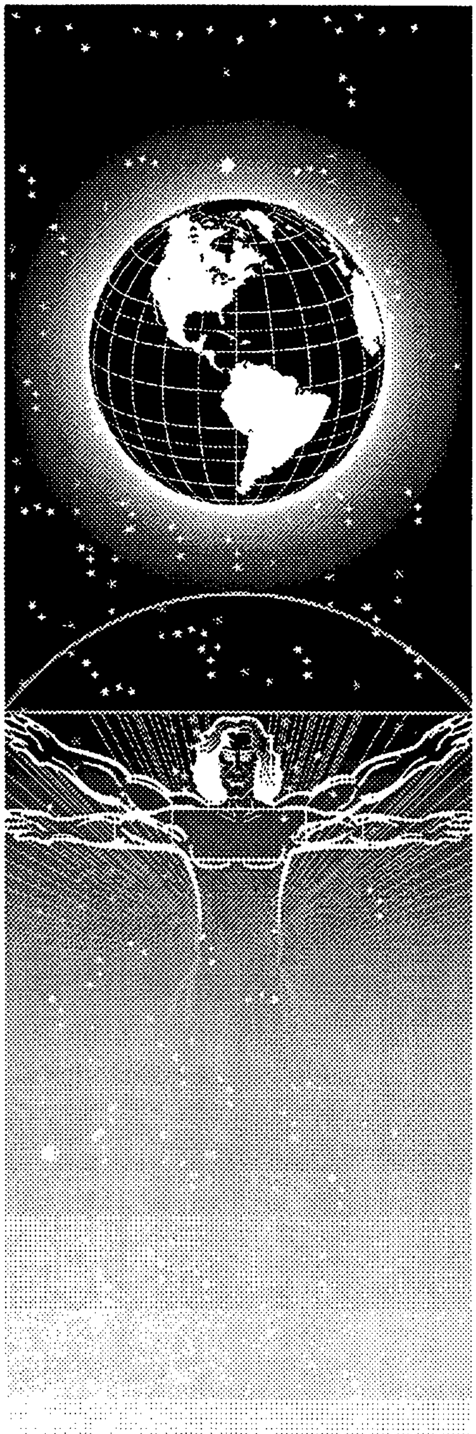
AL/CF-TR-1997-0171



UNITED STATES AIR FORCE ARMSTRONG LABORATORY

AUTOMATIC TARGET CUEING AND OPERATOR PERFORMANCE WITH ENHANCED APG-70 SYNTHETIC APERTURE RADAR IMAGERY

Judi E. See
Iris Davis

LOGICON TECHNICAL SERVICES, INC.
P.O. BOX 317258
DAYTON OH 45437-7258

Gilbert G. Kuperman

CREW SYSTEMS DIRECTORATE
HUMAN ENGINEERING DIVISION
WRIGHT-PATTERSON AFB OH 45433-7022

NOVEMBER 1997

INTERIM REPORT FOR THE PERIOD JUNE 1996 TO NOVEMBER 1997

Crew Systems Directorate
Human Engineering Division
2255 H Street
Wright-Patterson AFB OH 45433-7022

NOTICES

TECHNICAL REVIEW AND APPROVAL

AL/CF-TR-1997-0171

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

The voluntary informed consent of the subjects used in this research was obtained as required by Air Force Instruction 40-402.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

JOHN F. KENT, COL, USAF, BSC
Deputy Chief
Crew System Interface Division

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
November 1997

3. REPORT TYPE AND DATES COVERED
Interim, Jun 96 - Nov 97

4. TITLE AND SUBTITLE
Automatic Target Cueing and Operator Performance with Enhanced APG-70 Synthetic Aperture Radar Imagery

5. FUNDING NUMBERS
C: F41624-94-D-6000
PE: 62202F
PR: 7184
TA: 10
WU: 44

6. AUTHOR(S)
*Judi E. See; Iris Davis
Gilbert G. Kuperman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
*Logicon Technical Services, Inc.
P.O. Box 317258
Dayton, OH 45437-7258

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Armstrong Laboratory, Crew Systems Directorate
Human Engineering Division
Human Systems Center
Air Force Materiel Command
Wright-Patterson AFB OH 45433-7022

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AL/CF-TR-1997-0171

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The present study examined automatic target cueing (ATC) and target localization performance using 4 ft resolution synthetic aperture radar (SAR) imagery collected during the Theater Missile Defense Eagle Smart Sensor ATC (TESSA) program at Eglin Air Force Base, FL. Twelve observers viewed 360 SAR images that varied in terms of aiding (unaided versus aided) and the amount and type of background foliage ("open," "treeline," and "sparse" sites). All images contained the target to be detected, a transporter-erector-launcher (TEL), as well as two support vehicles. In the aided condition, each image was presented with four cue boxes to identify the ATC's target designations. The same set of 180 images was also presented in a separate unaided block without cues. After viewing each SAR patch map, participants identified the location of the TEL and rated their confidence in their decision. Overall, ATC cueing at the 4 ft resolution enhanced operators' confidence in their decision making but did not alter their localization accuracy, perceptual sensitivity (d'), or speed relative to the unaided condition. Further analysis revealed, however, that a critical determinant of operator performance and confidence was the "reliability" of the ATC. If all of the ATC's cues were false alarms, performance was worse than if no aiding had been provided at all. On the other hand, performance was most effective when the majority of the cues were centered over man-made vehicles. Finally, clutter in the form of background foliage degraded localization accuracy, sensitivity, reaction time, and operator confidence. As revealed by the present findings, when unaided performance is exacerbated by factors such as high background clutter, the assistance of a reliable ATC may be most beneficial.

14. SUBJECT TERMS
Aided Target Acquisition, Synthetic Aperture Radar (SAR), Theater Missile Defense (TMD), Operator Performance, Signal Detection Theory, Automatic Target Cueing.

15. NUMBER OF PAGES
53

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

This page intentionally left blank.

PREFACE

This effort was conducted by the Collaborative Systems Technology Branch of the Armstrong Laboratory (AL/CFHI), Wright-Patterson Air Force Base, Dayton, Ohio. The project was completed under Work Unit 71841044, "Crew-Centered Aiding for Advanced Reconnaissance, Surveillance, and Target Acquisition." Logicon Technical Services, Inc. (LTSI), Dayton, Ohio, provided support under contract F41624-94-D-6000, Delivery Order 0007. Mr. Donald Monk was the Contract Monitor.

The authors wish to acknowledge the Air Force Theater Missile Defense Attack Operations Program Office of the Aeronautical Systems Center (ASC/FBXT) and the Automatic Target Recognition Technology Branch, Wright Laboratory (WL/AARA), for providing the imagery used in this effort.

The authors also wish to acknowledge the extra efforts of their LTSI co-workers: Mr. Dave Robinow, for developing the software to control image presentation and data collection in the experiment, and Ms. Elisabeth Fitzhugh, for her thorough literature search and thoughtful input.

TABLE OF CONTENTS

TABLE OF CONTENTS, CONT'D.

LIST OF FIGURES

LIST OF TABLES

This page intentionally left blank.

INTRODUCTION

Background

Mobile ground targets such as mobile missile launchers, surface-to-air threats, and other vehicles that can relocate autonomously represent a continuing challenge for U. S. defense forces. As demonstrated by experience in the Persian Gulf War, where there was no evidence for the destruction of a single Iraqi mobile Scud launcher (Fulghum, 1994), detecting and attacking mobile ground targets can be difficult even in relatively open, barren terrain. While mobile missile launchers such as Scuds are unable to hide the launch event itself, they do have the capability to move within three to five minutes of a launch, exacerbating the task of re-acquisition. In an effort to ensure that we fully exploit this small post-launch window of opportunity, the Department of Defense has begun to emphasize the development of technologies that will improve the ability of U. S. armed forces to locate and identify mobile ground threats. Such developments will increase the efficiency of modern weapons, thereby reducing the risk to our armed forces while also reducing collateral damage.

One area currently receiving considerable attention is the development of Automatic Target Cuer/Recognition (ATC/ATR) technologies. Typically, the long term goals expressed for ATC/ATR systems have striven to develop completely automated systems. However, current system performance is limited by the inability of ATC/ATR algorithms to accommodate the natural unpredictability of background scenes in sensor data as well as the various orientations in which targets may be positioned within those backgrounds. The problem is further compounded by the fact that relaxing the algorithm's criteria can do more harm than good since it frequently results in unacceptably high false alarm rates for a completely automated system (Kuperman, Bryant, & Clark, 1991; Walters, 1993). Thus, it is expected that near-term systems will be human-in-the-loop systems designed to assist rather than replace the operator. ATC devices would aid crewmembers by filtering large volumes of imagery data and by cueing regions of interest in a scene that may represent targets, while ATR devices would provide information regarding target type, thereby enabling faster and more accurate targeting decisions.

As a step toward improving the ability of the Air Force to locate and identify mobile theater ballistic missile systems, the Wright Laboratories Avionics Directorate (WL/AA) is pursuing the use of ATC/ATRs for application with the F-15E's APG-70 radar and Low Altitude Navigation and Targeting Infrared for Night (LANTIRN) sensor systems. The APG-70 radar is a high frequency pulse-Doppler radar designed for air-to-air and air-to-ground (A/G) attack operations. For A/G operations, the radar provides various ground mapping modes to help the crew navigate, detect, and designate ground targets. The F-15E Synthetic Aperture Radar (SAR) in High Resolution Map (HRM) mode has the capability to produce patch maps at eight different coverage sizes, ranging from 80 nautical miles (nmi) to 0.67 nmi. The 0.67 nmi map corresponds to an image resolution of 8.5 ft/pixel, which is currently the highest level attainable on the F-15E (Nonnuclear Weapon Delivery Manual). In order to provide the type of higher resolution imagery needed to support the application of ATC/ATR technologies, the Theater Missile Defense (TMD) office (ASC/FBXT) sponsored the development and testing of a new high resolution capability with a "smart" sensor management system for use with the APG-70 radar. The APG-70 was modified to include a New High Resolution Mode (NHRM) with a patch size of 0.33 nmi, which corresponds to an image resolution of 4 ft x 6 ft (hereafter referred to simply as 4 ft). The smart sensor system will use the APG-70's SAR mode to search cued locations and locate likely targets using consecutive SAR patch maps of increasing resolution. The LANTIRN system will then be used to search for and detect the infrared signatures of possible targets within the area of the SAR maps using forward looking infrared (FLIR).

The flight demonstration of these capabilities took place at Eglin Air Force Base (AFB), Florida, from February through April 1995. The collection effort was designated as the *T*heater Missile Defense *E*agle *S*mart *S*ensor and *A*TC (TESSA) program. The main program objectives of the flight test included the collection of LANTIRN FLIR and APG-70 SAR data of mobile missile targets for use in the development of ATC/ATR algorithms and in TMD targeting simulations. An Air Force Development Test Center (AFDTC) F-15/188 sensor suite was configured to include the modified APG-70 air-to-ground radar with NHRM and a modified LANTIRN targeting pod.

During this phase of the TESSA program, missions were flown against three target sites at various times of the day and night. Weather conditions were generally good for all missions. Flight profiles were identical for each data collection mission and consisted of ten passes toward a target array that included a mobile missile launcher, a confuser vehicle, and a support vehicle. The flight profile was initiated at 40 nmi from the target array on the first pass and 20 nmi from the array on each subsequent pass. The approach angle to the target varied systematically with each pass to provide 180° coverage of the array at 22.5° intervals from a tail-on view to a head-on view. Both SAR and LANTIRN FLIR were collected for each pass. SAR data were collected from pass initiation to 10 nmi from the target array. A flight path was flown to produce a constant radar squint (off-nose) angle of 45°. Altitude for the radar portion of each pass was maintained at 17,000 ft. The APG-70 SAR data were recorded on high density data tapes via a Modular Airborne Recording System 2028 recorder. For the FLIR segment of the pass, the aircraft descended to 10,000 ft and data were collected from 10 nmi to target overflight (O'Byrne, 1995; Pryce, 1995).

The three target sites were selected to represent different degrees of clutter with respect to the capabilities of either the SAR or the infrared sensor. The three background conditions were referred to as "open," "treeline," and "sparse" clutter. The site designated as open was fairly level and consisted of approximately 50% low cut vegetation and grass, with the remainder being exposed sandy soil. This was selected to represent the lowest level of clutter with little chance of background foliage contributing to target confusion or obscuration. Background for the treeline site consisted of a well-defined treeline bordered by an unpaved road and an open field. Tree height ranged from 30 to 40 ft. Target vehicles were parked in the open on the road bordering the trees, and although the target vehicles were not obscured, the strong radar return that may be generated by natural clutter such as a treeline was expected to add to detection difficulty at this site. This site was selected to represent a medium level of clutter. Background for the sparse site consisted of low bushes and 10 to 20 ft high short-leaf pine trees randomly spaced from 5 to 50 ft apart. Again target vehicles were not placed so as to be obscured by the native vegetation. However, the radar returns and shadows generated by the trees and scrub represented the highest level of clutter of the three sites.

The target array consisted of three vehicles: a Scud-B mobile missile transporter-erector-launcher (TEL), a German MAN truck carrying a high pressure air compressor (HiPAC) unit, and a Zil-131 communications van. The Scud-B, which was the primary target for detection, was an authentic fully functional (except for the inert and unfueled stowed missile) specimen of a late 1960's Soviet TEL. The TEL had four drive axles, and its dimensions were 43 ft 6 in. long by 10 ft 8 in. wide by 11 ft 6 in. high. The MAN was intended as a confuser vehicle because it shared many cue features with the TEL, including size, number of axles, and engine location. Its dimensions were 29 ft 7 in. long by 8 ft 4 in. wide by 10 ft high. The Zil-131 vehicle (23 ft long by 8 ft 2 in. wide by 8 ft 2 in. high) was a three axle all wheel drive unit commonly used throughout the former Soviet block. It served as a vehicle that would be expected to accompany a TEL to an unprepared launch site. At each of the sites included in the TESSA program, placement of the unnetted vehicles varied from mission to mission; however, the vehicles were always facing true north and were generally spaced within 150 ft of one another.

Operator Performance with SAR Imagery

The present investigation involves use of the SAR imagery collected during the TESSA program. It was designed as a follow-on to a previous study that also used the TESSA SAR imagery (Davis, See, Shacklett, & Kuperman, 1996). That investigation indicated that operator target detection performance improved with the enhanced 4 ft (NHRM) resolution when compared to the 8.5 ft resolution. The focus of the current study was to evaluate the operator performance effects associated with the addition of target cueing to the 4 ft resolution images; therefore, cueing was the primary independent variable of interest. Since each site represented a different level of background clutter that might interfere with operator performance or modify the effectiveness of the cueing, the second variable of interest was site.

Cueing

A number of researchers have begun to examine the effects of automatic target cueing and automatic target recognition on human operator performance. In general, most studies support the notion that the addition of cueing to assist crewmembers engaged in locating or identifying targets does indeed enhance operator performance. For example, as part of the Radar Aided Mission/Aircrew Capability Exploration (RAM/ACE) program, Jauer, Quinn, Hockenberger, and Eggleston (1986) attempted to identify which types of SAR system enhancements would yield optimal performance. The conditions that were explored ranged from a manual low resolution SAR system to a higher resolution system incorporating a SAR autoscreener for automated target detection and classification. The probability of correct SAR target designation showed a clear improvement overall for conditions employing automatic SAR target screeners. However, the results further revealed that detection performance with the SAR autoscreener was dependent upon the number of false alarms associated with the automatic cueing. Specifically, detection performance with an autoscreener that produced three false alarms per image was no better than that in the manual condition where the operator performed without the assistance of cueing (P = 0.67, in each case). The probability of target acquisition improved significantly only when the number of false alarms was one ($P = 0.90$) or zero ($P = 0.93$) per frame. These two conditions differed from the manual condition, but not from each other. Thus, the cutoff between a beneficial autoscreener and one that provides no additional benefits above and beyond unaided performance appeared to be about three false alarms per image presentation. In fact, for optimal performance, the researchers recommended that system designers strive for an autoscreener that produces only one false alarm or fewer per screen.

In a more recent investigation, Becker, Hayes, and Gorman (1991) examined both ATR hits and false alarms in order to determine the optimal levels for user acceptance of the system. Their intent was to establish a cut-off in ATR performance below which the system would be viewed as having little tactical value. The ATR they examined used a symbolic overlay on top of infrared imagery to indicate the locations of possible targets. Observers viewed each ATR-processed image and then rated the ATR's apparent performance and tactical value. Following the ratings, participants were able to compare the ATR-processed image with an image showing the locations of actual targets in the scene. This comparison enabled them to determine the ATR's average level of hit and false alarm responses for the series of images. Four levels of hit rates (0.30, 0.50, 0.70, and 0.90) were combined factorially with four false alarm levels (1.5, 1.0, 0.5, and 0.167 false alarms/degree^2 of scene) to provide a total of 16 conditions in the experiment. These false alarm levels represented 2 to 8, 1 to 6, 0 to 3, and 0 to 1 false alarms per display, respectively. In all, observers viewed 16 sets of 30 displays, one set for each condition.

The results indicated that participants rated all 16 of the ATRs as having at least some degree of tactical value; however, none of the ATRs received a value rating above about 60 on a scale of 100. Judgments of tactical value increased as the ATR's hit rate increased and as the false alarm level decreased. The two systems that received tactical value ratings above 50 had hit rates of 0.70 and 0.90 and false alarm levels of 0.167 per square degree (0 to 1 false alarms per image). Thus, the more beneficial ATRs were those that achieved a fairly high hit rate while producing a minimal number of false alarms.

A second experiment was conducted in order to determine whether a false alarm level even lower than 0.167 per degree^2 of scene would be viewed as more beneficial. Three levels of hit rate (0.50, 0.70, and 0.90) were combined with three false alarm levels (1.0, 0.167, and 0.067 false alarms/degree^2 of scene) to provide nine conditions. As in the first experiment, judged tactical value increased with increases in the ATR's hit rate and decreases in the number of false alarms. However, contrary to what might be expected, the two lowest levels of false alarms (0.067 and 0.167) did not differ from each other. Overall, the results of the two experiments led the authors to conclude that an ATR system appears to have obvious tactical value only if it functions at a hit rate of 0.70 or better coupled with 0 to 1 false alarms per image or better. In addition, when the false alarms are already at an acceptably low level, improving the hit rate appears to be more advantageous than decreasing the false alarms even further.

These conclusions coincide with those reached by Fulkerson (1980) in an earlier study that examined target acquisition with SAR imagery. The targets to be detected included surface-to-air missile sites, depots, convoys, and armored battalions. During separate blocks of trials, observers were instructed to detect only one of the four types of targets. They completed the target acquisition task under three levels of autocueing as well as a no-autocueing baseline condition. In the autocueing conditions, the true target was always cued along with either one, four, or eight false alarms to provide high-, medium-, and low-power levels of autocueing. Significant improvement in target acquisition performance was achieved only with the highest level of autocueing (the level associated with 100% detections and only one false alarm per image). Observers responded more quickly and more accurately with the high-power cueing as compared to the no-autocueing baseline condition. In terms of both reaction time and accuracy, the high-power (one false alarm) and medium-power (four false alarms) autocueing conditions did not differ significantly. These outcomes led the authors to conclude that improvements in target acquisition performance are possible when the maximum number of false alarms is four but that no additional advantage is gained at lower false alarm levels. Further, in a post-experimental questionnaire, participants indicated not only that the presence of more than four false alarms was distracting but also that the occurrence of too few false alarms was disconcerting. They were most comfortable with the middle range of false alarms.

Weisgerber and Savage (1990) approached the study of ATR effectiveness by examining the effects of adding an ATR to FLIR imagery in a task requiring identification of seven ship classes. The percentage of correct identifications was used to index the operators' ability to discriminate among the seven target images. Ship images occurred at either close, medium, or distant ranges. The participants' task was to identify the ship in the image and also rate their confidence in their response. They completed 126 trials in an unaided condition and 126 trials in an aided condition where they received assistance from an ATR, which functioned at one of three levels of reliability (50%, 70%, or 90% accuracy). The ATR provided a recommended identification of the image, and the subjects were free to accept this identification if they believed it to be accurate or reject it and choose another that they judged to be more appropriate. Comparisons of aided and unaided performance revealed that the percentage of correct identifications was always greater in the aided condition, regardless of the quality of the ATR or the target distance. The greatest difference between aided and unaided performance occurred when the ATR functioned at a 90% level of reliability. The subjects not only performed more accurately in the aided condition but were also more confident of their responses when they received assistance from the ATR. Further, they were more confident when the ATR was highly reliable.

Kibbe and Weisgerber (1991) expanded upon the Weisgerber and Savage (1990) study by adding the variable of time constraint to the ship identification task. As in the earlier study, they examined the effects of range (near, medium, and distant) and ATR accuracy (50%, 70%, and 90%) on identification accuracy and on reaction time. Operators were given either 2, 4, 6, or 8 seconds to make their decision. Overall, operator accuracy was always significantly greater in aided conditions as compared to the unaided baseline condition. Accuracy improved progressively as the reliability of the ATR increased from 50% to 90%, although only the two highest levels differed significantly from unaided performance. Exposure time had no effect on operator accuracy. Further, while reaction time with the best ATR (90%) was significantly faster than reaction time with the worst ATR (50%), none of the aided conditions differed from the unaided condition. Use of the best ATR saved only about 300 milliseconds as compared to the unaided baseline condition.

In a similar study, Adams (1991) also compared ATR-aided human performance with unaided performance in order to determine the benefits of providing ATR assistance in a ship identification task. FLIR images of seven types of ships were presented at one of three distortion levels (low, medium, and high). The accuracy of the ATR was set at either 60% or 90%. Participants were required to identify the ship in each image and also rate their confidence in their decision. The results revealed that identification accuracy and decision-making confidence were both highest for ATR-aided performance versus unaided performance, regardless of the level of accuracy of the ATR or the level of distortion in the FLIR image. The greatest difference between the aided and unaided conditions occurred when the ATR was highly accurate *and* the imagery was highly distorted.

Entin, Entin, and Serfaty (1996) explored not only ATR accuracy but also the type of supporting information provided by the ATR. They presented simulated scenes composed of objects randomly selected from sets of target and nontarget objects. The objects appeared as black silhouettes against a white background scene, which was distorted in order to simulate the uncertainty and noise associated with realistic sensor systems. In the experiment, scenes comprised of seven to ten objects (on average half targets and half nontargets) were presented under both unaided and aided conditions. Participants were required to examine each image and detect as many targets as possible. When the ATR was available, it designated target objects in the scene with a black square. For each of these designated objects, participants could obtain additional information. The *Alpha* ATR presented the closest matching target. The *Beta* ATR provided the closest match and a confidence rating. The *Delta* ATR provided the three closest matches, rank ordered in terms of the degree of fit. The accuracy of each ATR was either high (90% hits and 10% false alarms) or low (90% hits and 40% false alarms). The results indicated that aided detection performance was always superior to unaided performance. Further, detection performance with the high accuracy ATR exceeded that with the low accuracy system. The effectiveness of the type of ATR depended on its accuracy. With the high accuracy ATR, the highest level of aided performance occurred with the *Delta* ATR, but this system yielded the lowest level of aided performance when the accuracy of the ATR was low. Following the experiment, most participants indicated they would prefer the *Beta* ATR (the one providing the closest match and a confidence rating) in high time pressure situations.

Following up on these outcomes, Entin and Entin (1997) conducted another study using the same apparatus. The ATR in this experiment duplicated the *Beta* ATR, providing the closest match and a confidence rating. As before, the accuracy of the ATR was either high (90% hits and 10% false alarms) or low (90% hits and 40% false alarms). The results indicated that aided performance accuracy was always superior to unaided performance; further, operators' hits were higher and their false alarms were lower with the highly reliable ATR as compared to the low accuracy system. However, this improvement in performance accuracy came at a cost to reaction time. Decision time was approximately 25% longer in both aided conditions as compared to the unaided baseline condition. This difference may simply be due to the extra time needed to elicit the ATR's supporting information, a step which would not be required under unaided viewing.

In summary, previous studies have indicated that operator performance accuracy may improve with the assistance of an ATC/ATR device, particularly if the device is reliable. Devices that achieve hit rates of 70% or better coupled with four or fewer false alarms per image appear to yield the highest levels of accuracy. Operator detection performance with ATRs that achieve hit rates of 50% or produce more than four false alarms per image tends not to differ from unaided manual performance. Differences in reaction time do not appear to be as consistent as those regarding performance accuracy. Reaction time may remain unchanged with the addition of automated aiding, it may increase, or it may decrease, depending on the task. Increases in reaction time may be attributable to the fact that the ATR presents more information for the operator to consider during decision-making.

Site

In addition to cueing, the effect of background clutter on performance accuracy was also examined in the present study. Toms and Kuperman (1991) describe background clutter as the busyness of a scene, either manmade or natural, in which a potential target may be embedded. High levels of background clutter may include geographical features such as a dense forest that can provide natural hiding places or the presence of large numbers of confusing objects such as decoys or other target-like objects. Low clutter would be characterized by terrain with low or minimal vegetation or the absence of confusing objects.

Previous studies investigating the effects of background clutter on the detection of relocatable targets defined clutter as the amount of vegetative coverage in a scene (Davis, See, Shacklett, & Kuperman, 1996; Kuperman, Wilson, & Perez, 1988; See & Kuperman, 1995). In the Kuperman et al. (1988) study, performance was lowest under the high clutter condition and improved as the background clutter decreased. See and Kuperman (1995) also found performance to be superior at the lowest level of clutter as compared to medium and high levels. Davis et al. (1996) found that both detection and localization performance degraded as the level of clutter increased. As described earlier, background clutter in the present study was defined by the three sites included in the TESSA program, which varied in terms of both the amount and type of clutter: the open site represented the lowest level of clutter; the treeline site represented a medium level; and the sparse site represented the highest level.

The Theory of Signal Detection

In order to examine the effects of cueing and site on performance accuracy in the current study, the techniques of the theory of signal detection (TSD) were applied. TSD is a model of perceptual processing that is frequently used to characterize performance effectiveness in target acquisition tasks (Gescheider, 1985; Green & Swets, 1966; Macmillan & Creelman, 1991; See & Kuperman, 1995; See, Riegler, Fitzhugh, & Kuperman, 1996; See, Warm, Dember, & Howe, 1997; Wilson, 1992). The application of TSD to a target detection task entails the derivation of two independent measures of performance: perceptual sensitivity (d') and response bias (c). The d' index of sensitivity is a perceptual measure that provides a bias-free estimate of the observer's ability to discriminate targets from nontargets. The index of response bias, c, provides an independent assessment of the operator's general willingness to make a detection ("target") response, which can vary on a continuum from conservative to lenient. Both measures are derived from observers' hits (correct detections) and false alarms (errors of commission) during the course of a task. A detection theory analysis is preferable to separate examinations of hits and false alarms because it permits performance to be characterized *independently* in terms of sensing abilities and decision making processes with measures that simultaneously take both the hits and false alarms into account, as reflected in the computing formulae for sensitivity and bias:

$$\mathbf{d'} = \mathbf{z}_{\mathbf{FA}} - \mathbf{z}_{\mathbf{H}} \qquad [1]$$

$$\mathbf{c} = .5\left(\mathbf{z}_{\mathbf{FA}} + \mathbf{z}_{\mathbf{H}}\right) \qquad [2]$$

In each formula, z represents the standard normal deviate associated with proportions of hits (H) and false alarms (FA), both of which enter directly into the derivation of each TSD index.

In many detection tasks in which TSD is applied, including target acquisition, observers may be required not only to detect the presence of a target but also to determine its location. Thus, once they have determined that a target is present, observers must decide which of several alternative "target-like" objects present in the scene has the greatest likelihood of being the target. The probability of correctly determining the target's location when it is present is derived from the joint probability of making both a correct detection of the target and a correct identification of its location. Similarly, in still other tasks, the objective may not be to determine whether or not a target is present but to decide where it is located, given that it is always present. Under these types of circumstances, it is still possible to apply TSD and obtain estimates of operator sensitivity, with some modification. The d' index of sensitivity for target localization can be interpreted as the operator's ability to differentiate the actual target from other alternative "target-like" objects that may be present. It is estimated, from either a computational formula or tables of d', on the basis of the number of alternatives available for designation and the operator's ensuing proportion of correct localization responses (Hacker & Ratcliff, 1979; Macmillan & Creelman, 1991). If desired, the index of bias can also be obtained in target localization tasks to provide a measure of the observer's degree of caution or conservatism in making the localization response. Its calculation is the same as that for target detection, with the proportion of correct localizations substituted for hits (Macmillan & Creelman, 1991).

The Present Study

The techniques of signal detection theory were applied in the present study in order to examine aided target localization performance using the 4 ft SAR imagery that had been collected as part of the TESSA program. Since the ultimate goal of the TESSA program is to produce data to facilitate the development of ATC/ATR algorithms, aided target acquisition performance was examined to evaluate the effectiveness of such a system in comparison to unaided performance. Thus, the primary purpose of the present study was to quantify the difference between aided and unaided operator performance within a single investigation using the same sample of human operators.

A second goal of this investigation was to use the performance accuracy results to provide inputs for a computer model of the target detection process known as ORION (Petersen, Fruchey, Rubin, & O'Rourke, 1995). This model was created to perform engagement effectiveness analyses on airborne systems during attack of relocatable, mobile, time critical, or imprecisely located targets. It supports the modeling of multiple, serial sensors such as SAR and FLIR and uses both correct detections and correct recognitions as well their concomitant false alarm rates. The primary performance inputs for ORION include estimates of perceptual sensitivity and associated distributions of false alarms. The present study was designed to provide baseline data regarding aided SAR target localization performance accuracy in a format that would also be suitable for ORION.

It was expected that performance accuracy in the present study would be higher with cued images. With regard to site, it was expected that performance accuracy would be highest at the open background site, lower at the treeline site, and lowest at the highly cluttered sparse site.

METHOD

Participants

Ten males and two females participated in the study. All participants were volunteers from various organizations at Wright-Patterson Air Force Base, OH. None of the individuals had prior operational experience with radar sensors, but nine had participated in previous studies employing SAR imagery in a target detection task. The remaining three were naive regarding the task of acquiring targets from SAR imagery, although they had varying experience with sensors and aircraft environments. Visual acuity was confirmed using both the Snellen visual acuity chart and the Vector Vision CSV 1000 contrast sensitivity test. All participants had normal or corrected-to-normal 20/20 vision. Their ages ranged from 23 to 48 years ($M = 33.6$, $SD = 7.2$).

Design

The basic design was a 2 (aiding) x 3 (site) within-subjects design. The two levels of aiding consisted of aided and unaided presentations of the imagery. In the aided condition, images contained cue boxes designating four regions of interest within the scene. In the unaided condition, the same imagery was presented without the cue boxes. The three levels of site included the open site (low clutter), the treeline site (medium clutter), and the sparse site (high clutter). Within each condition, the experimental trials consisted of images in which the three vehicles of the target array were always present. Consequently, the operator's task was not to detect whether or not the TEL was present but rather to determine its location in the scene.

Stimuli

The stimulus set consisted of Hughes modified APG-70 SAR imagery from the TESSA collection as previously described. Because the cueing algorithm was applied selectively to 4 ft images that contained the target array, only target imagery at the 4 ft resolution was available for use in the current study. The ATC system employed in this experiment was developed by Sverdrup Technologies, Incorporated, in support of a SAR study sponsored by the TMD program office and managed by Wright Labs. The ATC system has three main components: a constant-false-alarm-rate (CFAR) detector, a clustering routine, and a bright-region mean squared error

(MSE) classifier. These components are executed in sequence with the output from one component providing input for the next.

To begin, a wide area SAR image provides the input to the first component, the CFAR detector. The output from this process is a CFAR decision statistic image whose dimensions are the same as those of the original SAR image. Both images provide the input for the clustering routine, which extracts regions of interest (ROIs) from the original SAR image by applying a user-determined threshold to the CFAR statistic. Each ROI corresponds to a spatial cluster of pixels. These ROIs are then supplied to the third ATC component for classification (Dilsavor, 1995). However, output from the classifier was not available for this experiment because of a lack of training imagery for the algorithm. Rather, centroid pixel coordinates were obtained for the ROIs from the output of the clustering component, along with the CFAR decision statistic values for those coordinates. This information was available for 216 image files, and 87 of these were selected as the base stimulus set for the experiment. Image selection was aided by the use of information provided by Sverdrup Technology, Incorporated, and the ground truth report (Pryce, 1995).

Image files (480 pixels by 480 pixels) were received in a compressed format on 8 mm tape. CANTATA, a graphical programming environment for KHOROS hosted on a Silicon Graphics (SG) Indy system, was used to process each image file. All image files had to be processed due to an artifact of data collection/processing, which produced a large number of pixel intensity values at 0 with the remainder primarily distributed around 100. An image with this distribution would display an unusually large number of bright pixels. Consequently, in order to achieve a normal distribution for each image, histograms of the intensity of the pixels comprising the image were plotted and, based on the output, parameters for clipping the upper and lower extremes were selected. The resulting image was stretched to re-distribute the pixel values more evenly within the image range of 0 to 255 and enhance detail that would otherwise be lost. The processed images were inspected and additional fine adjustments were made to the clip parameters as needed.

Corner reflectors were removed from all scenes by replacing reflector pixels with background pixels in every target scene using the ColorIt 3.0 software hosted on a Macintosh Performa 636. To increase the number of "unique" scenes available for use in the study, eleven images were mirrored using the Flip function of KHOROS. An additional nine images, not included in the stimulus set, were selected as practice images and were processed using the same methods as described above.

Since the CFAR probability was applied across missions and not per image, the number of ROI reports per image varied considerably (from a minimum of 0 to a maximum of 51). Because previous studies have demonstrated operator intolerance for more than four false alarms (Becker, Hayes, & Gorman, 1991; Fulkerson, 1980; Jauer, Quinn, Hockenberger, & Eggleston, 1986), cues were presented for only four ROIs in each image. These ROIs were selected based on the magnitude of the CFAR decision statistic. Any ROIs associated with corner reflectors were automatically excluded from the selection process. Cue boxes (10 pixels by 10 pixels) were constructed and centered at the pixel coordinates provided by the algorithm for the four highest ROIs in each image. The cue boxes could be centered over any of the vehicles in the target array or over any feature of the background terrain. Specifically, the imagery was sufficient to support three different forms of cueing: (1) none of the vehicles in the target array cued, (2) the TEL and the MAN cued, or (3) the TEL, the MAN, and the Zil cued. Thus, the TEL target was cued in two thirds of the aided presentations. In one third of the imagery, all of the cues represented false alarms. Within each of the three background sites, 20 image trials were presented with each of the three types of cueing, providing a total of 180 aided images. These same images were also presented in a separate unaided block without the cues so that the differences between aided and unaided performance could be assessed within a single study. Thus, a total of 360 image trials was used in the experiment.

Apparatus

The research was conducted in the Crew-Aiding and Information Warfare Analysis Laboratory (CIWAL) located at Armstrong Laboratory, Wright-Patterson Air Force Base, OH. An SG Indy system with a high resolution 20-in. color monitor was used for stimulus presentation and data collection. To emulate aircraft radar displays as closely as possible, the intensities of the stimulus images were displayed entirely by using only the green gun of the monitor. The cues, when present, appeared as white boxes contrasting noticeably against the green background. Prior to the experiment, the brightness and contrast controls of the monitor were adjusted to clearly display each step of a 32 step gray-scale without overdriving the monitor. The gray-scale steps were verified prior to each data collection session. Luminance values were recorded using a TOPCON BM-7 luminance colorimeter. The BM-7 calibration was checked against a Hoffman LS-65B/HO standard luminance source and was found to be within the manufacturer's specifications (+/- 4%). Weekly measurements of the monitor displaying a 100% and 50% white field were taken. Participants were instructed not to adjust the brightness and contrast controls during data collection. A push button control pad and a computer mouse were used for response entry. Ambient lighting was provided by a desk lamp, which was positioned so as to minimize glare on the monitor.

Procedure

Each individual participated in one experimental session lasting approximately 1.5 hours, including the briefing and data collection. Upon arrival, participants were provided with a short description of the experimental procedures and were given time to read the consent form, ask questions, and sign the form if they wished to participate. They were then given a detailed briefing of the study objectives and imagery as well as more detailed instructions regarding task performance. The briefing also included specific information regarding the TEL, the MAN, and the Zil as well as their deployment. The TEL was identified as the target vehicle, and participants were instructed to locate the radar return they felt was most likely associated with the TEL. In the aided condition, they were not confined to selecting one of the ATC's cues but were free to choose any object in the image. Participants were also told that the best indicators for the presence of a target vehicle included the appearance of three bright (hard), relatively closely spaced returns in an image and that the most prominent distinguishing feature of the TEL at the 4 ft resolution was its length.

During the experimental session, participants were seated in a comfortable chair at a viewing distance approximately 30 in. from the screen. A practice session was conducted first to familiarize participants with the response apparatus and procedures and to provide clear examples of imagery representative of all sites and cueing conditions. Individuals were allowed to repeat the practice session until they were ready to start the experimental session. During both the practice and the actual data collection session, each trial began with the presentation of a *READY* prompt in the center of a black screen. Individuals began a stimulus presentation trial by pressing the *READY* button on the response keypad. For each image presentation, they were asked to indicate the location of the TEL. Target localization was accomplished by centering the mouse pointer over the radar return associated with the TEL and clicking the left mouse button to record the coordinates. After locating the target, participants were further prompted to rate their confidence in that decision by pressing one of six labeled buttons on the response panel (with 1 representing the lowest confidence and 6 representing the highest confidence). A diagram was provided above the control panel to remind them which button was associated with each of the six possible confidence rating responses. When the *READY* prompt reappeared on the display screen, participants initiated the next trial by again pressing the *READY* button on the response pad. They were instructed to respond as quickly and accurately as possible with neither speed nor accuracy receiving more emphasis. The sequence of events in each trial is depicted in Figure 1.

During the experimental session, the imagery was presented in two blocks: aided (180 images) and unaided (180 images). Within the aided block, the four cue boxes were visible upon image presentation; however, a button on the control panel allowed participants to turn the cues on and off as often as they wished. Target localization could also be made with the cue boxes on or off. The same images used in the aided block were presented in a separate unaided block without the cues. Block order was balanced across participants. Within each block, images were presented in a unique random order for each participant.

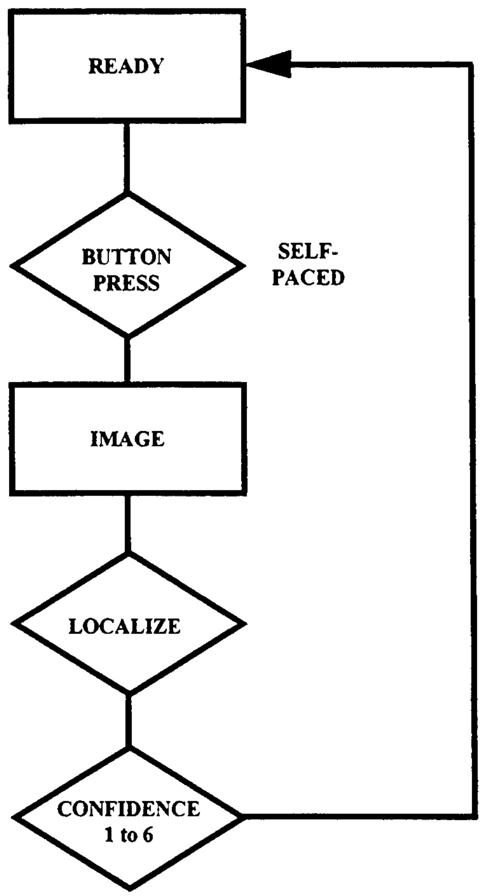


Figure 1. The sequence of events during each trial.

RESULTS

We examined performance effectiveness via four primary dependent variables: the percentage of correct localizations, perceptual sensitivity (*d'*), reaction time (RT) for correct localizations, and operators' confidence ratings. In determining whether an operator's localization response was correct or incorrect, an error tolerance of 10 pixels was used. This value was selected to represent the approximate length of the target at the 4 ft resolution. Thus, if the operator's localization point lay within 10 pixels of the center of the TEL target to be detected, it was considered correct. Percentages of correct localizations were then used to derive the *d'* index of perceptual sensitivity for each individual in the various experimental conditions by consulting the appropriate tables of *d'* for localization (Hacker & Ratcliff, 1979; Macmillan & Creelman, 1991). For localization tasks such as ours, the primary determinant of perceptual sensitivity, in conjunction with the percentage of correct localizations, is the number of alternative items that could be selected as the target. As in our previous study (Davis, See, Shacklett, & Kuperman, 1996), we operated under the assumption that the three vehicles in the target array represented the three alternatives that were available for possible selection as the TEL target. Before determining *d'*, percentages of 0 and 100 were first adjusted by means of the procedure recommended by Snodgrass and Corwin (1988) to permit the calculation of perceptual sensitivity when such values are encountered.

Data analysis proceeded in three phases. First, we examined the overall differences between the aided and unaided conditions. Second, within the aided condition, we sought to determine whether there were any performance effects due to the type of aiding that was provided. Although four cue boxes always appeared for every image in the aided condition, the cueing itself could take one of three different forms: (1) none of the vehicles in the target array cued, (2) the TEL and the MAN cued, or (3) the TEL, the MAN, and the Zil cued. Finally, the results of the second phase of the data analysis led us to conduct a global analysis designed to look at the differences among the four "conditions" of aiding *simultaneously*: Unaided, Aided/No vehicles cued; Aided/TEL and MAN cued; and Aided/TEL, MAN, and Zil cued.

Phase I: Aided vs. Unaided

Percentage of correct localizations

Mean percentages of correct localizations at each site for the unaided and aided conditions are presented in Table 1. The most salient feature of the table is the observation that the mean percentage of correct localizations was identical in the unaided and aided conditions. Thus, contrary to what might be expected when the operator is assisted by an automatic detection device that helps locate potential targets, performance accuracy did not improve relative to the situation in which the operator received no such assistance. The means in the table further indicate that the percentage of correct localizations was highest in the open site, moderate in the treeline site, and lowest in the sparse site. This was true for both the unaided and aided conditions.

Table 1

Mean Percentage of Correct Localizations (Standard Deviations in Parentheses) at Each Site for the Unaided and Aided Conditions

	Site			
	Open	Treeline	Sparse	Mean
Unaided	95	80	63	79
	(3)	(8)	(10)	(15)
Aided	94	83	62	79
	(2)	(8)	(11)	(15)
Mean	94	81	63	79
	(3)	(8)	(10)	(15)

A 2 (aiding) x 3 (site) repeated measures analysis of variance (ANOVA) was conducted to test the statistical significance of the means in Table 1. The alpha level for this and subsequent ANOVAs was set at .05. Probabilities for any effect containing three or more levels (e.g., site) were obtained via the Huynh-Feldt epsilon adjustment (Huynh & Feldt, 1970, 1976). The results of the ANOVA confirmed expectations gained by visual inspection of the means in Table 1. Namely, only the effect for site was statistically significant, $F(2,22) = 76.11, p < .0001$.

Post hoc correlated *t*-tests were conducted next to determine which sites differed significantly from one another. The overall alpha for the set of three comparisons was set at .10; thus, the alpha for each individual test was .033. The results of these comparisons indicated that the percentage of correct localizations differed significantly among all three sites.

<u>Perceptual sensitivity</u>

Mean values of perceptual sensitivity at each site for the unaided and aided conditions are presented in Table 2. As with the percentages of correct localizations, the perceptual sensitivity scores reveal that the provision of automated cueing did not enhance operator sensitivity to target localization. The overall mean *d'* was 1.8 in both the unaided and aided conditions, a value that is nearly identical to the mean localization sensitivity of 1.7 that was observed at the 4 ft resolution in our previous study (Davis, See, Shacklett, & Kuperman, 1996). The primary factor affecting operator sensitivity was background site. Perceptual sensitivity was highest in the open site and lowest in the sparse site. A 2 (aiding) x 3 (site) repeated measures analysis of variance (ANOVA) of the *d'* scores revealed only a significant main effect for background site, $F(2,22) = 91.41, p < .0001$. Post hoc correlated *t*-tests further indicated that operator sensitivity differed significantly among all three sites.

Table 2

Mean Perceptual Sensitivity (Standard Deviations in Parentheses) at Each Site for the Unaided and Aided Conditions

	Site			
	Open	Treeline	Sparse	Mean
Unaided	2.8	1.7	1.0	1.8
	(0.3)	(0.5)	(0.4)	(0.8)
Aided	2.6	1.8	1.0	1.8
	(0.3)	(0.5)	(0.4)	(0.8)
Mean	2.7	1.8	1.0	1.8
	(0.3)	(0.5)	(0.4)	(0.8)

RT for correct localizations

Mean RTs (in seconds) for correct localizations at each site in the unaided and aided conditions appear in Table 3. As with the percentage of correct localizations and perceptual sensitivity, the RT data indicate that the unaided and aided conditions did not differ. Thus, overall, operators responded neither more accurately nor more quickly when working with the assistance of the ATC to guide their search of the SAR map. However, as can be seen in Table 3, response time did increase progressively as the level of clutter in the background increased from the open site to the treeline and sparse sites. Further, the aiding did appear to stabilize the RT in the open and treeline sites more so than in the unaided condition, which exhibited a marked increase in RT from the open to the treeline site.

Table 3

Mean RT for Correct Localizations (Standard Deviations in Parentheses) at Each Site for the Unaided and Aided Conditions

	Site			
	Open	Treeline	Sparse	Mean
Unaided	3.2	4.8	8.4	5.4
	(1.4)	(3.6)	(5.0)	(4.2)
Aided	4.3	4.6	7.6	5.5
	(2.0)	(2.2)	(4.0)	(3.2)
Mean	3.7	4.7	8.0	5.5
	(1.8)	(2.9)	(4.4)	(3.7)

A 2 (aiding) x 3 (site) repeated measures ANOVA revealed a significant main effect for site, $F(2,22) = 19.94, p < .0002$. In addition, the Aiding x Site interaction was statistically significant, $F(2,22) = 4.82, p < .0300$. Post hoc correlated *t*-tests revealed that RT in the sparse site was significantly slower than in the open and treeline sites, which themselves did not differ. With respect to the interaction, which is portrayed graphically in Figure 2, the RT increased progressively with clutter in the unaided condition but did not increase significantly in the aided condition until the clutter had reached its highest level (i.e., the sparse site). Post hoc tests indicated that the magnitude of the increase in RT from the open to the treeline site was significantly greater in the unaided condition than in the aided condition. However, the extent of the difference in RT from the treeline to the sparse site did not differ by condition.

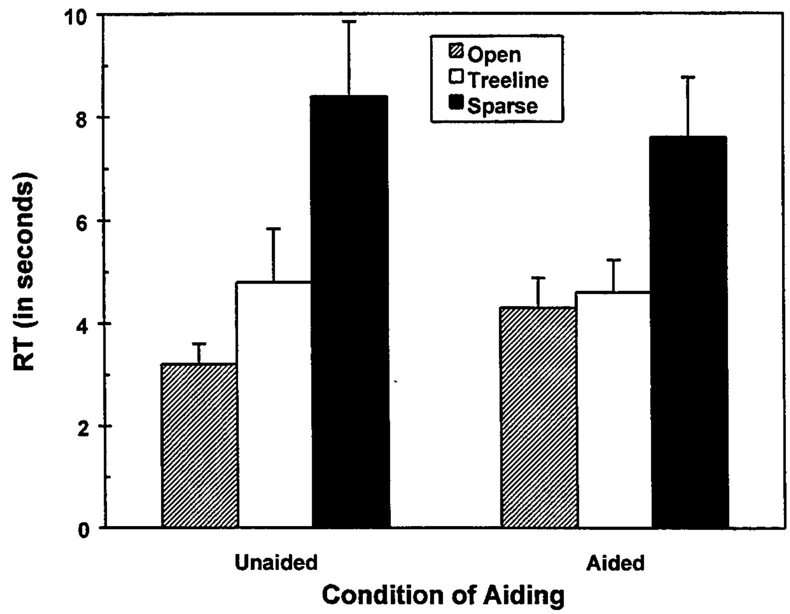


Figure 2. Mean RT in seconds at each site for the unaided and aided conditions.

Confidence rating

In addition to localizing the TEL in each image, observers were also asked to provide a rating of confidence in their response on an integer scale ranging from 1 (low confidence) to 6 (high confidence). Mean confidence ratings at each site in the unaided and aided conditions are tabulated in Table 4. In contrast to the three previously discussed dependent variables, which showed no differences by condition of aiding, the mean confidence rating was higher in the aided condition as compared to the unaided condition. In addition, confidence decreased progressively as clutter increased, though less sharply in the aided condition relative to the unaided.

A 2 (aiding) x 3 (site) repeated measures ANOVA revealed significant main effects for aiding, $F(1,11) = 9.16, p < .0115$, and for site, $F(2,22) = 36.47, p < .0001$. The Aiding x Site interaction was also statistically significant, $F(2,22) = 10.52, p < .0037$. Post hoc correlated *t*-tests indicated that the confidence ratings differed significantly among all three sites. As depicted in Figure 3, the interaction revealed that observers' confidence tended to drop off considerably from the open site to the treeline and sparse sites in the unaided condition. In the aided condition, observers remained more confident in their decision-making despite the growth in clutter. Post hoc testing further indicated that the drop in confidence from the open to the treeline site was significantly larger in the unaided condition than in the aided condition. However, the extent of the decline from the treeline to the sparse site did not differ by condition.

Table 4

Mean Confidence Rating (Standard Deviations in Parentheses) at Each Site for the Unaided and Aided Conditions

	Site			
	Open	Treeline	Sparse	Mean
Unaided	5.1	4.2	3.1	4.1
	(0.8)	(0.9)	(1.0)	(1.2)
Aided	5.0	4.5	3.8	4.4
	(0.7)	(0.7)	(0.6)	(0.8)
Mean	5.0	4.4	3.5	4.3
	(0.7)	(0.8)	(0.9)	(1.0)

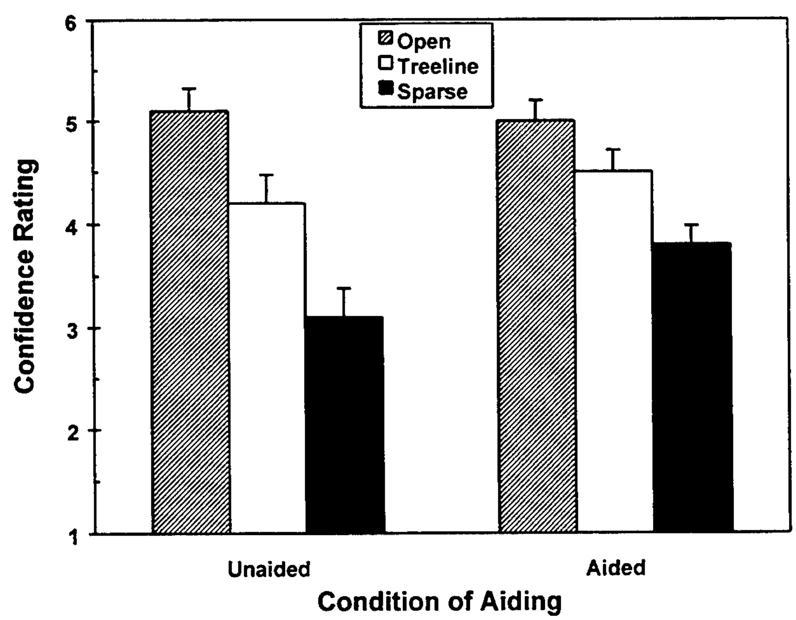


Figure 3. Mean confidence rating at each site in the unaided and aided conditions.

Summary of Phase I results

In summary, the results of the statistical analyses in Phase I revealed no benefit for the aided condition in terms of the percentage of correct localizations, perceptual sensitivity, and the RT for correct localizations. That is, observers were neither more accurate nor faster when they received assistance from the ATC as opposed to when they performed the target localization task on their own. However, the results did reveal a significant impact of aiding on observers' confidence in their responses. Thus, while they attained the same level of performance when assisted by the ATC as they did without it, observers did feel more confident that they had correctly localized the TEL.

Phase II: Type of Aiding

Although the results of the first phase of data analysis indicated no effects of aiding on performance speed and accuracy, the question still remained of whether there might be performance differences within the aided condition itself due to the pattern of aiding employed. Specifically, we surmised that there might be variations in performance depending upon whether (1) none of the four cue boxes was centered over any vehicle in the target array, (2) two of the four cue boxes were over the TEL and MAN, or (3) three of the four cue boxes were over the TEL, MAN, and Zil. In order to verify the validity of this supposition, 3 (type of aiding) x 3 (site) repeated measures ANOVAs were completed for the percentage of correct localizations, perceptual sensitivity, the RT for correct localizations, and the confidence ratings. For all four dependent variables, the main effects for type of aiding and site were statistically significant as was the two-way interaction ($p < .0003$ in all cases).

Inspection of the data indicated that the condition in which no vehicles were cued appeared to be the primary contributor to the main and interactive effects involving type of aiding. Specifically, the mean percentage of correct localizations and mean sensitivity were lower when no vehicles were cued than when two or more vehicles were cued. In addition, the RT was considerably slower and observers' confidence ratings were markedly lower when no vehicles were cued. Furthermore, a comparison of the means in the condition where no vehicles were cued with the means in the unaided condition from Phase I revealed that observers did *worse* in this condition of aiding than they did when they received no assistance whatsoever!

Phase III: Conditions of Aiding

The outcomes just described led us to conduct a third phase of data analysis that we had not originally planned to complete. In Phase III, we conducted separate analyses of the four dependent variables in which we examined the effects of each condition of aiding simultaneously. That is, condition of aiding encompassed the unaided condition as well as the three aided conditions (No Vehicles cued, TEL/MAN cued, and TEL/MAN/Zil cued). With site included, the resulting analysis for each dependent variable was a 4 (condition of aiding) x 3 (site) repeated measures ANOVA.

In essence, this type of analysis comprises an examination of the effects of ATC accuracy on operator effectiveness, with the unaided condition providing a baseline for comparison. In the condition where no vehicles were cued, four cue boxes were present but none captured the TEL target or any of the other vehicles that were present in the image. Hence, for these images, all cues represented false alarms, reducing the ATC's reliability to 0%. In the two conditions where the TEL was cued, one of the four cue boxes was placed correctly over the target, giving the ATC a hit rate of 100% (i.e., it always detected the target) with only three false alarms per image.

Percentage of correct localizations

The mean percentages of correct localizations at each site for the four conditions of aiding are presented in Table 5. First, the means in the table indicate that there were differences among the four conditions of aiding. Specifically, the worst performance occurred in the Aiding/No Vehicles cued condition, and the best performance occurred in the condition where all three vehicles were cued. Second, with respect to site, performance accuracy declined as the background clutter increased. However, the pattern of the decline appeared to differ within each condition of aiding.

The ANOVA of the means in Table 5 revealed significant main effects for condition, F $(3,33) = 19.53, p < .0002$, and for site, $F(2,22) = 72.54, p < .0001$. The Condition x Site interaction was also significant, $F(6,66) = 17.97, p < .0001$. For both condition and site, post hoc correlated t-tests were used to determine where the significant differences lay. The overall alpha for each set of tests was set at .10. The resulting alpha for each individual test was determined by the number of comparisons within the set. When six comparisons were made (for condition), the individual alpha was .017; for three comparisons (site), .033. The results of the post hoc analyses appear in Tables 6 and 7. As can be seen in Table 6, the performance degradation in the No Vehicles condition as compared to the unaided condition was significant. Further, the percentage of correct localizations in the TEL/MAN/Zil condition was significantly different from all other conditions. With respect to Table 7, the percentage of correct localizations differed significantly among all three sites.

Table 5

Mean Percentage of Correct Localizations (Standard Deviations in Parentheses)at Each Site for Four Conditions of Aiding

	Site			
	Open	Treeline	Sparse	Mean
Unaided	95	80	63	79
	(3)	(8)	(10)	(15)
Aided/No Vehicles Cued	100	61	51	71
	(0)	(19)	(29)	(29)
Aided/TEL and MAN Cued	95	91	53	80
	(7)	(5)	(6)	(20)
Aided/TEL, MAN, Zil Cued	86	96	82	88
	(3)	(6)	(8)	(8)
Mean	94	82	62	79
	(6)	(17)	(20)	(20)

Table 6

Results of Post Hoc Correlated t-Tests of Correct Localizations for Condition of Aiding

	No Vehicles	TEL/MAN	TEL/MAN/Zil
Unaided	*	--	*
No Vehicles		--	*
TEL/MAN			*

Table 7

Results of Post Hoc Correlated t-Tests of Correct Localizations for Site

	Treeline	Sparse
Open	*	*
Treeline		*

The two-way interaction between condition and site is depicted in Figure 4. As can be seen in the figure, the percentage of correct localizations tended to decline progressively as the level of background clutter increased in the unaided condition and in the Aided/No Vehicles cued condition. In particular, the percentage declined sharply from the open to the treeline site. The effect of site was less prominent in the remaining two conditions. Although the percentages of correct localizations did decline from the treeline to the sparse site in the TEL/MAN condition, they remained more or less stable in the condition where all three vehicles were cued.

In order to assess the statistical significance of the effects of site within each condition, post hoc correlated *t*-tests were conducted. Specifically, differences in the percentages of correct localizations among the three sites were compared separately within each condition. The overall alpha for each set of tests was set at .10; thus, the alpha for each individual comparison was .033. The results of these analyses can be seen in Figure 4. Within each condition, two sites labeled with the same letter did not differ significantly, and vice versa for the occurrence of different letters. Thus, in the unaided condition, the percentage of correct localizations differed among all three sites. In the No Vehicles cued condition, the percentage declined from the open to the treeline site but remained stable thereafter. When two vehicles were cued, the open and treeline sites did not differ; when all three vehicles were cued, the open and sparse sites did not differ.

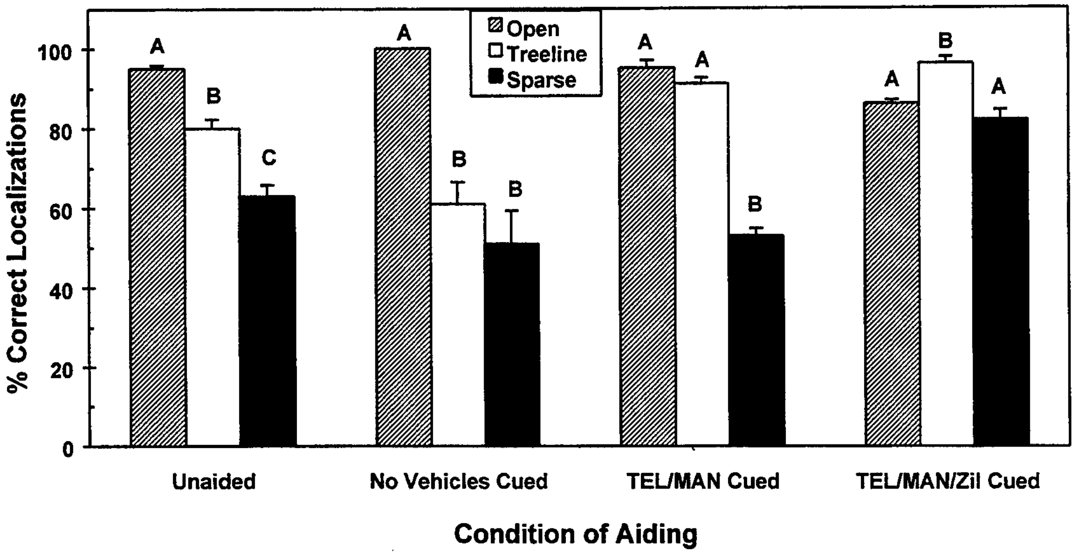


Figure 4. Mean percentage of correct localizations at each site for four conditions of aiding.

Perceptual sensitivity

Mean *d'* scores at each site for the four conditions of aiding appear in Table 8. With respect to the condition of aiding, the means in Table 8 indicate that operator sensitivity was lowest in the Aided/No Vehicles cued condition and highest in the Aided/TEL, MAN, Zil cued condition. With respect to site, mean perceptual sensitivity decreased progressively as the level of background clutter increased. However, the nature of the decline appeared to differ depending on condition. Sensitivity decreased considerably when no vehicles were cued and remained most stable when all three vehicles were cued.

The ANOVA of the *d'* scores revealed significant main effects for condition of aiding, $F(3,33) = 12.93, p < .0002$, and for site, $F(2,22) = 101.73, p < .0001$. The Condition x Site interaction was also significant, $F(6,66) = 23.02, p < .0001$. Post hoc tests for condition indicated that the TEL/MAN/Zil condition differed significantly from each of the three remaining conditions, which themselves did not differ from one another. Post hoc tests for site indicated that the sensitivity differed significantly among all three sites.

Table 8

Mean Perceptual Sensitivity (Standard Deviations in Parentheses) at Each Site for Four Conditions of Aiding

	Site			
	Open	Treeline	Sparse	Mean
Unaided	2.8	1.7	1.0	1.8
	(0.3)	(0.5)	(0.4)	(0.8)
Aided/No Vehicles Cued	3.2	1.0	0.6	1.6
	(0.0)	(0.8)	(1.3)	(1.4)
Aided/TEL and MAN Cued	2.8	2.4	0.7	1.9
	(0.6)	(0.4)	(0.2)	(1.0)
Aided/TEL, MAN, Zil Cued	2.0	2.8	1.8	2.2
	(0.2)	(0.5)	(0.4)	(0.6)
Mean	2.7	2.0	1.0	1.9
	(0.6)	(0.9)	(0.8)	(1.0)

The nature of the Condition x Site interaction is portrayed graphically in Figure 5. As can be seen in the figure, the decline in sensitivity as background clutter increased was most pronounced in the unaided and No Vehicles cued conditions. In the TEL/MAN cued condition, sensitivity declined noticeably only in the most highly cluttered background site. In the TEL/MAN/Zil condition, sensitivity was paradoxically highest in the treeline site. As in the case of correct localizations, the statistical significance of the effects of site within each condition was assessed by means of post hoc correlated *t*-tests, with an overall alpha of .10 for each set of comparisons. The results of these analyses can be seen in Figure 5.

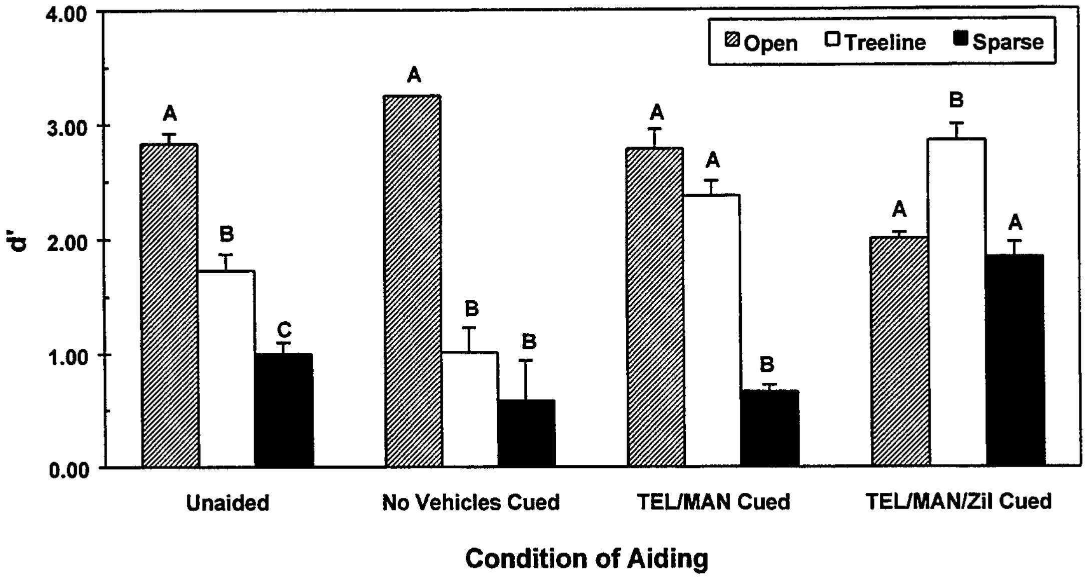


Figure 5. Mean perceptual sensitivity at each site for four conditions of aiding.

RT for correct localizations

The mean RT scores at each site for the four conditions of aiding appear in Table 9. With respect to the condition of aiding, the means in Table 9 indicate that observers were the slowest in the Aided/No Vehicles cued condition and fastest in the Aided/TEL, MAN, Zil cued condition. With respect to site, the mean RT increased progressively as the level of background clutter increased. However, the nature of the increase appeared to differ depending on condition. RT increased most precipitously when no vehicles were cued and remained most stable when all three vehicles were cued.

The ANOVA of the RT scores revealed significant main effects for condition of aiding, $F\,(3,30) = 12.88, p < .0001$, and for site, $F\,(2,20) = 23.17, p < .0001$. The Condition x Site interaction was also significant, $F\,(6,60) = 19.96, p < .0001$. Post hoc tests for condition indicated that the No Vehicles condition differed significantly from each of the three remaining conditions, which themselves did not differ from one another. Post hoc tests for site indicated that the RT was significantly slower in the sparse site than in the open and treeline sites; however, RT in the latter two sites did not differ.

Table 9

Mean RT for Correct Localizations (Standard Deviations in Parentheses) at Each Site for Four Conditions of Aiding

	Site			
	Open	Treeline	Sparse	Mean
Unaided	3.2	4.8	8.6	5.5
	(1.4)	(3.8)	(5.2)	(4.3)
Aided/No Vehicles Cued	4.2	6.2	14.5	8.3
	(2.6)	(3.2)	(7.3)	(6.5)
Aided/TEL and MAN Cued	5.1	4.1	5.7	5.0
	(1.9)	(1.7)	(3.0)	(2.3)
Aided/TEL, MAN, Zil Cued	4.0	4.4	4.4	4.2
	(2.0)	(2.1)	(1.9)	(1.9)
Mean	4.1	4.9	8.3	5.8
	(2.0)	(2.8)	(6.1)	(4.4)

The nature of the Condition x Site interaction is portrayed graphically in Figure 6. As can be seen in the figure, RT showed progressive increases as the background clutter increased, but only for the unaided and No Vehicles cued conditions. When at least two vehicles were cued, the RT remained more or less stable. The figure further indicates that the sharpest increase in RT with background clutter occurred in the No Vehicles condition, where observers required an average of 14.5 seconds to make a localization response in the highly cluttered sparse site.

As in the case of correct localizations and *d'*, the statistical significance of the effects of site within each condition was assessed by means of post hoc correlated *t*-tests, with an overall alpha of .10 for each set of comparisons. The results of these analyses can be seen in Figure 6. As expected on the basis of visual inspection of the figure, RT differed by site only in the unaided and No Vehicles cued conditions. The RT increased significantly from the treeline to the sparse site in both cases. In the two remaining conditions where at least two vehicles were cued, RT was fairly low and did not differ by site.

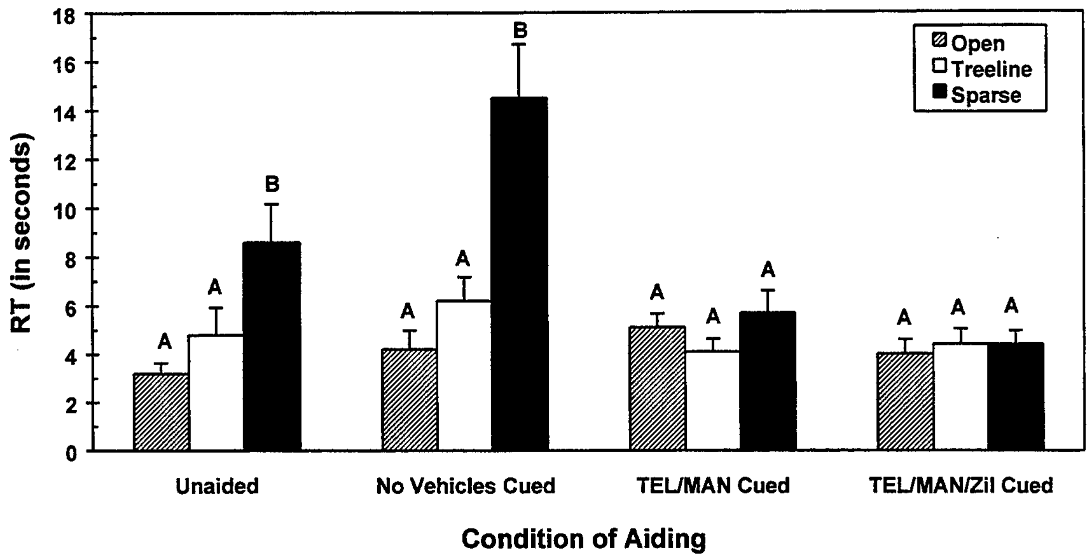


Figure 6. Mean RT in seconds at each site for four conditions of aiding.

Confidence rating

Mean confidence ratings at each site for the four conditions of aiding are presented in Table 10. As can be seen in the table, there appeared to be a disparity in confidence between the unaided and No Vehicles conditions on the one hand and the TEL/MAN and TEL/MAN/Zil conditions on the other. Observers reported having more confidence in their responses when at least two of the vehicles in the target array were cued. Further, as expected, confidence decreased as the background became more and more cluttered. This effect was most prominent in the No Vehicles condition and least prominent in the TEL/MAN/Zil condition.

The ANOVA of the confidence ratings indicated significant main effects for condition of aiding, $F(3,33) = 36.58, p < .0001$, and for site, $F(2,22) = 36.91, p < .0001$. The Condition x Site interaction was also significant, $F(6,66) = 36.28, p < .0001$. The results of post hoc tests for condition and site appear in Tables 11 and 12, respectively. As expected on the basis of the means in Table 10, the unaided condition did not differ from the No Vehicles condition, and the TEL/MAN condition did not differ from TEL/MAN/Zil; however, all remaining comparisons were significantly different. With respect to site, the confidence ratings for all three sites were significantly different.

Table 10

Mean Confidence Rating (Standard Deviations in Parentheses)at Each Site for Four Conditions of Aiding

	Site			
	Open	Treeline	Sparse	Mean
Unaided	5.1	4.2	3.1	4.1
	(0.8)	(0.9)	(1.0)	(1.2)
Aided/No Vehicles Cued	5.1	4.0	2.5	3.9
	(0.8)	(0.8)	(0.6)	(1.3)
Aided/TEL and MAN Cued	4.9	4.9	4.3	4.7
	(0.6)	(0.7)	(0.7)	(0.7)
Aided/TEL, MAN, Zil Cued	5.0	4.6	4.6	4.7
	(0.7)	(0.8)	(0.7)	(0.7)
Mean	5.0	4.4	3.6	4.4
	(0.7)	(0.8)	(1.1)	(1.1)

Table 11

Results of Post Hoc Correlated t-Tests of Confidence Ratings for Condition of Aiding

	No Vehicles	TEL/MAN	TEL/MAN/Zil
Unaided	--	*	*
No Vehicles		*	*
TEL/MAN			--

Table 12

Results of Post Hoc Correlated t-Tests of Confidence Ratings for Site

	Treeline	Sparse
Open	*	*
Treeline		*

The interaction between condition and site is portrayed in Figure 7. As shown in the figure, observers' confidence ratings gradually declined from the open site to the treeline and sparse sites in the unaided and No Vehicles conditions. However, in the two conditions where the vehicles were cued, observers' confidence remained relatively stable despite degradations in background clutter. The results of post hoc correlated *t*-tests, which are also portrayed in Figure 7, indicated that observers' confidence ratings did decline significantly from site to site when no aiding was provided and when aiding was available but no vehicles were cued.

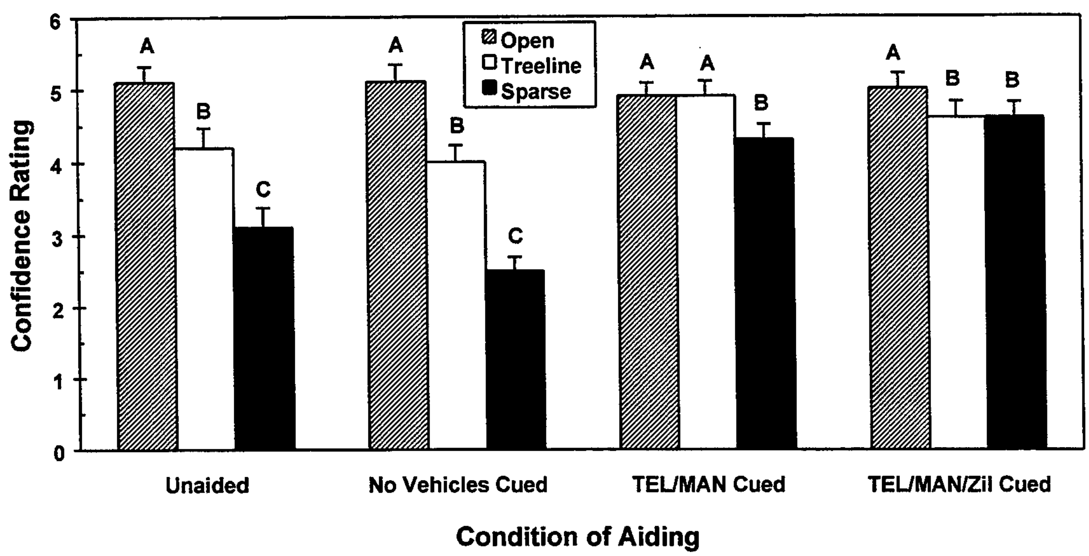


Figure 7. Mean confidence rating at each site for four conditions of aiding.

<u>Summary of Phase III results</u>

In essence, the outcomes of Phase III analyses revealed that not all aiding is created equal. Localizations, *d'*, RT, and confidence all differed significantly depending upon the type of aiding that was provided. When cue boxes were present but no vehicles were cued, both localizations and RT were significantly *worse* than if no aiding had been present at all. Conversely, when all three vehicles in the target array were cued, observers' localization responses and perceptual sensitivity as well as their confidence in their decision-making were significantly better than in the unaided condition and in the Aided/No Vehicles cued condition.

DISCUSSION

The present study was designed as a follow-up to a previous study (Davis, See, Shacklett, & Kuperman, 1996) in which unaided target detection and localization performance were examined via the same set of TESSA imagery used here. In the first study, we demonstrated that the 4 ft resolution yields an improvement in performance accuracy over the 8.5 ft resolution, the highest level that is currently available on the APG-70 radar of the F-15E. The primary goal of the current investigation was to determine whether performance at the 4 ft resolution could be further enhanced by the assistance of an ATC whose cues identify potential target locations. The results of this investigation can be discussed in terms of three critical outcomes regarding the performance effects of aiding, the type of aiding that is provided, and the impact of background site.

Performance Effects of Aiding

Our global analysis of the effects of aiding was very clear in indicating that aiding in and of itself did not enhance performance accuracy, sensitivity, or speed. However, aiding did enhance observers' confidence. Hence, whereas the same level of performance was achieved in both the unaided and aided conditions, higher decision-making confidence was observed when operators were assisted by the ATC. The ATC cues did not enable them to perform any more accurately, but it gave them a boost in confidence, as if the ATC provided confirmation in their decisions (e.g., "I think the TEL is here and so did the ATC, so this must be it" versus "I think the TEL is here, but I'm not sure"). The observation that cueing was able to enhance confidence is not trivial since operator confidence plays a large role in overall performance. All things being equal, operators who exhibit confidence in their work are preferred over those who can achieve at the same level but have no faith in their abilities. Along these lines, our results further indicated that the aiding tended to keep confidence elevated across variations in background clutter, which can also be beneficial to overall performance. Operators must be able to maintain confidence in their decision-making even under difficult circumstances such as those characterized by attempting to locate a target in a highly cluttered background.

The absence of a global effect of aiding on localization accuracy, sensitivity, and reaction time contradicts our hypothesis that aiding would enhance performance accuracy or speed. On the surface, it also appears to contradict many findings in the literature which indicate that aided performance is more effective than unaided (Adams, 1991; Entin & Entin, 1997; Entin, Entin, & Serfaty, 1996; Fulkerson, 1980; Jauer, Quinn, Hockenberger, & Eggleston, 1986; Kibbe & Weisgerber, 1991; Weisgerber & Savage, 1990). As reviewed in the Introduction, however, these studies also demonstrated that aided performance is most effective primarily when the cuer achieves a hit rate of 70% or better coupled with four or fewer false alarms per image. Overall, the ATC's hit rate for the imagery presented in the current study was only 67%, with three to four false alarms per image. Thus, the ATC's hit rate may have been too low to produce a performance advantage over unaided performance. This supposition is further corroborated by subsequent analyses which indicated that differences between aided and unaided performance became apparent only when the aided condition was further subdivided according to the specific type of aiding that was provided.

Type of Aiding

This analysis of the type of aiding that was provided represents a crude assessment of the impact of ATC reliability. Here we broke the aided condition down into its various types and compared those to the unaided condition. The three types of aiding were as follows: (1) none of the vehicles in the target array were cued by the ATC, (2) the TEL and the MAN were cued by the ATC, and (3) all three vehicles in the target array were cued by the ATC. This type of analysis provides a rough evaluation of ATC reliability in that all of the cues for the first type of aiding were false alarms; hence, the ATC's reliability would be 0%. For the remaining two types, at least one of the cues captured the target vehicle, giving it a 100% hit rate with 3 false alarms. Thus, one would expect performance to be better when the "reliability" of the ATC is higher.

The results of this phase of our analyses indicated that the type of aiding does indeed have a significant impact on all aspects of performance: localization, perceptual sensitivity, RT, and confidence. There were two salient findings here. First was the observation that operators were actually worse off when aiding was provided but no vehicles were cued relative to the situation in which they received no assistance at all. Second, aiding enhanced performance most effectively when all three vehicles in the target array were cued. In general, performance in this condition was better than in all other conditions, aided or unaided. The advice stemming from these outcomes is relatively straightforward: if you are going to use an ATC, use one that is reliable; otherwise, performance may be worse than it would be under normal unaided circumstances.

Three additional outcomes of importance in this phase of our analyses relate to the effectiveness of the third type of aiding where all three vehicles in the target array were cued. First was the observation that this type of aiding enhanced performance the most in the more highly cluttered treeline and sparse sites. Thus, whereas performance in the remaining unaided and aided conditions deteriorated as clutter increased, particularly in the sparse site, it remained more stable in cases where all three vehicles were cued. This outcome suggests that reliable aiding may be most useful in situations where unaided performance is marginal. Its effectiveness may be minimal when unaided performance is already at an acceptable level, in which case the ATC may not be needed at all. These results concur with those obtained by Adams (1991), who also found that the greatest performance advantage occurred when an accurate ATR was used to identify objects in highly distorted imagery.

Second, in terms of correct localizations and perceptual sensitivity, performance was significantly better when all three vehicles were cued than when only the TEL and MAN were cued. In both conditions, the ATC correctly located the target, giving it a hit rate of 100% with 3 false alarms. The difference between the two conditions therefore lies within the nature of the false alarms. When all three vehicles were cued, only one of the false alarms was centered over background clutter. When only two vehicles were cued, the remaining two cues were clutter false alarms. The fact that this difference between the two conditions impacted performance effectiveness suggests that operators might be sensitive to the ATC's tendency to cue clutter versus vehicles. Operators may have more confidence in ATCs that cue vehicles as opposed to those that cue clutter. In the former case, the ATC may be seen as more capable of differentiating between vehicles and clutter and may therefore be viewed as somewhat more reliable.

Finally, although previous studies have indicated that aided performance accuracy invariably exceeds unaided target acquisition performance, they have not been as clear regarding the effects of aiding on reaction time. In particular, some studies have shown that aided response times may be faster than unaided RT (Fulkerson, 1980); others have indicated that RT may be slower with aiding (Entin & Entin, 1997); and still others have reported no difference between aided and unaided response times (Kibbe & Weisgerber, 1991). In the current investigation, there was no improvement in RT in the condition which exhibited the most improvement in localization accuracy and sensitivity (i.e., the Aided/TEL, MAN, and Zil cued condition). However, operators were significantly slower in the Aided/No Vehicles cued condition than in the unaided condition. Thus, these results suggest that whereas response time may not be any faster with reliable devices that benefit target acquisition performance itself, it may be slowed considerably with an unreliable ATC that generates a low hit rate and four or more false alarms per image.

Background Site

The results of all analyses that we conducted consistently indicated that background site was a significant determinant of performance effectiveness. Localization accuracy, sensitivity, and confidence declined and RT increased as the background became more highly cluttered. Performance was most effective in the open site and worst in the sparse background. The mean values of d' in the sparse site, which ranged from 0.6 to 1.8, indicated that the task of target localization was exceedingly difficult when the background was highly cluttered. In fact, sensitivity for target localization reached its highest level only when aiding was provided and only when all three vehicles were cued. In general, operators were also less confident in their decision making when background clutter was high. Further, on average, it took them twice as long to locate a target in the sparse site as compared to the open site. These outcomes signify not only that the effects of background site must be taken into consideration when assessing performance but also that the assistance of an accurate ATC may be most beneficial when the background is highly cluttered.

CONCLUSIONS

1. ATC cueing enhances operators' confidence in their decision making.

2. The type of aiding provided by the ATC is a critical determinant of operator localization accuracy, perceptual sensitivity, RT, and confidence. If all cues are false alarms, performance is worse than if no aiding had been provided at all. Performance is most effective when the majority of the cues are centered over man-made vehicles.

3. Clutter in the form of background foliage degrades localization accuracy, sensitivity, RT, and operator confidence.

4. The assistance of an accurate ATC may be most beneficial when unaided performance is exacerbated by such factors as high background clutter.

REFERENCES

Adams, S. R. (1991). *Presentation format as a determinant of targeting task performance* (Report No. NWC TP 7125). China Lake, CA: Naval Weapons Center.

Becker, C. A., Hayes, B. C., & Gorman, P. C. (1991). *User acceptance of intelligent avionics: A study of automatic-aided target recognition* (Report No. NASA CR-177583). Moffett Field, CA: Ames Research Center.

Davis I., See, J. E., Shacklett, S. L., & Kuperman, G. G. (1996). *Unaided target acquisition performance with enhanced APG-70 synthetic aperture radar: A signal detection theory analysis* (Report No. AL/CF-TR-1996-0104). Wright-Patterson Air Force Base, OH: Armstrong Laboratory.

Dilsavor, R. L. (1995). *Overview of ATC system implementation.* Unpublished manuscript.

Entin, E. B., & Entin, E. E. (1997). The effects of decision time and decision-aid accuracy on workload and performance. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting,* (pp. 177-181). Santa Monica, CA: The Human Factors Society.

Entin, E. B., Entin, E. E., & Serfaty, D. (1996). Optimizing aided target-recognition performance. In *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting,* (pp. 233-237). Santa Monica, CA: The Human Factors Society.

Fulghum, D. A. (1994). Scud hunting may drop under 10-minute mark. *Aviation Week and Space Technology, 140 (8)*, 90.

Fulkerson, D. C. (1980). Required performance levels for autocueing devices. In *Proceedings of the Human Factors Society 24th Annual Meeting,* (pp. 511-515). Santa Monica, CA: The Human Factors Society.

Geschieder, G. A. (1985). *Psychophysics: Method, theory, and application* (2nd ed.). Hillsdale, NJ: Erlbaum.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Hacker, M. J., & Ratcliff, R. (1979). A revised table of d′ for M-alternative forced choice. *Perception & Psychophysics, 26*, 168-170.

Huynh, H., & Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. *Journal of the American Statistical Association, 65*, 1582-1589.

Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics, 1*, 69-82.

Jauer, R. A., Quinn, T. J., Hockenberger, R. L., & Eggleston, R. G. (1986). *Radar aided mission/aircraft capability exploration (RAM/ACE): Full task simulation* (Report No. AAMRL-TR-86-015). Wright-Patterson Air Force Base, OH: Armstrong Laboratory.

Kibbe, M. P., & Weisgerber (1991). *Operator use of automated target recognition (ATR) systems under time constrained conditions* (Report No. NWC TP 7136). China Lake, CA: Naval Weapons Center.

Kuperman, G. G., Bryant, M. L., & Clark, L. G. (1991). Man-machine interfaces for target recognition systems. In *Proceedings of the 4th National Symposium on Sensor Fusion: Volume 1*. Ann Arbor, MI: ERIM.

Kuperman, G. G., Wilson, D. L., & Perez, W. A. (1988). *Relocatable target acquisition performance with simulated synthetic aperture radar imagery* (Report No. AAMRL-TR-88-025). Wright-Patterson Air Force Base, OH: Armstrong Aerospace Research Laboratory.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge: Cambridge University Press.

Nonnuclear Weapon Delivery Manual, TO 1F-15E-34-1-1, USAF Series F-15E Aircraft McDonnell Douglas Aerospace. D 15 Jan 94; Change 2 1 Sept. 94.

O'Byrne, B. (1995). *TMD Eagle Smart Sensor and ATC (TESSA)* (Letter Report Test Directive 921AEA98).

Petersen, J., Fruchey, D., Rubin, R., & O'Rourke, P. (1995). *ORION sortie effectiveness model: User's manual*. Pico Rivera, CA: Northrop Grumman Corporation.

Pryce, J. D. (1995). *TESSA SAR/FLIR Data Collection Report: Volume 1* (Report No. OMI-560). Dayton, OH: Optimetrics, Inc.

See, J. E., & Kuperman, G. G. (1995). *The effects of resolution and other factors on target acquisition performance using simulated synthetic aperture radar imagery* (Report No. AL/CF-TR-1995-0149). Wright-Patterson Air Force Base, OH: Armstrong Laboratory.

See, J. E., Riegler, J. T., Fitzhugh, E., & Kuperman, G. G. (1996). *Unaided target acquisition performance with first-generation forward looking infrared imagery: A signal detection theory analysis* (Report No. AL/CF-TR-1996-0094). Wright-Patterson Air Force Base, OH: Armstrong Laboratory.

See, J. E., Warm, J. S., Dember, W. N., & Howe, S. R. (1997). Vigilance and signal detection theory: An empirical evaluation of five measures of response bias. *Human Factors, 39*, 14-29.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34-50.

Toms, M. L., & Kuperman, G. G. (1991). *Sensor fusion: A human factors perspective* (Report No. AL-TR-1991-0152). Wright-Patterson Air Force Base, OH: Armstrong Laboratory.

Walters, C. P. (1993). Human perception testing and automatic target recognition: Performance against a common imagery set. In *Proceedings of the Third Automatic Target Recognizer Systems and Technology Conference: Volume 1* (U) (Report No. GACIAC PR 93-01). Chicago, IL: GACIAC.

Weisgerber, S. A., & Savage, S. F. (1990). *Operator ship classification using an automatic target recognition (ATR) system in conjunction with forward looking infrared (FLIR) imagery* (Report No. NWC TP 7101). China Lake, CA: Naval Weapons Center.

Wilson, D. L. (1992). *Theory of signal detection and its application to visual target acquisition: A review of the literature* (Report No. AL-TR-1992-0083). Wright-Patterson Air Force Base, OH: Armstrong Laboratory.

GLOSSARY

AFB	Air Force Base
AFDTC	Air Force Development Test Center
A/G	Air-To-Ground
ANOVA	Analysis of Variance
ASC/FBXT	U.S. Air Force Aeronautical Systems Center/Theater Missile Defense Integrated Product Team
ATC	Automatic Target Cuer
ATR	Automatic Target Recognizer
c	Response bias index
CFAR	Constant-False-Alarm-Rate
CIWAL	Crew-Aiding and Information Warfare Analysis Laboratory
d'	Perceptual sensitivity
FA	False Alarm
FLIR	Forward Looking Infrared
H	Hit
HiPAC	High Pressure Air Compressor
HRM	High Resolution Map
LANTIRN	Low Altitude Navigation and Targeting Infrared for Night
MSE	Mean Squared Error
NHRM	New High Resolution Mode
nmi	Nautical Mile
RAM/ACE	Radar Aided Mission/Aircrew Capability Exploration
ROI	Region of Interest
RT	Reaction Time
SAR	Synthetic Aperture Radar
SG	Silicon Graphics
TEL	Transporter/Erector/Launcher
TESSA	TMD Eagle Smart Sensor and ATR
TMD	Theater Missile Defense
TSD	Theory of Signal Detection
WL/AA	Wright Laboratories/Avionics Directorate