Conflict Resolution Automation and Pilot Situation Awareness

Arik-Quang V. Dao[1], Summer L. Brandt[1], L. Paige Bacon[2], Joshua M. Kraut[2], Jimmy Nguyen[2], Katsumi Minakata[2], Hamzah Raza[2], and Walter W. Johnson[3]

[1] San Jose State University, Moffett Field, CA 94035, USA
[2] California State University Long Beach, Dept of Psychology,
Long Beach, CA 90840, USA
[3] NASA Ames Research Center, Moffett Field, CA 94035, USA
{quang.v.dao,summer.l.brandt}@nasa.gov,
{lauren.bacon,josh.kraut,jimmy.nguyen,
hamzah.raza}@student.csulb.edu,
kminakata@gmail.com, walter.w.johnson@nasa.gov

Abstract. This study compared pilot situation awareness across three traffic management concepts that varied traffic separation responsibility between the pilots, air-traffic controllers, and an automation system. In Concept 1, the flight deck was equipped with conflict resolution tools that enable them to perform the tasks of weather avoidance and self-separation from surrounding traffic. In Concept 2, air-traffic controllers were responsible for traffic separation, but pilots were provided tools for weather and traffic avoidance. In Concept 3, a ground based automation was used for conflict detection and resolution, and the flight deck tools allowed pilots to deviate for weather, but not detect conflicts. Results showed that pilot 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 pilot situation awareness on conflict resolution tasks can be improved by keeping them in the decision-making loop.

Keywords: situation awareness, flight deck, automation, NextGen, SAGAT, SPAM.

1 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 [1]. To meet the capacity demands of the future air transportation system, as well as meet or improve safety and efficiency standards, human controller tasks such as 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 [2]. However, there may be trade-offs related to situation awareness when deploying

G. Salvendy, M.J. Smith (Eds.): Human Interface, Part II, HCII 2011, LNCS 6772, pp. 473–482, 2011.

automation [3]. 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.

2 Situation Awareness

Situation awareness (SA) has many definitions. The most widely used definition of SA 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" [4]. This definition implies that SA is mostly internal. An alternative definition suggests that SA can be distributed between the operator and his/her task environment (e.g., information located on a display of traffic) [5].

Endsley [6] developed an off-line probe technique, called the Situation Awareness Global Assessment Technique (SAGAT) to assess SA in experimental contexts and simulations. With this technique, task analysis is first used to identify critical information requirements. Probe questions are then developed to capture the operator's awareness of this information. 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 [7].

Alternatively, Durso, Bleckley, and Dattel [8] 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 in the surrounding environment. Knowing where to find critical information should yield better SA, 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. In order for the operator to be 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, the ready prompt will be accepted by pressing a button or by saying "ready." The probe question is then administered right after the ready prompt has been accepted. This procedure yields three 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 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. That is, the lower the workload, the shorter the ready latencies. 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.

3 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 completely out-of-the-loop leading to complacency [3]. When complacent, the operator no longer proactively seeks to maintain awareness of task 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 imposes high workload. When workload is high, cognitive tunneling can occur where the operator is forced to selectively and narrowly attend to the primary task, reducing the cognitive resources needed to monitor or process other task relevant components [9]. 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 [10].

Dao et al. [10] 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 provide better support for SA than at fully automated levels.

Because Dao et al.'s [10] study examined short, 2-minute conflict scenarios, it is not clear whether the same effect of automation would be observed when pilots must fly longer scenarios that involve different phases of flight as well and where they have additional responsibilities. Thus, the present study expands on Dao et al.'s findings by examining pilot SA when separation responsibility is distributed between pilots, controllers, and automation in longer, 80-minute scenarios.

4 Current Study

Pilots and controllers engaged in real-time simulations focused on trajectory-based en route and arrival operations into Louisville International-Standiford Field Airport (SDF). In trajectory-based operations controllers and pilots attempt to maintain complete trajectories at all times, modifying complete trajectories rather than using temporary vectors. Although both pilot and controller SA was a focus in this study, this paper will only focus on the pilot's SA.

Situation awareness for pilots was examined under three concepts of operations and under high en route traffic density (three times normal). In all three concepts the pilots were responsible for engaging in an interval management task (often referred to as merging and spacing). In ***Concept 1,*** experimental pilots had onboard conflict detection and resolution tools (CD&R) and were responsible for interval management, for autonomous weather avoidance, and conflict resolution/separation assurance (they did not have to obtain concurrence from ATC). In ***Concept 2***, experimental pilots again had CD&R tools and were responsible for interval management. Pilots were also responsible for generating conflict free weather avoidance route modifications but, unlike in Concept 1, they had to downlink proposed routes for concurrence from the ATC (who was responsible for separation assurance). ***Concept 3,*** was similar to Concept 2, but without flight deck CD&R. As a result pilots often could not see traffic conflicts on their proposed routes, requiring the ATC to modify them.

Based on results from Dao et al. [10], 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.

5 Method

5.1 Participants

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

5.2 Apparatus

Pilots in the simulation managed a desktop simulator that included the Cockpit Situation Display (CSD), a PC-based interactive 3-D volumetric display developed by the NASA Ames Flight Deck Display Research Laboratory (see Fig. 1). The CSD provides pilots with the location of surrounding traffic and weather, plus the ability to view planned 4-D trajectories [11]. Although both standard airborne weather radar depictions, and advanced 3D weather depictions were examined, results were not presented as part of the present report. The CSD contained logic that detected and highlighted simulated conflicts and 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, graphically depicting ownship proximity to traffic along the planned route.

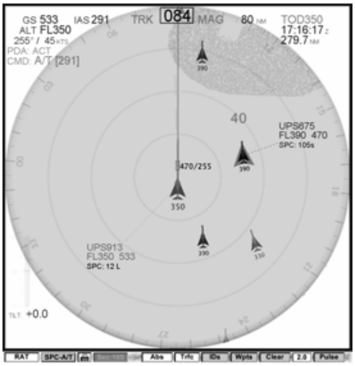

Fig. 1. Plan View of 3-D Cockpit Situation Display (3-D CSD) with Weather Radar

Pilots modified the flight path of ownship for weather and traffic avoidance using the Route Assessment Tool (RAT) [12], a graphical tool that permitted them to move, insert and delete waypoints. In Concepts 1 and 2, the RAT was linked to conflict detection software allowing pilots to find conflict-free paths. In Concept 3 conflict detection was disabled.

The interval management task was implemented using the NASA Langley ASTAR algorithms [13]. When engaged, ASTAR calculated speed adjustments designed to achieve a time-in-trail of the leading aircraft at 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.

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

The Multi-Aircraft Control System (MACS) [14] provided an underlying 747 aircraft simulation, plus a display of flight deck instruments and controls (Fig. 2). These included a primary flight display (PFD), a mode control panel (MCP), a 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. Uplinked route modifications from the controller appeared in the data link window from which they were loaded, visually examined by the pilot on the CSD, and, if acceptable, directly loaded into the flight management system. MACS, a highly versatile piece of software, also provided the interface for controllers and pseudo-pilots. However, since the controllers' activities are not specifically germane to the present report on pilot SA, the reader is directed to a separate book chapter [15] for details of their tasks.

Workload and situation awareness probe questions were administered using a separate touch screen tablet computer. All probes required a yes/no or multiple choice response (equal number of yes/no and multi-choice questions per pilot per scenario).

5.3 Design and Procedure

The independent variable was Concept [Concept 1: Pilot Responsible with Flight Deck CD&R, Concept 2: Ground (Controller) Responsible with Flight Deck CD&R, Concept 3: Ground (Auto-Resolver Agent) Responsible without Flight Deck CD&R] 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.

Experimental pilots started each scenario in an en-route phase of flight during which they initiated an interval management operation (also known as merging and spacing) that continued through the arrival into SDF. Approximately 2 minutes into each scenario the pilots received and implemented their interval management instructions, which included spacing interval and lead aircraft, from an air traffic control (ATC) ground scheduling station. Subsequently pilots also needed to use the RAT to make, or request, a route modification to avoid hazardous weather. In Concepts 1 pilots were responsible for avoiding and resolving all traffic conflicts, and in Concept 2 for generating weather deviation requests that were conflict free. Thus, pilots in all concepts adjusted their route relative to the weather based on their own safety criteria and, in concepts 1 and 2, with respect to the constraints imposed by surrounding traffic. In addition to experimental pilots, confederate "pseudo-pilots" were used to manage the background traffic. This traffic was set at three times today's level to be consistent with the future traffic levels expected when the concepts being explored in this simulation may be implemented.

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?

6 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.

6.1 Timeouts

A repeated measures ANOVA of timeouts as a function of Concept showed no significant effect ($F(2, 12) = 1.99$, $p = .18$). Although not significant, the data pattern suggested that pilots attended to the probe questions more when they were responsible for traffic separation. 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. [16].

6.2 Analyses of Ready Response Latency

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. The p-values were adjusted using Greenhouse-Geisser for violations of Sphericity where appropriate.

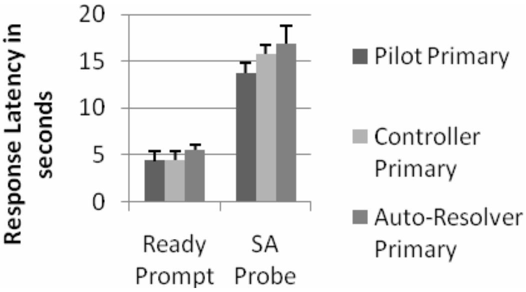

Fig. 3. Response Latencies by Concept

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 less time to respond to the ready prompt in Concept 1 (pilot primary) and Concept 2 (controller primary) than in Concept 3 (auto-resolver primary), see Fig. 3. The pattern of results was also consistent with workload findings from the same study reported in Ligda et al. [16].

6.3 Analyses of Probe Response Latency

Probe response latencies (in seconds) were 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), see Fig. 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.

6.4 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.

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 sec) than questions they answered incorrectly (*M*=18.13 sec). This is consistent with Durso's [8] proposition that shorter response times suggest better SA.

7 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. [10]. Although not significant, the pattern of the results observed for timeouts and response latencies to the probe questions suggest that an intermediate level of automation introduced in Concept 1 can be implemented to help reduce workload. Furthermore, the improved SA scores in the Concept 1 condition where pilots remained involved in the conflict resolution task showed that reduced workload can be achieved without a high cost to SA.

The presence of diminished pilot situation awareness under conditions where the automation carried greater responsibility for air traffic separation and where pilots were not involved in the decision-making suggests that automation mistrust or complacency factors could play a greater role in influencing pilot situation awareness

[3]. Also under these same high automation conditions, mistrust in the automation may have lead to over-monitoring of system behavior and subsequently increasing workload – as shown by higher workload patterns in the Concept 3 condition [9].

SA probe latencies with the online probe technique were found to be a more sensitive measure of SA than probe accuracy (see also [7]). 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.

Findings from this study demonstrate that automated decision support tools can be introduced to the flight deck without significant loss of SA, and that it is possible to keep the operator in the decision-making loop without the burden of high workload. Thus future flight deck system designs should focus on designs that support interaction between the operator and automation. In addition, future studies may implement the SA and workload probe techniques described in this study to examine how to optimally distribute roles and responsibilities between the human operator and automation.

Acknowledgement. This study was supported by the NASA Concepts and Technology Development Project, and in collaboration with NASA cooperative agreement (NNA06CN30A) researchers. These researchers, located at Cal State University Long Beach, Cal State University Northridge, and Purdue University, provided pseudopilots and controllers as part of a distributed simulation network. All participant pilots were tested at NASA Ames FDDRL. We thank Kim-Phuong Vu, Vernol Battiste, and Tom Strybel for their comments on prior versions of this paper.

References

1. Joint Planning and Development Office. Concept of operations for the Next Generation Air Transportation System, Version 2.0. Washington, DC (2007)
2. Prevot, T., Homola, J., Mercer, J., Mainini, M., Cabrall, C.: Initial Evaluation Of Air/Ground Operations With Ground-Based Automated Separation Assurance. In: Proceedings of the 8th USA/Europe Air Traffic Management Research and Development Seminar, Napa, CA (2009)
3. Parasuraman, R., Sheridan, T.B., Wickens, C.D.: 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 (2000)
4. European Air Traffic Management Programme. The Development of Situation Awareness Measures in ATM Systems. HRS/HSP-005-REP-01 (2003)
5. Durso, F., Rawson, K., Girotto, S.: Comprehension and Situation Awareness. In: Durso, F., Nickerson, R., Dumais, S., Lewandowsky, S., Perfect, T. (eds.) Handbook of Applied Cognition, 2nd edn. Wiley, Hoboken (2007)
6. Endsley, M.R.: Measurement Of Situation Awareness In Dynamic Systems. Human Factors 37(1), 65–84 (1995)
7. Pierce, R., Strybel, T., Vu, K.-P.L.: Comparing Situation Awareness Measurement Techniques In A Low Fidelty Air Traffic Control Simuluation. In: Proceedings of the 26th International Congress of the Aeronautical Sciences (ICAS), Anchorage, AS (2008)
8. Durso, F.T., Bleckley, M.K., Dattle, A.R.: Does Situation Awareness Add To The Validity Of Cognitive Tests? Human Factors, 721–733 (2006)

9. Parasuraman, R., Wickens, C.D.: Humans: Still Vital After All These Years Of Automation. Human Factors 3, 511–520 (2008)
10. Dao, A.-Q.V., Brandt, S.L., Battiste, V., Vu, K.-P.L., Strybel, T., Johnson, W.W.: The Impact Of Automation Assisted Aircraft Separation On Situation Awareness. In: Salvendy, G., Smith, M.J. (eds.) HCI International 2009. LNCS, vol. 5618, pp. 738–747. Springer, Heidelberg (2009)
11. Granada, S., Dao, A.Q., Wong, D., Johnson, W.W., Battiste, V.: Development And Integration Of A Human-Centered Volumetric Cockpit Display For Distributed Air-Ground Operations. In: Proceedings of the 12th International Symposium on Aviation Psychology, Oklahoma City, OK (2005)
12. Canton, R., Refai, M., Johnson, W., Battiste, V.: 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 (2005)
13. Abbott, T.S.: Speed Control Law For Precision Terminal Area. NASA Technical Memorandum 2002-211742. National Aeronautics and Space Administration, Hamptom (2002)
14. Prevot, T.: Exploring The Many Perspectives Of Distributed Air Traffic Management: The Multi Aircraft Control System MACS. In: International Conference on Human-Computer Interaction in Aeronautics, pp. 23–25 (2002)
15. Vu, K.-P.L., Strybel, T.Z., Battiste, V., Johnson, W.W.: Factors Influencing The Decisions And Actions Of Pilots And Controllers In Three Plausible NextGen Environments. In: Proctor, R.W., Nof, S., Yih, Y. (eds.) Cultural Factors in Systems Design: Decision Making and Action.. CRC Press, Boca Raton (in press)
16. Ligda, S.V., Dao, A.-Q.V., Strybel, T.Z., Vu, K.-P., Battiste, V., Johnson, W.: Impact Of Phase Of Flight On Operator Workload In A Distributed Air Traffic Management System. In: 54th Annual Meeting of the Human Factors and Ergonomics Society, San Francisco, CA (2010)