Conflict Resolution Automation and Pilot Situation Awareness

Arik-Quang V. Dao and Summer L. Brandt
NASA/San Jose State University Research Foundation

Paige Bacon, Josh Kraut, Jimmy Nguyen, Katsumi Minakata, and Hamzah Raza
California State University, Long Beach

David Rozovski
Purdue University

Walter W. Johnson
NASA Ames Research Center

This study compared pilot situation awareness across three traffic management concepts. The Concepts varied in terms of the allocation of traffic avoidance responsibility between the pilot on the flight deck, the air traffic controllers, and a conflict resolution automation system. In Concept 1, the flight deck was equipped with conflict resolution tools that enable them to fully handle the responsibility of weather avoidance and maintaining separation between ownship and surrounding traffic. In Concept 2, pilots were not responsible for traffic separation, but were provided tools for weather and traffic avoidance. In Concept 3, flight deck tools allowed pilots to deviate for weather, but conflict detection tools were disabled. In this concept pilots were dependent on ground based automation for conflict detection and resolution. Situation awareness of the pilots was measured using online probes. Results showed that individual situation awareness was highest in Concept 1, where the pilots were most engaged, and lowest in Concept 3, where automation was heavily used. These findings suggest that for conflict resolution tasks, situation awareness is improved when pilots remain in the decision-making loop.

INTRODUCTION

Rapid increase in air traffic density will exceed the ability of the human controller to successfully manage operations in the national air space using existing traffic management concepts and technology (Joint Planning and Development Office, 2007). To meet the capacity demands of the future air transportation system, as well as meet or improve safety and efficiency standards, human controller tasks like air traffic conflict detection and resolution must be supported by, or shared with, humans in the flight deck and/or new automation technologies. Studies conducted at NASA Ames Research Center have shown that controller performance on conflict avoidance tasks decreases when traffic load increases, but this decrement can lessen when the controller is assisted by automation (Prevot, Homola, Mercer, Mainini, & Cabrall, 2009). However, there may be trade-offs related to situation awareness when deploying automation (Parasuraman, Sheridan, & Wickens, 2000). The focus of this paper is to assess these trade-offs with respect to pilot situation awareness under conditions where traffic separation responsibility is shared between the flight crew, controllers, and automation.

Situation Awareness

Situation awareness (SA) has many definitions. For the purpose of this study, situation awareness is "the operator's understanding of the state of the relevant environment and his or her ability to anticipate future changes and developments in the environment" (European Air Traffic Programme, 2003).

Endsley (1995) developed an off-line probe technique, called the Situation Awareness Global Assessment Technique (SAGAT) to asses SA. With this technique, task analysis is used to identify critical information requirements. Probe questions are then developed to capture the operator's awareness of these items. During a simulation, the scenario is paused, the screen blanked, and the operator is presented with the probe questions. Higher accuracy scores on the questions are indicative of higher SA. However, SAGAT has been criticized for being too heavily reliant on working memory and the process of freezing and resuming a scenario interrupts the operator's primary task (Pierce, Strybel, & Vu, 2008). If performance measures are of interest, then this task interruption can negatively impact results of a study.

Alternatively, Durso, Bleckley, and Dattel (2006) proposed that SA can be measured based on the operator's understanding of the task environment. That is, the operator may not have the information needed to answer the probe question in working memory, but may know the location of SA relevant information on a display. Knowing where to find critical information should yield better situation awareness, thus allowing operators access to the display can then improve their accuracy for these events. Therefore, SA information normally available from the display should be available when the operator is being probed. Because the operator is probed without stopping the task, SA probes are administered in a two-step process. First, a ready prompt is presented. This prompt informs operators that a probe question is ready to be

presented. If the operator's workload is not too high, and s/he has the resources to answer the probe question, then s/he will accept the ready prompt (by hitting a button or saying "ready".) The probe question is then administered right after the ready prompt has been accepted. Durso et al. (2006) noted that this procedure yields three important measures: a ready latency (response latency between the appearance of a ready prompt and when the operator indicates that s/he is ready), a probe response latency (response latency between the presentation of the question and operator response), and a probe accuracy score. The ready latency is considered to be a measurement of workload because the operator should be able to indicate that s/he is ready more quickly when s/he is not busy compared to when s/he is busy. That is, the lower the workload, the shorter the ready latencies should be. The probe response latency can be used as an indicator of SA because the operator should take less time to answer questions when the information needed to answer the question is easily accessible (either in his/her working memory or s/he knows where to look for the information). In other words, shorter response times suggest better SA than longer response times.

Automation Affects Situation Awareness

The implementation of automation can vary in terms of degree, with each level of increasing automation having an impact on situation awareness. In cases where automation is completely responsible for undertaking a task, humans may be thrown out-of-the-loop leading to complacency (Parasuraman, et al., 2000). When complacent, the operator no longer proactively seeks to maintain relevant information in the environment leading to diminished SA. SA can also be diminished when the level of automation provided does not adequately support the task or impose high workload. When workload is high, cognitive tunneling can occur where the operator is forced to selectively attend to the primary task, reducing the cognitive resources needed to monitor or process other task relevant components (Parasuraman & Wickens, 2008). However, a performance benefit can be gained from reduced workload without trading off SA, if the human operator is kept "in the loop" by interacting with automation to complete tasks (Dao et al., 2009).

Dao et al. (2009) examined the impact of varying levels of automation on individual pilot SA. Pilots were asked to perform a traffic conflict avoidance task with and without the support of automation. On manual trials, pilots were given a null resolution (no change to route) which they had to modify in order to resolve the conflict. On automated trials, pilots were given a resolution proposed by an automated system, which they could evaluate to ensure that it does solve the conflict, but could not modify it for efficiency or other preferences. On interactive trials, pilots were given an automation-proposed resolution that they could accept as is or revise to improve it based on his/her preference. Pilots were probed for SA at the end of each trial, when the scenario was frozen, but all displays were still active and in sight. Results showed that pilot SA was lowest in the automated condition when compared to the manual and interactive conditions;

there were no differences between the manual and interactive conditions. Low SA in the automated condition suggests that factors such as automation complacency had a significant impact on SA. Additionally, comparable SA found in the interactive and manual conditions suggest that an interactive, human-in-the-loop implementation of automation would fare better than at fully automated levels.

Because Dao et al.'s (2009) study examined short, 2-minute conflict scenarios, it is not clear whether the same affect of automation would be observed when pilots must fly longer scenarios that involve different phases of flight as well and additional responsibilities. Thus, the present study expands on Dao et al.'s findings by examining pilot SA under different levels of automation allocation in more realistic, 80-minute scenarios.

Current Study

Pilots and controllers engaged in real-time simulations focused on en route and arrival operations into Louisville International-Standiford Field Airport (SDF). Experimental pilots started the scenario in an en-route phase of flight and were asked to engage in company spacing during the arrival into SDF. At about 2 minutes into each scenario the pilots received their spacing interval and lead aircraft from an air traffic control (ATC) ground scheduling station. In addition, they were to make route modification to avoid hazardous weather, and in some conditions, to avoid traffic conflicts. Controllers managed 2 high altitude en route sectors, where they assisted with weather avoidance, managed en route, arrival and departure traffic and provided resolutions to aircraft in conflict based on each concept. Situation awareness for pilots was examined in three concepts of operations, which varied based on the distribution of traffic conflict responsibility across human controllers, pilots, and automation. Although both pilot and controller SA was a focus in this study, this paper will only focus on the pilot's SA across the 3 operational concepts.

In **Concept 1,** pilots had onboard conflict detection and resolution tools (CD&R) and were responsible for merging and spacing, weather avoidance, and conflict resolution (3/4 of conflicts) between ownship and other aircraft, and the controller for the remaining (1/4 of conflicts). In **Concept 2**, pilots also had CD&R tools and were responsible for merging and spacing and weather avoidance, but not responsible for conflict identification and resolution, which was managed by air traffic controllers (3/4 of conflicts) and automation (1/4 of conflicts). **Concept 3**, was similar to Concept 2, except that the majority of conflict resolution responsibility was allocated to an automation system (3/4 of conflicts) rather than to a controller (1/4 of conflicts). In concept 3, unlike 1 and 2 the pilot conflict detection was disabled, thus they could only vet route changes for weather, but not for traffic.

Based on results from Dao et al. (2009), it was predicted that pilot SA would be greatest when operators were involved in the decision making process. Therefore, better pilot SA scores were predicted for Concept 1 and 2 than for Concept 3.

METHOD

Participants

Eight commercial airline pilots with glass cockpit experience and two controllers were recruited for this experiment. They were compensated $25/hr for their participation.

Apparatus

Pilots in the simulation managed a desk top simulator that included an interactive 3-D Cockpit Situation Display (CSD) developed by the NASA Ames's Flight Deck Display Research Laboratory (see Figure 1). The CSD is a PC-based 3-D volumetric display that provides pilots with the location of surrounding traffic, plus the ability to view planned 4-D trajectories (Granada, Dao, Wong, Johnson, & Battiste, 2005). Embedded within the CSD was logic that detected and highlighted conflicts. Pilots were told that conflict detection tool was 100% reliable. In addition, the CSD had pulse predictors that emitted synchronous bullets of light that traveled along the displayed flight plans at a speed proportional to the speeds of the associated aircraft. Using these functions (conflict detection and pulse), a prediction of up to 20 minutes into the future could be made, thus providing graphical confirmation of ownship proximity to traffic along the planned route.
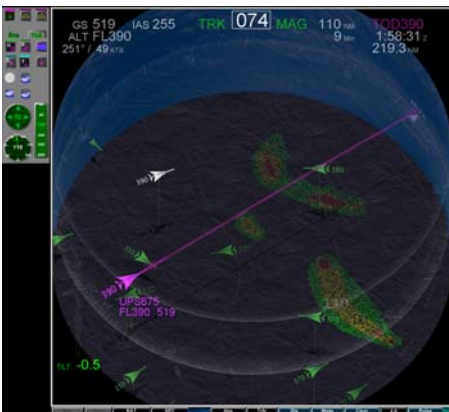

Figure 1. 3D Cockpit Situation Display (3D CSD)

Pilots modified the flight path of ownship for weather and traffic avoidance using a Route Assessment Tool (RAT; Canton, Refai, Johnson, & Battiste, 2005). In Concepts 1 and 2, the RAT was linked to conflict detection software allowing the pilots to find conflict-free paths. In Concept 3, however, the conflict detection was disabled. Proposed resolutions were color coded in gray to distinguish them from the current route (amber if in conflict; magenta in nominal conditions).

Pilots were also asked to merge and space behind an assigned lead aircraft. This task was supported by an automated merging and spacing tool, which was based on the NASA Langley ASTAR algorithms (Abbott, 2002). When engaged, the merging and spacing tool calculated if the aircraft would achieve its assigned spacing by the runway. A spacing error time, how early or late the aircraft was expected to be at the runway, was displayed in seconds. When coupled with the auto throttles, the spacing tool gradually modified the aircraft speed to achieve the assigned spacing interval.

In addition to the CSD, pilots were provided with the Multi-Aircraft Control System (MACS; Prevot, 2002) which was displayed on a separate monitor. MACS provides an interactive display that simulates a 747 flight deck controls (Figure 2). These flight deck controls included the primary flight display (PFD), the mode control panel (MCP), data link display and controls for sending/receiving data link messages and new routes from the ground or automation, as well as flaps and gears for landing procedures.

Figure 2. Multi-Aircraft Control System (MACS).

In addition to the pilot interface described briefly above, MACS was implemented in controller mode, allowing air traffic controllers to track and manage aircraft in their assigned sector. Built into MACS was a conflict resolution algorithm called the Auto-Resolver (Erzberger, 2006). In all concepts, the controller was able to use the Auto-Resolver as a tool and request an automated resolution at the controller's discretion. In Concepts 2 and 3, the Auto-Resolver was implemented as an agent designed to independently manage a percentage of the traffic in the scenario without consent from the human controller. The Auto-Resolver agent was disabled in Concept 1. In addition to the Auto-Resolver tool, controllers used a software based tool called the Trial Planner to search for new conflict free routes. Controllers sent new routes to the flight deck stations via data link. Pilots loaded the new routes through the flight deck data link window, and visually examined the new routes on the CSD.

A separate touch screen tablet computer was also used to measure workload and situation awareness. All probe questions required a yes/no or multiple choice response (equal number of yes/no and multi-choice questions per pilot per scenario).

Design and Procedure

The independent variable was Concept (Concept 1: Pilot Responsible, Concept 2: Controller Responsible, Concept 3: Auto-Resolver Agent Responsible) and the dependent measures were the three metrics obtained from the probes (ready latencies, probe latencies, and probe accuracy). Participants completed three blocks of four trials over four data collection days. Each block was grouped by Concept and was presented once per day. Two trials were repeated on the fourth data collection day due to software malfunctions. Each trial lasted approximately 80 minutes. Participants received classroom training prior to data collection days and were provided three practice trials, one with each concept level.

Pilots were asked to fly an arrival route with a continuous descent approach (CDA) into Louisville, Kentucky airport. All pilots flew within the same scenario in real time and were assigned a spacing interval and lead aircraft two minutes after the start of the trial by an automated ground station. Additionally, pilots were responsible for weather avoidance in all conditions and were trained to maneuver using the RAT. Pilots adjusted their route relative to the weather based on their own safety criteria and constraints imposed by surrounding traffic.

In Concept 1, pilots independently managed weather and traffic separation using the RAT and executed the modified flight plan without prior approval from ground. In Concepts 2 and 3, pilots were required to data link their route modification request to ground stations for approval. Approvals from the ground came from the Auto-Resolver agent when the proposed route was clear of conflicts. Proposed routes that were not clear were forwarded to the human controller for review and final approval. Onboard conflict alerting was provided in Concepts 1 and 2, which allowed pilots to vet their proposed route for traffic, whereas in Concept 3, conflict alerting was disabled.

Pilots received a ready prompt for one SA question every 3 minutes from the start of each trial. Pilots were instructed to press the "ready" button on the touch screen panel to reveal the question. The simulation did not stop while they were answering the questions, and they were allowed to reference information on the displays (see Table 1 for example SA questions). The display timed out after one minute of non-responsiveness for both the ready prompt and the probe question. Pilots completed a trial when they landed in SDF.

Table 1: Sample Situation Awareness Questions

In the next 5 minutes how many aircraft will be within 10nm and 2000ft of ownship?
What is the heading of the aircraft closest to you?
How many times did ownship change speed more than 5 knots in the last five minutes?
Is the difference in heading between ownship and lead less than 10 degrees?

RESULTS

One participant's data was removed from analyses due to non-compliance with probe procedures. Timeouts, or when the participant did not respond to either the ready button or select their response (presumably because workload was too

high to attend to the probe questions) occurred 9% of the time. A repeated measures ANOVA of timeouts as a function of Concept showed no significant effect ($F(2,12) = 1.99$, $p = .18$). In Concept 1 (Pilot Responsible), pilots timed out on 3.8% of the ready prompts compared to 4.6% in Concept 2 (Controller Responsible) and 7.9% in Concept 3 (Auto-Resolver Agent Responsible). Although not significant, the pattern suggests pilots attended to the probe questions more when they were responsible for traffic separation. This pattern is consistent with workload findings reported in Ligda et al. (submitted).

A natural log transformation was performed on all response latency data, given the non-normal distribution. A repeated measures ANOVA was performed for each SA probe measure with Concept as a factor (see Table 2 for means and standard deviations). The p-values were adjusted using Greenhouse-Geisser for violations

Table 2: Means & Standard Deviations of Situation Awareness Probes

	Ready Prompt Latency (sec)	Probe Response Latency (sec)	% Correct
Overall	4.73 (7.54)	15.40 (13.61)	81 (39)
Concept			
Pilot Primary	4.32 (6.57)	13.67 (13.19)	84 (36)
Controller Primary	4.38 (7.05)	15.78 (13.81)	81 (39)
Auto-Resolver Primary	5.51 (8.81)	16.81 (13.68)	79 (41)

n = 7

Analyses of Ready Response Latency

A repeated measure ANOVA for ready response latencies (in seconds) was performed, with Concept as a factor. There were no significant differences in ready prompt latencies by Concept ($F(1.15,6.87) = 2.36$, $p = .17$). Pilots took an average of 4.36s to respond to the ready prompt in Concept 1, compared to 4.43s in Concept 2 and 5.52s in Concept 3. The pattern of results was consistent with workload findings reported in Ligda et al. (submitted).

Analyses of Probe Response Latency

The natural log of probe response latencies (in seconds) was submitted to a repeated-measures ANOVA, with Concept as a factor. A significant effect of Concept on probe response latency was found, ($F(2, 12) = 4.01$, $p = .046$). In Concept 1, the probe response time was 13.77s, whereas response time was 15.80s for Concept 2 and 16.86s for Concept 3 (see Figure 3). Post-hoc comparisons indicated that pilots were faster answering the SA questions in Concept 1 compared to Concept 3, $p = .05$. Again, the pattern of the means was in the same direction as hypothesized. This suggests that when pilots were responsible for separation, they had the lowest probe response latency, implying they had the best SA.

Analyses of Percent Correct Responses to SA Probes

The percent correct responses to the SA questions were analyzed in a similar manner. There was no effect of Concept,

$F(1.06, 6.33) = 2.34$, $p = .18$; however, the direction of the means was consistent with the probe response latency findings. In Concept 1, pilots correctly answered 84% of the SA probes compared to 81% in Concept 2 and 79% in Concept 3. Again, this pattern suggests that pilots have better SA when they are responsible for separation.
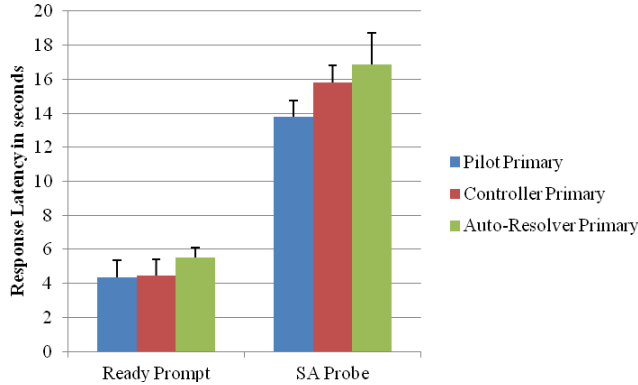

Figure 3: Response Latencies by Concept

An additional analysis was performed that examined probe response latencies as a function of probe accuracy. There was a significant difference between probe response latency for correct versus incorrect SA questions ($t=(1,6) = 2.95$, $p = .03$). Overall, pilots responded quicker to questions they answered correctly ($M=14.85$, $SD=5.29$) than questions they answered incorrectly ($M=18.13$, $SD=7.63$). This is consistent with Durso's (2006) proposition that shorter response times suggest better SA.

CONCLUSION

Pilot situation awareness in the conflict avoidance task was improved when they remained in the decision-making loop. This finding is consistent with that obtained by Dao et al. (2009). The presence of diminished pilot situation awareness under conditions where the Auto-Resolver agent carried greater responsibility for air traffic separation and where pilots did not have onboard CD&R suggests that automation mistrust or complacency factors could play a greater role in influencing pilot situation awareness.

SA probe latencies with the online probe technique was found to be a more sensitive measure of SA than probe accuracy (see also Pierce et al., 2008). The fact that the SA probe latencies were able to distinguish between levels of automation suggest that they are good tools that can be used in the evaluation of operator SA in future ATM concepts.

ACKNOWLEDGEMENT

REFERENCES

Abbott, T. S. (2002). Speed control law for precision terminal area. *NASA Technical Memorandum 2002-211742.* Hamptom, VA: National Aeronautics and Space Administration.

Canton, R., Refai, M., Johnson, W., & Battiste, V. (2005). Development and integration of human-centered conflict detection and resolution tools for airborne autonomous operations. In: *Proc.12th International Symposium on Aviation Psychology*, Oklahoma City, OK.

Dao, A.-Q. V., Brandt, S. L., Battiste, V., Vu, K.-P. L., Strybel, T., & Johnson, W. W., (2009). The impact of automation assisted aircraft separation on situation awareness. *Human Interface, Part II, HCII 2009*, LNCS 5618, 738-747.

Durso, F. T., Bleckley, M. K., & Dattle, A. R. (2006). Does situation awareness add to the validity of cognitive tests? *Human Factors*, 721–733.

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

Erzberger, H. (2006). Automated Conflict Resolution for Air Traffic Control. *Proceedings of the 25th International Congress of the Aeronautical Sciences (ICAS),* Germany.

European Air Traffic Management Programme (2003). The Development of Situation Awareness Measures in ATM Systems. HRS/HSP-005-REP-01.

Granada, S., Dao, A. Q., Wong, D., Johnson, W. W., & Battiste, V. (2005) Development and integration of a human-centered volumetric cockpit display for distributed air-ground operations. *Proceedings of the 12th International Symposium on Aviation Psychology*, Oklahoma City, OK.

Joint Planning and Development Office. (2007). *Concept of operations for the Next Generation Air Transportation System, Version 2.0.* Washington, DC: Author.

Ligda, S. V., Dao, A.-Q. V., Strybel, T. Z., Vu, K.-P., Battiste, V., & Johnson, W. (submitted). Impact of phase of flight on operator workload in a distributed air traffic management system. *54th Annual Meeting of the Human Factors and Ergonomics Society*, San Francisco, CA.

Pierce, R., Strybel, T., & Vu, K.-P. L. (2008). Comparing situation awareness measurement techniques in a low fidelty air traffic control simuluation. *Proceedings of the 26th International Congress of the Aeronautical Sciences (ICAS).* Anchorage, AS.

Prevot, T. (2002). Exploring the many perspectives of distributed air traffic management: the multi Aircraft Control System MACS. *International Conference on Human-Computer Interaction in Aeronautics*, 23-25.

Prevot , T., Homola , J., Mercer, J., Mainini, M., & Cabrall, C. (2009). Initial evaluation of air/ground operations with ground-based automated separation assurance. *Proceedings of the 8th USA/Europe Air Traffic Management Research and Development Seminar,* Napa, CA.

Parasuraman, R., & Wickens, C. D. (2008). Humans: Still vital after all these years of automation. *Human Factors, 3,* 511-520.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans, 3,* 286-297.