Endsley, M. R. and Garland D. J (Eds.) (2000) *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

DIRECT MEASUREMENT OF SITUATION AWARENESS: VALIDITY AND USE OF SAGAT

Mica R. Endsley
SA Technologies

Introduction

The Situation Awareness Global Assessment Technique (SAGAT), is a global tool developed to assess SA across all of its elements based on a comprehensive assessment of operator SA requirements (Endsley, 1987b; 1988b; 1990c). Using SAGAT, a simulation employing a system of interest is frozen at randomly selected times and operators are queried as to their perceptions of the situation at that time. The system displays are blanked and the simulation is suspended while subjects quickly answer questions about their current perceptions of the situation. As a global measure, SAGAT includes queries about all operator SA requirements, including Level 1 (perception of data), Level 2 (comprehension of meaning) and Level 3 (projection of the near future) components. This includes a consideration of system functioning and status as well as relevant features of the external environment.

SAGAT queries allow for detailed information about subject SA to be collected on an element by element basis that can be evaluated against reality, thus providing an objective assessment of operator SA. This type of assessment is a direct measure of SA — it taps into the operator's perceptions rather than infers them from behaviors that may be influenced by many other factors besides SA. Furthermore it does not require subjects or observers to make judgments about situation knowledge on the basis of incomplete information, as subjective assessments do. By collecting samples of SA data in this manner, situation perceptions can be collected immediately (while fresh in the operators' minds), reducing numerous problems incurred when collecting data on mental events after the fact, but not incurring intrusiveness problems associated with on-line questioning. Multiple "snapshots" of operators' SA can be acquired in this way, providing an index of the quality of SA provided by a particular design. By including queries across the full spectrum of an operator's SA requirements, this approach minimizes possible biasing of attention, as subjects cannot prepare for the queries in advance since they could be queried over almost every aspect of the situation to which they would normally attend. The method is not without some costs, however, as a detailed analysis of SA requirements is required in order to develop the battery of queries to be administered.

Development of Queries

Probably one of the most important issues that must be addressed when using SAGAT is that of determining the queries to use for a particular experimental setting. Asking queries that are relevant to the operator's SA during the freeze is a prime determinant of the utility of the technique. These queries should also be posed in a cognitively compatible manner. That is, they should be phrased to be as similar as possible to how the person thinks about the information, and should not require extra transformations or decisions on the part of the operator.

SA Requirements Analysis

The problem of determining what aspects of the situation are important for a particular operator's SA has frequently been approached using a form of cognitive task analysis called a goal-directed task analysis, illustrated in Table 1. In such analysis, the major goals of a particular job class are identified, along with the major subgoals necessary for meeting each of these goals. Associated with each subgoal, the major decisions that need to be made are then identified. The situation awareness needed for making these decisions and carrying out each sub-goal are identified. These SA requirements focus not only what data the operator needs, but also on how that information is integrated or combined to address each decision. In this analysis process, SA requirements are defined as those dynamic information needs associated with the major goals or sub-goals of the operator in performing his or her job (as opposed to more static knowledge such as rules, procedures and general system knowledge).

Table 1 Format of Goal-Directed Task Analysis

Goal
  Subgoal
    Decision
      Projection (SA Level 3)
      Comprehension (SA Level 2)
      Data (SA Level 1)

Conducting such an analysis is usually carried out using a combination of cognitive engineering procedures. Expert elicitation, observation of operator performance of tasks, verbal protocols, analysis of written materials and documentation, and formal questionnaires have formed the basis for the analyses. In general, the analysis has been conducted with a number of operators, who are interviewed, observed and recorded individually, with the resultant analyses pooled and then validated overall by a larger number of operators.

An example of the output of this process is shown in Table 2. This example shows the SA requirements analysis for the subgoal "Maintain Aircraft Conformance" for the major goal "Avoid Conflictions" for an air traffic controller (Endsley & Rodgers, 1994). In this example, the sub-goal is even further divided into lower level sub-goals prior to the decisions and SA requirements being listed. In some cases, addressing a particular sub-goal occurs through reference to another sub-goal in other parts of the analysis, such as the need to re-address aircraft separation in this example. This shows the degree to which a particular operator's goals and resultant SA needs may be very inter-related. The example in Table 2 shows just one major subgoal out of four that are relevant for the major goal of "Avoid Conflictions", which is just one of three major goals for an air traffic controller.

This analysis systematically defines the SA requirements (at all three levels of SA) that are needed to effectively make the decisions required by the operator's goals. Many of the same SA requirements appear throughout the analysis. In this manner, the way in which pieces of data are used together and combined to form what the operator really wants to know is determined.

Although the analysis will typically include many goals and sub-goals, they many all be active at once. In practice, at any given time more than one goal or subgoal may be operational, although they will not always have the same prioritization. The analysis does not indicate any prioritization among the goals (which can vary over time), or that each subgoal within a goal will always be active. Unless particular events are triggered, for instance the subgoal of assuring aircraft conformance in this example, may not be active for a given controller.

This type of analysis is based on goals or objectives, not tasks (as a traditional task analysis might). This is because goals form the basis for decision making in many complex environments.

Table 2 Example of Goal-Directed Task Analysis for En-route Air Traffic Control (Endsley & Rodgers, 1994)

1.3Maintain aircraft conformance

1.3.1 Assess aircraft conformance to assigned parameters

- *aircraft at/proceeding to assigned altitude?*
- *aircraft proceeding to assigned altitude fast enough?*
  - time until aircraft reaches assigned altitude
  - amount of altitude deviation
  - climb/descent
    - altitude (current)
    - altitude (assigned)
    - altitude rate of change (ascending/descending)
- *aircraft at/proceeding to assigned airspeed?*
- *aircraft proceeding to assigned airspeed fast enough?*
  - time until aircraft reaches assigned airspeed
  - amount of airspeed deviation
    - airspeed (indicated)
    - airspeed (assigned)
    - groundspeed
- *aircraft on /proceeding to assigned route?*
- *aircraft proceeding to assigned route fast enough?*
- *aircraft turning?*
  - time until aircraft reaches assigned route/heading
  - amount of route deviation
    - aircraft position (current)
    - aircraft heading (current)
    - route/heading (assigned)
  - aircraft turn rate (current)
    - aircraft heading (current)
    - aircraft heading (past)
    - aircraft turn capabilities
      - aircraft type
      - altitude
    - aircraft groundspeed
    - weather
    - winds (direction, magnitude)

1.3.2 Resolve non-conformance

- *Reason for non-conformance?*
  - Verify data
    - *Is presented altitude correct?*
      - Aircraft altimeter setting
      - Aircraft altitude (indicated)
    - *Is presented airspeed correct?*
      - Aircraft airspeed (indicated)
      - groundspeed
      - winds (magnitude, direction)
    - *Is presented position/heading correct?*
      - Fix distance to Nav aid
      - range/bearing to Fix
      - track code
- *Will current behavior cause a problem?*
  - **Assess aircraft separation (1.1.1)**
  - **Assess aircraft/airspace separation (1.2.1)**
  - **Assure minimum altitude requirements (1.4)**
- *Action to bring into conformance?*

- **Provide clearance (2.2)**

Furthermore, tasks tend to be technology dependent. Completely different tasks may be carried out to perform the same goal in two different systems. For example, navigation may be done very differently in an automated cockpit as compared to a non-automated cockpit. Yet, the SA needs associated with navigation are essentially the same (e.g. location, deviation from desired course, …).

The analysis strives to be as technology free as possible. How the information is acquired is not addressed, as this can vary considerably from person to person, from system to system, and from time to time. In some cases it may be through system displays, verbal communications, other operators, or internally generated from within the operator. Many of the higher level SA requirements fall into this category. The way in which information is acquired can vary widely between individuals, over time and between system designs.

The analysis seeks to determine what operators would ideally like to know to meet each goal. It is recognized that they often must operate on the basis of incomplete information and that some desired information may not be available at all with today's system. However for purposes of design and evaluation of systems, we need to set the yardstick to measure against what they ideally need to know, so that artificial ceiling effects based on today's technology are not induced in the process. Finally, it should be noted that static knowledge, such as procedures or rules for performing tasks, is outside the bounds of an SA requirements analysis. The analysis focuses only on the dynamic situational information that affects what the operators do.

These analyses tend to be very long and complex and can take as much as a person-year of effort to complete for a given domain. On the positive side, once completed for a major class of operators, it need not be repeated unless the goals and objectives of the operator change radically. To date, these analyses have been completed for many domains of common concern including en route air traffic control (Endsley & Rodgers, 1994), TRACON air traffic control (Endsley & Jones, 1995), fighter pilots (Endsley, 1993), bomber pilots (Endsley, 1989a), commercial transport pilots (Endsley, Farley, Jones, Midkiff, & Hansman, 1998), aircraft mechanics (Endsley & Robertson, 1996), and airway facilities maintenance (Endsley, 1994). A similar process was employed by Hogg, Torralba and Volden (1993) to determine appropriate queries for a nuclear reactor domain.

SA requirements analyses have formed the basis for the development of appropriate SA queries in each of these domains which can be used and adapted for a wide range of experimental testing in these areas. For the reader who needs to develop queries for a new domain, the methodology described can be readily applied to derive the SA requirements for operators in that area. While this requires a considerable investment of energy, the resultant analysis also provides a highly useful foundation for directing design efforts and is well worth the effort expended.

Query Format

Based on the goal-directed task analysis, a list of the SA requirements for that domain can be constructed either as a whole, or for just particular goals and subgoals of the operator. These SA requirements form the basis for determining the queries to be used in that domain. Table 3 lists the SA queries that were developed for air traffic control, based on the analysis done by Endsley and Rodgers (1994) and Endsley and Jones (1995). In each case the query is presented and categorical response options are normally provided so that the operator's job in responding to the queries is minimized.

Table 3 SAGAT Queries for Air Traffic Control (Endsley & Kiris, 1995b)

1. Enter the location of all aircraft (on the provided sector map)
   aircraft in track control
   other aircraft in sector
   aircraft will be in track control in next 2 minutes
2. Enter aircraft callsign (for aircraft highlighted of those entered in query 1)
3. Enter aircraft altitude (for aircraft highlighted of those entered in query 1)
4. Enter aircraft groundspeed (for aircraft highlighted of those entered in query 1)
5. Enter aircraft heading (for aircraft highlighted of those entered in query 1)
6. Enter aircraft's next sector (for aircraft highlighted of those entered in query 1)
7. Enter aircraft's current direction of change in each column (for aircraft highlighted of those entered in query 1)

Altitude change	Turn
climbing	right turn
descending	left turn
level	straight

8. Enter the aircraft type (for aircraft highlighted of those entered in query 1)
9. Enter aircraft's activity in this sector (for aircraft highlighted of those entered in query 1)
   enroute, inbound to airport, outbound from airport
10. Which pairs of aircraft have lost or will lose separation if they stay on their current (assigned) courses?
11. Which aircraft have been issued assignments (clearances) that have not been completed?
12. Did the aircraft receive its assignment correctly?
13. Which aircraft are currently conforming to their assignments?
14. Which aircraft must be handed off to another sector/facility within the next 2 minutes?
15. Enter the aircraft that are experiencing a malfunction or emergency that is effecting operations.
16. Enter the aircraft which are not in communication with you.
17. Enter the aircraft that will violate special airspace separation standards if they stay on their current (assigned) path.
18. Which aircraft are weather currently an impact on or will be an impact on in the next 5 minutes along their current course?
19. Which aircraft will need a new clearance to achieve landing requirements?
20. Enter all aircraft that will violate minimum altitude requirements in the next two minutes if they stay on their current (assigned) paths?
21. Enter the aircraft that are not conforming to their flight plan.
22. Enter the aircraft and runway for this aircraft.

An example of a SAGAT query is shown in Figure 1. The operator serving as subject in the experiment need only click with a cursor on the desired answer, and then click on the next button to continue to the next question. This makes it fairly fast and easy to complete the battery of questions at each freeze point. Another example for a military pilot is shown in Figure 2. Note that in each of these examples an appropriate reference map is provided on which the location of

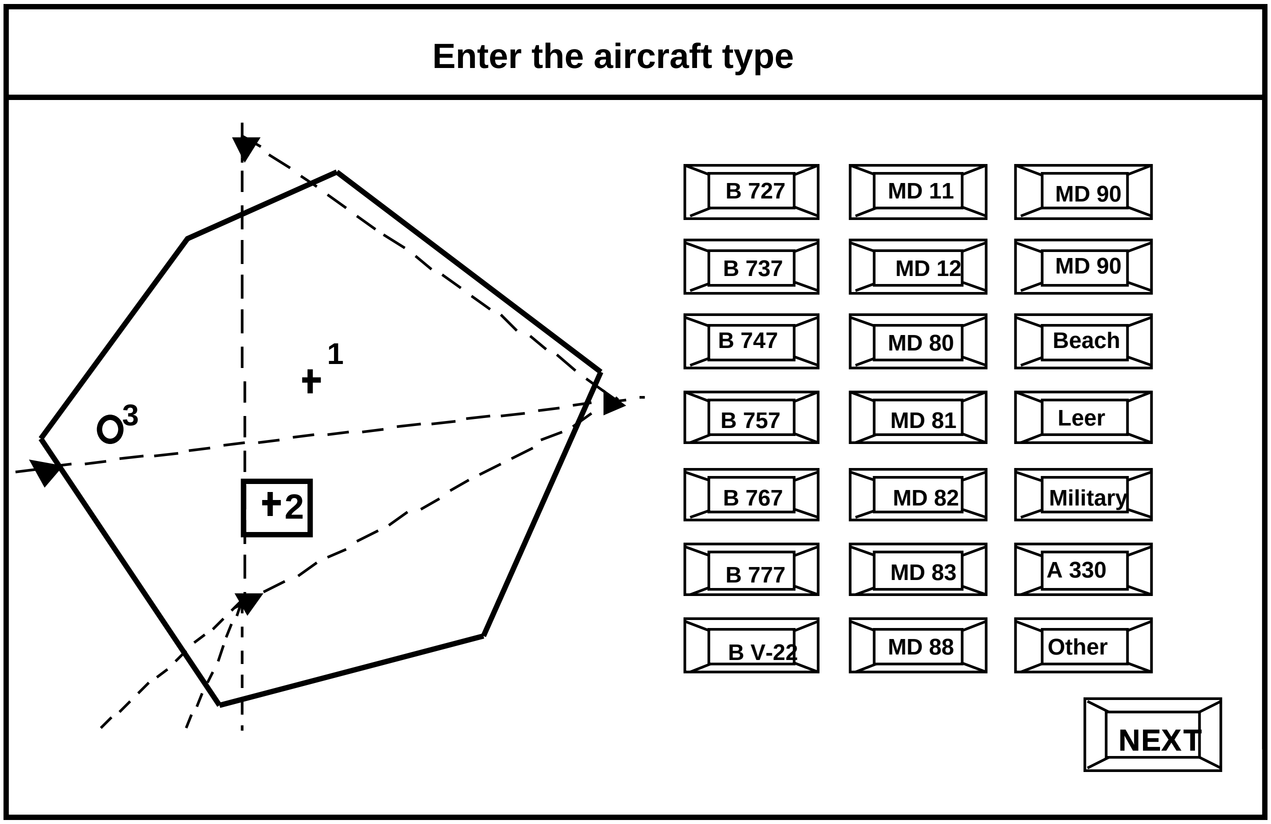

Figure 1 SAGAT Query Example for ATC: Aircraft Type (Endsley & Kiris, 1995b)

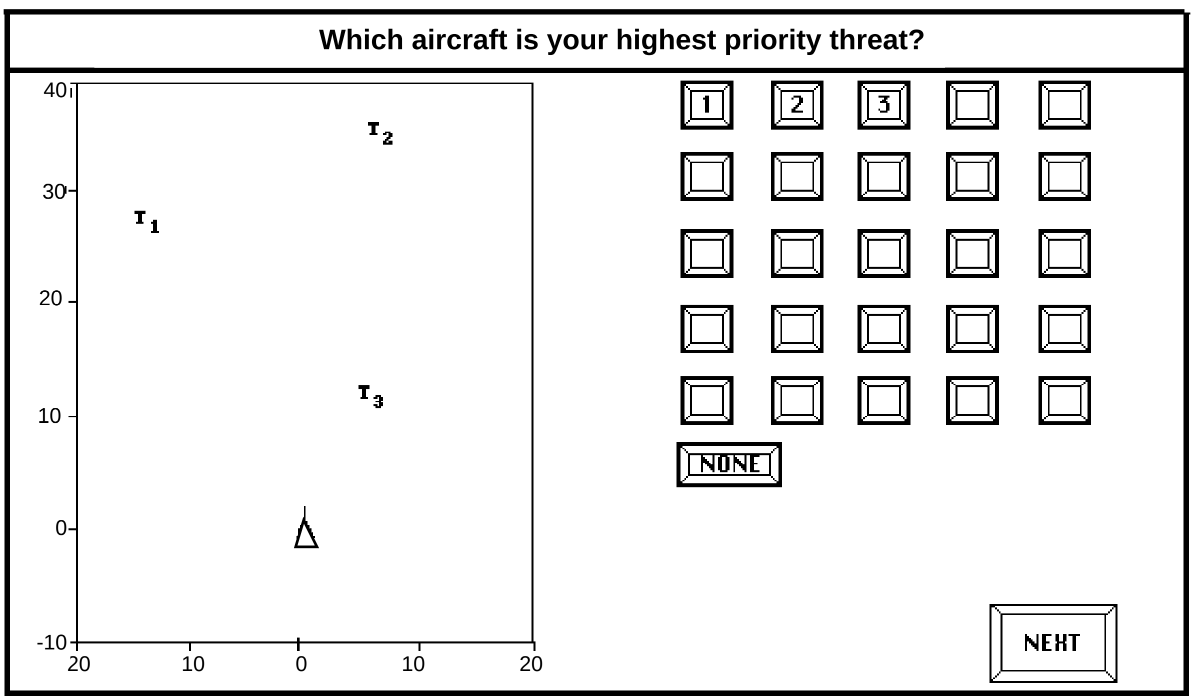

Figure 2 SAGAT Query Example for Military Pilot: Aircraft Priority

relevant aircraft are marked. These maps are derived from the first question that is provided to the operator in each case. An appropriate blank map is provided (with relevant boundaries and reference points marked) and the operator is asked to enter the aircraft that he or she knows about at the time in the appropriate location on the map. This completed map then serves as an organizing point for the remainder of the questions. This removes the problem of aircraft callsign confusion (which controllers are very poor at remembering) and which pilots may not know about at all. The questions are therefore self-referenced in a way which is compatible with how the operator pictures the situation at the time.

For a given testing situation, not all of the queries may be required. For instance, if the simulation does not employ aircraft emergencies, query number 16 in Table 3 may be omitted. Other queries should only be asked at times when they are appropriate. For example, query 12 and 13 dealing with proper reception and conformance to clearances should only be asked if the controller indicates in query 11 that there are aircraft that have been issued a new clearance that is not yet completed. Query 21 should only be provided if the controller indicates in query 9 that the aircraft will be landing in the sector. Query 22 is only relevant for landing aircraft and is generally not provided in a test involving en route air traffic control.

A determination of which queries should be provided for a given test can be made based on the SA requirements analysis, the capabilities of the simulation and the simulation scenarios, and the objectives of the test. It should be noted that the queries should not be too narrowly focused on just one or two items of interest, such as just measuring altitude when evaluating a new head-up display, for instance. This is because if only a few probes are used, subjects are much more likely to shift their attention to the factors they know they will be questioned on, affecting SA artificially. It has also been found that changes to one part of a system may inadvertently affect SA on other issues (Endsley, 1995a). A wide range of queries is needed to detect these changes. Vidulich (Chapter 13) finds that narrowly focused queries are less sensitive than a broad based range of queries for detecting changes in SA associated with a design. For this reason it is recommended that a broad range of the queries be used, with the exception of those which are clearly not relevant to SA in a particular simulation or situation.

For ease of administration of the queries and analysis of the data SAGAT queries have frequently been computerized (e.g. Endsley & Kiris, 1995b). The computerized battery insures that certain queries are not asked when they are not appropriate. While it is not required that administration be computerized (queries can be administered verbally or via paper and pencil) it has been found that administration in this manner greatly reduces the burden of analysis. The computer can collect the subject’s answers to each query into a database which can be compared to the correct answers from the simulation’s computer database. Relational database programs for correlating and scoring this data are a great benefit in the analysis process.

Examples of SAGAT Results

The rich data provided by SAGAT provides some justification for the effort involved in using it. It has been used to examine changes in SA induced by changes in display formats, display hardware, and automation concepts; individual differences in SA among operators and factors related to SA; and major changes in the rules of operation.

Evaluation of Sensor Hardware

Figure 3 shows an example of SAGAT results from a study examining two different avionics concepts for a fighter aircraft system (Endsley, 1988b). In this case, a new avionics system was evaluated which pilots subjectively felt was superior to the old system. Simulation testing did not show an improvement in mission performance, however. SAGAT testing revealed that the new system did indeed provide better SA as evidenced by better knowledge of where enemy aircraft were located, as well as improved knowledge on several other factors important for SA in this domain. As shown in Figure 3, pilots had significant improvement in SA with the new system for aircraft at various ranges, with the greatest improvement for aircraft at the farthest range (> 70 miles). From these results, it was concluded that the new system was providing the desired benefit to SA. The lack of improvement in mission performance was most likely due to the fact that the pilots had not yet discovered how to modify their tactics to take advantage of the improved SA, and that insufficient testing had been conducted to see improvements in the less sensitive performance metrics available for fighter combat scenarios. As an aid for making design decisions, this sort of discrimination is highly useful.

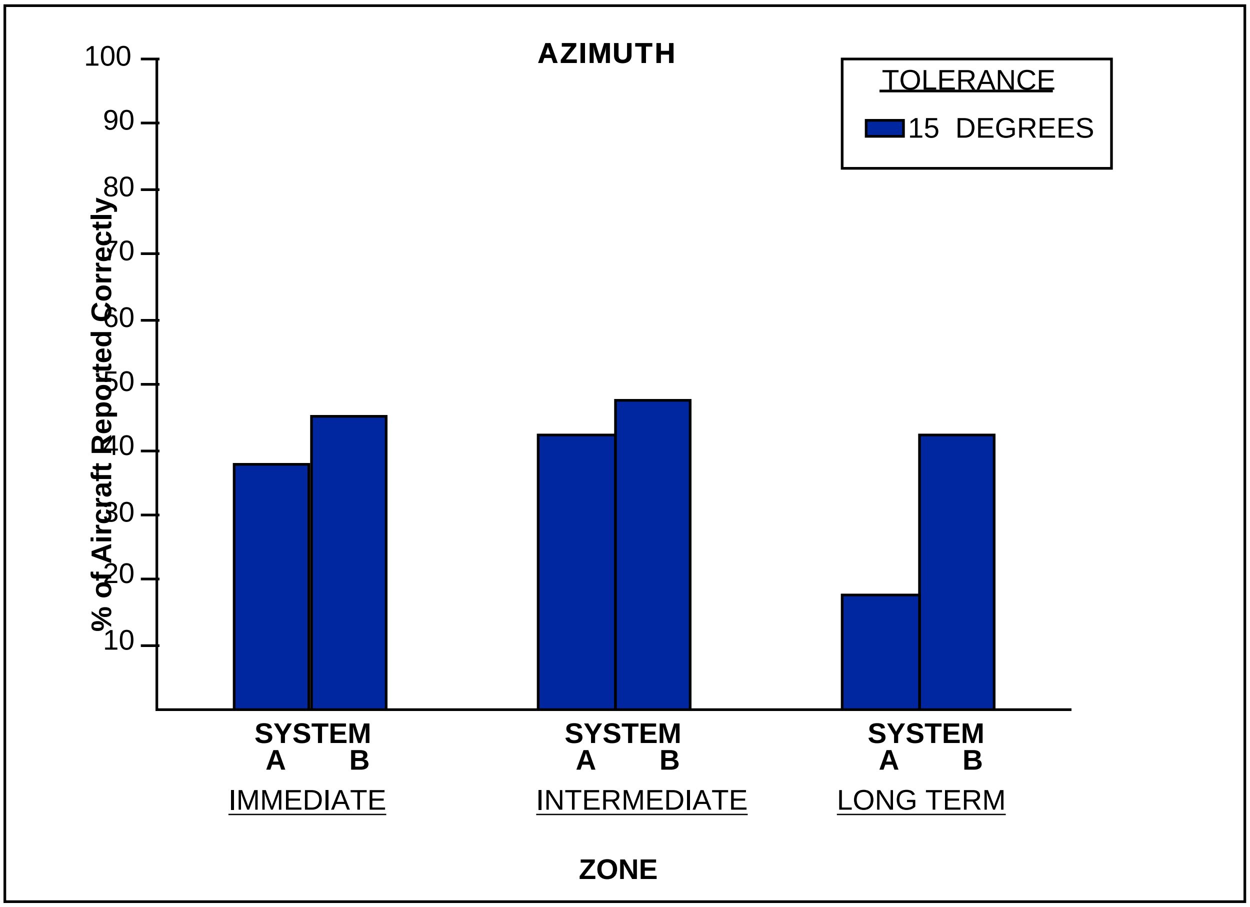

Figure 3 Example of SAGAT results: Knowledge of Aircraft Location (by Azimuth) for Two Avionics Systems (Endsley, 1988b)

Evaluation of Free Flight

In another example, a totally new form of distributing roles and responsibilities between pilots and air traffic controllers was examined. Termed “free flight”, this concept was originally described to incorporate major changes in the operation of the national airspace. It may include aircraft filing direct routes to destinations rather than along pre-defined fixed airways, and the authority for the pilot to deviate from that route, either with the air traffic controllers permission or perhaps even fully autonomously (RTCA, 1995). As it was felt that such changes could have a marked effect on the ability of the controller to keep up as monitor in such a new system, a study was conducted to examine this possibility (Endsley, Mogford, Allendoerfer, Snyder, & Stein, 1997).

Results showed a trend towards poorer controller performance in detecting and intervening in aircraft separation errors with these changes in the operational concept and poorer subjective ratings of performance. Finding statistically significant changes in separation errors during ATC simulation testing is quite rare however. More detailed analysis of the SAGAT results provided more diagnostic detail as well as backing up this finding. As shown in Figure 4,

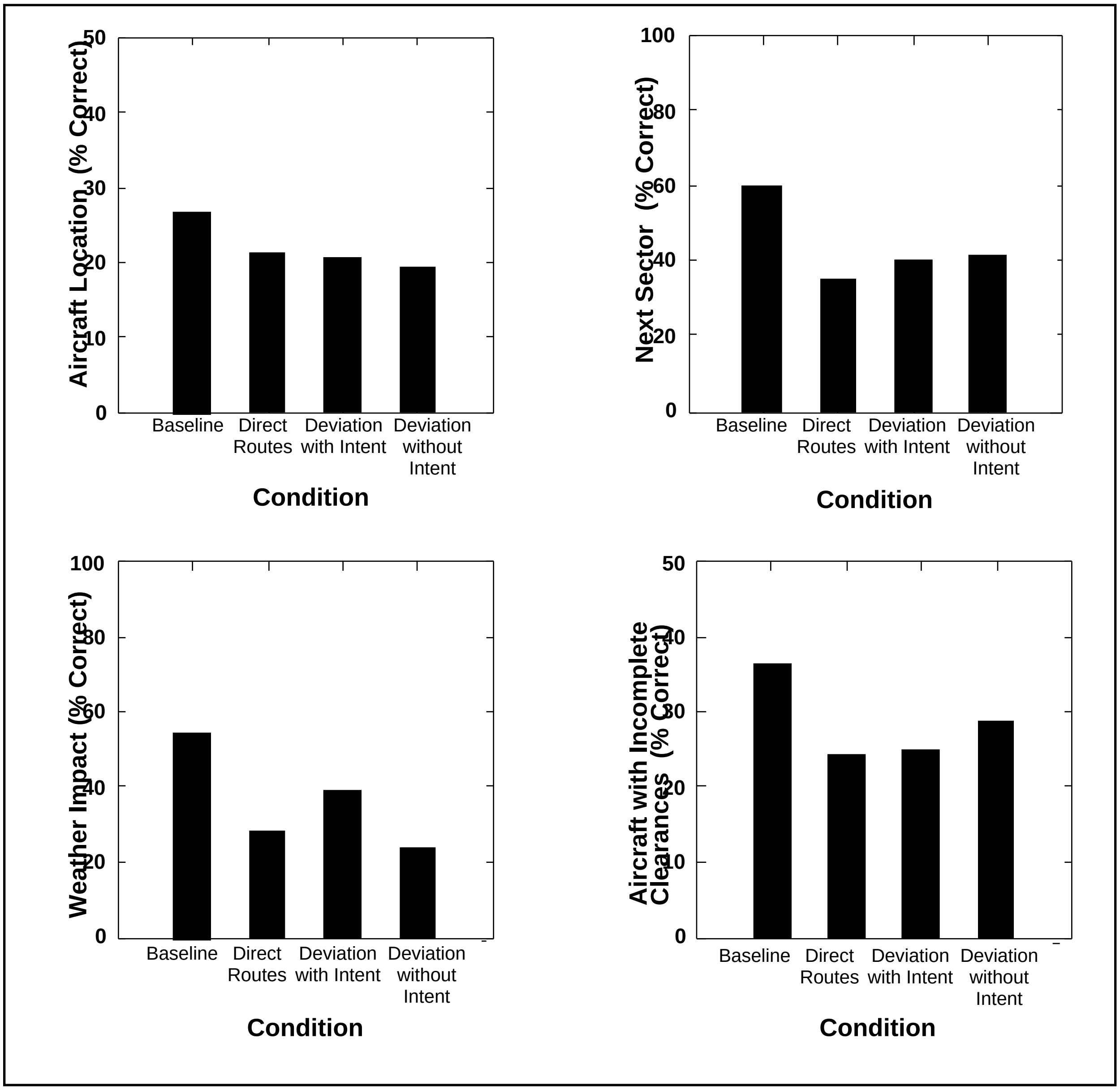

Figure 4 Example of SAGAT Results: Free Flight Implementations (Endsley et al, 1997)

controllers were aware of significantly fewer of the aircraft in the simulation under free flight conditions. Attending to fewer aircraft under higher workload has also been found in other studies

(Endsley & Rodgers, 1998). In addition to reduced Level 1 SA, however, controllers also had a significantly reduced understanding (Level 2 SA) of what was happening in the traffic situation, as evidenced by lower SA regarding which aircraft weather would impact on and a reduced awareness of those aircraft that were in a transitionary state. They were less aware of which aircraft had not yet completed a clearance, and for those aircraft whether it was received correctly and whether they were conforming. Controllers also demonstrated lower Level 3 SA with flight. Their knowledge of where the aircraft was going to (next sector) was significantly lower under free flight conditions.

These findings were useful in pinpointing whether concerns over this new and very different concept were justified, or whether they merely represented resistance to change. The SAGAT results showed not only that the new concept did indeed induce problems for controller SA that would prevent them from performing effectively as monitors to back-up pilots with separation assistance, and also showed in what ways these problems were manifested. This information is very useful diagnostically in that it allows one to determine what sorts of aids might be needed for operators to assist them in overcoming these deficiencies. For instance, in this example, a display which provides enhanced information on flight paths for aircraft in transitionary states may be recommended as a way of compensating for the lower SA observed. Far from just providing a thumbs-up or thumbs-down input on a concept under evaluation, this rich source of data is very useful in developing iterative design modifications and making tradeoffs decisions.

Evaluation of Automation Concepts

SAGAT has also been used to provide information on questions of more scientific interest. For instance, it has long been observed that operators who act as monitors of automated systems may become “out-of-the-loop” and less able to take-over manual control when needed. This was hypothesized to be due partially to a lessening of manual skills, but primarily due to a loss of situation awareness through both decreased vigilance and lower understanding associated with passive monitoring (Endsley, 1987a; 1995a). Applying SAGAT at the end of short simulation involving an expert system for aiding in automobile navigation, Endsley and Kiris (1995a) found that Level 2 SA was significantly lower when subjects were operating under full automation, as compared to several intermediate levels of automation and purely manual control, as shown in Figure 5. This lower level of SA corresponded to the predicted increase in time required to take-over manual performance following an automation failure. Porter (1996) found a similar result with lower SA under high automation than low automation in a higher fidelity simulation involving an air traffic control tool called Center TRACON Automation System (CTAS). These studies demonstrate that SAGAT can help address basic issues involving questions about observed changes in human performance with certain types of system implementations. They also may direct us towards solutions to these problems, such as the level of automation manipulation examined in this study

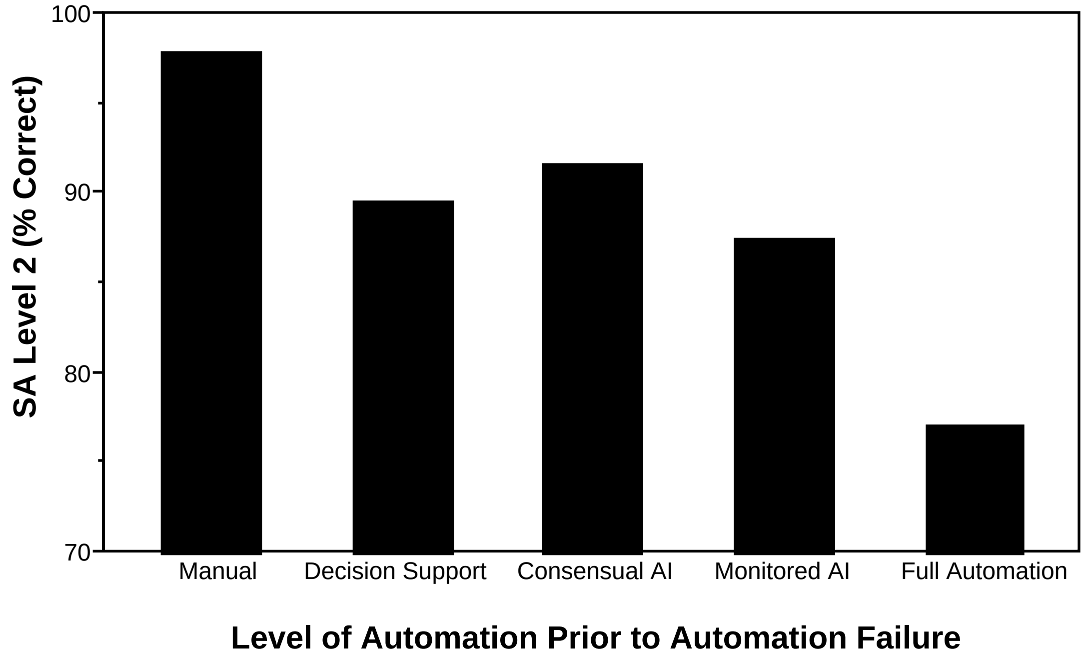

Figure 5 Example of SAGAT Results: Level of Automation (Endsley and Kiris, 1995a)

Investigations of Individual Differences in SA

In some cases, SA itself is the subject of the investigation, rather than merely an index of merit for some other factor being investigated. In these studies a direct measure of SA is highly desirable. For instance, Endsley and Bolstad (1994) used mean SAGAT scores across a number of simulation trials as an index of individual pilot SA ability. This measure provided useful index of SA that was independent of simple experience measures, which team the individual flew on (the effect of team-mates) and which aircraft system they had been provided. It showed a ten to one ratio in mean SA score across the group of experienced fighter pilots studied. While the number of subjects was limited (21), this study was able to find significant correlates to SA ability as measured by SAGAT, including attention sharing, spatial skills, perceptual speed and pattern matching ability. In another experiment, Collier and Folleso (1995) reported on the use of a technique adapted after SAGAT for use in a nuclear power plant. They found consistent differences in the level of SA measured in this manner between two operators with very different levels of expertise. Studies such as these demonstrate the utility of measuring SA objectively when performing studies that seek to investigate the relationship between SA and other cognitive measures of interest.

Issues of Validity and Reliability of SAGAT

The SAGAT technique has thus far been shown to have a high degree of validity, sensitivity and reliability for measuring SA.

Sensitivity

The examples referenced here show that SAGAT has good sensitivity to system manipulations, automation manipulations, expertise differences and operational concepts in a variety of system domains. Vidulich (Chapter 13) found good sensitivity for the technique across a wide range of studies, as long as a broad range of queries is used.

Criterion Validity

SAGAT has also been shown to have predictive validity, with SAGAT scores indicative of pilot performance in a combat simulation (Endsley, 1990b). This study found that fighter pilots who were able to report on an enemy aircraft's existence via SAGAT were three times more likely to latter kill that target in the simulation.

Responses to SA queries also appear to be sensitive as expected to features of the task situation. Fracker (1990) reported that SA measured using a freeze methodology was sensitive to the importance of the aircraft (degree of threat) and the number of enemy aircraft present in a low-fidelity task. Gronlund , et. al. (in press) report that memory probes are sensitive to the importance of the aircraft in a high fidelity air traffic control task. Endsley and Smith (1996) similarly found that recall accuracy for aircraft location on a tactical situation display by experienced fighter pilots was significantly related to the aircraft's importance to their decision tasks. Gugerty (1997) found that the percentage of cars correctly reported decreased with increases in load (number of cars) in a driving task, with attention focused on a subset based on the car's importance to the driver. Endsley and Rodgers (1998) found that the percentage of aircraft reported by experienced controllers via SAGAT significantly decreased with increases in the number of aircraft in an air traffic control simulation. For those aircraft controllers were aware of, they found that controllers were significantly less accurate on most other factors related to the aircraft as the number of aircraft present increased; however controllers interestingly preserved their knowledge of aircraft separation under load. These findings indicate that this type of measure is sensitive to changes in taskload and to factors that effect operator attention.

Reliability

Measurement reliability has been demonstrated in a study that found high reliability (test-retest scores of .98, .99, .99 and .92) of mean SAGAT scores for four fighter pilots who participated in two sets of simulation trials (Endsley & Bolstad, 1994). Fracker (1991) reported low reliability for his measure of aircraft location, however, this may be reflective of the absolute error score used in his test, as well as the low-fidelity of the simulation and use of in-experienced subjects. Collier and Folleso (1995) reported good reliability for their measure involving two experienced nuclear power plant operators. Gugerty (1997) also reported good reliability for the percentage of cars recalled, recall error and composite recall error (even-odd reliabilities of .93, .92 and .96) in his study involving a driving task. In general, these results support the reliability of the measure.

Construct Validity

Probably the biggest two issues of concern involving SAGAT have centered around the perceived intrusiveness of freezes in a simulation to collect SAGAT data, and the degree to which it reflects memory and as such is limited. Each of these issues will be addressed separately.

Intrusiveness

Several studies have shown that a temporary freeze in a simulation to collect SAGAT data does not impact performance. Endsley (1990a; 1995a) reported on a study in which one, two or three freezes were introduced to collect data for durations of 30 seconds, one minute or two minutes to collect SAGAT data. This study found no impact of the stops on subject performance measures. Other studies have used SAGAT in some trials, but not others in order to examine whether SAGAT interfered with subject performance (Bolstad & Endsley, 1990; Endsley, 1989b; Northrop, 1988). No effect was found in any of these studies. Subjectively, the pilots in the studies appeared to adjust to the technique quite well and were able to freeze and return to the action fairly readily. Hogg, Folleso, Torralba and Volden (1993) also reported that power plant operators in their study subjectively reported no effect from the freezes and considered it similar to their training exercises.

Despite the conflicting evidence of these studies, a concern over the possibility of intrusiveness has been voiced (Sarter & Woods, 1991), which continues to be repeated. In an effort to continue to dispel these concerns, another study was conducted on the issue of intrusiveness. This study investigated whether operator performance could be effected by merely the threat of a stop to collect SAGAT data. That is, are operators somehow altering their behavior during simulation trials in which they feel they may be stopped and tested on their SA? To answer this question, a study was conducted so that performance on trials in which subjects were told that only performance would be measured could be compared to trials in which subjects were told that a stop to collect SAGAT data might occur. In the later case, SAGAT stops occurred only half of the time. Any effect of the actual SAGAT stop could therefore be differentiated from merely the threat of the stop, and compared to trials in which subjects knew they would not be stopped.

A set of trials was conducted of an air-to-air fighter sweep mission. The subject, flying as the pilot of single aircraft, was to penetrate enemy territory, maximizing kills of enemy fighters while maintaining a high degree of survivability. Four computer controlled aircraft were the adversaries in these engagements. Subject instructions were manipulated during the test. In one-third of the trials, subjects were told that only performance would be measured. In the other two-thirds of the trials, subjects were told that there might be a stop to collect SAGAT data in addition to performance measurement. Half of these trials actually were stopped once at a random point in the trial for two minutes to collect SAGAT data. Half were not stopped. Each of six subjects completed five trials in each of the three conditions: no stop/none expected, no stop/stop expected, stop/stop expected. The conditions were presented in a random order. A total of 90 trials were completed. Pilot performance in terms of kills and losses was collected as the dependent measure.

The test was conducted using a medium fidelity mission simulation on a Silicon Graphics 4D-220 computer. The system had a high-resolution, 19" color display monitor and realistic stick and throttle controls. A simulated head-up display, tactical situation display, vertical situation display, fuel gage and thrust gage were provided. Six subjects participated in the test. The subjects were all experienced former military fighter pilots. The mean subject age was 43.6 years (range of 33 to 57). They had an average of 2803 hours (range of 1500 to 3850) and an average of 15.2 years (range of 7 to 25) of military flight experience. Two of the six subjects had combat experience.

Analysis of variance was used to evaluate the effect of the test condition (no stop/not expected, no stop/stop expected, and stop/stop expected) on each of the two performance measures: aircraft kills and losses. The test condition had no significant impact on either performance measure, $F(2, 87) = .15$, $p = .861$, $F(2, 87) = 1.53$, $p = .223$, shown in Figure 6. In viewing the data, it can be seen that the number of kills was almost identical, independent of whether subjects expected a stop or not and independent of whether they actually experienced a stop. While the subject died slightly more often in the trials where they expected a stop but did not receive one, this difference was not significant. This data supports the null hypothesis, indicating that a stop or even the threat of a stop to collect SAGAT data does not have a significant impact on performance.

The results of this study confirm previous findings which have not found a demonstrable effect on performance of freezes in a simulation to collect SAGAT data. It furthermore expands on these studies to reveal that even the threat of a stop does not significantly impact performance. Subjective comments by the subjects after the study confirm this. They reported that the information about whether to expect a SAGAT stop was irrelevant to them. At least on a conscious level, they were not preparing in any way for the SAGAT test. The results of this study indicate that they were not doing so unconsciously either. These results are also useful in that the “opponents” in the trial were digital aircraft (computer controlled). This eliminates the possibility that performance effects may have been masked by the freezes affecting both the red team and the blue team equally.

Overall, the results of this study indicate that using SAGAT to collect data on situation awareness is not intrusive on subject performance, and therefore provides an additional indication of the validity of the method for directly measuring subject SA during simulations. While it is never possible to "prove" the null hypothesis, that SAGAT does not affect performance, all of the studies collected so far indicate that it does not appear to significantly affect performance, as long as the stops are unpredictable to the subject (Endsley, 1988b). Based on the utility that has been found with the technique, it is believed that any such risk is well worth the information that is collected.

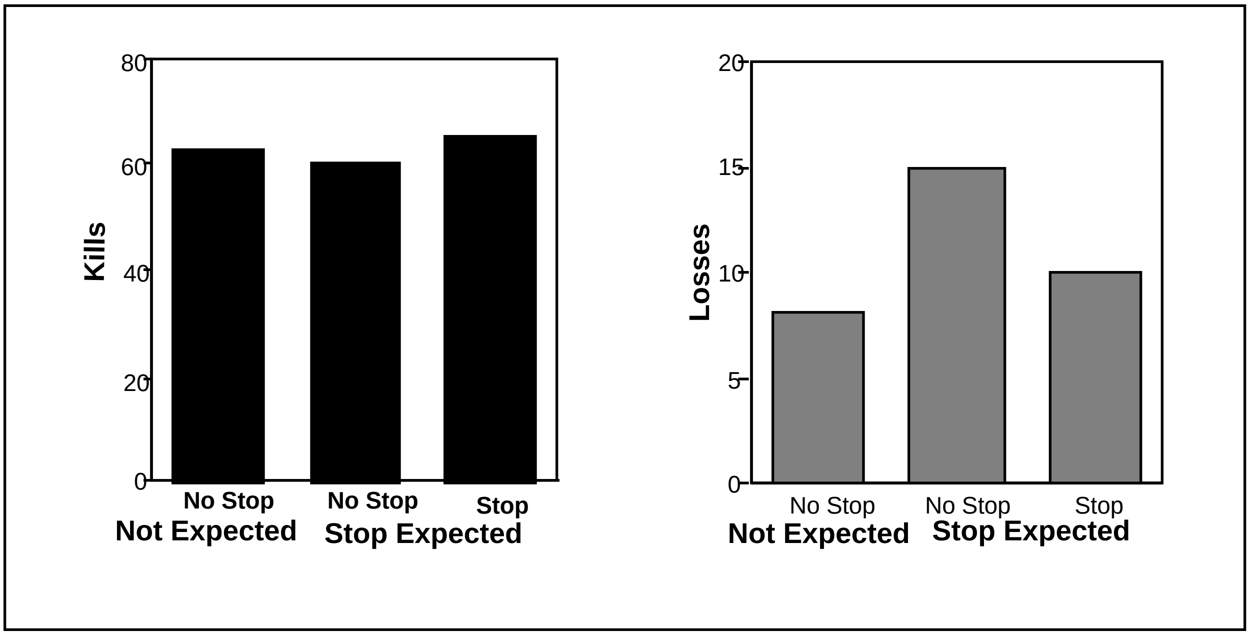

Figure 6 Pilot Performance in Simulations as a Function of SAGAT Freezes and Expectations

Memory

Retrospective versus concurrent memory. A second issue relevant to the use of SAGAT pertains to whether it provides a good representation of the operator's SA or whether it is hindered by being dependent on memory. Addressing this issue depends on carefully examining the construct of memory. First it has been widely reported that retrospective memory for past cognitive events is rather poor (Nisbett & Wilson, 1977). A technique which is dependent on retrospective memory would clearly represent a problem for SA measurement. SAGAT, however, seeks to tap into concurrent memory by placing the queries immediately following the events occurrence.

In addition to thinking of memory as a retrospective phenomenon, the concept of working memory is also pertinent to the active manipulation and use of information. Baddeley's (1986) model of working memory, for instance, includes a central executive for processing (comparing, manipulating, combining) information along with a visual-spatial scratchpad and a verbal-auditory loop for retaining information. In this sense, working memory is where dynamic information from the environment would be resident. It would make sense therefore that the contents of working memory are exactly what needs to be tapped by an SA measure. This information would need to be obtained immediately, however, as memory decay can be an issue.

Access to working memory and long-term memory. In addition, experiments have shown that SA is not solely dependent on working memory. Endsley (1988a; 1995b) and Fracker (1988) have both emphasized the importance of long-term memory stores for SA, particularly for experienced operators. Hartman and Secrist (1991) talk about a concept of skilled memory, positing access to long term memory stores relevant to SA. Endsley (1990a; 1995a) found that pilots' ability to report their SA via SAGAT was unaffected by how long after the freeze the question was asked, testing intervals from around 20 seconds to up to 6 minutes. This manipulation was obtained by providing the queries in a random order. Thus pilots were answering SAGAT questions during the entire interval, however, the timing of a particular query could be altered and compared to accuracy on the same query provided at different intervals across subjects and freezes. This study showed that SA information was available for quite some time after a freeze (when other conflicting tasks or information are not introduced as interference), indicating long term memory has a role in SA as well as working memory.

Two explanations are offered for these findings. First, this study used expert subjects performing a realistic task. Most laboratory studies of working memory employ stimuli with little or no inherent meaning to the subjects. The storage and utilization of task relevant information has been found to be very different from irrelevant information (Chase & Simon, 1973). Information has been found to be more retrievable when it is processed effortfully and with awareness (Cowan, 1988), as would be case with a meaningful task. It is also likely that experienced operators have long-term memory stores (such as schema and mental models) that serve to organize information, having an effect on its availability for a measure such as SAGAT. By contrast, Fracker (1991) found very poor performance on SA queries in his tests which probed student subjects performing a very artificial computer task.

These results might also be most parsimoniously explained by a more integrated model of working memory and long-term memory. Cowan (1988) proposed such a model in which working memory is depicted as an activated subset of long-term memory. The focus of attention is represented within that activated memory, controlled by the central executive. Information can become activated either effortfully through that focus of attention or habituated information may by-pass that focus. Spontaneous activation of long-term memory based on associations may also bring information into activation. Durso and Gronlund (in press) also conclude that an integrated model of working memory and long-term memory might best explain research on SA and memory. They draw on a model by Ericsson and Kintsch (1995) in which working memory contains pointers to information in long-term memory. Durso and Gronlund conclude that working memory may be a constraint on the processing of information to form SA (via the central executive for instance), but not to the storage of SA.

Explicit versus implicit memory. In terms of SAGAT and the measurement of SA via queries, these results indicate that SA information should be retrievable for a short period of time during a simulation freeze via some combination of working and long term memory stores. The situational representation of the operator should be reportable. This forms an explicit measure of memory. Other research has posited the idea of implicit memory — information which is available to affect performance, but which operators cannot report (Lewandowsky, Dunn, & Kirsner, 1989). Gugerty (1997) specifically addressed this issue in a driving task. He compared subject responses to recall probes regarding other cars (a measure of explicit SA) to subject performance on crash avoidance and detection of blocking cars (measures of implicit SA). He found that both the recall measures and the performance-based measures were reliable and significantly correlated with each other. Gugerty concluded that the two measures tapped the same knowledge base, supporting their validity, and agreeing with researchers who have found the distinction between implicit and explicit memory to be artificial and based on insensitive measurement.

A final concern, revolves around whether information even needs to be in the subject's memory (or mental representation) in the first place. Perhaps it is sufficient that they merely reference the information on their displays when needed (Durso, Hackworth, Truitt, Crutchfield, Nikolic, & Manning, 1998). If this were the case, SA, as a dynamic mental representation of the environment, would not be needed. To address this question, I will leave an experiment to the reader. The next time you are pulled over by a policeman for speeding (you may have to break the law to carry out this experiment), and he (or she) asks “Do you know how fast you were going?”, your reply should be “No, but it doesn’t matter because I can look at my speedometer any time I want”. My hypothesis is that the policeman will not be amused by this argument. Getting information to a display is not tantamount to providing an operator with SA (or with effective performance). One of the most frequent causal factors for SA errors involves situations in which all the information is available to the operator, however it is not attended to for various reasons (Jones & Endsley, 1996). The real challenge in today’s systems is the ability of the operator to dynamically locate and integrate needed information from the plenitude that is available in the system. For this reason, SA, as a measure of the operator’s ability to form the needed understanding of the status a dynamic system under the constraints of many opposing demands, is the measure of interest.

In summary, these studies indicate that SA measures such as SAGAT can reliably tap into memory stores (either working memory or long-term memory) as an index of SA. They do not appear to be hampered by problems of retrospective memory recall or implicit memory constraints.

IMPLEMENTATION RECOMMENDATIONS

Several recommendations for SAGAT administration have been made based on previous experience in using the procedure (Endsley, 1995a).

Training

An explanation of SAGAT procedures and detailed instructions for answering each query should be provided to subjects before testing. The idea of using a “surprise” to collect this data will probably only work once, and may be hampered by misunderstandings regarding the questions themselves. For this reason, it is usually better to thoroughly brief participants. Several training trials should be conducted in which the simulator is halted frequently to allow subjects ample opportunity to practice responding to the SAGAT queries. Usually three to five samplings are adequate for a subject to become comfortable with the procedure and to clear up any uncertainties in how to answer the queries.

Test Design

SAGAT requires no special test considerations. The same principles of experimental design and administration apply to SAGAT as to any other dependent measure. Measures of subject performance and workload may be collected concurrently with SAGAT, as no ill effect from the insertion of breaks has been shown. To be cautious, however, half of the trials may be conducted without any breaks for SAGAT so that a check is provided for this contingency.

Procedures

Subjects should be instructed to attend to their tasks as they normally would, with the SAGAT queries considered as secondary. No displays or other visual aids should be visible while subjects are answering the queries. If subjects do not know or are uncertain about the answer to a given query, they should be encouraged to make their best guess. There is no penalty for guessing, allowing for consideration of the default values and other wisdom gained from experience that subjects normally use in decision making (e.g. embedded schema information). By giving subjects “credit” for this type of knowledge, the true benefits of systems which seek to supplant that knowledge can be more fairly assessed. If subjects do not feel comfortable enough to make a guess, they may go on to the next question. Talking or sharing of information between subjects should not be permitted. If multiple subjects are involved in the same simulation, all subjects should be queried simultaneously and the simulation resumed for all subjects at the same time. (Later comparisons of team members' SAGAT data can also be used to assess the degree to which they possessed a common picture of the situation — shared SA.)

Which Queries to Use

Random selection.

As it may be impossible to query subjects about all of their SA requirements in a given stop due to time constraints, a portion of the SA queries may be randomly selected and asked each time. A random sampling provides consistency and statistical validity, thus allowing SA scores to be easily compared across trials, subjects, systems and scenarios.

Due to attentional narrowing or a lack of information, certain questions may seem unimportant to a subject at the time of a given stop. It is important to stress that they should attempt to answer all queries anyway. This is because (a) even though they think it unimportant, the information may have at least secondary importance, (b) they may not be aware of information that makes a question very important (e.g. the presence of a pop-up aircraft), and (c) if only questions of the highest priority were asked, subjects might be inadvertently provided with artificial cues about the situation that will direct their attention when the simulation is resumed. Therefore a random selection from a constant set of queries is recommended at each stop.

Experimenter controlled.

In certain tests it may be desirable to have some queries omitted, due to limitations of the simulation or characteristics of the scenarios. For instance if the simulation does not incorporate aircraft malfunctions, the query related to this issue may be omitted. In addition, with particular test designs it may be desirable to insure that certain queries are presented every time. When this occurs, it is important that subjects also be queried on a random sampling from all SA requirements and not just on those related to a specific area of interest to the evaluation being conducted. This is due to the ability of subjects to shift attention to the information they know they will be tested on. What may appear to be an improvement in SA in one area may just be a shift of attention from one aspect of the situation to another. When the SAGAT queries cover all of the SA requirements, no such artificial cueing can occur and attention shifts can be detected.

When to Collect SAGAT Data

It is recommended that the timing of each freeze for SAGAT administration be randomly determined and unpredictable enough so that subjects can not prepare for them in advance. If the freeze occurrence is associated with the occurrence of specific events, or at specific times across trials, prior studies have shown that the subjects will be able to figure this out (Endsley, 1988b), allowing them to prepare for them or actually improve SA through the artificiality of the freeze cues. An informal rule has been to insure that no freezes occur earlier than three to five minutes into a trial to allow subjects to build up a picture of the situation and that no two freezes occur within one minute of each other.

The result of this approach is that the activities occurring at the time of the stops will be randomly selected. Some stops may occur during very important activities that are of interest to the experimenter, others when no critical activities are occurring. This gives a good sampling of the subjects' SA in a variety of situations During analysis the experimenter may want to stratify the data to take these variations into account.

How Much SAGAT Data to Collect

The number of trials necessary will depend upon the variability present in the dependent variables being collected and the number of data samples taken during a trial. This will vary with different subjects and designs, but between 30 and 60 samplings per SA query (across subjects and trials) with each design option have previously been adequate in a within subjects test design.

Multiple SAGAT stops may be taken within each trial. There is no known limit to the number of times the simulator can be frozen during a given trial. No ill effects have been found for as many as three stops during a 15 minute trial (Endsley, 1995a). In general, it is recommended that a freeze last until a certain amount of time has elapsed and then the trial is resumed, regardless of how may questions have been answered. Freezes as long as two minutes in duration were used with no undue difficulty or effect on subsequent performance. Freezes as long as five to six minutes have been shown to allow subjects access to SA information without memory decay in experiment one (Endsley, 1995a).

Data Collection

The simulator computer should be programmed to collect objective data corresponding to the queries at the time of each freeze. Since some queries will pertain to higher level SA requirements that may be unavailable from the computer, an expert judgment of the correct answer may be made by an experienced observer who is privy to all information, reflecting the SA of a person with perfect knowledge. A comparison of the subjects' perceptions of the situation (as input into SAGAT) to the actual status of each variable (as collected per the simulator computer and expert judgment) results in an objective measure of subject SA. Questions asked of the subject but not answered should be considered incorrect. No evaluation should be made of questions not asked during a given stop.

It is recommended that answers to each query be scored as correct or incorrect based upon whether it falls into an acceptable tolerance band around the actual value. For example, it may be acceptable for a subject to be 10 MPH off of actual groundspeed. This method of scoring poses less difficulty than dealing with absolute error (see Marshak, Kuperman, Ramsey, & Wilson, 1987). A tabulation of the frequency of correctness can then be made within each test condition for each SA element. As data scored as correct or incorrect are binomial, the conditions for analysis of variance are violated. A correction factor (Y’ = arcsine (Y)) can be applied to binomial data, however, which allows analysis of variance to be used. In addition, a chi-square, Cochran's Q, or binomial t-test (depending on the test design) can be used to evaluate the statistical significance of differences in SA between test conditions.

Finally, it is recommended that each query be evaluated separately, rather than in some combined form. In general, analysis has not supported the combination of SAGAT queries into a single measure, or into groups by SA level. SA on each query can be independent of the others, as reflections of shifts in operator attention in association with various display manipulations and as dictated by the demands of their tasks. Combinations of queries are likely to reduce sensitivity of the metric by losing these important distinctions.

Limitations and Applicability for Use

This technique has primarily been used within the confines of high-fidelity and medium-fidelity part-task simulations. This provides experimenter control over freezes and data collection without any danger to the subject or processes involved in the domain. It may be possible to use the technique during actual task performance if multiple operators are present to insure safety. For example, it might be possible to verbally query one pilot in flight while another assumes flight control. Such an endeavor should be undertaken with extreme caution, however, and may not be appropriate for certain domains.

Sheehy (1993) employed an adaptation of this technique by making video-tapes of an ongoing situation in a nuclear power plant control room. These tapes were then replayed to naive subjects with freezes for SAGAT queries employed. It is not known how different the SA of subjects passively viewing a situation may be from subjects actually engaged in task performance, however this approach may yield some useful data. Endsley and Rodgers (1998), for instance, used computer generated replays of air traffic control errors to study controller SA. Similar to Sheehy, their study involved passive viewing of the displays. None-the-less, they were able to examine shifts in controller SA that corresponded to workload changes and which were indicative of SA problems as experienced in real time operations.

Conclusions

To date, SAGAT has been used for a wide variety of system evaluations and construct investigations. in such diverse applications as aircraft, ATC, driving, nuclear power operations and teleoperations. Perhaps the widely tested measure of SA, it has been shown to have good levels of sensitivity, reliability and predictive validity.

References

Baddeley, A. D. (1986). Human memory. Oxford: Clarendon Press.

Bolstad, C. A., & Endsley, M. R. (1990). Single versus dual scale range display investigation (NOR DOC 90-90). Hawthorne, CA: Northrop Corporation.

Chase, W. G., & Simon, H. A. (1973). Perceptions in chess. Cognitive Psychology, 4, 55-81.

Collier, S. G., & Folleso, K. (1995). SACRI: A measure of situation awareness for nuclear power plant control rooms. In D. J. Garland & M. R. Endsley (Eds.), Experimental analysis and measurement of situation awareness (pp. 115-122). Daytona Beach, FL: Embry-Riddle University Press.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. Psychological Bulletin, 104(2), 163-191.

Durso, F. T., & Gronlund, S. D. (in press). Situation awareness. In F. T. Durso, R. Nickerson, R. Schvaneveldt, S. Dumais, M. Chi, & S. Lindsay (Eds.), Handbook of applied cognition. New York: Wiley.

Durso, F. T., Hackworth, C. A., Truitt, T. R., Crutchfield, J., Nikolic, D., & Manning, C. A. (1998). Situation awareness as a predictor of performance for en route air traffic controllers. Air Traffic Control Quarterly, 6(1), 1-20.

Endsley, M. (1987a). The application of human factors to the development of expert systems for advanced cockpits. In Proceedings of the Human Factors Society 31st Annual Meeting (pp. 1388-1392). Santa Monica, CA: Human Factors Society.

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

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

Endsley, M. R. (1988b). Situation Awareness Global Assessment Technique (SAGAT). In Proceedings of the National Aerospace and Electronics Conference (NAECON) (pp. 789-795). New York: IEEE.

Endsley, M. R. (1989a). Final report: Situation awareness in an advanced strategic mission (NOR DOC 89-32). Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (1989b). Tactical simulation 3 test report: Addendum 1 situation awareness evaluations (81203033R). Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (1990a). A methodology for the objective measurement of situation awareness. In Situational Awareness in Aerospace Operations (AGARD-CP-478) (pp. 1/1 - 1/9). Neuilly Sur Seine, France: NATO - AGARD.

Endsley, M. R. (1990b). Predictive utility of an objective measure of situation awareness. In Proceedings of the Human Factors Society 34th Annual Meeting (pp. 41-45). Santa Monica, CA: Human Factors Society.

Endsley, M. R. (1990c). Situation awareness in dynamic human decision making: Theory and measurement. Unpublished doctoral dissertation, University of Southern California, Los Angeles, CA.

Endsley, M. R. (1993). A survey of situation awareness requirements in air-to-air combat fighters. International Journal of Aviation Psychology, 3(2), 157-168.

Endsley, M. R. (1994). Situation awareness in FAA Airway Facilities Maintenance Control Centers (MCC): Final Report. Lubbock, TX: Texas Tech University.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. Human Factors, 37(1), 65-84.

Endsley, M. R. (1995b). Toward a theory of situation awareness. Human Factors, 37(1), 32-64.

Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot situation awareness. International Journal of Aviation Psychology, 4(3), 241-264.

Endsley, M. R., Farley, T. C., Jones, W. M., Midkiff, A. H., & Hansman, R. J. (1998). Situation awareness information requirements for commercial airline pilots (ICAT-98-1). Cambridge, MA: Massachusetts Institute of Technology International Center for Air Transportation.

Endsley, M. R., & Jones, D. G. (1995). Situation awareness requirements analysis for TRACON air traffic control (TTU-IE-95-01). Lubbock, TX: Texas Tech University.

Endsley, M. R., & Kiris, E. O. (1995a). The out-of-the-loop performance problem and level of control in automation. Human Factors, 37(2), 381-394.

Endsley, M. R., & Kiris, E. O. (1995b). Situation awareness global assessment technique (SAGAT) TRACON air traffic control version user guide. Lubbock, TX: Texas Tech University.

Endsley, M. R., Mogford, R., Allendoerfer, K., Snyder, M. D., & Stein, E. S. (1997). Effect of free flight conditions on controller performance, workload and situation awareness: A preliminary investigation of changes in locus of control using existing technology (DOT/FAA/CT-TN 97/12). Atlantic City, NJ: Federal Aviation Administration William J. Hughes Technical Center.

Endsley, M. R., & Robertson, M. M. (1996). Team situation awareness in aircraft maintenance. Lubbock, TX: Texas Tech University.

Endsley, M. R., & 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.

Endsley, M. R., & Rodgers, M. D. (1998). Distribution of attention, situation awareness, and workload in a passive air traffic control task: Implications for operational errors and automation. Air Traffic Control Quarterly, 6(1), 21-44.

Endsley, M. R., & Smith, R. P. (1996). Attention distribution and decision making in tactical air combat. Human Factors, 38(2), 232-249.

Ericsson, K. A., & Kintsch, W. (1995). Long term working memory. Psychological Review, 102, 211-245.

Fracker, M. L. (1988). A theory of situation assessment: Implications for measuring situation awareness. In Proceedings of the Human Factors Society 32nd Annual Meeting (pp. 102-106). Santa Monica, Ca: Human Factors Society.

Fracker, M. L. (1990). Attention gradients in situation awareness. In Situational Awareness in Aerospace Operations (AGARD-CP-478) (Conference Proceedings #478) (pp. 6/1-6/10). Neuilly Sur Seine, France: NATO - AGARD.

Fracker, M. L. (1991). Measures of situation awareness: an experimental evaluation (AL-TR-1991-0127). Wright-Patterson AFB, OH: Armstrong Laboratory, U. S. Air Force.

Gronlund, S. D., Ohrt, D. D., Dougherty, M. R. P., Perry, J. L., & Manning, C. A. (1998). Role of memory in air traffic control. Journal of Experimental Psychology: Applied), 4, 263-280.

Gugerty, L. J. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. Journal of Experimental Psychology: Applied, 3, 42-66.

Hartman, B. O., & Secrist, G. E. (1991). Situational awareness is more than exceptional vision. Aviation, Space and Environmental Medicine, 62, 1084-9.

Hogg, D. N., Torralba, B., & Volden, F. S. (1993). A situation awareness methodology for the evaluation of process control systems: Studies of feasibility and the implication of use (1993-03-05). Storefjell, Norway: OECD Halden Reactor Project.

Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. Aviation, Space and Environmental Medicine, 67(6), 507-512.

Lewandowsky, S., Dunn, J. C., & Kirsner, K. (Ed.). (1989). Implicit memory: theoretical issues. Hillsdale, NJ: LEA.

Marshak, W. P., Kuperman, G., Ramsey, E. G., & Wilson, D. (1987). Situational awareness in map displays. In Proceedings of the Human Factors Society 31st Annual Meeting (pp. 533-535). Santa Monica, CA: Human Factors Society.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84(3), 231-259.

Northrop Corporation. (1988). Tactical simulation 2 test report: Addendum 1 situation awareness test results. Hawthorne, CA: Author.

Porter, A. W. (1996). Investigating the effects of automation on situation awareness (DRA/LS(LSC4)/CHCI/CD299/1.0). Malvern, UK: Defense Research Agency.

RTCA (1995). Report of the RTCA Board of Directors select committee on free flight. Washington, D.C.: Author.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. The International Journal of Aviation Psychology, 1(1), 45-57.

Sheehy, E. J., Davey, E. C., Fiegel, T. T., & Guo, K. Q. (1993, April). Usability benchmark for CANDU annunciation - lessons learned. Paper presented at the ANS Topical Meeting on Nuclear Plant Instrumentation, Control and Man-Machine Interface Technology, Oak Ridge, TN.