Third International Command and Control Research and Technology Symposium, National Defense University, June 17-20, 1997

Eye Tracking in Tactical Decision Making Environments: Implications for Decision Support Evaluation*

Jeffrey G. Morrison
NCCOSC - RDT&E Div., Code D44210
San Diego, CA 92152-5001, E-Mail: JMorriso@nosc.mil

Sandra P. Marshall
Department of Psychology, San Diego State University
San Diego, CA 92182, E-Mail: SMarshal@sciences.sdsu.edu

Richard T. Kelly and Ronald A. Moore
Pacific Science & Engineering Group, Inc.
6310 Greenwich Drive, Suite 200, San Diego, CA 92122

Abstract

This paper discusses the use of eye movement monitoring systems in evaluating visual displays to support tactical decision making. A variety of analytic approaches are considered from a conceptual and practical standpoint and applied to preliminary data collected using the TADMUS DSS display. Eye movement data were collected using a head-mounted eye tracking system worn by command-level decision makers (Commanding and Tactical Action Officers) in tactical scenarios using a decision support system in a simulation of an Aegis cruiser Combat Information Center. While results are preliminary and not definitive, prospects for the use of eye movement monitoring systems as tools for evaluating decision support displays are considered good.

1 Introduction

The Tactical Decision Making Under Stress (TADMUS) program began with the premise that environmental and emotional stressors could have subtle but significant effects on tactical decision making. This premise was suggested by the congressional investigation of the USS Vincennes incident in 1988, wherein an Iranian Airbus was shot down by an Aegis cruiser because relevant tactical data could not be accessed, integrated and interpreted in the short time available to assess the aircraft's status as a threat. The inquiry concluded that the quality of decision making within the combat information center (CIC) during the incident may have been negatively impacted by the stress induced by an on-going surface engagement in which the USS Vincennes was engaged. The inquiry also concluded that the scientific community and system designers did not know enough about these phenomena to address this issue. With this in mind, ONR established the TADMUS program to study command decision making and the effects stress might have on it. As a result of this effort, we have learned much about complex decision making and how to mitigate the ef-

* Funding for the research cited in this paper was received from the Cognitive and Neural Science Technology Division of the Office of Naval Research.

fects of stress on it. The project has developed a number of training and decision support, human-computer interface (HCI) interventions which continue to be refined and analyzed. The study of these interventions and validation of their impact has, however, been problematic, specifically with regard to the effects of HCI interventions, because appropriately diagnostic metrics and methods for evaluating them and their effects on performance do not exist.

The design of HCI display formats is still as much an art as it is a science. There is no definitive means for assessing the efficacy of a design, or for that matter developing a theory of how displays are used by decision makers. Designers of man-machine systems have long wanted to be able to determine quantitatively how displays and controls are actually used by their human operators, so as to facilitate the design of better displays. If we cannot know how the operator processes information, we would like to be able to treat the operator as a “black box” where we can measure what goes in and what comes out, so we can infer how to increase throughput of information. In order to evaluate the use of a display, the researcher needs to know how a display is being used in the context in which it is to be used, i.e., what information the operator is looking for and where he is looking to get it. This is very difficult to achieve using traditional evaluation tools, and satisfactory measurement tools are needed.

The problem becomes even greater in HCI displays, such as those found in the TADMUS Decision Support System (DSS) because they are designed to *integrate* data into meaningful information to solve specific decision making problems. While the DSS and similar displays are typically designed as a series of modules, the modules are in fact highly interrelated and often present similar data in several different ways to support different decision making tasks. Further, the modules are collocated on a single visual display surface, (e.g., a video monitor). This makes assessing how individual modules are being used to solve various decision making tasks which occur in natural decision making extremely problematic.

A number of innovative methods have been attempted to evaluate how displays are being used and the efficacy of alternative display formats. Among the most common methods are those that attempt to assess the individual modules by including / excluding modules in a systematic manner and seeing what the effect is on task performance metrics. In effect, this approach asks “how does the module(s) affect task performance?” There are several problems with this approach. First, it requires alternative formats to be generated to provide comparable information so that the task can be completed, and allow the realtive merits of formats to be compared. This can be resource intensive to design and build, and if the alternatives are not carefully designed to differ from each other in theoretically meaningful ways, can lead to results that are very difficult to interpret. The problem becomes increasingly difficult as the display or task grows more complicated or the number of modules within the display increases beyond two or three. Further, such an approach creates problems because it, by definition, destroys the integrated properties of a well designed display. Nonetheless, if the application designer is interested in comparing several alternative displays in terms of their effects on performance, a suitable method is required. This approach can be very useful in evaluating alternative designs from a practical perspective, but it is not particularly diagnostic in terms of understanding why one display is better than another, or how the data are used in decision making. People, and particularly expert decision makers, are very adept at compensating for the deficiencies of a display. Further, a variety of factors can interact in affecting the performance seen from different display formats. These include: legibility of the displays, display placement relative to each other and those for competing tasks; and information processing issues such as: the interpretability of the displays, the success with which the display components have been integrated to solve specific decision making problems, the level of cognitive workload experienced by the decision maker in using the display, and the strategies used by the decision makers while performing the tasks. Therefore, a more diagnostic means of evaluating the information utility of display formats is highly desirable.

The use of an eye movement monitoring system is conceptually able to solve a number of problems in evaluating how a display is, used because it allows data to be collected regarding where the decision maker is looking at any given time within the context in which the display is designed to be used, rather than an artificial one contrived on the basis of a research paradigm and experimental control. This is not to say that sophisticated analytic tools are not required to interpret eye movement data; however, it may be argued that incorporation of eye movement data into an system evaluation offers the potential for significant diagnosticity into how displays are used and the underlying decision making than is available from other tools and techniques. The following section will survey some of these techniques and their potential application to the study of tactical decision making displays, as illustrated with an on-going evaluation of the TADMUS DSS.

2 TADMUS Decision Support System

The TADMUS DSS was developed with the objectives of: (1) minimizing the mismatches between cognitive processes and the information available in the CIC to facilitate decision making; (2) mitigating the shortcomings 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 to facilitate the interpretation of spatial data. The design goal of the DSS was to take the data in the system and present them as meaningful information, (i.e., when, where, and in the form needed), relative to the decision making tasks being performed based on a theoretical understanding of human decision making. For a detailed discussion of the DSS and how it was developed, see [Hutchins, *et al.*, 1996a], [Hutchins, *et al.*, 1996b], and [Morrison, *et al.*, 1996].

The current generation DSS was designed expressly for the evaluation of display elements to support feature matching, story generation (viz., Explanation-Based Reasoning (EBR)), and Recognition-Primed Decision making (RPD) 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 is implemented on a personal computer which may operate independent of, synchronized with, or linked to a scenario driver simulation.

Figure 1 shows the first DSS prototype display. The DSS is a composite of several display modules, which 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, *et al.*, 1996]. Nevertheless, three modules will be discussed here 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.

2.1 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 rapidly: (1) see where the target track is relative to own-ship, (2) see what the track has been doing over time, (3) recognize whether the target can shoot you, and (4) recognize whether you could shoot the target. Knowledge engineeri ng showed that these issues were preeminent every time a decision maker considered a track. An important aspect of this display is that it shows a historical plot of what the target has done in space and time (the history is replayed each time the target is selected). This greatly reduces the short term memory requirements on the CO and TAO for interpreting the significance of the selected target. 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 to account for its actions. In addition, the profiles show own-ship weapon and target 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.

2.2 Response Manager

The response manager is located immediately below the track profile and is tied to it via a line indicating the target'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 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.

2.3 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 decision makers focus on: potential threat, non-threat, or unknown. The decision maker selects the hypothesis he 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 target 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 target 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 target 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 necessary to "buy into" the selected assessment. The basis for assessment module is expected to be helpful in avoiding "Blue-on-Blue" and "Blue-on-White" engagements.

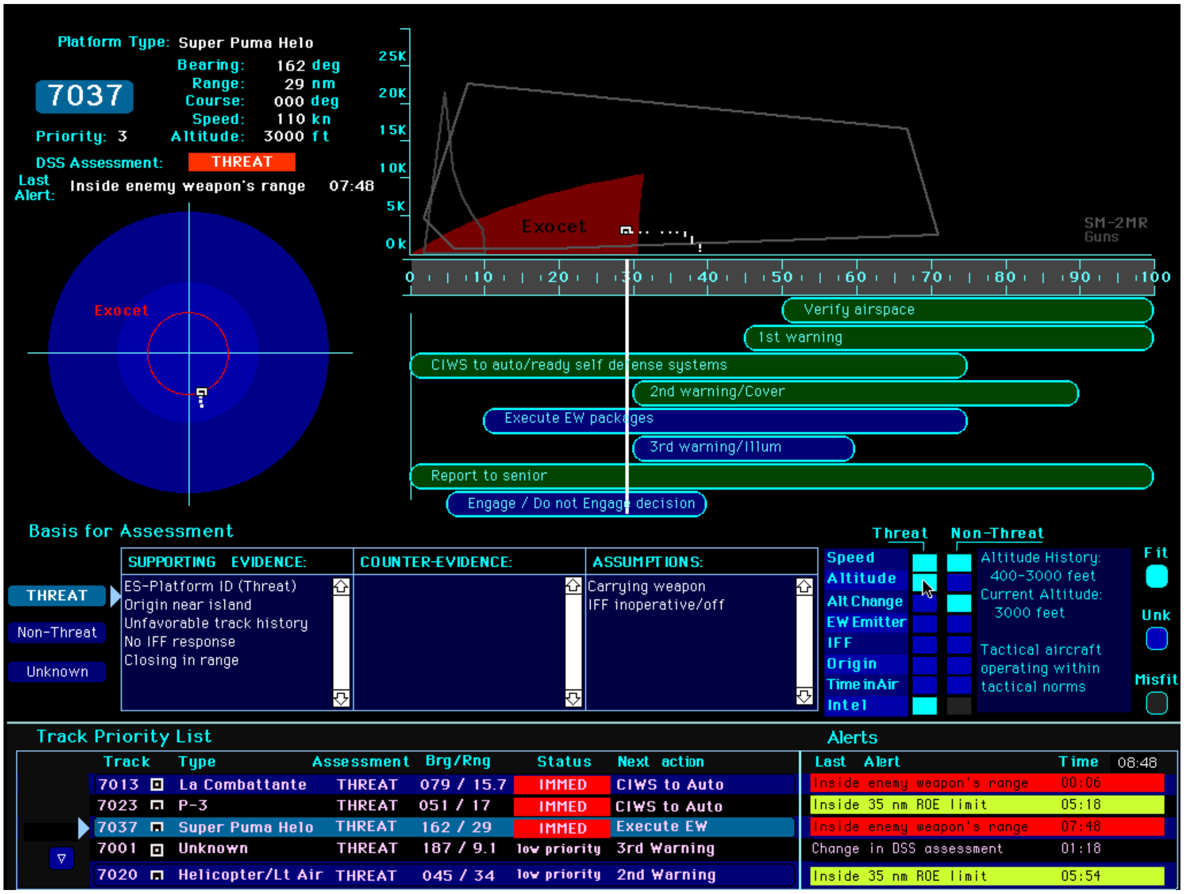

Figure 1. TADMUS DSS-1

3 DSS Evaluation Using Eye Movements

One of the purposes of this study was to explore technical and methodological issues associated with collecting eye movement data in an applied decision making laboratory setting. Various difficulties, such as electrical interference from nearby equipment, have complicated analysis of the data collected. Thus, the results reported here are preliminary and based on the data that have been reduced and verified so far. Subsequent reports will provide a more complete presentation of the results from this study. Nevertheless, the results presented here are illuminating and demonstrate the potential of eye movement data for display design and evaluation.

The study was conducted in two parts. The first entailed bringing qualified tactical decision makers into the NRaD Decision making Evaluation Facility for Tactical Teams (DEFTT) laboratory, and having them serve as the CO and TAO in air warfare (AW) simulations in peace-keeping missions. Subjects were then trained in the use of the DEFTT systems and the DSS. Following this, the eye tracking systems were donned and data were collected during the course of two tactical scenarios. This procedure was followed by a "directed questioning" procedure described below. Raw eye movement data were processed and reduced into fixations and dwell times in areas of interest. For a full description see the companion paper in these proceedings [Marshall, *et al.*, 1997].[1]

[1] A more thorough discusion of derivative eye movement measures is provided in [Harris, *et al.*, 1982], [Harris, *et al.*, 1986) and [Spady, *et al.*, 1982].

3.1 Overall DSS Usage

Eye tracking time histories were available for seven officers (4 CO, 3 TAO) each of whom performed two test scenarios. The two test scenarios involved similar types of tracks, tactical situations, and tactical decisions. They did, however, differ substantially in decision making load as determined by the density of tracks and by the number of concurrent, time-dependent tasks. The time histories indicated whether the officers were looking at the DSS screen at each moment during the scenario. Using these time histories, three important performance indices were computed: mean dwell time, mean fixation rate, and overall dwell percentage. A dwell consists of all visual fixations within a specific area of interest. Areas of interest in these analyses were defined as either the entire DSS display or each of the individual DSS modules, depending on the analysis.

3.1.1 Mean Dwell Time

Mean dwell time in this analysis reflects the average amount of time that officers spent looking at the DSS screen during any single eye fixation or gaze. Figure 2 shows the mean dwell time for the COs and TAOs in both the low and high load scenarios. Since the scenarios differed in total duration, dwell time has been normalized as average seconds per minute.

Figure 2 shows that COs tended to spend more time looking at the DSS on average than did TAOs whenever they looked at the DSS. Both officers tended to spend somewhat more time with the DSS in the high load scenario. The variability in mean dwell time across officers also increased for the high load scenario.

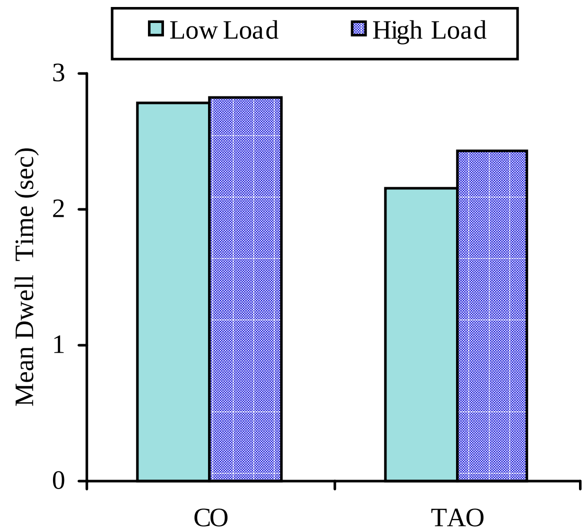

Figure 2. Scan Duration by Position and Scenario

3.1.1 Mean Dwell Time

While mean dwell time provides an indication of the amount of time required to obtain specific information about a track and therefore provides a sense of the rate in which information is extracted from the DSS, there is another important component to consider. The mean fixation rate is the average number of times that officers looked at the DSS. Mean dwell time and mean fixation rate are often considered to be compensatory. That is, a decision maker could, logically, acquire the same amount of information from a display either by taking one long look or by taking many brief looks to get the necessary information to solve a decision making problem. Thus, one would expect these measures to be negatively correlated, and our data confirm this inverse relationship (r = -.487).

Figure 3 presents the mean fixation rate for the COs and TAOs in both the low and high load scenarios. It can be seen that the mean fixation rate is approximately 5.5 per minute regardless of decision making position or scenario load. There are, however, substantial differences across officers' scanning patterns, and the differences between them increase markedly for the high load scenario.

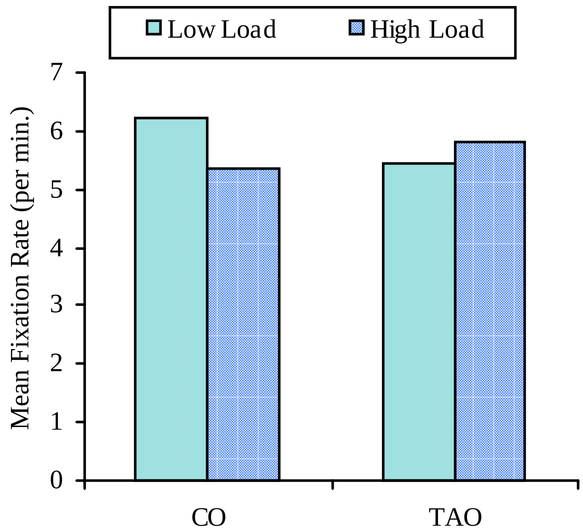

Figure 3. Fixation Rate by Position and Scenario

From Figures 2 and 3 we can see that officers tended to consult the DSS at a similar rate but that COs tended to spend more time looking at it per fixation. Also, officers spent more time looking at the DSS per fixation during the more demanding scenario.

3.1.3 Total Dwell Percentage

The total dwell percentage reflects the proportion of the overall scenario duration spent looking at the DSS. That is, this measure combines both the mean dwell time and the mean fixation rate in order to provide an overall index of time spent looking at the DSS.

Figure 4 shows the total dwell percentage for the COs and TAOs in both the low and high load scenarios. The most notable finding is the greater overall use of the DSS by the COs. As discussed in previous empirical studies with the DSS [Kelly, *et al.*, 1996], this finding is consistent with the differences in roles among these tactical decision makers. Typically, the CO is heavily involved in exploring patterns, forecasting likely events, and maintaining overall awareness of the tactical situation. These tasks are supported by the DSS. The TAO, on the other hand, is generally more heavily involved with the detailed management of multiple tracks. These tasks are supported by the DSS but rely heavily on a separate display, the geo-plot.

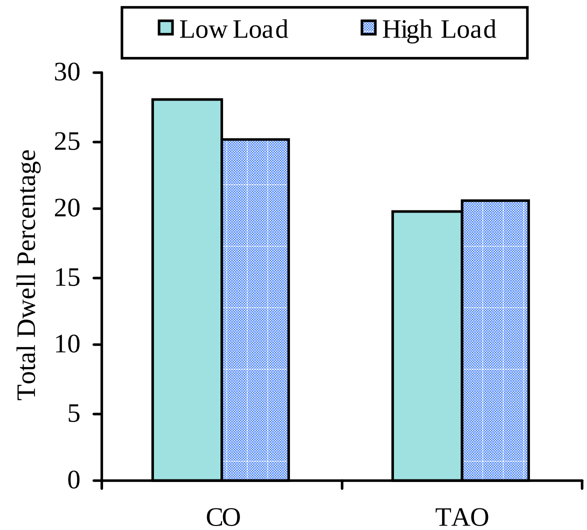

Figure 4. Percent of Total Time on DSS by Position and Scenario

Again, it should be noted that there are large individual differences in the total dwell percentage and that the variance increased for the high load scenario. This finding underlines the importance of further research and data analysis to explain why different decision makers exhibit such different scanning patterns, particularly as their task becomes more demanding. One promising approach might be to examine decision makers' eye movements around pre-defined critical events in the scenarios. In this way, we could determine how the DSS was used at particular times when events in the tactical situation demand the commanders' attention.

3.2 Usage of Individual DSS Modules

Limited eye movement data were available that permit us to examine which DSS modules were used while performing certain key tactical decision making tasks. To assess this issue, a "directed questioning" technique was developed to ensure we would know when decision makers were looking for specific information. Officers were presented with static "storyboard" DSS screens populated with scenario data and were asked to use them to answer tactical questions. These questions were selected to tap cognitive processes that are important for tactical decision making based on knowledge engineering performed in developing the DSS. They in fact represent distinct decision making tasks and will serve as a basis for understanding how scan sequences correlate to tactical decision making.

- Track Location – "Where is (Track X) relative to own ship?"
- Track Identification – "What is the identity of (Track X)?"
- Track Prioritization – "What priority would you assign to (Track X)?"
- Threat Assessment – "What has (Track X) been doing that suggests it is a threat/non-threat?"
- Critical Thinking – "Why might you be wrong in your assessment of (Track X)?"
- Response Management – "Which actions should be considered next with regard to (Track X)?"

Each of the six questions was asked five times. Specific track numbers were inserted into the questions to provide appropriate context for the storyboard, and the track numbers differed for each of the five replications. Prior to each question, officers were asked to fix their gaze on a dot on the display screen in order to verify calibration of the eye movement monitoring system. A question was then presented, followed by the DSS display. The officers then looked around the DSS as necessary to gather the information that they needed to answer the question. When they felt that they had determined the answer, they pressed a button which removed the DSS display and restored the screen with the reference dot. This procedure not only confined the data analyses to a single, specific decision making task, but also enabled calibration of eye tracking measures immediately before and after each trial.

Data from one officer (TAO) have been analyzed for the five replications of each of the six tactical decision making tasks. These were compared with independent predictions by subject matter experts about how the DSS modules would be used for the six decision tasks. The absolute deviation between the predicted and observed use of the DSS, compiled across modules, is shown in Table 1. It may be seen that for some tasks (e.g., track prioritization), there was quite good agreement, while for other tasks (e.g., critical thinking) predicted DSS usage patterns were not observed as consistently.

Table 1
Mean Absolute Deviation between Predicted and Observed Use of DSS Modules for Six Decision Making Tasks

Decision Making Task	Absolute Deviation
Track Location	15.8%
Track Identification	22.9%
Track Prioritization	10.3%
Threat Assessment	33.5%
Critical Thinking	31.7%
Response Management	17.5%

In order to interpret these differences, we will need to examine the predicted and observed use of individual modules. For discussion purposes, Table 2 presents these data for this subject in the track prioritization task. The concordance of the predictions and observations may be seen across all of the DSS modules. A marked division in the utility of DSS modules for this task can also be noted. The decision maker tended to rely on the upper-half of the DSS and on the Track Priority List. These findings suggest that the DSS layout was effective in supporting track prioritization decisions and that decision makers were quickly able to learn how to use the DSS for this task.

Table 2
Predicted and Observed Use of DSS Modules for Track Prioritization

DSS Module	Predicted	Observed
Track Summary	78%	100%
Track Profile (range x altitude)	89%	100%
Track Profile (range x bearing)	78%	80%
Response Manager	61%	60%
Threat Assessment	38%	20%
Comparison to Norms	31%	20%
Track Priority List	97%	80%
Alerts List	41%	40%

By contrast, improvements to the DSS may be needed to support other tactical decision making tasks, like critical thinking. Table 3 shows the predicted and observed usage of DSS modules for critical thinking. It was gratifying that the modules that were predicted to be most relevant for this task (i.e., Track Profile, Threat Assessment, and Comparison to Norms) were consulted frequently. It can also be seen, however, that nearly all of the DSS modules were consulted during the critical thinking task. In particular, the Response Manager, Track Priority List, and Alerts List were consulted much more often than predicted. Perhaps because critical thinking is a complex task requiring integration of data from several modules, the decision maker tended to scan the entire DSS for relevant information.

Table 3
Predicted and Observed Use of DSS Modules
for Critical Thinking

DSS Module	Predicted	Observed
Track Summary	57%	40%
Track Profile (range x altitude)	75%	80%
Track Profile (range x bearing)	64%	100%
Response Manager	36%	100%
Threat Assessment	92%	100%
Comparison to Norms	75%	60%
Track Priority List	33%	100%
Alerts List	39%	80%

This pattern of findings suggests that the DSS design may not support critical thinking as effectively as it could. Redesigning the display to concentrate information that supports critical thinking in particular modules and to allow decision makers to acquire this information by "quick-looks" should improve the DSS for critical thinking. This improvement, in turn, should be confirmed by subsequent eye movement data. Collaborative efforts are currently underway between NRaD, Cognitive Technologies Inc., and NAWC-TSD to make changes to the DSS that would promote more effective support for critical thinking tasks. It is hoped that further research with the eye movement system will allow this hypothesis to be explored empirically.

3.3 Scanning Sequences

Eye movements during the performance of the directed tasks also provided useful data on the decision maker's scanning behavior. This enables us to determine the order in which he consulted the DSS modules and to define regular scan patterns. These analyses could serve as the basis for mathematical models of scanning / information acquisition, which may be useful in realistic, dynamic tactical situations.

An eye movement transition matrix was computed for each of the six tactical decision making tasks. Each cell of the matrix represented the probability that decision makers would move from a particular DSS module (e.g., Track Summary) to another specific module (e.g., Alerts List).

For purposes of this paper, the DSS display was subdivided into three regions based on the intended uses in designing the various modules in the DSS. The first region involves the upper-half of the DSS (viz., Track Summary, Track Profile, and Response Manager). This region was designed primarily to support track identification and management, relying heavily on recognition-primed decision (RPD) processes. The second region includes the middle of the DSS (viz., Threat Assessment and Comparison to Norms). It supports explanation-based reasoning (EBR) and was designed to support threat assessment decisions. The third region includes modules designed to support multi-track management and uses the lower-part of the DSS (viz., Track Priority List and Alerts List). Generally, this region may be thought of as promoting overall situation awareness (SA).

Figure 5 shows these three abstracted DSS regions, spatially separated for ease of presentation. The connecting lines show the transition percentages observed for the track prioritization task. It can be seen that although the decision maker does seek information from all regions of the DSS, the majority of his scans concentrate in to upper, RPD region. By contrast, the scan pattern for the critical thinking task (see Figure 6) is more diffuse.

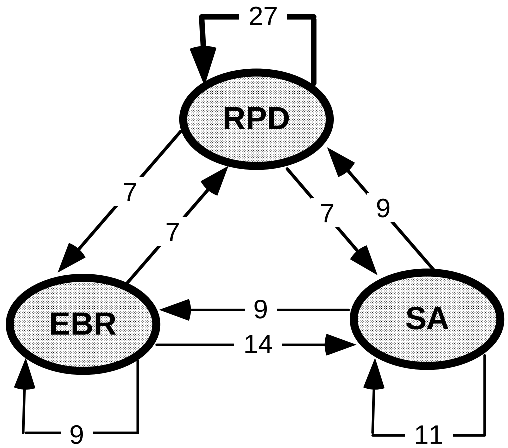

Figure 6. Observed Scanning Pattern for Critical Thinking

Note: Values are the scan transition percentages between recognition-primed decision (RPD), explanation-based reasoning (EBR), and situation awareness (SA) regions of the DSS display. Mean number of dwells was 8.8.

Analysis of scanning sequences can contribute not only to an evaluation of the DSS display design but also to our understanding of how information is acquired and used for decision making tasks. The fact that the DSS was designed on the basis of underlying theories of decision making may prove critical in developing a meaningful analysis of eye movement data. In this sense, these data can simultaneously support both applied work in display design and basic research in cognitive processes.

4 Future Directions

The above analyses reflect preliminary data and fairly basic analytic techniques. It is our intention to continue collecting data and developing additional metrics and methods associated with the use of eye movement data and decision support systems. We are very encouraged by the results of the scanning patterns and transition diagrams, and we are eager to pursue further analyses. Unfortunately, our present methodology requires that they be used only with the short presentation of static storyboards, rather than in the context of a more dynamic, realistic decision making task. If we are able to isolate a distinct set of scan patterns associated with recognized decision making tasks, it may be possible to use these patterns as templates, and build an automated algorithm for detecting them in the context of real-world decision tasks. Such algorithms could prove invaluable in understanding tactical deci-

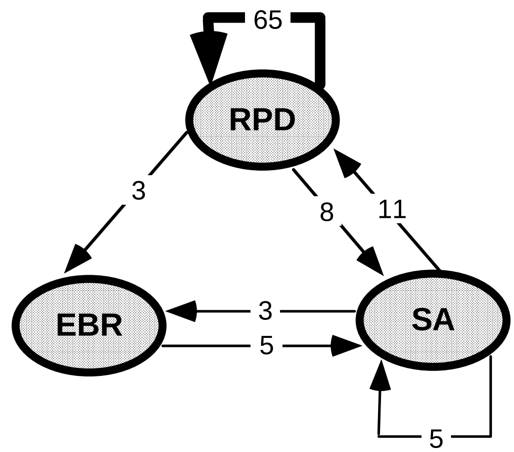

Figure 5. Observed Scanning Pattern for Track Prioritization

Note: Values are the scan transition percentages between recognition-primed decision (RPD), explanation-based reasoning (EBR), and situation awareness (SA) regions of the DSS display. Mean number of dwells was 7.0.

sion making and in developing more sophisticated models of it.

Another possibility we would like to pursue is the use of information metrics as a means for evaluating the usefulness of alternative information displays in achieving criterion performance on a task. This issue has been quite problematic because there is no way to know which portions of a given display will provide useful information for a particular decision maker at a particular moment in time. In effect, what constitutes information depends on what the decision maker knows and how he uses it. As there is no way to know this for certain, traditional information metrics have been difficult to implement in evaluating HCI displays. However, measuring the relative entropy, or randomness, in alternate displays for performing the same task in the same context would allow the designer to assess the relative goodness of the displays in terms of their efficiency in reducing (or increasing) randomness.

Acknowledgments

The authors gratefully acknowledge the continued support and guidance that has been provided by Gerald S. Malecki as well as the many technical contributions by Steve Francis, Brent Hardy, Connie O'Leary, Mike Quinn, and Will Rogers.

References

[Harris, *et al.*, 1982] Harris, R. L., Toll, J. R., Ephrath, A. R., & Stevens, A. T. How a new instrument affects pilots' mental workload, (pp. 1010-1013). In *Proceedings of the Human Factors Society 26th Annual Meeting*, Santa Monica, CA: Human Factors Society, 1982.

[Harris, *et al.*, 1986] Harris, R. L., Glover, B. J., & Spady, A. A. *Analytic Techniques of Pilot Scanning Behavior and Their Application.* (NASA Technical Report 2525) Hampton, VA: NASA-Langley Research Center, 1986.

[Hutchins, *et al.*, 1996a] Hutchins, S. G., Morrison, J. G., & Kelly, R. T. Principles for aiding complex military decision making, (pp. 186-203). In *Proceedings of the Second International Symposium on Command and Control Research and Technology*. Washington, DC: National Defense University, 1996.

[Hutchins, *et al.*, 1996b] Hutchins, S. G., Kelly, R. T., & Morrison, J. G. Decision support for tactical decision making under stress, (pp. 204-215). In *Proceedings of the Second International Symposium on Command and Control Research and Technology*. Washington, DC: National Defense University, 1996.

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

[Marshall, *et al.*, 1997] Marshall, S. P., Morrison, J. G., Allred, L. E., Gillikin, S., & McAllister, J. A. Eye tracking in tactical decision making situations: Implementation and analysis, (this volume). In *Proceedings of the Third International Symposium on Command and Control Research and Technology*. Washington, DC: National Defense University, 1997.

[Moore, *et al.*, 1996] Moore, R. A., Quinn, M. L., & Morrison, J. G. A tactical decision support system based on naturalistic cognitive processes, (p. 868). In *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society, 1996.

[Morrison, *et al.*, 1996] Morrison, J. G., Kelly, R. T., & Hutchins, S. G. Impact of naturalistic decision support on tactical situation awareness, (pp. 199-203). In *Proceedings of the Human*

Factors and Ergonomics Society 40th Annual Meeting. Santa Monica, CA: Human Factors and Ergonomics Society, 1996.

[Spady, *et al.*, 1982] Spady, A. A., Jones, D. H., Coates, G. D., & Kirby, R. H. The effectiveness of using real-time eye scanning information for pilot training, (pp. 1014-1017). In *Proceedings of the Human Factors Society 26th Annual Meeting.* Santa Monica, CA: Human Factors Society, 1982.