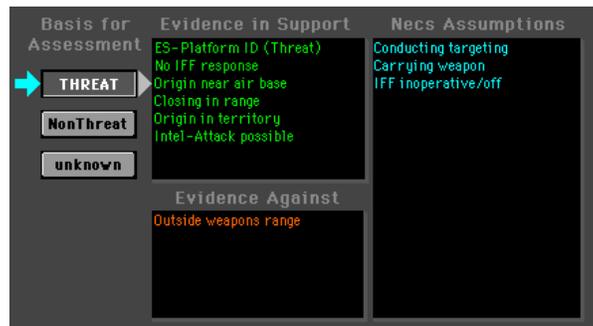

Figure 8. Basis for Assessment

DSS EVALUATION EXPERIMENT

The ultimate goal of any display design is to positively influence the performance of the person-machine system of which it is a part. Therefore, a study was performed to examine how the DSS impacted the decision making of COs and TAOs relative to performance in a traditional CIC in a medium-fidelity simulation. Although the contributions of individual display modules could not be assessed objectively due to resource limitations, overall effects of the DSS on decision performance were examined in terms of a variety of performance criteria.

The results reported here address DSS-1 design issues. Other findings from this experiment have been reported elsewhere (Kelly, Hutchins, & Morrison, 1996; Morrison, Kelly, & Hutchins, 1996). Again, the DSS-1 did not feature an integrated geo-plot, but it relied on a geo-plot similar to that used in current tactical systems, provided as part of the DEFTT simulator. The DSS-1 also had a comparison to norms module, which provided color-coded squares to show how well a set of critical parameters for the selected track fit a template for known threats and support pattern matching. This module was not well liked and was not used as had been intended. It was, therefore, dropped from the DSS-2. The DSS-2 is a refined version of the one tested, and it was refined based on the results of this study.

As previously discussed, there is substantial evidence that experienced tactical decision makers employ feature matching and, to a lesser degree, story generation strategies. Moreover, various errors observed during tactical scenarios and exercises have been linked with basic cognitive limitations (memory, attention, etc.). To build on the naturalistic decision strategies that experienced COs and TAOs use and to help overcome their cognitive limitations, a series of decision support modules were developed. Since these decision support modules were developed with a user-centered design perspective, we expected that they would be effective in reducing decision maker errors. Also, we expected that COs and TAOs would consider these modules to be useful and easy to use, since they were consistent with the strategies that they use in processing information and making decisions.

Method

Participants and Support Team

Sixteen active-duty U. S. Navy officers participated in this study as eight CO-TAO teams. These officers were highly experienced in air warfare tactical decision making, and several were actual shipboard CO-TAO teams. The participants had completed the necessary training courses to be TAO-qualified and had extensive shipboard experience standing watch as a TAO. Those participating as the CO were of appropriate ranks for this role (two Captains and six Commanders), had an average of 20.4 years in the Navy, and had been part of an average of 5.9 deployments. Similarly, those participating as the TAO held appropriate ranks (five Lieutenant Commanders and three Lieutenants), had served in the Navy for an average of 12.8 years, and had deployed an average of 4.0 times, mostly to the Western Pacific and Persian Gulf.

A standard team of enlisted personnel served as the support staff for each CO-TAO team. These personnel served as Electronic Warfare Supervisor, Identification Supervisor, Tactical Information Coordinator, and Anti-Air Warfare Coordinator. They worked at consoles adjacent to the CO and TAO consoles within the DEFTT Lab. The support staff were trained to use their displays and to provide information in a consistent manner across CO-TAO teams, controlling a major source of extraneous variability.

Materials

This study was performed in the DEFTT Laboratory at NRaD using dual-screen workstations for officers participating as the CO and the TAO. For each, one screen presented a standard geo-plot, and the other screen presented the DSS-1 display. Two large screen displays, comparable to those in many shipboard CICs, were also available.

Two training scenarios and four test scenarios were used. All scenarios were set in the Persian Gulf and involved a similar mix of fixed-wing aircraft, helicopters, and surface ships. The test scenarios, in particular, were constructed to simulate peace keeping missions with a very high number of targets to be dealt with in a short period of time (i.e., were time compressed) and with a significant number of highly ambiguous tracks regarding assessment and intent. Analyses of the test scenarios indicated that while they differed in many details, they were roughly equivalent in decision workload overall. Each of the test scenarios had a duration of approximately 20 minutes.

Procedures

Upon arriving at the DEFTT Laboratory, participants were given a pre-briefing on the objectives and procedures for this study, were oriented to the laboratory equipment and staff, and were provided with necessary reference materials. Criterion-referenced training with the baseline DEFTT display system and with the DSS was then provided, and two practice scenarios were run prior to beginning the test session.

Participants were given appropriate geo-political and intelligence briefings prior to each test run and were encouraged to take as much time as necessary in order to set-up their displays the way that they wanted and to familiarize themselves with the tactical situation. This was done in order to reduce the artificiality of feeling as if they were dropped into the middle of a tactical situation which would normally have developed over several hours. During the scenario runs, the activity and communications from both the support team and various external sources was scripted. At various points in each scenario, CO-TAO teams were requested to report their primary tracks of interest. Subjective workload was assessed immediately following each test scenario using the NASA Task Load Index (TLX; Hart & Staveland, 1988). Short breaks were provided between each scenario.

At the conclusion of the last test scenario, each participant was asked to complete a questionnaire that involved ratings of the usage, utility, and usability of each of the modules and of the DSS overall. Then, a brief structured interview was used to solicit comments about the strengths and weaknesses of the DSS, including suggestions for changing the displays and the information provided within them.

Experimental Design and Performance Measures

A within-subject factorial design was employed across the four test scenarios such that each team performed two scenarios with the DSS and two scenarios without it. In all scenarios, the CO-TAO team had the use of geo-plot and CRO displays comparable to those in use in

current ship CICs. The order of the scenarios and DSS conditions was counterbalanced using a Latin Square procedure.

In addition to collecting objective data on tactical actions, display usage, control inputs, and voice communications, subjective assessments, via questionnaires and a structured interview, were solicited from each CO and TAO at the conclusion of the test session. The voice communications for each test scenario run were transcribed and analyzed in order to identify decision making anomalies and communication patterns.

Results

Several classes of research questions were examined as part of this study. These concern DSS utility, situation awareness, team communications, and DSS usability. While the data discussed below were subjected to various statistical analyses, care should still be taken in their interpretation. The small sample size and the simulated test environment used in this field research make it difficult to draw definitive conclusions about tactical decision making performance. Nevertheless, these findings do provide a strong indication of the potential utility of DSS displays for tactical decision making, particularly in littoral air warfare situations.

DSS Utility

If COs and TAOs considered the DSS to be useful for tactical decision making, we would expect them to make use of it during the test scenarios when it was available. Similarly, we would expect them to report that the information provided by it was useful for their decision processes.

At 1 minute intervals throughout the test scenarios, subject matter experts recorded whether or not the COs and TAOs were attending to the DSS. On the average, participants were observed to be attending to the DSS at 66% of the time samples. At the remainder of the time samples, the COs and TAOs were observed attending to the geo-plot, which was on a separate display. In many cases, it was possible to determine to which DSS module(s) they were attending, and these data are shown in Figure 9. It can be seen that the COs tended to use the DSS somewhat more than the TAOs, particularly for purposes of maintaining awareness of individual track's behavior, responses completed and pending, and the relative priority of active tracks. TAOs, on the other hand, tended to spend more of their time using the DSS to acquire quantitative track status and sensor data. Both COs and TAOs tended to use the upper half of the DSS (i.e., Track Summary, Track Profile, and Response Manager modules) most often. Self-reports of module usage from the participants corroborate these findings.
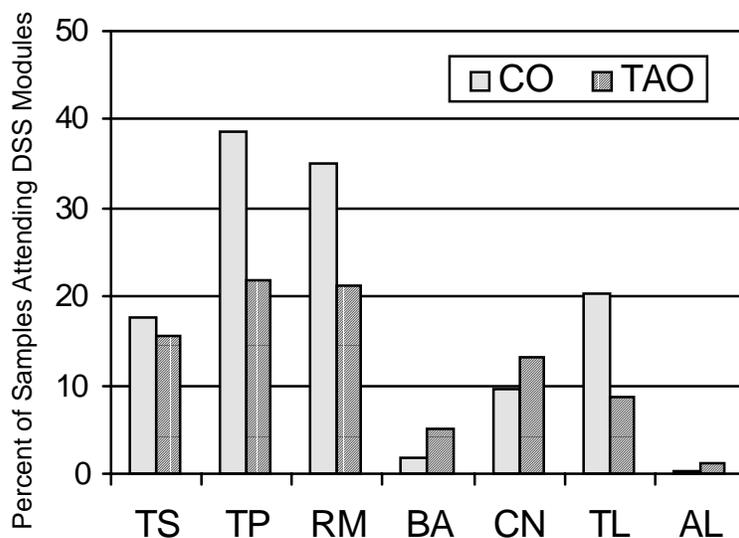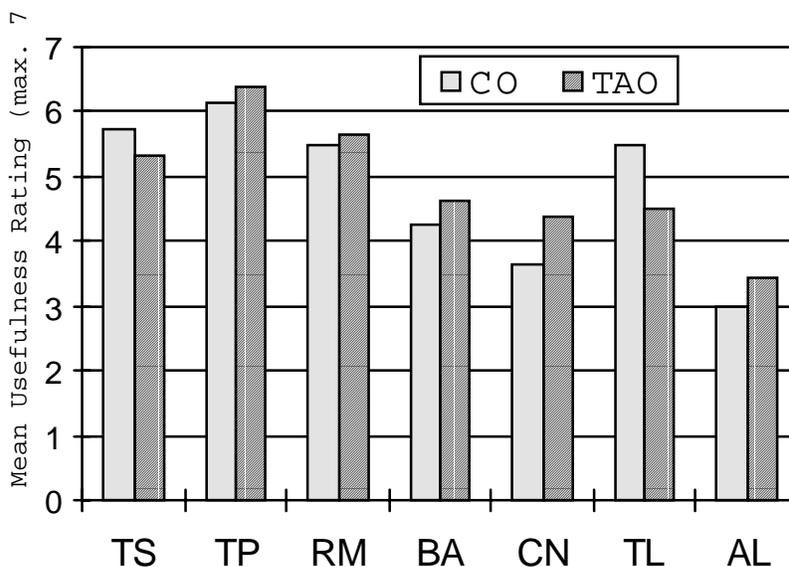
Figure 9. Mean Percent of Time where the CO and TAO Attended to DSS Modules

Note: TS = Track Summary, TP = Track Profile, RM = Response Manager, BA = Basis for Assessment, CN = Comparison to Norms, TL = Track Priority List, AL = Alerts List.

At the completion of the test session, participants were asked to complete a questionnaire that called for a variety of ratings of the DSS. Figure 10 shows the average ratings (on a 7-point scale) of how useful the COs and TAOs felt the information provided in the DSS modules was. Ratings indicate that most modules were considered to be quite useful for tactical decision making, particularly those parts of the DSS designed to support quick decision making. COs and TAOs noted that these modules enabled them to extract key information rapidly and to visualize track behavior easily. Participants considered that the DSS overall offered high utility (average rating = 5.97 of 7-points) for tactical decision making in littoral warfare situations.

Figure 10. Mean Rating of the Usefulness of DSS Modules

Feedback from the CO-TAO teams who participated in this experiment indicated that the DSS provided them an excellent summary of the overall tactical situation as well as of key data for individual tracks. In particular, COs and TAOs considered that both the Track Profile and the Basis for Assessment modules provided important information not readily available in present day systems. Since the Track Profile module supported feature matching, which is the most commonly used decision strategy, its high rating was anticipated. Yet, when the track data are conflicting or ambiguous and when the decision maker has time available, the Basis for Assessment module was rated as helping substantially. Note that by encouraging decision makers to consider the full range of available evidence along with various explanations for it, this module reduces the likelihood of mistakenly engaging friendly or neutral tracks, and was rated highly with regard to avoiding Blue-on-Blue and Blue-on-White engagements.

Situation Awareness

Awareness of the tactical situation was examined via several performance measures. Specifically, it was predicted that if the CO-TAO team was more aware of the tactical situation in a peace-keeping mission, they would:

- identify the critical contacts earlier and more accurately;

- take more of the defensive tactical actions required by the rules of engagement earlier, and take more of the provocative (offensive) actions later; that is to say the window of time during which decision makers had to evaluate and handle a track would increase; and

- ask fewer questions to clarify previously reported track data and the relative locations of tracks.

Critical contacts. During the scenario runs, the CO-TAO team was probed at pre-speci-fied times to identify the tracks that were considered to be of greatest tactical interest at that time. Their responses were contrasted with those of an independent group of five subject matter experts. As shown in Figure 11, significantly more of the critical contacts were identified when the DSS was available. Significant differences ($p < .05$) were noted at both the early and mid-scenario probes; performance was comparable at the late probe, however. Late in the scenario the critical tracks may become more obvious even without the DSS. Nevertheless, recognition of critical tracks earlier in the scenario affords decision makers a broader array of response options and permits more effective coordination of response actions.
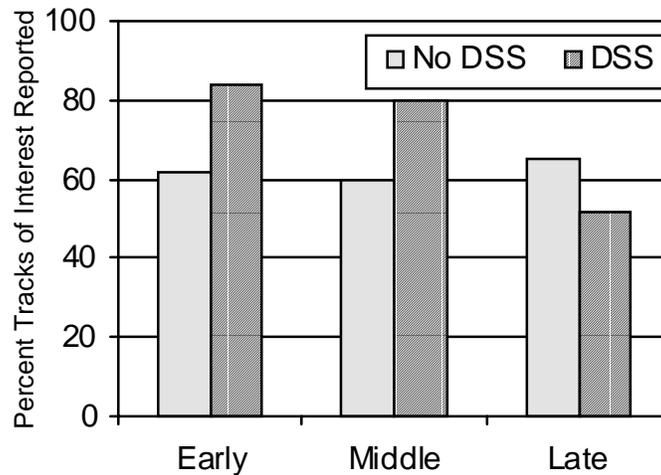
Figure 11. Percent of Critical Contacts Reported as Tracks of Interest

Tactical actions. Using the rules of engagement as a benchmark for decision performance in the scenarios, a group of subject matter experts assessed whether the CO-TAO teams warned and/or illuminated threat tracks at specified times and took appropriate defensive actions. A modified form of the Anti-Air Warfare Team Performance Index (ATPI) was used for scoring tactical performance (after Dwyer, 1992), and these data are summarized in Figure 12. In scenarios when the DSS was available, CO-TAO teams were significantly more likely to take defensive actions in a timely manner against imminent threats ($p < .05$). This indicates that the DSS promoted an earlier recognition of the emerging risks of the tactical situation. By contrast, no difference was observed in the number of tracks that were warned or illuminated (i.e., provocative actions) when the DSS was available.   However, several subject matter experts contended that warnings and illuminations may not be diagnostic performance indices in these scenarios since they represent provocative tactical actions that commanders may consider to be inappropriate against certain tracks in a littoral situation. Not taking provocative actions would be appropriate and expected if commanders had assessed that the track was not an imminent threat, and felt comfortable with prolonging those actions because they had a good tactical picture - as would be expected if the DSS was being effective in meeting its design objectives.
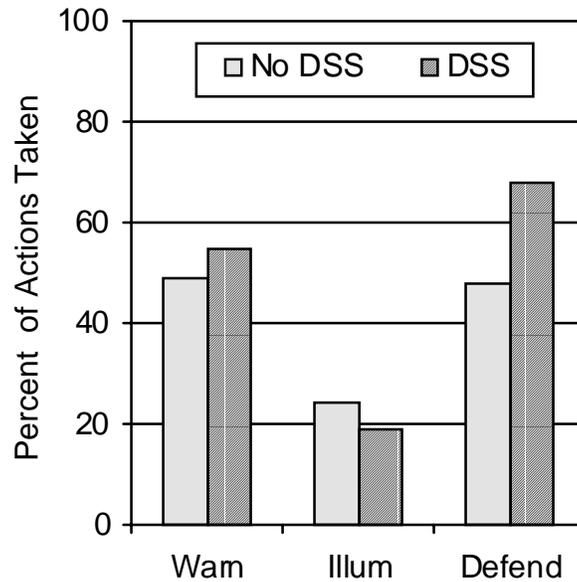
Figure 12. Team Performance of Tactical Actions required by the Rules of Engagement

Team Communications

Although the DSS was primarily designed to support an individual decision maker, the information that it provides could be expected to influence the team's collective decision process. One way that influence could be observed is by changes in the team's communications. Therefore, the communications rate, the pattern of communication, and the content of the communications were compared with and without the DSS.

Communications rate and pattern. It was hypothesized that when using the DSS, teams would have less need to exchange data verbally and would, thus, communicate less often. To test this, all voice communications that requested or provided information were tabulated for each of the 32 test runs. Since the length of the scripted test scenarios differed, the total number of voice communications observed was divided by the scenario duration to give a communications rate.

Figure 13 shows the mean rate of communications originating with the CO, TAO, other members of their team, and others external to the ship's combat center (e.g., the battle group commander, the bridge). A general decrease in communications rate with the DSS was observed. This decrease remains fairly consistent regardless of who originated the communication. In fact, the pattern of communications was unaffected by the presence of the DSS. About 40% of the communications occurred between the CO and TAO, and another 35% occurred between the TAO and the team. Each of the remaining links accounts for about 5% or less of the total communications. The decrease in communications across positions suggests that the DSS supported the entire team by providing basic data about tracks, thereby reducing their need to request or provide such data verbally.
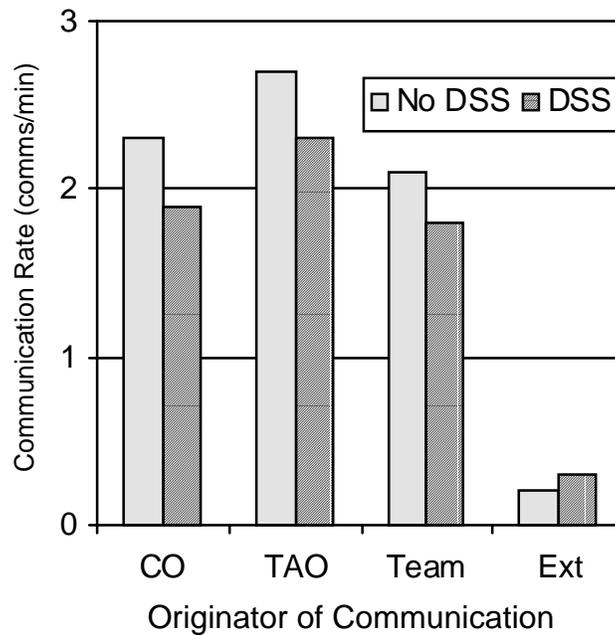
Figure 13. Voice Communications Rate by Message Originator and DSS Condition

Communications content. While the pattern of communications was not found to be affected by the presence of the DSS, the content of the teams' communications may be altered by the DSS. That is, without the DSS, teams might need to spend more time exchanging basic track data while those with the DSS might spend the bulk of their time assessing track intent or evaluating alternate courses of action. To explore the possibility of a qualitative trade-off in communications, voice communications were coded by their message content according to the following scheme:

- *Information* – exchange of sensor-based data;

- *Status* – exchange of procedure-based data;

- *Clarification* – redundant communication to elucidate, interpret, or correct other communications;

- *Correlation* – association of two or more data;

- *Assessment* – discussion of expected track behavior, likely intent, or future actions;

- *Orders* – commands to perform an action.

Figure 14 shows the overall average proportion of communications observed for each of these content categories. The largest proportion involved Information communications, in which sensor-based data were exchanged. This, of course, is not surprising since these data effectively drive the decision processes. The *rate* of these communications, however, was found to be significantly lower when the DSS was available ($p < .05$). Since the DSS provides much of these data, there was less need for verbal exchanges among the team. Similarly, fewer Correlation communications were observed when using the DSS. Although decision makers were less likely

to ask about or report correlation data with the DSS (since much of it is displayed automatically), they were somewhat *more* likely to talk about correlations in the data that they observed on the DSS. This effect is consistent with the intended purpose of the DSS in supporting situation assessment and track evaluation. No differences between DSS and No DSS runs were observed in the other communications content categories.
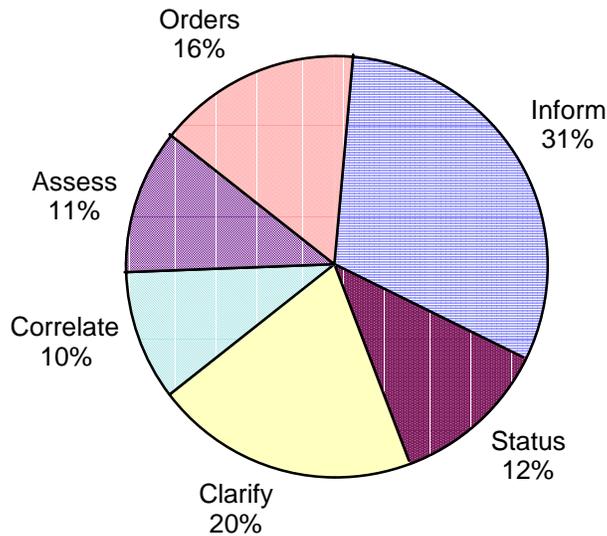

Figure 14. Mean Proportion of Communications by Content Category

Clarification communications. Overall, about 20% of the communications were for clarification purposes, reflecting uncertainty about track location, kinematics, identification, status, or priority. Figure 15 shows the type of information that was discussed during clarification communications. The relative percentage of each type of clarifying communication is shown when the DSS was available and when it was not available. Thus, it can be noted that clarification communications about Track Location (e.g., locating the symbol on the geographic display that corresponds to the track of interest) and Track Status (e.g., response to warnings) were equally likely whether or not the DSS was available. In contrast, clarification communications about Track Kinematics (e.g., speed or altitude), EW Information (e.g., IFF and emitter signature), and Tactical Picture (e.g., track identity and relative position) were less likely when the DSS was available. Clarifications regarding Ambiguous Orders (e.g., incorrect track number) were somewhat more likely to occur when the DSS was available.

The pattern of these findings is actually quite revealing about the information in the DSS that was used by tactical decision makers. Since the multi-track geo-plot display and the DSS were not linked in this study, the DSS by itself provided little help in locating tracks in relation to each other. Similarly, the DSS provided no information to decision makers about the status of actions taken or about tracks' responses to warnings. Thus, no difference between DSS and No DSS conditions was expected for these types of clarifications, and the data confirm this prediction.

22

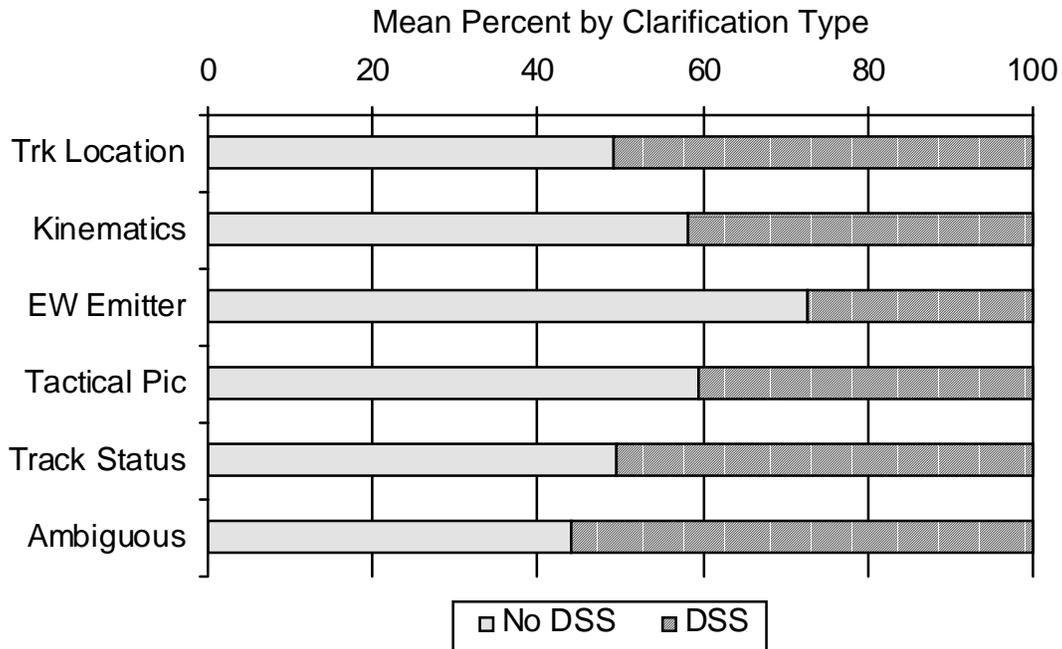## Mean Percent by Clarification Type

Figure 15. Relative Percent of Clarification Communications With and Without the DSS.

In contrast, the DSS Track Summary module showed available track kinematic data, the supplementary data read-out in the Comparison to Norms module summarized EW information, and the Track Profile and Track Priority List modules helped decision makers prioritize and maintain awareness of the overall tactical picture. Therefore, the DSS was expected to reduce the need to clarify these types of information. Again, these predictions were confirmed, especially for the EW information that was consulted frequently in making threat assessment decisions.

Although not statistically significant, the tendency for there to be more clarification communications about ambiguous orders when using the DSS was interesting. In fact, we observed a tendency for decision makers to be more precise in referencing tracks when using the DSS. This increased precision thereby encouraged the team to ask for clarification about which track was being referenced when ambiguous orders occurred.

This detailed analysis of clarification communications confirmed that decision makers did indeed use the information that was displayed in the DSS, even with only very limited experience with it. This analysis also revealed ways in which the DSS might be enhanced to further reduce the burden of clarification communications. For example, linking the geo-plot display with the DSS such that selecting a track on either display would highlight it on the other would probably reduce the communications to clarify track location. This enhancement was incorporated into the DSS-2 design, as previously described. Similarly, enhancements to the Response Manager module to show the status of actions and responses by tracks would be likely to reduce the need for decision makers to repeatedly ask about such track status information. Current work is exploring these and other enhancements to the Response Manager module.

Communications about critical contacts. The tracks to which the teams' communications referred were also examined under the DSS and the No DSS conditions. The hypothesis was that the DSS would enable teams to focus on the critical contacts more quickly, resulting in a greater proportion of their communications about those tracks. The average proportion of communications about the critical contacts was slightly greater (but not significant) when using the DSS. It is not particularly surprising that these teams concentrated the bulk of their communications on the critical contacts regardless of whether or not they were using the DSS. After all, these were highly experienced tactical decision makers who are accustomed to functioning effectively with their current (non-DSS) systems. Thus, greater effects might be obtained with less experienced decision makers.

DSS Usability

Even if the DSS provides useful information, promotes better situation awareness, and facilitates team communications, it must also be usable. That is, decision makers should consider it easy to learn, easy to understand, and easy to use.

Participants were asked to complete a questionnaire following the test session in which they rated the usability of the DSS and its modules. Comments about the DSS interface were also solicited in the questionnaire and during a follow-up interview. The overall rating of DSS usability was high (average rating of 4.16 of 5-points). Similarly, most modules were considered easy to use, as shown in Figure 16. The modules that promoted "quick-look" assessments of track status, location, and priority were rated as more usable. The modules that were predominately text-based, particularly the Alerts List, were rated as less easy to use.

The COs and TAOs offered many valuable suggestions for improving the DSS to make it more useful and usable for tactical decision making and were incorporated into the DSS-2 described above. Frequently heard suggestions included the need to:

(a) integrate the DSS display with the geo-plot to simplify track selection procedures and to display multiple tracks,

(b) allow user-customizable display areas and content, particularly for control window size, range scale, and Response Manager actions,

(c) allow command-override of track priorities and threat assessments, and

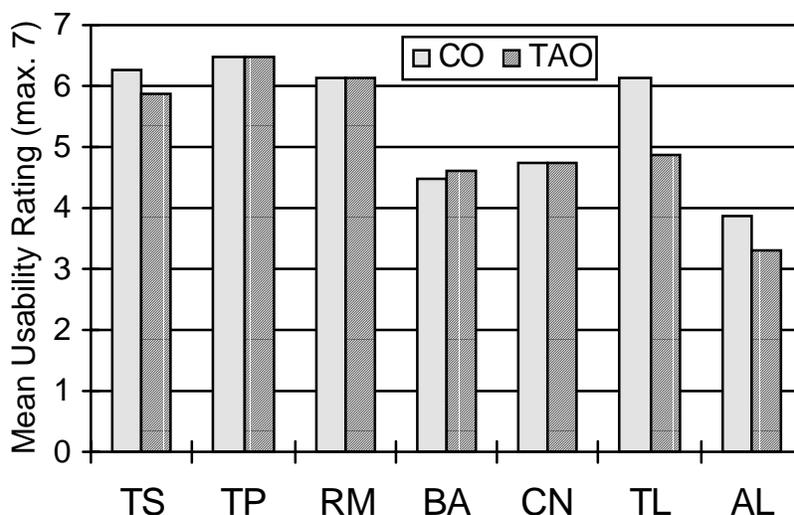(d) provide an expanded Track Priority List that shows more tracks in a graphic/spatial format.

Figure 16. Mean Rating of the Usability of DSS Modules.

Note: TS = Track Summary, TP = Track Profile, RM = Response Manager, BA = Basis for Assessment,
CN = Comparison to Norms, TL = Track Priority List, AL = Alerts List.

While several of the usability suggestions by experienced tactical decision makers concern improvements to the "look and feel" of the DSS interface, others have substantial implications for our understanding of their underlying naturalistic decision processes. Their call to integrate the DSS and geo-plot displays and to provide more tracks in the Track Priority List reminds us that feature matching decision strategies involve evaluating tracks within the context of other related tracks and events. Decision makers' requests to permit them to customize their displays and to override system defaults have implications for their story generation / explanation-based reasoning strategies. Namely, decision makers seem to use a "stepping stone" approach whereby they use available data to reach an intermediate conclusion about a track, say its priority. Based on that intermediate conclusion, they then continue to use other data to explore further implications for that track, given its priority. To support such a process, it is clear that decision makers would want to have the ability to override and customize their DSS displays.

CONCLUSIONS

Operational decision making predominantly relies on feature matching strategies. To a lesser extent when faced with conflicting or ambiguous data, decision makers employ story generation or explanation based reasoning strategies. Displays that are consistent with these naturalistic decision making strategies provide the most useful support to commanders, facilitating the rapid development of an accurate assessment of the situation. Displays that support both feature matching and explanation based reasoning are recommended for complex decision making tasks. While the feature matching displays will likely be used far more often, the explanation based reasoning display is of substantial value under certain circumstances, particularly with less experienced decision makers.

The DSS was developed for application to Navy tactical decision making on a single ship in support of Air Warfare in dense, fast-paced littoral settings. With some adaptation, it could support other military decision situations, including concurrent decisions involving other warfare areas, higher-level, supervisory decisions involving multi-ship battle groups, and even collaboration among tactical decision makers in joint service or multi-national (coalition) operations. Several new research projects are underway to explore these applications. In addition to these direct applications to support military decision making, the decision support and display principles identified through this effort are relevant to other complex decision making settings, such as nuclear power control, flight control, process control, and disaster relief planning. Further, additional work is looking at developing derivative displays that reflect emerging theories of decision making, extension of the DSS concepts to other workstations within the CIC, as well as better integration of DSS modules with shipboard data processing systems.

Lessons Learned

1. Experienced tactical decision makers' reactions and performance indicated that the DSS provided them useful information in a readily understood form. An advantage of the DSS over current systems is that it supports "quick-look" decision processes (i.e., RPD), which are typically used in tactical situations. Naturalistic decision processes are further supported by presenting data in an operationally relevant context, which involves historical, geo-political, and tactical doctrine components.

Implications. A user-centered approach to decision support and display design is important for achieving user acceptance and ecologically relevant performance enhancements. A thorough understanding of the users' information requirements and decision processes provides a critical foundation for successful design efforts. This allows data to be viewed in relation to a specific decision situation (e.g., embedding a track's kinematic data in graphical theat templates).

2. User reactions to the DSS indicated that tactical decision makers appreciated access to data that were not heavily filtered or preprocessed. Those algorithms that were implemented (e.g., the track prioritization algorithm) were kept simple and understandable. Thus, the DSS makes use of tactical data already available in the CIC without fusion into more complex, abstract concepts. This gives the COs and TAOs greater confidence in their decisions, since they can maintain conceptual links to basic physical data (e.g., a contact's kinematics - range, speed, altitude). The DSS organizes these data around key tasks that tactical decision makers need to perform. This display organization makes it easier for experienced decision makers to get the information that they need quickly and to note inconsistencies or ambiguities in the data. Because the display is structured around specific operational tasks, decision makers were able to navigate the display efficiently and extract the required information. For less experienced decision makers, the DSS structures their information search so that they are able to better understand the relationship among the data and how it could be used. As a result, less experienced decision makers function more like experts.

Implications. Decision support and display design efforts should carefully consider how automation is incorporated. Human decision makers require access to the underlying data to enable situation assessment based on the patterns detected in these data. A promising design approach is to organize archival sensor data into modules that support critical decision making tasks.

3. DSS display modules that used graphics and that integrated information from several sources were used most often and were considered the easiest to learn and use.

Implications. Operational decision making is characterized by time stress, multiple concurrent task requirements, and situation uncertainty. Displays that enable decision makers to recognize data patterns quickly provide substantial support under these conditions. Whenever possible, designers should make use of graphic displays and selectable overlays, with color, shape, size, and position coding schemes.

4. Early TADMUS research suggested that the user interface to existing tactical display systems required the decision makers to spend excessive time interacting with displays rather than making decisions. Field studies and analyses have revealed that with some current tactical systems, the user interface dialog places an enormous burden on operators, requiring as much as a thousand control actions an hour (Osga, 1989). It was concluded that time spent interacting with the systems detracts from time available for making tactical decisions.

Implications. Decision support systems should be designed to minimize the amount of interaction required of the decision maker to extract information. Interfaces that require the user to take overt actions in order to obtain information (e.g., selecting between overlapping windows, activating pop-up windows, using pull-down menus, etc.) interfere with the user's ability to process information and should be minimized. The use of tiled windows, with distinct modules organized around specific decision making problems is preferred.

5. Development of the DSS did not follow a simple or straight path. The design was the product of extensive discussions among staff from many different disciplines and backgrounds. In addition, many forms of testing and evaluation of prototype displays were employed – from table-top critiques to formal evaluation experiments.

Implications. The design of decision support displays needs to draw on the unique skills of professionals in many complementary fields, particularly human factors engineers, subject matter experts, and cognitive scientists. Iterative testing with representative users needs to be conducted, in different forms, throughout development. In the early phases, testing of concepts can employ storyboards and static mock-ups. As the design matures, structured feedback from users can be obtained via scripted scenarios with a dynamic mock-up. Finally, more formal laboratory and field studies can be performed with prototype systems to evaluate its effectiveness.

Implications for Future Training Research

In addition to its utility as a real-time decision support tool, the DSS may have substantial value for training the complex cognitive skills that characterize expert tactical decision makers. Since the DSS organizes and presents tactical data in a form that is consistent with experts' usage patterns, it guides decision makers through the huge amount of tactical data available in the CIC. Essentially, the DSS is expected to help an intermediate decision maker quickly access relevant data, identify important patterns, and develop a higher level of expertise. In this regard, the DSS may be useful not only for initial training but also for refresher or recurrency training.

In a training mode, we envision that the DSS could be central to teaching the application of tactical decision making skills and developing expertise. Because the DSS has been designed to support expert decision makers, it is reasonable to expect that it would prove useful in shaping the decision making strategies of novice decision makers and developing expert decision making skills. There is the implication, however, that there would be a need to call-up particular scenarios from a library that are appropriate for the tactical decision making skills to be trained. There would need to be a comprehensive training strategy for developing expert decision making skills as well as appropriate training objectives and performance criterion. The scenarios would presumably be sequenced to fit within an appropriate curriculum, leading trainees from more basic skills to more advanced skills under more demanding tactical situations (i.e., greater workload, more ambiguity). Nevertheless, the system could be flexible enough to enable it to detect a trainee's level of expertise and particular training needs and then adjust the curriculum accordingly. During training, the decision maker could be guided through the use of tactical data shown on the DSS as the scenario unfolded. Intelligent agents, wizards, balloon help, or other such HCI tools could "pop-up" at appropriate points in the scenario to advise trainees on data relationships, requirements to shift attention, or specific applications of other key tactical skills. Then, at the conclusion of each training scenario run, key indicators could be calculated in order to assess decision making performance and provide rapid feedback to the trainee. The DSS would also need to include other standard features that support debriefing and evaluation following training exercises, such as the ability to replay selected parts of the scenario, to view selected points in time, and to annotate the scenario with comments.

To extend the DSS to support this type of tactical decision making training, several research and development activities are required, and are being actively pursued as part of the on-going TADMUS project. These include:

(a) identification of key tactical decision making skills, knowledge, and abilities,

(b) development of a training curriculum for producing (or sharpening) these skills,

(c) development of a comprehensive library of scenarios that exercise these skills under various workload demands,

(d) design and testing of intelligent agents, wizards, and other HCI tools that could be incorporated into the DSS as "over-the-shoulder" advisors or synthetic expert instructors, and

(e) investigation of diagnostic decision performance measures that could be compiled automatically to indicate whether the trainee mastered the skills.

ACKNOWLEDGMENTS

REFERENCES

Dwyer, D. J. (1992). An index for measuring naval team performance. *Proceedings of the Human Factors Society 36th Annual Meeting*, (pp. 1356-1360). Santa Monica, CA: Human Factors Society.

Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload*. Amsterdam: North Holland.

Hutchins, S. G. and Kowalski, J. T. (1993). Tactical decision making under stress: Preliminary results and lessons learned. *Proceedings of the 10th Annual Conference on Command and Control Decision Aids*. Washington, DC: National Defense University.

Hutchins, S. G., Morrison, J. G., and Kelly, R. T. (1996). Principles for aiding complex military decision making. *Proceedings of the Second International Command and Control Research and Technology Symposium*. Washington, DC: National Defense University.

Jacobsen, A. R. (1986). The effects of background luminance on color recognition and discrimination. *Society for Information Display Technical Digest*, 78-81.

Kaempf, G. L., Wolf, S., and Miller, T. E. (1993). Decision making in the Aegis combat information center. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, (pp. 1107-1111). Santa Monica, CA: Human Factors and Ergonomics Society.

Kelly, R. T., Hutchins, S. G., and Morrison, J. G. (1996). Decision process and team communications with a decision support system. *Proceedings of the Second International Symposium on Command and Control Research and Technology*, (pp. 216-221). Washington, DC: National Defense University.

Klein, G. A. (1989). Recognition-primed decisions. In W. R. Rouse (Ed.), *Advances in Man-Machine Systems Research*, vol. 5 (pp. 47-92). JAI Press, Inc.

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 Action: Models and Methods* (pp. 138-147). Norwood, NJ: Ablex Publishing Corporation.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.) *Decision Making in Action: Models and Methods*. Norwood, NJ: Ablex Publishing Corporation.

Moore, R. A., Quinn, M. L., and Morrison, J. G. (1996). A tactical decision support system based on naturalistic cognitive processes. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting*, this volume. Santa Monica, CA: Human Factors and Ergonomics Society.

Morrison, J. G., Kelly, R. T., and Hutchins, S. G. (1996). Impact of naturalistic decision support on tactical situation awareness. *Proceedings of the Human Factors Society 40th Annual Meeting*, (pp. 199-203). Santa Monica, CA: Human Factors and Ergonomics Society.

Nugent, W. A. (September 1996). Comparison of variable coded symbology to a conventional tactical situation display method. *Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society*, (pp. 1174-1178). Santa Monica, CA: Human Factors and Ergonomics Society.

Osga, G. A. (March 1989). *Measurement, modeling, and analysis of human performance with Combat Information Center consoles*. (NOSC Technical Document 1465). San Diego, CA: Naval Ocean Systems Center.

Osga, G. A. and Keating, R. L. (March 1994). *Usability study of variable coding methods for tactical information display visual filtering*. (NRaD Technical Document 2628). San Diego, CA: Naval Command, Control and Ocean Surveillance Center, Research Development Test and Evaluation Division.

Rausche, D. P. (March 1995). *Variable coded symbology and its effects on performance*. (Masters Thesis). Monterey, CA: Naval Postgraduate School.

Van Orden, K. F. and Benoit, S. (March 1994). *Color recommendations for prototype maritime symbolic displays*. (NSMRL Technical Report 1192). Groton, CT: Naval Submarine Medical Research Laboratory.