

Real-Time Assessment of Situation Awareness of Air Traffic Control Specialists On Operational Host Computer System and Display System Replacement Hardware

Ben F. Willems, M.A.,
FAA ACT-530, Pomona, NJ, USA Ben.Willems@tc.faa.gov

Michele Heiney, Ph.D.,
Federal Data Corporation, Egg Harbor Township, NJ, USA Michele.CTR.Heiney@tc.faa.gov

Abstract

Improving the current Air Traffic Control (ATC) system with new technologies such as decision support systems (DSTs) is necessary to accommodate continued growth in air traffic. These new technologies will potentially move ATC Specialists (ATCSs) from tactical to more strategic control. How this change will affect ATCS situation awareness (SA) is unknown. To assess SA real-time on operational hardware and software is both a technical and human factors challenge.

This paper describes the measurement of SA, including a new instrument, in an experiment that investigated the effect of traffic level, automation, and ATCS position. Sixteen ATCSs from Air Route Traffic Control Centers (ARTCCs) within the United States volunteered. We attached the Situation Awareness Verification and ANalysis Tool (SAVANT) to hardware used in operational facilities and probed ATCS real-time.

The results of the experiment showed that the implementation of an online SA assessment instrument on operational hardware is feasible. ATCS position, load, and level of automation had an effect on ATCS SA. The results have implications for the introduction of new automation tools and the assessment of their effect on ATCS SA.

Introduction

The introduction of new technologies into the National Airspace System (NAS) will alter the role of the Air Traffic Control Specialist (ATCS). As the Federal Aviation Administration (FAA) moves forward with National Airspace System (NAS) modernization and the implementation of Free Flight concepts, the ATCS will move from a controlling role to a collabora-

rating role[1]. With increases of up to 60% expected in air traffic over the next decade, automation will play a significant role in supporting the NAS and the ATCS[1], the anticipated benefits from the use of decision support tools (DSTs) include increases in safety, efficiency, and ATCS productivity [2].

To gain a better understanding of the impact of DSTs on ATCSs' situational awareness (SA), we compared two levels of automation, limited and full, to a no-automation condition under high - and low-traffic loads. This paper will discuss a new tool to assess the effect of increases in air traffic and automation on ATCSs' SA in an operational Host Computer System with the Display System Replacement hardware.

Air Traffic Control Automation

A primary goal of Air Traffic Control (ATC) is safety. Operating under high workloads, tight regulations, and challenging conditions, the ATCS's job is demanding [3]. Improving the current ATC system with new technologies is necessary because of the projected increase in air traffic in the coming decades. These new technologies include automated tools for all phases of flight [1], [3], [5]. In this study, we used a prototype en route DST that is currently used at two ARTCCs in the United States. The introduction of automated DSTs in the en route environment may alter ATCSs' SA.

The literature discusses the technical and practical considerations of automation [6], [7], but often does not address how it produces new problems for human performance[8]. As we continue to rely on the support of automation and errors involving automated systems tend to be more costly [9], it is essential to better understand its impact on human performance.

Automation tends to increase system complexity. The increased probability of system failure related to the increased number of systems adds to the complexity of the user's job [9]. Complex systems may negatively affect SA[10].

Situation Awareness Assessment

Previously used measures to assess how operators develop and maintain SA include physiological measures such as eye movements [11], [12], verbal protocol analysis [13], [14], retrospective recall[15], rating techniques [16], [17], memory probes [18], [19], and on-line queries [20]. Most of these techniques have demonstrated some degree of validity and usefulness. In this study, we assessed SA through four instruments. We used (1) self-ratings of several aspects of SA obtained from the post scenario questionnaires, (2) over-the-shoulder ratings of several aspects of SA obtained from the observations by our subject matter expert, (3) SA assessment based on the Situation Awareness Global Assessment Technique (SAGAT), and (4) SA assessment through a new online tool developed for this study.

Objectives

The purpose of the current study was to explore the effects of traffic load and automation on ATCSs in an ATC en route environment.

1.1 Method

In this study, we implemented an online SA assessment tool that uses the simulation environment to probe ATCSs. The instrument was implemented on operational Host computer and Display System Replacement hardware.

1.2 Participants

Sixteen DSR certified, non-supervisory, full-time ATCSs participated. None of the participants was on medical waiver or in a staff position at the time of the experiment. Eleven participants had normal vision and five had corrected-to-normal vision. The mean age of the participants was 38.7 years. They had actively controlled traffic for a mean of 15.1 years. None of the ATCSs had previously received training on a DST.

1.3 Airspace

We used a generic en route ARTCC to make our findings easier to generalize and to increase the size of our participant pool [21]. During the simulation, the weather conditions required instrumented flight rules to be in effect.

1.4 Scenarios

The results discussed in this paper are based on six different scenarios under the following conditions. The scenarios included two load levels (High and Low) and three levels of automation (No Automation, Limited Automation, and Full Automation). The ATCSs worked in teams of two, consisting of a Radar- (R-) and Data- (D-) side.

1.5 Training

The training environment for the DST consisted of an instructor position and two student positions. The DST itself ran on a DEC Alpha system with remote displays. The instructor and student positions consisted of SUN workstations with 19" displays, a keyboard, and a three-button mouse.

For training on the airspace, we used two standalone ATCoach simulators [22], and six simulation pilots. Participants trained on six scenarios of moderate load.

1.6 Experimental Sessions

For the experimental sessions, we used an integrated system including the Target Generation Facility (TGF), the Host computer, the Display System Replacement (DSR) system, and the DST. We used the TGF to generate targets and air space. The air space in this study used the concept of generic air space from previous studies but required adaptation of the TGF, the Host, the DSR, and the DST [21]. Because the DST required a 20-minute look-ahead, we created an airspace several sectors wide. Only one of the sectors was active, whereas the other sectors functioned as ghosts.

1.7 ATCS Experimental Environment

The ATCS environment included a fully functional sector position and a new tool that allowed us to probe SA on line. In the following sections, we describe this environment.

1.7.1 Sector Position

For the R side, a 2,000 by 2,000 pixel, 29" video display unit (Sony, Japan) presented the radar-scope. The workstation had a DSR flight strip bay, an en route keyboard, and a trackball with three buttons. A landline allowed interfacility and intrafacility communications. The D-side ATCS had access to the DST on a 29" video display.

1.7.2 Implementation of an Online Situation Awareness Probe

The Situation Awareness Verification and Analysis Tool (SAVANT) ran on separate multiple network SUN workstations programmed in C++, Java, and ODS Toolbox [23], one for the D side and one for the R side ATCSs. Approximately every 3 minutes, we probed the perceived workload with the Air Traffic Workload Input Technique (ATWIT). Twenty seconds after ATWIT went off, the ATCSs' monitors displayed the SAVANT query for 3 seconds, and then the replication of the DSR screen appeared for 9 seconds while they responded to the query.

SAVANT uses the ODS toolbox as the human-computer interface. SAVANT obtained part of the data for the replication of the DSR screen from the TGF. SAVANT needed other data to correctly place data blocks. Those data included the orientation and length of the leader line relative to the raw radar return, the history trail, and the vector line. A token ring sniffer listened in on two of the three DSR token rings and presented the DSR messages on a socket. Middleware software grabbed the appropriate messages and presented it to DESIREE.

The SAVANT computer display presented a replication of the airspace including targets and data blocks. SAVANT indicated targets of interest on the screen by highlighting them in a separate color and with portions of their data blocks hidden. Participants responded to queries by clicking on the target that has the highest values for the queried variable. When SAVANT started, the computer screen was blank. Each session lasted 12 seconds (3 seconds to display the question and 9 seconds to respond).

We used two types of queries. The first type was a forced choice question that asked about the relationship between one aircraft and either another aircraft or airspace. For this type of query, visual information was present except for the information about which the query was asking. The second type question was open ended with multiple responses possible. The only information available for this question type was the aircraft position. We recorded the ATCS's responses by having them click on the aircraft representing the correct answer.

1.8 Design and Procedure

Our study was a 2 x 2 x 3 design and contained two levels of position (D-side and R-side), two

levels of load (Low and High), and three levels of automation (No Automation, Limited Automation, and Full Automation).

Position: ATCSs worked either as R or a D side.

Load: Low-load scenario had a level of complexity at which a first line supervisor would be about to remove the assistance of a data ATCS. High-load scenario had a level of complexity at which a first-line supervisor would be about to move from a two- to three-person team.

Automation: The No Automation (baseline) featured paper flight strips. The Limited Automation or second level of automation featured electronic flight strips with conflict indication but no trial flight planning. The Full Automation or third level featured electronic flight strips with flight planning. We collected the ATCS response and response time for each question

Dependent variables: In this experiment, we used four SA assessment techniques; SAVANT provided objective online assessment of SA for gist-type information; SAGAT provided a single objective assessment per scenario of Endsley's three levels of SA; and our PSQ and OTS form queried ATCSs and SMEs respectively about perceived ATCS SA.

2. Results

2.1 SAVANT

Correct Responses Only – Aircraft-Pair Questions. As an R-side, ATCSs responded faster to the current SAVANT questions than as a D-side [$F(1,15)=5.86, p < .05$, Figure 1].

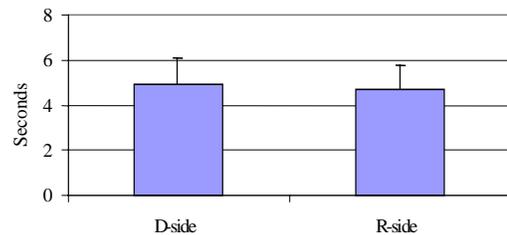


Figure 1. Response time for current correct answers by position.

All Responses Answered – Aircraft-Pair Questions. ATCSs responded faster in the full automation condition than in the limited automation condition. There were no differences between full and no automation or limited and no automation (Figure 2).

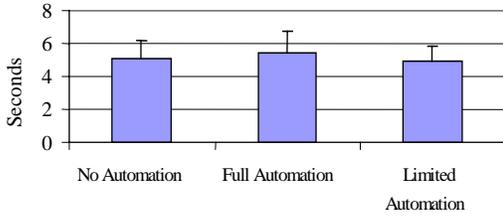


Figure 2. Response time for all responses of current SA by automation.

ATCSs responded faster to queries when load was low [F(1,15)=7.57, p<.05, Figure 3].

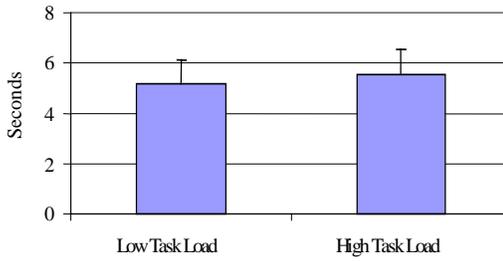


Figure 3. Response time for all responses of future SA by load.

Sector-Based Queries. As an R-side, ATCSs answered a significantly higher percentage of questions correctly than as a D-side [F(1,15)=13.40, p<.01, Figure 4]. A trend for automation showed that, as automation increased, the percentage of correct responses decreased (Figure 5).

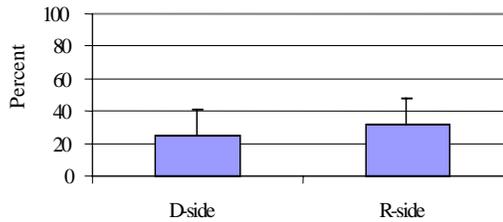


Figure 4. Percent of correct responses for Sector Based questions by position.

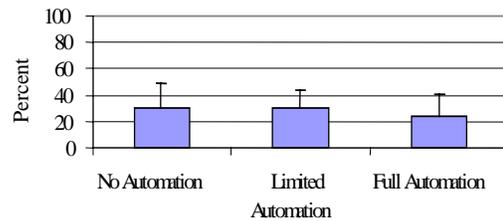


Figure 5. Percent of correct responses for Sector Based questions by automation.

Automation tended to have an effect for R-side ATCSs but not for D-side ATCSs. R-side ATCSs' percentage of correct responses was higher when they did not use any automation.

2.2 SAGAT

We compared ATCSs' perceptions of the traffic situation as reported on the SAGAT queries to the actual state of the traffic situation at the end of the scenario.

Level 1 SA: When working on the R-side position, ATCS awareness of aircraft location was higher than when on the D-side (Figure 6). ATCSs identified more callsign letters correctly under low load conditions (Figure 7). Under full automation and high load, ATCSs' awareness of headings was higher than full automation and low load (Figure 8). Awareness for aircraft type was highest in the high load, limited automation condition.

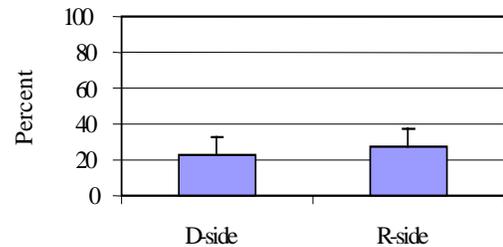


Figure 6. Percent correct for aircraft position by position.

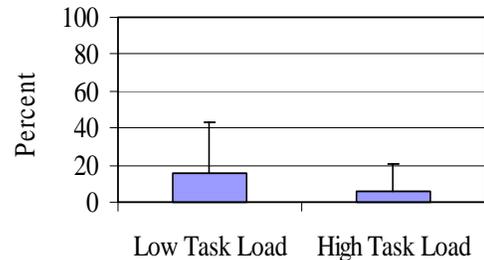


Figure 7. Percent correct for awareness of aircraft callsign letter by load.

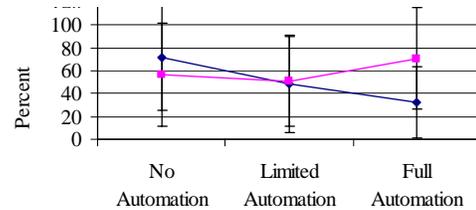


Figure 8. Percent correct for awareness for headings by load and automation.

Level 2 SA. Awareness of the next fix was lowest under limited automation. When on the R-side, ATCSs' awareness of aircraft's flight profiles was higher than when on the D-side. ATCSs had a higher awareness of separation is-

sues under low load conditions. Under high load conditions, ATCSs were more aware of aircraft clearances received. Awareness for clearance conformance was lowest under the low load, no automation condition.

Level 3 SA. We found no significant effects, nor trends for the variables assessing Level 3 SA.

2.3 Post-Scenario Questionnaire

Position and load affected perceived SA for current aircraft location (SA_{cal}) (Figure 9). When working as an R-side, ATCSs rated their SA_{cal} higher than when they were in the D-side position [F(1,15)=11.33, $p<.01$]. Further, when controlling traffic in a low load scenario, ATCSs perceived they had higher SA_{cal} than in high traffic load scenarios [F(1,15)=11.91, $p<.01$].

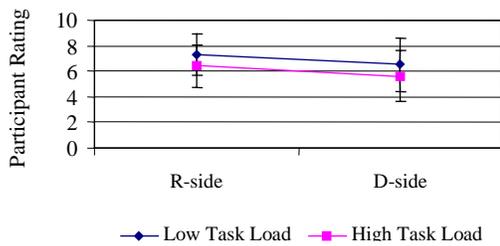


Figure 9. SA for current aircraft locations by load.

ATCSs perceived they had higher SA for projected aircraft locations when controlling traffic in a low load scenario compared to a high load scenario [F(1,15)=12.76, $p<.01$, Figure 10]. .

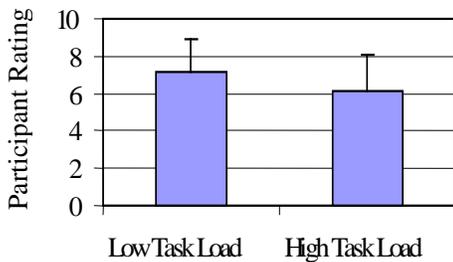


Figure 10. SA for projected aircraft locations by load.

When asked to rate their SA for potential violations, ATCSs reported higher levels of SA under low load scenarios in comparison to high load scenarios [F(1,15)=22.21, $p<.001$]. The univariate results showed a secondary trend for the interaction between load and automation. As shown in Figure 11, the secondary trend indicated load had the largest effect under the full automation condition. When ATCSs used full

automation, their SA for potential violations was lower when the load was high compared to low.

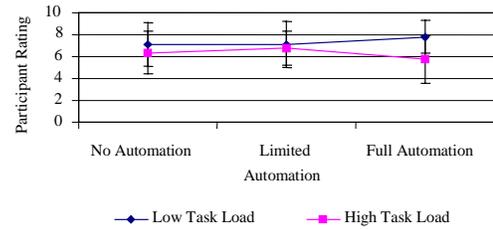


Figure 11. SA for potential violations by load and automation.

2.4 Over-the-Shoulder

The SME felt that ATCSs maintained awareness of aircraft positions, ensured positive control, and corrected their own errors in a timely manner better when the load was low [F(1,7)=40.69, $p<.001$, F(1,7)=56.52, $p<.001$, and F(1,7)=20.05, $p<.01$, respectively, Figure 12]. A trend in the data suggests the ATCSs also detected pilot deviations from control instructions better under low load conditions.

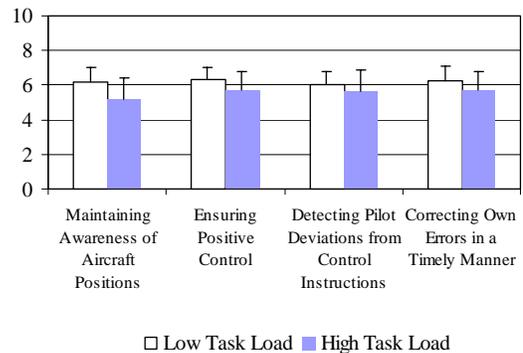


Figure 12. Maintaining attention and SA by load.

3. Discussion

We will first discuss our interpretation of the results for each of the instruments. Based on the instrument specific discussion, we will draw conclusion about the overall effect of automation, traffic load, and ATCS position on SA. Finally, we will discuss some of the advantages and disadvantages of SAVANT.

SAVANT. With limited automation, ATCSs were less involved with the traffic. In contrast, full automation provided a graphical representation of aircraft positions and conflict information, allowing the D-side ATCS to stay involved. Even without automation, ATCSs had better SA than with limited automation. In the field,

ATCSs training includes building the “picture” from flight strips and maintaining that picture without a graphical representation. Therefore, by giving them flight progress strips, they may have been better able to stay involved than with the limited automation where ATCSs were not familiar with the format of displayed aircraft information.

ATCSs had better awareness of the current situation when working on the R-side. R-side ATCSs had the advantage that they were in a more active role related to the traffic compared to the D-side who is in more of a monitoring role. This compliments findings that SA is lower when passively processing information instead of actively processing it [24], [25].

ATCSs had better awareness for future situations when load was low. The fewer the aircraft, the easier it was to project them into the future, increasing SA. Because the Aircraft Pair questions assessed SA relevant to both positions, load did not affect ATCS SA relative to their control position.

SA was higher for these action-oriented items when ATCSs were on the R-side. The R-side needs to have an overall comprehension of whether the aircraft are conforming to clearances; In contrast, the D-side ATCS needs to know the state of the aircraft. In addition, a trend showed lower SA under increasing levels of automation. This coincides to findings that SA decreases under automation [25], [26]. R-side ATCSs’ SA was higher when no automation was present. When no automation was present, the D-side ATCS could devote his or her full attention to the DSR screen and assist the R-side ATCS which freed up some cognitive resources of the R-side ATCS.

SAGAT. With limited automation, ATCSs only had electronic flight strips and could not pull up the graphical plans display. This appeared to hamper their ability to stay up with next fix information. As mentioned earlier, ATCSs have traditionally trained to use paper flight strips to help build their picture, and, in the no automation condition, they may have used this to stay involved. Similarly, they may have used the graphical display provided in the full automation to maintain the picture.

Even though ATCS scores for aircraft positions were low, we did find that R-side ATCSs had better awareness of aircraft location than D-side ATCSs. R-side ATCSs spend more time viewing

the radar display and actively controlling traffic, and the D-side ATCSs assist and are in a more monitoring role. The active control of traffic may increase SA as reflected in this item and supported findings [24] that SA was higher for active ATCSs than monitoring ATCSs. We had a similar finding for awareness of flight profile. The R-side ATCSs were better aware of aircraft flight profiles – en route, outbound from an airport, or inbound to an airport in the sector – than the D-side ATCSs. This corresponds well with the notion that one has better memory for things that you do yourself (the R-side radioed the aircraft and discussed changes in flight profile with the pilots) than things that are done for you (the D-side received second-hand information after the R-side initiated the control action).

In SAGAT, we queried ATCSs only on those aircraft that they initially placed on the SAGAT map. Under low load, ATCSs may be able to recall more aircraft than the ones that required their attention. Under high load, ATCSs most likely had so many aircraft that required their attention, that they could not recall all of them. Therefore, when queried about whether an aircraft that they recalled had received a clearance, under low load, they may have remembered less information correctly because the recalled aircraft included unimportant aircraft as well. Under high load, all recalled aircraft may have been important and whether or not one of these aircraft had received a clearance now was important to the ATCS, resulting in higher scores.

ATCSs were more aware of headings when load was high. With full automation, ATCSs had conflict indications and reached out more to resolve those conflicts, bringing the headings for those aircraft to their attention. In addition, these aircraft may have been important because all recalled aircraft under high load had the attention of ATCSs. ATCSs had higher SA for aircraft type under limited automation and high load conditions. In limited automation, ATCSs could easily access aircraft type from the aircraft list. Under high load, this information seems to have more importance to the ATCSs. Awareness of aircraft conformance to clearances was lowest under low load, no automation. Under low load, ATCSs may have remembered less information correctly because the recalled aircraft included unimportant aircraft as well important aircraft. Further, they did not have access to conflict indications and were not as concerned with possible conflicts because the likelihood of them was low. In contrast, under high load, recalled aircraft in-

cluded only important aircraft and the potential for possible conflicts was large, therefore, even under no automation conditions, ATCSs paid great attention to aircraft conformance.

First, the discrepancy in size between the DSR screen and the computer monitor may physically create a scaling problem, making it difficult for ATCSs to accurately place aircraft at positions coinciding with the DSR position. We have not established that a change in scale does not distort ATCS recall for aircraft positions. The probe itself also may not access the relevant information within the ATCS knowledge. ATCSs may have the knowledge of aircraft positions relative to one another but cannot express them at the exact location as required by the format provided by the probe. Or, perhaps, the ATCSs really do not have the information stored in memory but instead use the radar display as an external memory. If this is the case, a good indication of SA may be the speed at which they can find the information from the external memory. For instance, awareness of aircraft's next fix and aircraft type, although not available directly from the radar display, can be accessed from flight plan information. Even though information is not in memory, knowing where to find it can imply good SA [20]. Helbing [28] suggested that ATCSs use aircraft positions as memory pegs (i.e. they store relevant information about aircraft by aircraft position and are not necessarily able to retrieve aircraft information when probed using cues such as callsign). The current research findings suggest that ATCSs do not have good memory for exact aircraft locations. Therefore, if Helbing's model is accurate, the spatial representation that ATCSs use to store aircraft information is not simply a scaled version of the actual displayed information

The results show we cannot expect ATCSs to recall verbatim information. This does not mean that they do not have that information stored in memory; it simply means that the information is not accessible to us for probing.

Although, in general, awareness for Level 1 SA items was relatively low, there did not seem to be an adverse effect for Level 2 SA. One could use the SAGAT results to infer that the information processing model proposed by Endsley and Smolensky [30] is not correct (i.e., an ATCS would not need Level 1 SA before achieving Level 2 SA). However, the results from this experiment do not indicate whether ATCSs have processed information from Level 1 SA into

other variables or formats or that the raw information is still present but cannot be directly measured by Level 1 SA queries.

We did not have significant findings for Level 3 SA, the projection of elements into future situations. Even with the DST that should have moved ATCSs to a more strategic orientation, our ATCSs still controlled traffic in a mostly tactical fashion.

Post scenario questionnaire Under conditions that involved the DST, the ATCS had to constantly monitor it to fully use its capabilities. Constant monitoring of automation induces higher workload and can result in lower levels of SA. Under low load conditions, the ATCS may be able to compensate for this, but, when load is high, the ATCS is less able to maintain the picture and maintain SA.

Increasing load led to decreases in perceived SA for current aircraft locations, projected aircraft locations, and potential violations. When load was low, ATCSs were able to better focus on the big picture and displayed higher SA.

Position affected SA for current aircraft locations. R-side ATCSs rated their SA higher along this dimension than D-side ATCSs. This item reflects the differences in the responsibilities of the R- and D-side ATCSs. The R-side ATCS actively scanned the DSR screen and directed traffic, but the D-side ATCS had other responsibilities that did not focus on the scope leading to lower SA for aircraft locations.

Over-the-shoulder ratings. We expected that automation would have an effect on SA items from the OTS, but automation did not have an effect on the SME's ratings for SA. These results imply that the change in automation levels did not affect ATCS SA in a manner that the SME could observe. Due to limitations in resources, we used only one SME to complete OTS forms. The SME predominantly focused on the R-side ATCS. We assumed that any changes in the team should be shown on both sides. The R-side functioned as if everything was normal, but the D-side was using the automation. The automation did not directly affect the R-side as it did the D-side, as viewed by the SME. Therefore, an automation effect was not found.

Load affected the SME's ratings of ATCSs' SA. High loads contributed to lower SA ratings. Once again, high loads made it difficult for the ATCS to maintain the big picture. This led to ATCSs losing SA.

Automation Related Observations. We have seen in most of our data sets that with increased levels of automation, D-side ATCSs pulled away from the radar display, leading to a further reduction of SA for D-side ATCSs.

With limited automation, the D-side lost some SA, possibly because of the lack of either the graphical representation of the DST or the familiar flight progress strips. The R-side ATCS may have received less effective assistance under limited automation conditions. This in turn could lead to a reduction in SA for the R-side. When full automation was available, although the R-side may have lost some of the D-side assistance on radar tasks, he or she gained assistance of the D-side on strategically resolving conflicts. The benefit that the strategic solution provides under low traffic load conditions no longer offsets the loss of D-side assistance on radar tasks under high load conditions.

Load Related Observations. The ATCSs themselves and the OTS raters felt that an increase in load decreased SA. The objective SAGAT measure also showed decreases in SA, although this was not true for Level 2 SA. With more aircraft in the scenario, ATCSs may have needed more cognitive resources than available, leading to a loss of SA. With the anticipated increase in traffic in the next decade, ATCSs will have increasing difficulty maintaining good SA. Developing tools that support ATCSs to focus their attention on aircraft that are at risk to get into potential conflicts or otherwise complex ATC situations may allow ATCSs to make most efficient use of their cognitive resources.

We found that ATCSs had very low levels of Level 3 SA (i.e., they were not able to project future situations). Although ATCSs in the current system work mostly tactically (i.e., within a timeframe of 3-5 minutes), we had expected that the use of the DST would extend this horizon, resulting in differences between the non-automated and automated conditions. One would expect that with more experience in the use of the DST, ATCSs would extend their time-horizon and become more strategic, resulting in higher Level 3 SA. Although the data may be available on the use of the prototypes that are in operation in the field, there are no studies available that show this shift in ATCS SA.

ATCS Position Related Observations. Without providing a change in the roles and responsibilities of the D-side ATCS when providing automation tools, the D-side will revert to established

procedures. These established procedures will pull the D-side back into tactical control of aircraft, leaving the DST not used to its full potential when it would be able to assist ATCSs most. This is not to say that the R-side ATCS does not need the D-side assistance. Under the high load conditions that we tested here, the R-side needed someone to separate data blocks and an extra pair of eyes.

In the field, traffic levels fluctuate and are relatively low before a push in traffic occurs. Therefore, one can bring in a D-side ATCS before an anticipated push in traffic. The D-side can then communicate longer with adjacent sectors and prepare for or resolve potential conflicts. One way of changing the role of the D-side may be to implement traffic load indices that indicate transitions (e.g., relative low traffic load) with a transition to a push expected in 30 minutes. During those 30 minutes, a D-side ATCS could focus on resolving potential conflicts and be more strategic. This could reduce the peak complexity of the push. During the push itself, the D-side ATCS pulls back from a strategic into a more tactical mode and fully assists the R-side ATCS.

Our participants were radar-certified Certified Professional Controllers) CPCs. They felt they could have assisted better when at the radar display (i.e., more actively involved in controlling traffic). We confirmed this feeling of being less able to directly influence the traffic situation by objective data from SAGAT. SAGAT results indicated that for a given ATCS, their SA was lower when working on the D-side SA compared to working on the R-side. More research in the roles and responsibilities and the workload of D-side ATCSs is necessary to determine if that position can be of more assistance during high traffic situations through a change in roles and responsibilities.

The R-side ATCS actively controls traffic, whereas the D-side ATCS has other responsibilities besides focusing on the radar screen. This gives R-side ATCSs an advantage because they are in a more active role compared to the D-side, who is in more of a monitoring role. This coincides with Willems and Truitt's [24] findings that SA was higher for ATCSs actively controlling traffic than monitoring traffic.

The open-ended questions were really more for the tactical, R-side ATCS. SA was higher for these action-oriented items when ATCSs were on the R-side. The R-side needs to have an over-

all comprehension of whether the aircraft are conforming to a clearance.

In contrast, the D-side ATCS needs to know the state of the aircraft. A trend showed lower SA under increasing levels of automation. R-side ATCSs' SA was higher when no automation was present. These results indicate that the experimental manipulations affected R-side ATCSs more. When no automation was present, the D-side ATCS could devote his or her full attention to the DSR screen and assist the R-side ATCS, which freed up some cognitive resources of the R-side ATCS. When no automation was present, the D-side ATCS could manipulate the data blocks allowing for easier viewing and no overlapping information.

Evaluation of SAVANT. The fact that the SAVANT results mostly point in the same direction as the findings from our other measures of SA is promising. Although SAVANT does not freeze a simulation, the ATCSs felt that the brief time that they could not see the traffic was awkward. Because SAVANT is tightly integrated into our simulator, queries can be quite specific. A drawback of that approach occurs when an ATCS changes the situation such that the query is no longer relevant. Additional techniques that would allow the queries to be more dynamic and would adapt to the current situation may be necessary to adapt to ad hoc changes in the simulation. SAVANT is currently an instrument that becomes part of the simulation that is easy to administer for a given set of queries about anticipated ATC situation

References

- [1] Office of Air Traffic System Development. (1997). *AUA program master plan (Volume 2): Automation strategic planning*: Washington DC: Federal Aviation Administration.
- [2] Kirk, D. B., Heagy, W. S., & Yablonski, M. J. (2000). *Problem resolution: Support for free flight operations*. (MITRE No. 02001204-21). McLean, VA: Center for Advanced Aviation System Development.
- [3] Finkelman, J. M., & Kirschner, C. (1980). An information-processing interpretation of air traffic control stress. *Human Factors*, 22(5), 561-567.
- [4] Gosling, G. D. (1992, June). *Artificial intelligence in air traffic control*. Paper presented at the NASA/FAA Workshop on Artificial Intelligence and Human Factors in Air Traffic Control and Aviation Maintenance, Daytona Beach, FL.
- [5] Zweben, M. (1992, June). *Artificial intelligence in aerospace systems: An overview of AI research at NASA Ames*. Paper presented at the NASA/FAA Workshop on Artificial Intelligence and Human Factors in Air Traffic Control and Aviation Maintenance, Daytona Beach, FL.
- [6] Ignizio, J. P. (1991). *An introduction to expert systems*. New York: McGraw-Hill.
- [7] Waterman, D. A. (1986). *A guide to expert systems*. Reading, MA: Addison-Wesley.
- [8] Endsley, M. R. (1993). SA: A fundamental factor underlying the successful implementation of AI in the air traffic control system. In D. J. Garland & J. A. Wise (Eds.), *Human factors and advanced automation technologies* (pp. 117-122). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.
- [9] Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: Harper Collins.
- [10] Endsley, M. R. (1999). Situation awareness in aviation systems. In D. J. Garland, J. A. Wise, & V. D. Hopkin (Eds.), *Handbook of aviation human factors* (pp. 257-276). Mahwah, NJ: Lawrence Erlbaum.
- [11] Moray, N., & Rotenberg, I. (1989). Fault management in process control: Eye movements and action. Special Issue: Current methods in cognitive ergonomics. *Ergonomics*, 32, 1319-1342.
- [12] Wierwille, W., & Eggemeier, F. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors*, 35, 236-282.
- [13] Ohnemus, K., & Biers, D. (1993). Retrospective versus concurrent thinking-out-loud in usability testing. In *Proceedings of the 37th Annual Meeting of the Human Factors Society*, (pp. 1127-1131). Santa Monica, CA: Human Factors and Ergonomics Society.
- [14] Sullivan, C., & Blackman, H. S. (1991). *Insights into pilot SA using verbal protocol analysis*. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 57-61). Santa Monica, CA: Human Factors Society.
- [15] Kibbe, M. (1988). *Information transfer from intelligent EW displays*. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 107-110). Santa Monica, CA: Human Factors Society.
- [16] Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 185-218). Amsterdam: North-Holland.

- [17] Taylor, R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational Awareness in aerospace operations* (AGARD-CP-478) (pp. 3/1 - 3/17). Neuilly Sur Seine, France: NATO - AGARD.
- [18] Endsley, M. R. (1988a). *Design and evaluation for SA enhancement*. In Proceedings of the Human Factors Society 32nd Annual Meeting (pp. 97-101). Santa Monica, CA: Human Factors Society.
- [19] 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.
- [20] Durso, F. T., Truitt, T. R., Hackworth, C. A., Crutchfield, J. M., Ohrt, D. D., Hamic, J. M., & Manning, C. A. (1995). Factors characterizing en route operational errors: Do they tell us anything about situation awareness? In D. J. Garland and M. R. Endsley (Eds.), *Experimental analysis and measurement of situation awareness* (pp. 189-195). Daytona Beach: Embry Riddle Aeronautical University Press.
- [21] Guttman, J.A., Stein, E. S., & Gromelski, S. (1995). *The influence of generic airspace on air traffic controller performance*. (DOT/FAA/CT-TN95/38). Atlantic City International Airport, NJ: Federal Aviation Administration.
- [22] ATCoach (Version 7.0) [Computer Software]. (1996) Lexington, MA: UFA, Inc.
- [23] ODS Toolbox (Version 5.1) [Computer Software]. (1999). Bremen, FRG: Orthagon GMBH.
- [24] Willems, B. F., & Truitt, T. R. (1999). *Implications of reduced involvement in en route air traffic control* (DOT/FAA/CT-TN99/220). Atlantic City Airport, NJ: Federal Aviation Administration Technical Center.
- [25] Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, 37(2), 381-394.
- [26] Carmody, M. A., & Gluckman, J. P. (1993). Task specific effects of automation and automation failure on performance, workload and situational awareness. In R. S. Jensen & D. Neumeister (Eds.), *Proceedings of the Seventh International Symposium on Aviation Psychology* (pp. 167-171). Columbus, OH: Department of Aviation, The Ohio State University.
- [27] Endsley, M. R., Sollenberger, R., Nakata, A., Hough, D., & Stein, E. (2000). *Situation Awareness in Air Traffic Control: Enhanced displays for advanced operations* (DOT/FAA/CT-TN00/01). Atlantic City Airport, NJ: William J. Hughes Technical Center.
- [28] Helbing, H. (1997). A preliminary cognitive model of en-route Air Traffic Control. In R. S. Jensen, & L. Rakovan (Eds.), *Proceedings of the Ninth International Symposium on Aviation Psychology*, (pp. 391-396). Columbus, OH: The Ohio State University.
- [29] Albright, C. A., & Lewandowsky, S. (1995). Momentum in a complex monitoring task. In R. S. Jensen, & L. A. Rakovan (Eds.), *Proceedings of the Eighth International Symposium on Aviation Psychology* (pp. 364-368). Columbus, OH: The Ohio State University.
- [30] Endsley, M. R., & Smolensky, M. W. (1998). Situation awareness in Air Traffic Control: The picture. In M. W. Smolensky & E. S. Stein (Eds.), *Human Factors in Air Traffic Control* (pp. 115-154). San Diego, CA: Academic Press.
- [31] Endsley, M. R., & Rodgers, M. D. (1998). Distribution of attention, SA, and workload in a passive Air Traffic Control task: Implications for operational errors and automation. *ATC Quarterly*, 6(1), 21-44.
- [32] Moray, J. (1986). Monitoring behavior and supervisory control. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Volume II, Cognitive processes and performance* (pp. 43-1 - 43-32). New York: Wiley-Interscience.
- [33] Wiener, E. (1988). Field studies in automation. In S. D. Norman & H. W. Orlady (Eds.), *Flight deck automation: Promises and realities* (pp. 37-55). Moffett Field, CA: NASA-Ames Research Center.

Author Biographies

Ben Willems is an Engineering Research Psychologist for the Federal Aviation Administration. He conducts human-in-the-loop simulations for the National Air Space Human Factors Branch (ACT-530) at the William J. Hughes Technical Center in Pomona, New Jersey. He has been involved in real-time Air Traffic Control simulations since 1995 and joined the Federal Aviation Administration in 1998. His areas of expertise are air traffic controller visual scanning, workload, and situation awareness.

Dr. Michele Heiney is an Engineering Psychologist for Logicon/Federal Data Corporation since 1999. She received her PhD in Experimental/Social Psychology from Virginia Commonwealth University in 1998. She also holds a Commercial Pilot Certificate with an Instrument Rating.