Proceedings of the 2004 Winter Simulation Conference
R .G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, eds.

MEASUREMENT OF AIR TRAFFIC CONTROL SITUATIONAL AWARENESS ENHANCEMENT THROUGH RADAR SUPPORT TOWARD OPERATING ENVELOPE EXPANSION OF AN UNMANNED AERIAL VEHICLE

Maj James Denford

Directorate of Land Synthetic Environments
PO Box 17000 Station Forces
Kingston, Ontario, Canada K7K 7B4

John A. Steele
Roger Roy
Eugenia Kalantzis

Directorate General Land Combat Development
Operational Research Team
PO Box 17000 Station Forces
Kingston, Ontario, Canada K7K 7B4

ABSTRACT

Current airspace restrictions in Kabul limit the potential capability of the Tactical Unmanned Air Vehicle (TUAV) within the area of operations of the Kabul Multinational Brigade. An experiment was conducted using the OneSAF Testbed Baseline and a range of virtual simulations to examine the impact of five different radar options and three different information displays on the level of airspace situational awareness (SA) of the air traffic control officer (ATCO). The use of SAGAT, SART and NASA-TLX techniques were effective in determining differences in workload, situational awareness and understanding. Simultaneous data capture through shared EXCEL workbooks and VBA macros permitted near real time analysis. The Mann-Whitney U test, used due to the nature and limited size of the data sets, showed that any of the radars examined in this experiment would assist in the establishment of positive control over TUAV operations in the controlled airspace over Kabul.

1 INTRODUCTION

Canadian Forces (CF) recently purchased a set of four Sperwer (Sparrowhawk) Tactical Unmanned Aerial Vehicles (TUAVs) to provide the commander of Kabul Multi-National Brigade with a tactical, beyond line-of-site, rapidly deployable, day/night intelligence, surveillance, target acquisition and reconnaissance capability for deployment to Afghanistan as part of Canada's contribution to the International Stabilization Force. The TUAVs were procured to enhance situational awareness, battlefield management and force protection. Though operated by CF personnel, the TUAV system would operate in a coalition environment as a Brigade asset and could be tasked to support any of the coalition assets within the KMNB Area of Operations (AO).

Prior to deployment, TUAVs were restricted in their operations in the vicinity of the airfield, as depicted in Figure 1. They were not allowed to fly inside an 8 km by 20 km zone immediately surrounding the airfield. A second 35 km by 14 km zone was defined around the runway within which all UAVs were restricted to a height of 500 ft. However, if the commander deemed it mission critical, air traffic could be restricted and the TUAV would be given the freedom of the sky. It was proposed that the use of radar systems to improve the air picture in Kabul could eliminate some restrictions on the employment of a TUAV. This experiment was designed to evaluate this proposal and provide substantiation for it if warranted. Denford *et al* (2003) is the original report to the sponsors.

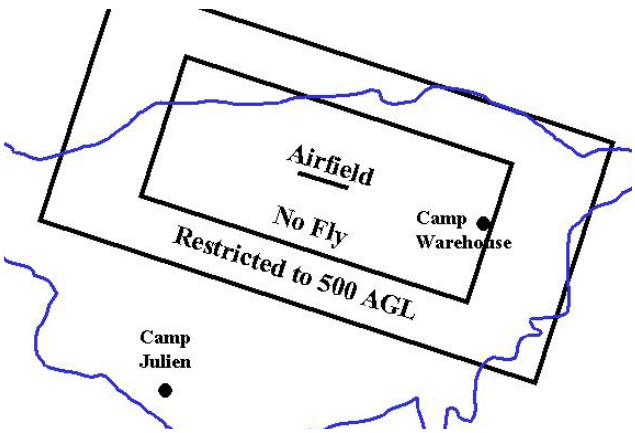

Figure 1: TUAV Airspace Restrictions over Kabul

Three candidate sites for radar installations were the Kabul International Airport (KIA) for civilian ATC radars, and Camp Warehouse or Camp Julien for military radar assets. The six radar options investigated were: 1) Baseline (no radar); 2) Quad Air Traffic Control (ATC) Radar at KIA; 3) MPN-25 ATC Radar at KIA; 4) Air Defence/Anti Tank Systems (ADATS) at Camps Julien and Warehouse; 5) Skyguard Fire Control Units (FCUs) at Camps Julien and Warehouse; and 6) Quad ATC Radar at KIA and Skyguard FCU at Camp Julien. It was also decided to examine different types of information displays presented within the ATC tower, focusing on a TUAV moving map display, the MPN-25 radar situation display and the Air Defence System Integrator (ADSI) display.

The location of the radar sites considered is illustrated in Figure 2, below, a satellite image of the AO. The range of peaks that can be seen snaking through the center of the AO limits the capability of radar on each side of the range to detect low flying aircraft on the other side.

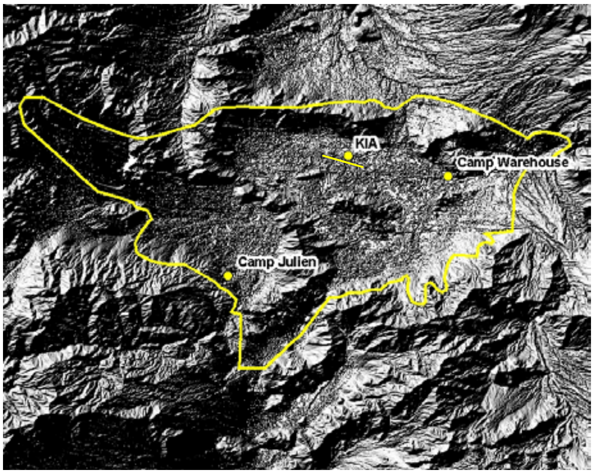

Figure 2: A Satellite Image Showing the Topography of the KMNB AO with Camps Julien and Warehouse and the Runway at KIA

2 THEORY

While several definitions of situational awareness (SA) have been offered, the most generally applicable definition is that provided by Endsley (1988). SA is "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future." Alternatively, it is a combination of a monitoring function, where the decision-maker tracks key system variables as he works toward a particular decision, and a control function, which includes the generation and evaluation of alternatives. The common thread among these definitions is that the achievement of SA requires conscious attention.

The primary method for measuring SA during the experiment was the Situational Awareness Global Assessment Technique (SAGAT) (Endsley 1987). When using SAGAT, the simulation is frozen at random intervals and data collected on all three levels of SA.

The SAGAT depends upon a comprehensive assessment of operator SA requirements. Such a study of ATC was completed by Endsley and Rogers (1994) and formed the basis for the tools developed for this experiment. Briefly, Air Traffic Control Officers (ATCOs) are called upon to sort-out and project the paths of an ever-fluctuating number of aircraft in order to ensure goals of minimum aircraft separation and safe, efficient take-off, en route and landing operations. The success of the ATCO in this task depends upon his or her awareness of the rapidly changing location of each aircraft (in three-dimensions) and its projected future location relative to every other, along with other pertinent aircraft parameters (destination, speed, fuel, altitude, etc…).

Figure 3 depicts a model by Endsley (1995) of levels of SA and their role in dynamic decision making. According to this model, there are three levels of SA (perception, comprehension and prediction,) all of which were evaluated in this experiment.

The first level involves perceiving the current status, attributes, and dynamics of relevant elements in the environment and is called Level 1 SA. In this study, it included aircraft type, call sign, location, altitude, altitude change, airspeed, heading, heading change, intention, and emergency status. For each item, criteria were developed for assigning to each response a score between 0 and 5. Subject matter expert participants in the experiment provided their rating of the impact of each of these items on air safety. The results were used to derive weight factors which were then used to combine the ten assigned scores into a single SAGAT Level 1 SA score out of 50.

Level 2 SA goes beyond current facts to include the comprehension of their significance in light of the controller's goals. Since the focus of this experiment was determining whether radar supported SA was adequate to ensure safety when a remotely piloted vehicle was inside controlled airspace, the measure of Level 2 SA was the accuracy of knowledge of which three aircraft were currently closest to the TUAV. Again, criteria for assigning scores out of 10 were developed rewarding the identification of the correct aircraft in the correct order.

Level 3 SA is the projection of the future actions of the elements of the environment in the near term. The corresponding measure used in this experiment was the answer to the Level 2 SA question in 2 and 5 minutes time and was scored in the same manner.

A second more subjective technique used was the Situation Awareness Rating Technique (SART) (Taylor, 1989) in which the ATCO assesses ten different dimensions of the quality of his own SA on a Likert scale from one to seven. These dimensions can be grouped and interpreted as indications of supply of attention, demands upon attention, and degree of situational understanding.

A useful overview of issues surrounding SA is given by Uhlarik and Comerford (2002).

To assess any change in the ATCO workload introduced by radar support, the NASA Task Load Index (TLX) (Hart 1986) was used. This technique requires participants to assign a Likert scale rating from 1 to 7 to each of the six different dimensions of the burden involved in the task at

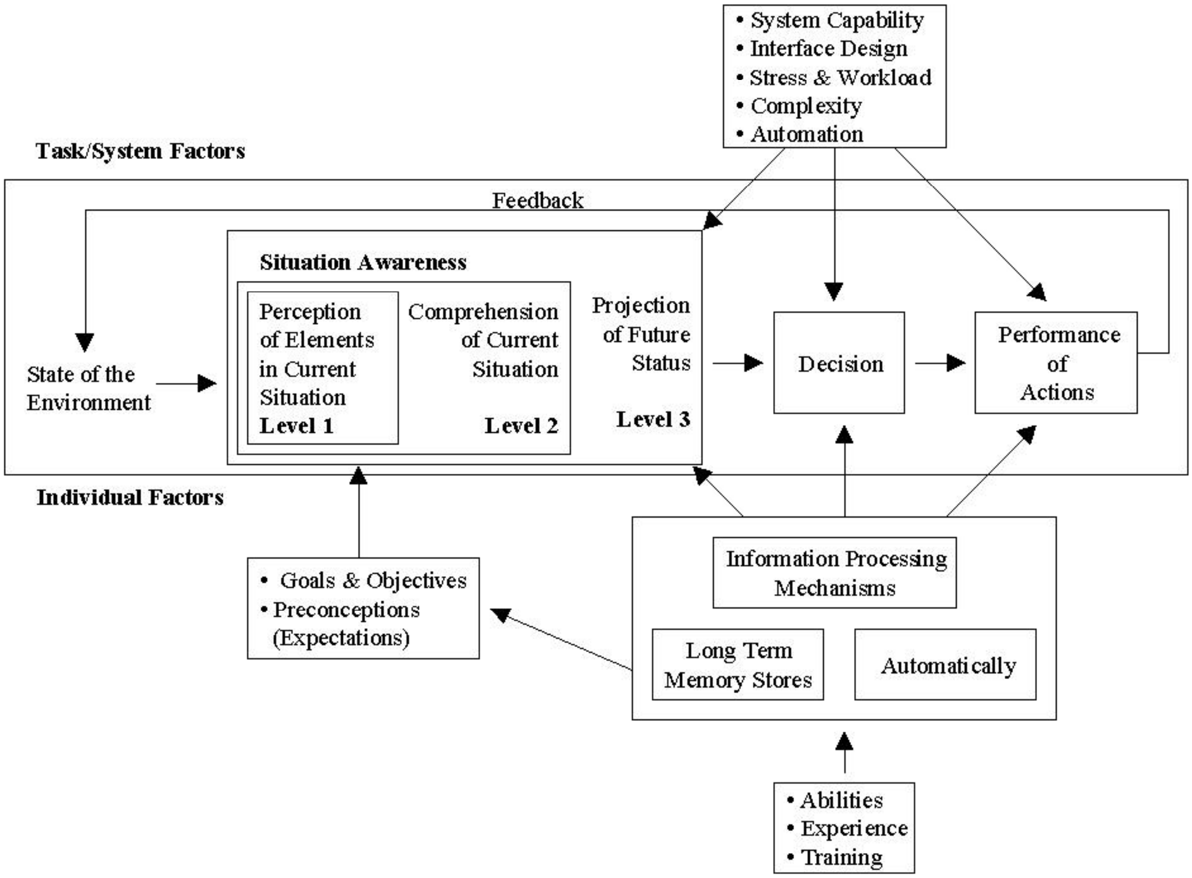

Figure 3: A Conceptual Map from Endsley (1995) of the Levels of SA and Their Role in Dynamic Decision Making

hand and to complete all of the 15 possible pair-wise comparisons between different dimensions indicating which of the two makes a greater contribution to the burden of the task. This permits the calculation of a single TLX score. The dimensions rated include mental demand, physical demand, time pressure, effort required, pressure due to performance level, and level of frustration.

3 SYNTHETIC ENVIRONMENT

The intent of the synthetic environment was to provide sufficient contextual cues to stimulate expert performance in the experiment participants. Elements simulated included the ATC tower, the Quad Radar or AN/MPN-25, Land Force ADATS or Skyguard FCUs, a Sperwer TUAV air vehicle and Ground Control Station (GCS) and an exercise control (EXCON) element consisting of air traffic in the Kabul area.

The EXCON function was to drive the experimental scenarios and consisted of two OnSAF Testbed Baseline (OTB) (an entity-based constructive simulation) stations to control aircraft operating within the simulation and two stations using ModIOS software to provide simulated radio communications. Flight profiles were developed using the Jeppesen diagrams for KIA, augmented by recommendations from two CC-130 Hercules pilots with recent experience in Afghanistan. Two pilots were employed for the execution of the EXCON task. The employment of pilots qualified in both fixed and rotary wing aircraft was vital to the creation of the virtual environment as their application of voice procedures situated the ATCO in the ATC task.

The Sperwer TUAV simulation consisted of a workspace that simulates the Ground Control Station (GCS) including an OTB workstation to direct the flight path, a moving map display showing the TUAV position over the terrain, a nose mounted camera and a payload operator's display to view the orientable line-of-site payload feed. Radio communications were provided between the tower and the GCS. The constructive simulation permitted the TUAV controller to fly the TUAV through various legs at various altitudes and speeds, take off and land the UAV, and initiate appropriate immediate actions. The air vehicle and payload characteristics were modeled to specifications of the Statement of Operational Requirement for the system.

The tower simulation consisted of a workspace that included a full 360-degree view from the tower, represented by three ModIOS presentations across six screens. The Air Traffic Control Officer (ATCO) and Assistant were provided with a station to simulate use of binoculars from the tower position. During the excursions to examine different types of information presented to the ATCO in the tower, a station was provided for remote display of the MPN-25 PPI, for the Air Defence System Integrator (ADSI) or the moving map display from the UAV. The ADSI was simulated by providing an OTB screen showing all flight operations, representing full coverage of the area (with the assumption that radar coverage from Bagram, Kandahar and Kabul was integrated). Communications were provided to the ATC in the form of telephone to the air defense radars and radio to the GCS and aircraft.

The air defence radar assets simulation consisted of a workspace that included two plan position indicator displays (PPIs) and two electro-optical displays. Telephone communications were provided to the AD Battery Command Post and the radar operators maintained listening watch on the tower radio frequency. The radar and optics performance characteristics were modelled to specifications provided by the manufacturer.

When present, the ATC radar simulation consisted of a workstation that included an operator with a PPI that was in its own workspace only for Quad radar operations and remoted to the tower location for MPN-25 operations. Telephone communications were provided to the tower and the radar operator maintained listening watch on the tower radio frequency. The radars' performance characteristics were modeled for Primary Search Radar operations to specifications provided by the ATC squadron employing the systems.

4 THE EXPERIMENT

Elements incorporated in the experiment were manipulated to simulate the radars and aircraft interacting with the ATCO in the Kabul area. The aircraft in the simulation included a variety of domestic and international commercial, military and private flights. Boeing 707s, DC-3s and Hercules transports shared the airspace with Cessnas, Helicopters and the TUAV. Circumstances modeled included mechanical, weather and medical emergencies as well as aircraft unable or unwilling to communicate with the tower.

Each half-hour scenario included a total of 12 aircraft on a variety of inbound, outbound, en route and local flight paths. Air traffic tempo was regulated to maintain a moderate load of between three and six aircraft within the ATCO's control at all times during the simulation. It was determined that only about four iterations of each radar option could be completed in the allotted time. This resulted in a very tight schedule with four iterations completed in the morning, four (sometimes five) in the afternoon and two (sometimes three) in the evening. A total of 24 iterations plus three excursions were completed.

All of the data collection for each run was implemented in a shared Microsoft Excel workbook which contained four worksheets for the ATCO (three for SAGAT and one for SART and TLX) and three for EXCON (SAGAT only). In this way, almost all subsequent data entry was eliminated and the comparison of ATCO SA data with air truth data from EXCON automated within other spreadsheets through macros written in Visual Basic for Applications (VBA) for EXCEL. Each of the six configurations were run four times with three pauses each, yielding 12 SAGAT data points per configuration.

5 ANALYSIS

The analysis of the data involved several stages, each described in a subsection below. Steele *et al* (2004) provides a more detailed description of the methods used and their results

5.1 Mapping ATCO-Listed to EXCON-Listed Aircraft

Since the queries to the ATCO were not multiple choice, some decision had to be made about which of the aircraft listed by EXCON was being described by the data for each aircraft on the ATCOs inputs. The scores assigned to SAGAT data from the ATCO will rise or fall depending upon which correspondences are assumed.

To make these choices, a macro was written which tested every possible set of mappings from ATCO aircraft to EXCON aircraft. An objective function was evaluated for each set of mappings. This function was a weighted sum of scores assigned to 4 data items: 40% on each location and aircraft type and 10% on each of altitude and speed. For every aircraft listed by EXCON but not included by the ATC

5.2 SAGAT Level 1 SA Items

Criteria were developed to for each raw score level between 0 and 5 for each of the 10 level 1 SA information items requested of the ATCO during each pause in the simulation.

5.2.1 Call Sign

Though the creation of a careful string comparison macro that systematically compared the call sign listed by the ATCO with the actual call sign and penalized character transpositions and substitutions was attempted and somewhat successful, it was too computationally intensive for repetitive use and was abandoned in favor of a much more rudimentary scheme using the EXCEL vlookup function. The ATCO call sign response was inserted into its place in an alphabetized list of call signs and one point taken away for each position the response was away from the true call sign.

5.2.2 Location

Locations were indicated by the ATCO by placing a number for each aircraft in an EXCEL table according to a graphical overlay on the table showing the outline of the area of operations with the KIA runway and Camps Warehouse and Julien. Each cell of the table was scaled to represent a 2.5 km square. For every 5 km straight-line distance between the ATCO location for an aircraft and the EXCON location for the same aircraft, one point was taken from the location score so 0 points were assigned only if the error was at least 25 km.

5.2.3 Altitude

Differences between the EXCON and the ATCO values for an aircraft's altitude of 500 ft., 1000 ft., 1500 ft., 2000 ft. and 3000 ft. were scored as 4, 3, 2, 1, and 0, respectively.

5.2.4 Aircraft Type

The six types of aircraft in the simulation were grouped as transport (Boeing 707, DC-3 and Hercules), small (Cessna and Griffon helicopter) and UAV. The size of penalty was correlated with the seriousness of the consequences of the confusion. Full marks only came with correct aircraft identification. Transports confused with other transports were given 4 points. Confusion between two different small aircraft were given 3 points. Confusing a transport with a small aircraft and vise versa was given 2 points. Confusing a small aircraft with a TUAV and vice versa was given 1 point, and confusing a transport with a TUAV and vice versa was given 0 points.

5.2.5 Speed

Differences between the ATCO and EXCON values for aircraft speed of 25 kts, 50 kts, 75 kts, 100 kts and 150 kts were scored as 4, 3, 2, 1, and 0, respectively.

5.2.6 Heading

When scoring aircraft heading errors, care was taken to ensure that 355 degrees was interpreted as 10 degrees away from 5 degrees. Differences between the ATCO and EXCON values for aircraft heading of 15 deg, 30 deg, 45 deg, 60 deg and 90 deg were scored as 4, 3, 2, 1, and 0, respectively.

5.2.7 Heading Change

A score of 5 was given for correctly identifying left or right turning or straight flight. Three points were given for confusing a turn with straight flight or vice versa, and a score of 1 was given for confusing a left turn with a right turn and vice versa.

5.2.8 Altitude Change

As with the aircraft heading change score, correctly identifying a climbing, descending or level flying aircraft earned 5 points. Confusion between level flight and either climbing or descending was given 3 points. Indicating a climbing aircraft to be descending and vice versa earned only 1 point.

5.2.9 Activity

Correctly identifying whether the aircraft was enroute, inbound, outbound or local earned 5 points. Calling an inbound aircraft outbound and vice versa was incorrect but at least understood the runway was involved and earned 3 points. Confusing local aircraft with in- or outbound aircraft and vice versa was deemed to show more serious confusion and given a 2. Indicating that an en route aircraft was not en route resulted in 1 point, and believing an aircraft not en route to be en route was deemed most serious because interactions with other aircraft would not be anticipated and was given 0 points.

5.2.10 Emergency Status

While correctly identifying the presence or absence of an emergency in an aircraft was given 5 points, incorrectly ascribing an emergency was considered less serious than incorrectly ascribing no emergency. The former earned 1 point and the latter 0 points.

5.3 Defining a Global SAGAT Level 1 SA Score

Rather than simply averaging the Level 1 SA item scores out of 5 (falsely suggesting that all ten items were of equal importance in determining Level 1 SA), subject matter experts participating in the experiment were asked to indicate which items were more or less important by completing the 45 possible pair-wise comparisons, indicating which of each pair of items was more important for air safety. From these, individual priority ratings were determined between the ten items. These were then combined giving 50% influence to the person playing the ATCO and the rest divided between the remaining four players. The resulting weights were used to combine the ten Level 1 SA items to give a Global Level 1 SA score out of 50.

5.4 SAGAT Level 2 and 3 Scores

Six, three and one point were available as component scores for correct listing of the first, second and third closest aircraft, respectively. The portions of each of these components assigned is given in Table 1, below. Listing the three closest aircraft but in the wrong order assures at least 4 out of 10.

Table 1: SAGAT Level 2 and 3 Scoring Scheme

Data↓ Score→	6	3	2	1	0
Nearest aircraft	listed nearest	listed 2nd nearest		listed 3rd nearest	not listed
2nd nearest aircraft		listed 2nd nearest	listed nearest	listed 3rd nearest	not listed
3rd nearest aircraft				listed	not listed

5.5 SART Scores

The Situational Awareness Rating Technique (SART) asks the ATCO to rate ten different aspects of his own SA. Analysis on the basis of the ten different dimensions is known as 10-D SART. The responses can then be averaged within three groups to indicate the demands on attention, the supply of attention, and the degree of situational understanding. Analysis of these three aggregate measures is called 3-D SART. The 10-D items rated are shown in Table 2, below.

Table 2: The Components of the SART

3-D	10-D	Description
Demands on Attention	Instability of Situation	Situation's likeliness to change suddenly
	Complexity of Situation	Situation's degree of complication
	Variability of Situation	The number of factors changing
Supply of Attention	Arounsal of Situation	Degree of alertness/readiness for activity stimulated by the situation
	Concentration of Attention	Degree to which thoughts are brought to bear
	Division of Attention	Ability to spread or distribute focus of attention
	Spare Mental Capacity	Mental ability available for new variables
Situational Under-standing	Information Quantity	Amount of knowledge received and understood
	Information Quality	Goodness or value of knowledge communicated
	Familiarity	Degree of prior situation experience and knowledge

5.6 NASA TLX Scores

Once the dimensions of mental demand, physical demand, time pressure, effort required, pressure due to performance level and level of frustration have been rated from 1 to 7, these ratings are combined in a weighted sum to give a TLX value. The weights assigned are determined by the number of times each dimension was judged to be a greater contributor to task burden than others it was compared with. This number divided by 15 becomes the weight used to combine the rating values to obtain a TLX value between 1 and 7. In this experiment, pair-wise comparisons were completed twice, once on the first day and once on the last day. The weights obtained were averaged between the two days before being used to combine individual ratings into a TLX value.

5.7 Statistical Method Used

The null hypotheses for the experiment were the following:

1. a., b., and c. The ATCO global level 1, 2 and 3 SAGAT SA scores, respectively, with radar support are no better than those given without radar support.
2. a., b., and c. Each of the five radar support option results in ATCO global level 1, 2, and 3 SAGAT SA scores, respectively, which are no different than those of any other radar support option.
3. The ATCO 3-D SART ratings with radar support are no better than those given by the baseline.
4. Each of the five radar support options results in ATCO 3-D SART levels which are no different than those of any other radar support option.
5. The ATCO TLX scores with each radar support option are no different from those given by the baseline or any other radar support option.

The fact that the SAGAT, SART and TLX data did not come from equal interval scales, the small size of the data sets, and their demonstrably non-normal distribution ruled out the use of *t*-tests in data analysis. Instead, all analysis was done through pairwise Mann-Whitney U tests. These are essentially the same as ordering the data from two sets, transforming the data into rank numbers and doing parametric *t*-tests on the rank number distribution.

6 RESULTS

6.1 Learning Curve/Maturation

There were concerns that the data gathered during the test would show evidence of a learning effect in which SA scores in initial trials were worse than subsequent scores. However, a plot of SAGAT Level 1 SA score totals with linear fit, shown in Figure 4, below, shows no such correlation.

6.2 SAGAT Scores

One-tail Mann-Whitney U tests determined confidence levels that ATCO SA with each radar support was better than without. The results showed that all radar options resulted in significantly better Level 1 SA at the 90% confidence level, with no significant differences between different radar support options.

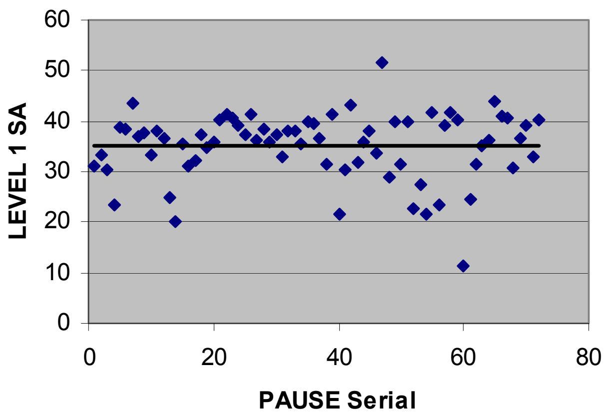

Figure 4: Sum of Level 1 SA Items by Pause with Line Fit

Table 3: Individually Valid Confidence Levels that Level 1 SA with Radar Support is Better than Without

Radar Support Option	Confidence Level
ADATS	96.3%
MPN	99.5%
QUAD	94.0%
QUAD/SKYGUARD	95.0%
SKYGUARD	99.1%

Only the MPN and Quad radar options resulted in significantly better Level 2 SA than without radar at the 95% confidence level. Level 2 SA scores also show that MPN is significantly different and better than the other radar support options. Only the MPN radar option shows statistically significant better Level 3 SA than without radar and only for the 2-minute future time horizon. The QUAD/Skyguard combination shows a significantly different and worse Level 3 SA than any other radar option for the 5-minute future time horizon, indicating that two different types of radars passing information in different formats may reduce Level 3 SA in the Tower. ATC radar operators use degrees and nautical miles while air defence radar operators pass information in mils and kilometers.

Table 4: Individually Valid Confidence Levels that Level 2 SA with Radar Support is Better than Without

Radar Support Option	Confidence Level
ADATS	60.2%
MPN	99.6%
QUAD	95.3%
QUAD/SKYGUARD	62.5%
SKYGUARD	79.9%

6.3 SART Scores

The 3-D SART showed no statistically significant differences in the demands on ATCO attention between the six variations in radar support. However, the MPN, Quad and QUAD/Skyguard options were all rated significantly better than the Baseline option for situational understanding. In addition, the MPN scores for situational understanding were significantly different (two-tail test) and better than all other radar support options at a greater than 90% confidence level. This may reflect the ATCO's comfort with a display in the tower and the terminology used by the supporting ATC-trained personnel operating the system.

Table 5: Individually Valid Confidence Levels that Level 3 SA (2 minutes) with Radar Support is Better than Without

Radar Support Option	Confidence Level
ADATS	69.8%
MPN	93.0%
QUAD	59.1%
QUAD/SKYGUARD	52.3%
SKYGUARD	62.5%

6.4 TLX Scores

An analysis of the 24 TLX data points from the ATCO indicate that the degree of burden for the ATCO was significantly different (Mann-Whitney two-tail test) and lower for the MPN radar than for any other option at better than 97% confidence level. All other radar options showed no significant differences between their TLX distributions from the baseline.

6.5 Excursions

Based on observation of increases in ATCO SA during the MPN-25 scenarios, excursions were developed to look at situation awareness aids in the tower. In three scenarios, the ATCO was provided with

1. a moving map display,
2. a moving map display and MPN-25 feed, or
3. an Air Defence System Integrator (ADSI) display.

The ATCO's confidence in correctly representing the location of the TUAV was increased by having an SA aid for reference. When paired with the MPN-25 feed, the moving map display was used less; it was mostly used when the MPN-25 lost the TUAV due to terrain masking but was referred to occasionally for confirmation. When the ADSI was employed, the moving map display was removed from the tower, as its information was superfluous. Due to familiarity with standard radar displays and terminology, the ATCO stated he was most comfortable with the MPN-25 display, however adapted quickly to the ADSI display and highly appreciated the utility of integrating of the various radar pictures. The benefit of the

ADSI was seen in the rapid updating of the ATCO's mental picture of the airspace (reducing errors and increasing confidence). This allowed the ATCO to give incoming aircraft relative bearings (clock face method) and distances with minimal cognitive effort compared to making calculations from his mental representation.

7 CONCLUSION

The employment of any radar in support of ATC operations significantly improved the Level 1 situational awareness of the ATCO in the experiment. Only the MPN-25 radar option shows statistically significant improvements in Level 2 and Level 3 ATCO SA. This seems to indicate that ATC radars are better than AD systems at providing the ATCO with SA information. The MPN-25 SART scores for situational understanding and NASA-TLX scores for workload were significantly better than all other radar support options. It appeared that the decrease in workload associated with the provision of an SA aid (display) allowed the ATCO to better leverage his SA into SU. Based on these findings, and considering the limitations of the simulation, the MPN-25 provided the best radar support option, but any of the radars examined in this experiment would assist in the establishment of positive control over TUAV operations in the KIA-controlled airspace.

The use of SAGAT, SART and NASA-TLX proved effective in determining differences in SA and task burden. The employment of simulation to create an immersive environment with sufficient contextual cues to elicit expert performance was found to be a highly effective and cost efficient means of gaining insight into fielding and employment issues for new mission equipment. The combination of both rigorous measurement techniques and quality synthetic environments exceeded expectations and provided useful feedback to commanders in the field.

Three aids in rapid report generation were the advance made in analysis methodologies, data capture techniques employing shared workbooks, and the automation of repetitive and labor-intensive analysis tasks. The advance preparation of both background material and analysis tools will be continued as methods of reducing experimentation cycle times.

The Mann-Whitney U test for statistical significance was useful even when there were only four points in each data set. However, more runs should be planned for any experiment if significance is in doubt.

REFERENCES

Denford, Maj J.S., Steele J., Kalantzis E., and Roy R.L. 2003. *Limited Objective Experiment 0301 Tactical UAV Operations in Controlled Airspace*, Report AEC-R 0303, October 2003.

Endsley, M. R. 1987. SAGAT: A methodology for the measurement of situation awareness. NOR DOC 87-83, Hawthorne, CA: Northrop Corporation.

Endsley, M. R. 1988. "Design and evaluation for situation awareness enhancement". In *Proceedings of the Human Factors Society 32nd Annual Meeting*, Santa Monica, CA: Human Factors Society, 97-101.

Endsley, M.R. 1995. "Measurement of Situation Awareness in Dynamic Systems," *Human Factors*, 37(1): 65-84.

Endsley, M. R. 2000. "Direct measurement of situation awareness: validity and use of SAGAT." In *Situation Awareness Analysis and Measurement*. eds. Endsley, M. R. and Garland D. J. Mahwah, NJ: Lawrence Erlbaum Associates.

Endsley, M. R., and Rodgers, M. D. 1994. *Situation awareness information requirements for en route air traffic control*. DOT/FAA/AM-94/27. Washington, D.C.: Federal Aviation Administration Office of Aviation Medicine.

Hart, S. G. 1986. *NASA Task Load Index (TLX) V1.0 Users Manual*. Hanover, MD: NASA Center for Aerospace Information (CASI).

Steele, J. A., Roy, R. L., Kalantzis, E, and Denford, J. S. 2004. Measuring Air Traffic Control Situational Awareness: Analysis of an Experiment Evaluating TUAV Radar Options in the Kabul Area of Operations. Defence Research and Development Canada – Centre for Operational Research and Analysis.

Taylor, R.M. 1989. "Situational Awareness Rating Technique (SART): The Development of a Tool for Aircrew Systems Design". In *Situational Awareness in Aerospace Operations*, AGARD-CP-478. 3/1 - 3/17. Seuilly-sur Seine, France: NATO-AGARD.

Uhlarik, J. and Comerford, D. A. 2002. "A Review of Situation Awareness Literature Relevant to Pilot Surveillance Functions." (DOT/FAA/AM-02/3) Washington, DC: Federal Aviation Administration, U.S. Department of Transportation.

AUTHOR BIOGRAPHIES

JAMES S. DENFORD is the head of the Concept Development and Experimentation section of the Canadian Army's Directorate of Land Synthetic Environments. A Major in the Canadian Army Signal Corps, he received a B.Eng in Engineering Management and a MBA with a concentration in Management Information Systems from the Royal Military College of Canada in 1992 and 2002 respectively. His research interests are in decision support systems and the role of computer simulation in command support. His e-mail address is `denford.js@forces.gc.ca`.

JOHN A. STEELE is a defence scientist serving the Canadian Army at the Kingston field office of Defence

R&D Canada – Centre for Operational Research and Analysis. He received a B.E. in 1995 and a Ph.D. in 2002, both in Mechanical Engineering from the University of Saskatchewan in Saskatoon. His research interests are automated data collection and analysis and weapon systems effectiveness. His email address is `steele.ja2@forces.gc.ca`.

ROGER L. ROY is a senior operational research analyst supporting the land staff of the Canadian Army in Kingston. He has served for more than 25 years in operational research with postings across Canada in support of all three services as well as NORAD in Colorado Springs. His email address is `roy.rl@forces.gc.ca`.

EUGENIA KALANTZIS is a defence scientist serving the Canadian Army at the Kingston field office of Defence R&D Canada - Centre for Operational Research and Analysis. She holds both a B.Eng (1994) and M.Eng (1997) from the Department of Civil Engineering and Applied Mechanics of McGill University in Montreal, and has recently completed an MBA (2003) with a concentration in Operations Management. Her email address is `kalantzis.e@forces.gc.ca`.