Return to Technical Session One

CAN_3

Improving the Usefulness and Acceptability of Automated Detection and Tracking Systems

Sharon McFadden, Melissa Lauz and Kirsten Stanger

Defence and Civil Institute of Environmental Medicine
1133 Sheppard Avenue West, Toronto, Ontario, Canada, M3M 3B9

Abstract

Automation is being introduced increasingly in an effort to improve the accuracy and timeliness of information processing within complex computer-based systems. One task that has received considerable attention in this respect is the detection and tracking of targets on sensor and tactical systems. However, initial experience with automated detection and tracking systems has been less than satisfactory. They are often seen as being unreliable and workload intensive. When they are perceived as being reliable, operators often fail to monitor them adequately. Most of the effort to improve the performance of these systems has focused on finding better algorithms. In contrast, the research presented in this paper has focused on understanding when and how operators make use of automated detection and tracking systems and the conditions under which they are likely to improve performance and reduce workload.

This paper summarizes the most pertinent results of our research examining human use of a simulated automated detection and tracking system as a function of reliability, workload, and experience. Our research shows that each of these factors impacts on system performance and on the perceived and actual usefulness of the automated system. Best performance tends to be associated with a moderately reliable tracking system with a low false detection rate while perceived reliability tends to be associated with low workload. Under low workload conditions, the operators' ability to detect targets improves, but their ability to detect automation induced errors does not. Suggestions for improving the usefulness of automated detection and tracking systems are offered.

Introduction

Improved sensors and processing capability have increased the amount of information that must be handled, while efforts to reduce manning have resulted in fewer operators to handle the available information. The only feasible solution to this increasing data overload is automation. Thus, we can expect automated systems to be an integral part of the operations room of the future and it is important to understand what they will and will not do and how best to employ them.

Experience with automated systems has often been less than satisfactory. In many cases, operators have under used them or rejected them entirely. This is especially true of automated detection and tracking systems. The response of system developers has been to try to improve the algorithms used by the automated systems. We have taken a somewhat different approach. We have tried to understand why operators might not be using these systems effectively and under what conditions they would.

An automated detection and tracking system (ADT) mimics an operator by scanning incoming signals to determine which ones are most likely to represent new targets (other ships submarines etc.) or the latest position of existing targets. The ADT detects targets by comparing the characteristics of the signal with a predefined template; it tracks the position of targets by comparing the location and strength of new signals with the location of existing targets. If the position of a signal is within a certain area surrounding an existing target, the automated tracker associates that signal with the target, updating its position. If the criterion used by the ADT is too stringent, real targets will be missed and tracks will be lost. On the other hand, the use of a lax criterion can result in false signals and signals from other targets being incorrectly associated with a target.

Given the variability in the strength, speed and paths of targets, it is probably inherently impossible to design a perfect automated detector and tracker. A more realistic goal is to design a useful automated system. A useful system may be defined as one that: improves overall system performance, reduces operator workload, and is used efficiently and effectively by the operator. Informal comments by operators of naval systems incorporating automated detection and tracking algorithms indicate that these goals are not currently being met despite extensive effort to develop improved algorithms for filtering the incoming data. As one operator recently put it, "the best thing about the automated detection and tracking system is the off button".

Unfortunately, there has been little research on human use of automated detection and tracking systems. However, research on other automated systems suggests reasons why operators may reject these systems. For example, Moray and his colleagues (Lee and Moray 1992, Muir and Moray, 1996) found that the use of an automated controller varied as a function of the operator's trust in it. Unless the automated controller was perceived as reliable, an operator did not use it to its full potential. Even when an aid is perceived as reliable, users may still use it suboptimally in order to maintain some feeling of control over the system (Morris, Rouse and Ward, 1988, Weisgerber and Savage, 1990). When users do rely on an automatic controller, they often fail to adequately monitor the automated system and miss errors (Parasuraman, Molloy & Singh 1993, Parasuraman, Molloy, Mouloua, & Hilburn 1996, Wiener 1985). These findings suggest that it may not be a simple task to develop an automated detection and tracking system that operators will use effectively and efficiently.

Previous research

Our initial studies looked at the use and usefulness of an Automated Tracker (AT) in a target detection and tracking task as a function of the perceived and actual reliability of the AT, task difficulty, and user experience (McFadden, Giesbrecht and Gula 1998, McFadden, Vimalachandran & Blackmore 1999). Providing only an automated tracker allowed the user the option of conducting the task manually or handing some or all of the targets over to the automated tracker. The results of these studies showed that people would make use of an AT if they perceived it as reducing their workload, even if it was not completely reliable. However, the extent to which the AT was used frequently depended on previous experience.

On average, participants who had extensive experience with a reliable AT tended to make less use of a moderately reliable AT (McFadden et al 1999). As well, participants who were able to carry out the task manually or with a moderately reliable AT tended to make less use of a reliable AT, at least under low workload conditions (McFadden et al 1998).

Without any kind of AT, most participants could track about 4 to 6 targets in the time available. With an AT which could track about 75% of the targets assigned to it, most participants were able to consistently monitor 90 to 100% of the strong targets. With the availability of a higher reliability AT, participants were more likely to detect and track weak targets as well. Moreover, workload, as measured by the time spent on interacting with the computer and perceived effort, declined substantially. However, percentage of targets tracked correctly did not improve beyond that found with the less reliable AT.

Ideally one would have expected performance to at least be equivalent to AT reliability if not better. Thus, if the AT could track 95 % of the detected targets, the percentage of targets tracked correctly should have been similar. At low levels of AT reliability, the increase in percentage of targets tracked came from a substantial reduction in lost targets. The AT relieved the participants of having to manually update each target. Freed from this time-consuming task, the participants had the time to search for weak targets. Thus, the percentage of missed targets also tended to decline as reliability increased, but primarily at higher levels of reliability. Unfortunately, as reliability increased a new type of error became more frequent – misassociations. Effectively, a decrease in missed targets was offset by an increase in misassociated targets.

Misassociations occurred in clearly defined situations where two targets passed close to one another. At that point, either the human or the AT could become confused as to which signal went with which target. If the targets were misassociated, the effect could often be seen in an unexpected deviation in the tracks of the misassociated targets on the display. A moderately reliable AT would sometimes fail to update the misassociated target. At that point, participants tended to re-add the target to the display instead of updating its position manually. This would get rid of the misassociation. A reliable AT was more consistent in updating targets. Thus, unless the participants noticed that the AT was associating the signals from one target with the marker originally assigned to a different target, the error continued until the target's path took it beyond the limit of the target display. Some participants were relatively good at handling misassociated targets; most were very poor.

Overall, our results suggest that providing an AT can both improve performance and reduce workload. Moreover, most participants will make use of the AT especially under high workload conditions. However, this use may be modified by previous experience. Thus, it would seem to be important to give users extensive experience with the automated system under the conditions in which it will be used. Finally, the benefits of an AT are not likely to be fully realized unless the percentage of automation induced errors such as misassociations is reduced.

Usefulness of an automated detection and tracking system

Having determined that even a moderately reliable AT would be used and could be effective in improving performance and reducing workload, the next step was to assess the effect of adding an automated detection function. In particular, would adding an automated detection function impact the percentage of misses and misassociations. In other tasks, automation induced errors have been reduced by keeping the operator more directly involved in the task. For example, Parasuraman et al (1993) found that automation induced errors were reduced when the reliability of the automated system was varied over time. Reducing AT reliability has somewhat the same effect. Some targets will continue to be tracked consistently while others will be lost from time to time. However, results with the AT showed that while this produce fewer misassociations, the number of missed targets and workload increased substantially. Thus, the overall benefit was nil. Reduced AT reliability may be partially compensated for by introducing an automated detection function. Strong targets that the AT fails to update will be re-added to the display. In addition, the detection function reduces the requirement to scan large quantities of data for new targets. On the other hand, adding a detection function changes the nature of the task. The operator no longer has the option of handling some targets manually. He or she becomes a passive monitor of the automated system whose primary role is to fix the errors introduced by it. Under such conditions, users may be even less effective monitors of the automated system.

To assess the potential of the detection function, performance and workload were assessed under two different detection criteria. With the first criterion, the automated detection function added only targets to the display. Participants received no assistance with detecting weak targets, but they did not have to deal with a large number of false alarms. With the second criterion, the automated function would potentially detect weak targets. However, it would also add non-targets. Detection threshold was investigated to determine if the number of missed targets could be reduce by increasing false alarm rate and the impact on workload of the lower threshold. As well, both a moderate and high reliability automated tracker were tested with the two levels of detection threshold to see if the addition of an automated detection function compensated for the increased workload and higher miss rate associated with a lower reliability AT. Targets that were lost by the automated tracker would, in many cases, be re-added to the display automatically thereby reducing the workload associated with coping with a moderately reliable tracker.

Automated Detection and Tracking Simulation System (ADTS)

The current experiment was carried out using the Automated Detection and Tracking Simulation (ADTS) experimental control software. The ADTS is a simulation of a target detection and tracking task for studying human use of an automated system. A schematic of the ADTS screen is shown in Figure 1. User tasks are performed through a series of tracking display (shown on the left of the figure), signal table (shown on the right), and function button (far right) selections. All selections are carried out via a mouse.

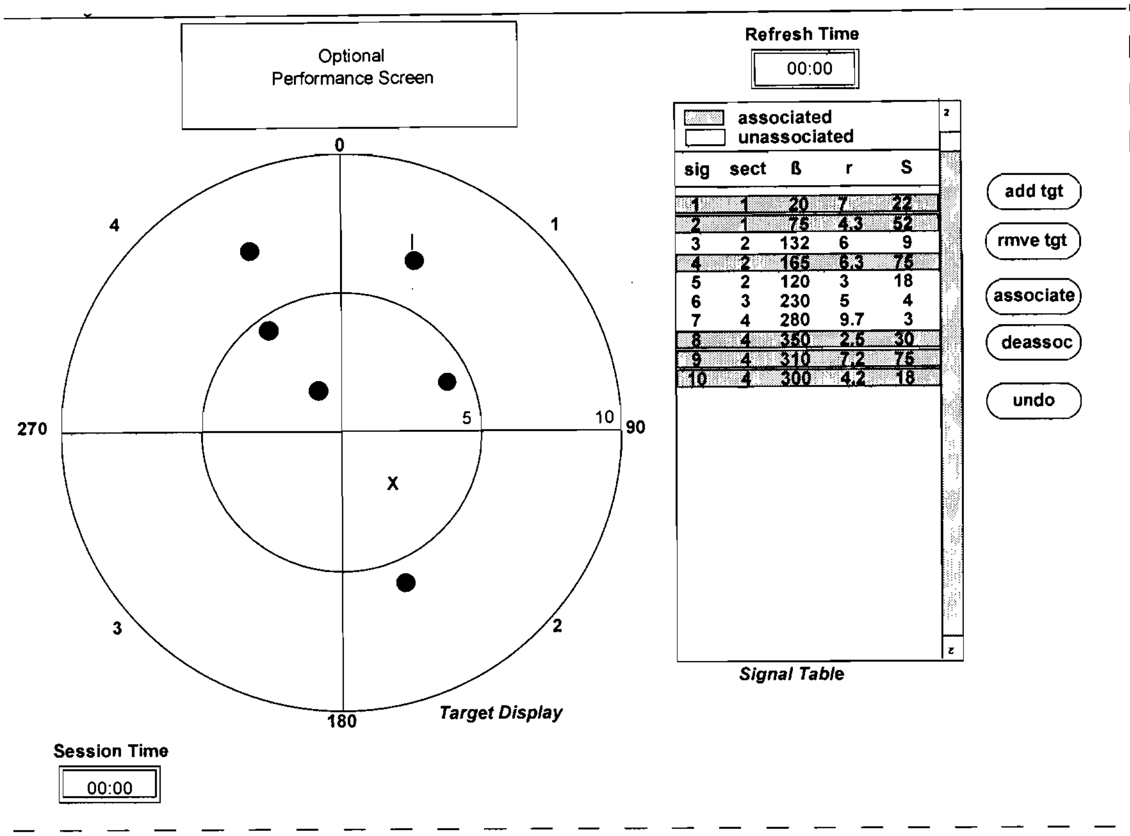

Figure 1: Schematic of the display for the Automated Detection and Tracking Simulation (ADTS) system. The "X" for the unassociated target marker appears as a white X in a black circle in the actual display. The dashed line in sector 1 of the tracking display shows the path traced out by that target. The solid line at the top of the same target shows the projected direction for that target. The numbers 1-4 are sector labels and the numbers 5 and 10 distance markers.

The ADTS is a modification of the ATS simulation used in our previous research. A detailed description of the Automated Tracking System (ATS) is available in McFadden et al (1998).

The task presented by the ADTS is to detect and then track the location of various targets (e.g. vessels) over time on the tracking display using information about the current location of a set of signals that is presented at regular intervals in the signal table. Some of the signals are due to the targets being tracked, some to targets that have not yet been detected, and the remainder are non-target signals. To add a new target, the participant selects the desired signal by positioning the cursor over it and clicking the mouse button and then selects the add function button in a similar way. To update the position of a target, the participant selects the desired signal and the existing target marker and then selects the associate button. Each time new information is presented, the associations between the target markers and the signals in the signal table are broken. Target markers that are not associated with a current signal have an X in them. The ADT act as preprocessors for the task. Each time new information is presented in the signal table, the ADT tries to update the position of all targets previously added to the tracking display as well as adding any new signals with a strength above a predefined threshold. The participant's task is to monitor the performance of the

ADT, adding missed targets, removing or updating targets that were not handled by the automated system (circles with an X), and correcting errors.

The set of target signals that the participant must track is called a scenario. Each scenario contains a selection of targets that follow either a straight line, zigzag, circular, or experimenter-defined path. In addition, several non-target signals are presented each time the participant receives new information about the position of the target signals. Non-target signals always have a strength of less than 20 and never last for more than one to two update periods. The strength of the target signals range from 0.1 to 100. At the beginning of a run, signals for new targets are added at a fixed rate until the maximum number of target signals for the condition under test has been reached. If a target's path takes it off the screen, the signal from a new target is added to the signal table.

Method

Four different combinations of detection threshold (10 and 20) and tracking reliability (approximately 87% and 98%) were run under two different levels of task difficulty (10 or 20 targets). With a detection threshold of 20, the ADT added only targets since all non-target signals had a strength of less than 20. With a threshold of 10, the ADT would also add weak targets and some non-target signals. Participants completed four scenarios under each combination of test conditions over a period of four days.

Prior to carrying out the test session, participants completed three training sessions. These introduced them to the task and gave them practice on simpler scenarios during which they had to detect and track either 6 or 8 targets. In the second and third training session, they had access to an ADT that tracked 98% of the detected targets and added all signals with a threshold of 20 or higher. New information was added to the signal table every 40 seconds.

Results and discussion

The percentages of targets detected and tracked under the various conditions are shown in Figure 2 along with the results under similar conditions from an earlier experiment (McFadden et al 1999) in which an automated detection function was not available (Threshold = 100). As can be seen, the percentage of targets tracked was very similar across all conditions. Adding an automated detection capability did not lead to improvements in performance over what had been achieved in previous experiments using just the automated tracker. Moreover, the same pattern and percentage of errors was found; misses declined and misassociations increased as tracker reliability increased.

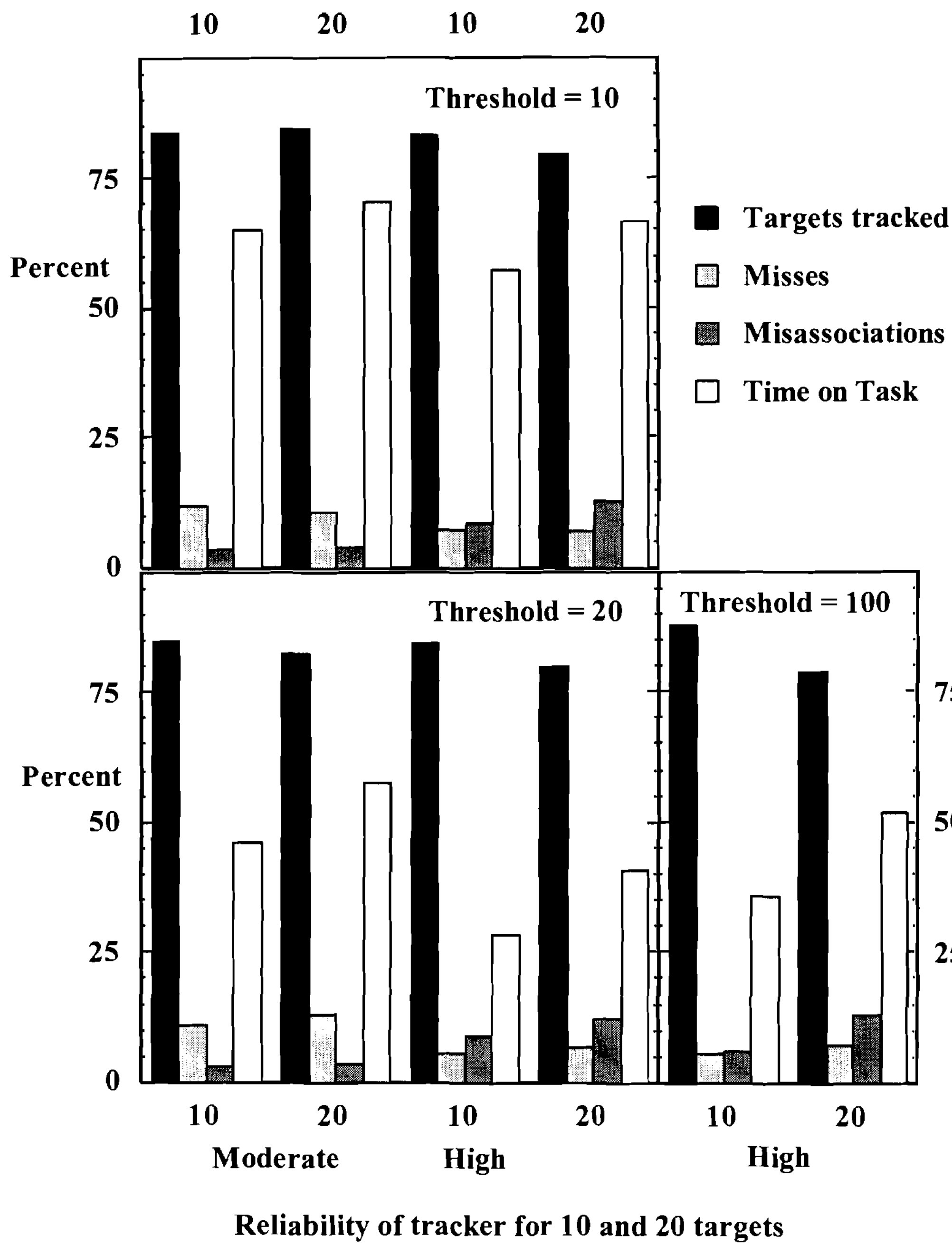

Figure 2: Percentage of targets tracked, misses, misassociations, and time taken as a function of the threshold of the automated detection function, the reliability of the tracker and the number of targets that the participant had to track. The results for threshold = 100 were collected in a previous experiment in which an automated detection function was not available.

The threshold for the auto detection function had little effect on targets tracked, but it did impact time taken or workload. It took approximately 20-25% longer to handle misses and errors when the threshold was set to 10 compared to when it was set to 20. Most of that extra time was taken up handling the larger number of false alarms in the lower threshold condition. With the threshold set to 20, time on task was about 10% lower than when an automated detection function was not available. However, in this study at least, that did not translate into either fewer missed targets or misassociations. One possible reason was insufficient experience under the different ADT conditions. In a previous experiment (McFadden et al 1999), it was found that extensive experience with a reliable AT was a necessary precursor to developing strategies for detecting weak targets. In the current experiment, only a quarter of the test runs involved the use of a reliable ADT with a detection criterion of 20.

Another possibility is that the reduction in misses with a highly reliable AT was not entirely due to the extra time available. Time on task was 10% lower under the high threshold, moderate AT condition than under the low threshold, high reliability condition, but the percentage of misses was lower in the latter condition. This suggests that the lower miss rate with the high reliability tracker was due to the fact that the AT tracked the weak targets and the participant did not have to keep reacquiring them.

Overall, it would appear that adding a detection function has little effect on percentage of targets tracked. Its primary advantage would be to further offload the operator by handling the detection of strong targets. This would be most advantageous in a multiple task environment or if there were specific task advantages to using a tracker with a stringent criterion. There is a distinct disadvantage, with which operators are all to familiar, to using an automated detection function to acquire weak targets.

Initially, it was thought that adding a detection function might impact the detection of misassociations. However, as Figure 2 indicates, this was not the case. The problem of automation induced errors is common to a wide range of automated systems and several different approaches have been suggested for reducing this type of error. With some tasks, these errors are lowered by extensive experience with the manual task (Wickens and Kessel 1979, Kessel and Wickens 1982, Endsley & Kiris 1995). Our research does not support this approach. Extensive training with the manual task had no effect on misassociations (McFadden et al 1999). Others have found that automation induced errors can be reduced by keeping the user involved in the task (Parasuraman et al 1993). This can be done by varying the reliability of the automated system over time or requiring manual intervention from time to time. Our results tend to support this approach in that misassociations are lower when a moderately reliable AT is used. The problem is to provide an AT that tracks weak targets without increasing misassociations. One possibility would be to employ a reliable tracker that alerted the operator when two tracks were in close proximity. However, providing alerts does not necessarily lead to improved performance. Often, it simply leads to the operator turning off the alert or increased workload. Thus, it would be important to evaluate the impact of such an alert on workload.

Conclusion

Considerable research has been carried out to develop automated detection and tracking systems that will improve the detection and tracking of targets on sensor and tactical systems. Our research in this area has concentrated on understanding why operators may not use these systems effectively and under what conditions they would. Based on the results of our experiments:

- Operators will make use of an automated tracker if they perceive it as reducing their workload and improving their capability to carry out the task, even if it is not completely reliable.

- An automated tracker can increase the number of targets an operator tracks and should reduce overall operator workload.
- However:
  - a moderately reliable AT does not tend to improve the operator's ability to handle weak targets,
  - with a highly reliable AT, misassociations impact negatively on overall performance.
- There appears to be little advantage to adding an automated detection function except possibly under high workload conditions with a large number of strong targets. Based on our results, an ADT should be used to detect and track strong targets in order to free up the operator to look for weaker targets. In order to facilitate this, the operator should be able to define the characteristics of the ADT and should be trained to use it for this purpose.
- Further research is required to determine a method for improving the user's ability to detect automation induced errors without increasing his or her workload substantially.

Recommendations

On the basis of the reported research, we recommend the following:

- The use of a moderately reliable automated tracker.
- Investigation of the usefulness of a detection function with user adjustable parameters so that it can be set to detect and track only strong targets.
- Investigation of alternate methods for alerting the user to the possibility of a misassociation.

References

Endsley, M. R., and Kiris, E. O. 1995, The out-of-the-loop performance problem and level of control in automation, *Human Factors,* **37**, 381-394.

Kessel, C. J., and Wickens, C. D. 1982, The transfer of failure-detection skills between monitoring and controlling dynamic systems, *Human Factors,* **24**, 49-60.

Lee, J. and Moray, N. 1992, Trust, control strategies and allocation of function in human-machine systems, *Ergonomics,* **35**, 1243-1270.

McFadden, S. M., Giesbrecht, B. L., and Gula, C. A. 1998, Use of an automatic tracker as a function of its reliability, *Ergonomics,* **41**, 512-536.

McFadden, S. M., Vimalachandran, A., and Blackmore, E. 1999, Factors Affecting Performance on a Target Monitoring Task in which Users had access to an Automatic Tracker, in preparation.

Morris, N. M., Rouse, W. B., and Ward, S. L. 1988, Studies of dynamic task allocation in an aerial search environment, *IEEE Transactions on Systems, Man, and Cybernetics,* **SMC-18**, 376-388.

Muir, B. M. and Moray, N. 1996, Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation, *Ergonomics,* **39**, 429-460.

Parasuraman, R., Molloy, R. and Singh, I. L. 1993, Performance consequences of automation-induced "complacency", *The International Journal of Aviation Psychology,* **3**, 1-23.

Parasuraman, R., Molloy, R., Mouloua, M., and Hilburn, B. 1996, Monitoring of automated systems, in R. Parasuraman and M. Mouloua (eds), *Automation and Human Performance: Theory and Application,* (Mahwah, NJ: Lawrence Erlbaum Associates), 91-115.

Weisgerber, S. A. and Savage, D. F. 1990, Operator ship classification using an Automatic Target Recognition (ATR) system in conjunction with Forward-Looking Infrared (FLIR) imagery, NWC TP 7101, Naval Weapons Center, China Lake, CA.

Wickens, C. D. and Kessel, C. 1979, The effects of participatory mode and task workload on the detection of dynamic system failures, *IEEE Transactions on Systems Man and Cybernetics,* **SMC-13,** 24-31.

Wiener, E. L. 1985, Cockpit automation: In need of a philosophy, In *Proceedings of the 1985 Aerospace Behavioral Engineering Technology Conference* (Warrendale, PA: Society of Automotive Engineers), 369-375.

Return to Technical Session One