

Pilot Confidence with ATC Automation using Cockpit Situation Display Tools in a Distributed Traffic Management Environment

Sarah V. Ligda², Nancy Johnson², Joel Lachter², Walter W. Johnson¹

Flightdeck Display Research Laboratory
NASA Ames Research Center¹
San Jose State University²
NASA Ames Research Center
Moffett Field, CA 94035, United State of America
{sarah.v.ligda, nancy.h.johnson, joel.b.lachter, walter.w.johnson}@nasa.gov

Abstract: NASA's Flight Deck Display Research Laboratory recently investigated air traffic automation designed to alleviate groundside workload in high traffic environments. This paper examines the data from post-experiment debriefings. We found that pilots are comfortable reviewing automated conflict resolutions, as well as modifying those resolutions before execution. The pilots were less comfortable with an automated system that had no pilot or controller human-in-the-loop review process. This traffic management concept will not be optimally achieved if pilots do not trust automation without a human review process in every conflict situation. While initial development of these systems should focus on ways to effectively enable such reviews of the automation, confidence can be expected to increase as pilots develop first-hand experience with the system.

Keywords: automation, conflict resolution, cockpit display of traffic information, CDTI, cockpit situation display, CSD.

1 Introduction

Air travel's increasing demand requires a change in the role of air traffic controllers (ATC) that reduces their per flight workload. To accomplish this, part of the responsibility for maintaining safe separation distances between aircraft must be transferred from ATC to the flight deck crew or to automation. NASA's Flight Deck Display Research Laboratory (FDDRL) recently investigated a form of air traffic automation designed to provide automated conflict resolutions. In this study, the FDDRL examined pilots' acceptance of resolutions produced by automation developed in the Aviation Systems Division at NASA Ames Research Center [1], and how these resolutions compare with pilot-created resolutions.

The automation system used in the present study is designed so that conflict detection and resolution (CD&R) functions would be performed by some combination of ground-based automation, air-based automation, and pilots - thereby substantially alleviating ATC of that particular function. In this paper, we discuss the results from

a debrief administered to pilots after completion of the study. The debrief assessed pilots' confidence and acceptability of the automated resolutions, the employment of these resolutions, and the flight deck tools used to support the pilot's activities. Other analyses from this study have been presented elsewhere [2][3]. This study also complements a similar study that examined ATC performance in a manual and interactive environment and ATC acceptability of automated resolutions [4].

A critical element of the concept under investigation is that all conflict detection responsibility and most conflict resolution responsibility is transferred from ATC to automation. The CD&R capabilities automatically detect a predicted loss of separation from Ownship (incursion of another aircraft into Ownship aircraft's protected zone - within 5 nm lateral and 1000 ft vertical), while the other aircraft is 8 or more minutes away from loss of separation. After detection, the CD&R automatically calculates a proposed resolution. Multiple concepts have been proposed for the use of this type of automated resolution system [4]. For example, if calculated on the ground, this resolution could be sent electronically (datalinked) to the aircraft for review and execution. Alternatively, if calculated on the flight deck, a resolution could be datalinked to the ground for approval prior to execution.

The primary flight deck tool for evaluating, modifying, and/or creating resolutions was the three-dimensional Cockpit Situation Display (3D CSD), a tool developed by the FDDRL to support flightdeck traffic awareness and flight path replanning [5]. The 3D CSD contains a conflict resolution tool that provides the pilot an automated resolution for evaluation, and a Route Assessment Tool (RAT) that can be used to modify the automated resolution or create a novel resolution. Additionally, the 3D CSD provides a three dimensional display of Ownship.

Three levels of automation were examined: *automated*: the automation resolver generated and displayed a conflict resolution on the 3D CSD that the pilot was not allowed to modify, only execute; *interactive*: the automation resolver generated and displayed a suggested conflict resolution on the 3D CSD which the pilot could execute or modify manually utilizing the RAT; *manual*: the automation resolver did not propose a conflict resolution on the 3D CSD and the pilot needed to generate a solution to resolve the conflict using the RAT alone.

In order to provide conflict detection and flight path information, we assume that future aircraft will have access to near real-time information about surrounding traffic (location, altitude, speed, heading, and flight plan). The present study assumes all aircraft are equipped with a version of Airborne Dependent Surveillance Broadcast (ADS-B) capable of transmitting this information to a range of 160 nm.

2 Methods

2.1 Participants

Seventeen air-transport rated (ATP) pilots participated in the study. The average glass-cockpit experience of the pilots was approximately 6000 hours, with a range of 1000 to 13,000+ hours. One pilot was retired (less than 5 years). Two subjects reported having previous FDDRL 3D CSD experience. All participants reported

owning personal computers. The study required the pilots to have a sufficiently high level of skill with the CSD tools to be able to easily assess the automated resolutions and generate unique resolutions. Five pilots failed to adequately master the CSD features during training and thus were excluded from data analysis, leaving twelve participants.

2.2 Equipment

Pilots were tested individually in a quiet, dimly-lit room. The testing station consisted of one 30" monitor on which the 3D CSD was displayed [see Fig. 1]. Pilots manipulated the CSD using a standard two-button mouse. The experimenter sat in another room and monitored the pilots using a digital video camera while the pilot's display was recorded using screen-capture software. The CD&R automation was composed of a conflict detection component and a resolution-generation component. The detection and resolution components utilized broadcast intent (flight plans) and a deterministic prediction logic to calculate conflicts and resolutions. Since no noise or uncertainty was modeled in this study, the automation provided error free predictions. However, it only issued alerts for predicted conflicts within 12 minutes of the anticipated loss of separation.

Conflict resolution responsibility was always assigned to Ownship (the pilot's aircraft). The automated resolution component was a version of a system developed by Erzberger [1] and was generated from identical conflicts found in the ATC study [4]. In addition to the automated resolution component, the pilots were provided with a Route Assessment Tool (RAT) in some conditions. The RAT allowed pilots to modify their flight plans by inserting/deleting waypoints, stretching lateral routes, or inserting/deleting climb and descent segments. Consistent with the constraints on the automated resolutions, no change in flight plans closer than 90 seconds from Ownship at the start of the trial were permitted. This was required to allow sufficient time for approval of plans by groundside automation. A more complete review of the present study's equipment, training, design and task is presented elsewhere [6].

2.3 Training

Each pilot was briefed on the flight concept before engaging in hands-on training. During the hands-on training phase, each pilot was taught the basic functions of the 3D CSD, the CD&R tool, and the RAT, which was used to manually make flight path modifications. Once trained, the pilots completed the experimental trials.

2.4 Design

The experimental design was a 3 (level of automation: automated [no RAT], interactive [automated suggestion plus RAT], manual [RAT only]) by 2 (time to loss of separation – LOS: near-term [8 minutes], far-term [12 minutes]) by 2 (automated resolution type: vertical, lateral) within-subjects factorial design.

In each scenario, there was three times normal traffic density in a double-sized sector composed of two sectors: ZKC 90 sector from the Kansas City center and ZID 91 sector from the Indianapolis center [4]. Forty-eight trials were presented to each pilot (16 trials with each automation level; within each automation level, there were 4 trials at each combination of time to LOS and automated resolution type). The



Fig. 1. The 3D Cockpit Situation Display presenting a conflict alert

automated condition was presented in the first block of trials, while the order of the interactive and manual conditions in the last two blocks was counterbalanced across the participants. Ownship was always responsible for resolving the conflict and was the aircraft which the automation selected (or would have selected) to maneuver to avoid the conflict [7].

After each trial, pilots rated the acceptability of the proposed and executed resolutions on the trials, using a five point scale ranging from unacceptable to excellent. The pilots also rated the complexity of the

conflict situation, using a five point scale ranging from very simple to very complex, and answered questions relating to their situation awareness [see 6].

2.5 Task

In the automated trials, pilots were required to review and execute the proposed conflict resolution, but could not modify the proposed resolution. In the interactive trials, pilots were required to review the proposed resolution and could (optionally) use the RAT to modify that resolution or create a new resolution before execution. In the manual trials, pilots were required to use the RAT to create their own resolution, then review and execute that resolution. Pilots were required to evaluate the acceptability of the proposed resolutions in the automated and interactive conditions, as well as the acceptability of the manually-created resolutions in the interactive and manual conditions. For all cases, the pilots had up to 90 seconds to view and/or modify the resolutions. In addition, pilots were asked to verbalize the rationale for their actions as they performed each task. The pilot ratings of the acceptability of the proposed automated resolutions will be analyzed in the following section.

2.6 Debrief

After completing the trials, each pilot was debriefed by a researcher. All pilots were given the same series of questions and asked to provide their answers verbally. They were questioned about their acceptance of the concept, resolution strategy, and trust in the automation. The questions asked in this debrief are described along with the responses in the following section.

3 Results & Discussion

3.1 Acceptability of Proposed Conflict Resolutions

In the automated and interactive conditions, pilots evaluated the acceptability of the proposed conflict resolutions generated by automation. Pilots rated the resolutions on a five point scale:

- 1) Unacceptable - ATC coordination required
You believe the resolution was unacceptable and would reject it because it compromises the safety of flight or you are unable to comply. ATC coordination is required to find a new resolution.
- 2) Poor - ATC coordination sought
You believe the resolution is poor and would definitely seek ATC coordination because a new resolution is highly desired.
- 3) Marginal - ATC coordination probably sought
You believe the resolution is marginal and would probably seek ATC coordination because a better resolution is possible.
- 4) Good - ATC coordination probably not sought
You believe that the resolution is good, although there might be a better one. You would probably not seek ATC coordination.
- 5) Excellent - ATC coordination unnecessary
You believe that the resolution is excellent and would not seek ATC coordination.

Of particular interest are the unacceptable, poor, and marginal ratings that the pilots gave to the resolutions generated by the automation. 384 automated resolutions were presented to the 12 pilots in the automated and interactive conditions; of those, 115 resolutions were rated 3 or below (30%). Pilots were asked to provide verbal comments explaining the rationale for their ratings. Pilot's explanations were generally within (but not limited to) 4 categories:

- 1) Safety (e.g., this resolution was less safe because... or could be safer if...)
- 2) Efficiency (e.g., this resolution was not efficient or could have been more efficient because...)
- 3) Comfort/severity of maneuver (e.g., not comfortable making the maneuver required or the maneuver required could have been less severe if...)
- 4) Preference (e.g., I would prefer to go vertically or laterally...)

Not all of the verbal comments fit in these four categories. Of the 115 comments that were rated 3 or below, 87 resolution comments fit into these categories. However,

because a few resolution comments fit into multiple categories, a total of 117 resolution comments were analyzed. Frequencies and percentages of resolution comments in the four categories are presented in Table 1.

Table 1. Frequencies and percentages of resolution comments rated marginal or below

	<i>Safety</i>	<i>Efficiency</i>	<i>Comfort/Severity</i>	<i>Preference</i>
<i>Frequency</i>	26 comments	57 comments	11 comments	23 comments
<i>Percentage</i>	22%	49%	9%	20%

As illustrated in Table 1, 31% of the low acceptability (marginal or lower) ratings related to safety or comfort/severity of maneuver. In the context of all of comments made in response to the automated resolutions, 7% of the comments are devoted to these two areas (5% safety, 2% comfort/severity). The majority (69%) of the low acceptability comments given to the automated resolutions occurred in the context of pilot preference and efficiency. More specifically, the pilots might have preferred a lateral maneuver rather than a vertical one to conserve fuel, or a maneuver that had less deviation from the original flight path. These findings suggest that future development of these automated systems should focus on methods to provide safer and more efficient proposed maneuvers.

3.2 Acceptability of Concepts

The first series of debrief questions were regarding pilot acceptance of several different air traffic management concepts, each providing the cockpit with conflict resolutions. The pilots were questioned on their level of comfort with four resolution concepts that differed in whether the pilot and/or ATC were allowed to review automated resolutions prior to their implementation:

- 1) An automated system that detects conflicts and generates resolutions which are reviewed by ATC before being datalinked to the aircraft where the pilot does a final review using 3D CSD.
- 2) An automated system that detects conflicts and generates resolutions which are NOT reviewed by ATC before being datalinked to the aircraft where the pilot does a final review using 3D CSD.
- 3) An automated system that detects conflicts and generates resolutions which are then reviewed by ATC before being datalinked to the aircraft where the pilot does a final review based on the datalinked route only (no flightdeck CD&R)
- 4) An automated system that detects conflicts and generates resolutions which are NOT reviewed by ATC before being datalinked to the aircraft where the pilot does a final review based on the datalinked route only (no flightdeck CD&R).

Almost all pilots reported being comfortable with the first concept, which allows for prior review of automated resolutions by both pilots and controllers (92% comfortable, 8% somewhat comfortable). In contrast, most pilots were uncomfortable with the last concept, which permits no prior human review of the automated

resolutions (17% comfortable, 25% somewhat comfortable, 57% not comfortable). The other two concepts fell in between, (pilot review only 42% comfortable, 50% somewhat comfortable, 8% not comfortable, and ATC review only 67% comfortable, 25% somewhat comfortable, 8% not comfortable).

Comfort ratings were submitted to a 2 (Pilot Review) x 2 (Controller Review) analysis of variance (ANOVA). The comfort ratings were dummy coded as 0, 1 and 2 for not comfortable, somewhat comfortable, and comfortable, respectively. This analysis yielded significant main effects for Pilot Review ($F(1,11) = 8.19, p < .05$) and Controller Review ($F(1,11) = 30.31, p < .05$), but no significant interaction ($p > .10$), see Fig.2.

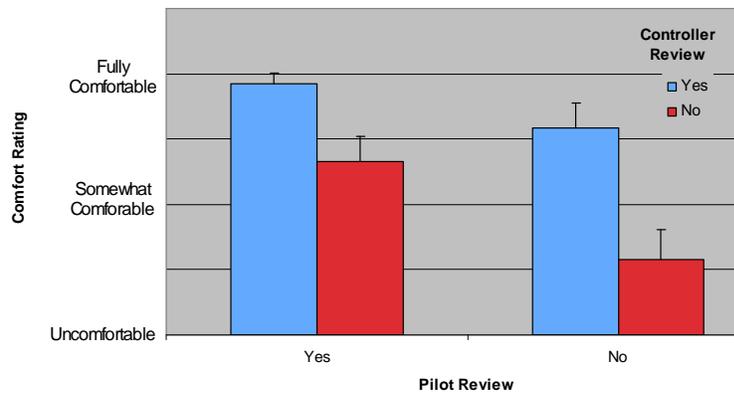


Fig. 2. Graph of pilot comfort with varying concept levels

Finally, the ratings of pilot only review were compared with those for ATC only review to see if pilots differentially valued one type of review over the other. Despite the small trend in favor of ATC review, this effect was not significant ($p > .10$). Thus, based on the ratings, the pilots appear to prefer human review of automated resolutions compared to not having a human operator in the loop. The pilots' comments also supported this conclusion.

During final debriefing, pilots stated that they were comfortable reviewing the resolutions created by the automation, as well as modifying those resolutions before execution. However, they stated that their comfort level and trust in the automation would be higher if ATC also reviewed the resolution before they considered it. The pilots were somewhat less comfortable with a system that had no human-in-the-loop review process of conflict resolutions before they were datalinked to the flight deck - although the pilot was given the option of final review at all times. This latter finding could be an artifact from the present-day system in which ATC provides most conflict detection and resolution services for the pilot (the exception being TCAS alerts). Alternatively, pilots might prefer a system that has a human review process in addition to only reviewing it themselves. Pilots stated that, although they would like to see the concept tested in the field before they were asked to use it, they would most likely accept a system that included automation as long as a human remains actively

involved in the process. As one pilot stated when asked if he felt comfortable with the first concept: *“Absolutely! Especially since it’s got the human element involved as a backup. Obviously it’s an automated system that doesn’t see all the human elements. As long as you have a human element in there to watch over it and make sure it doesn’t go awry, and doesn’t interfere too much that it defeats the purpose, yea, I’m very comfortable with that concept.”*

3.3 Preferred Resolutions

The pilots were also asked a series of questions regarding the characteristics of both the automated resolutions and the resolutions in their actual flights. None of the pilots had any ATC experience, and most reported that their flight school training (TCAS training in particular) was an important factor when learning how to resolve conflicts in the most efficient and effective manner. Overall, the pilots said that they preferred speed and vector changes (11 out of 12), depending on the situation. Furthermore, they were less enthusiastic about altitude changes. This might be due to a motivation to conserve fuel. When questioned about what factors they considered when making conflict resolutions, the pilots stated that their first concern was the overall safety of their flight, but that they also consider fuel and time to the gate to be major factors.

In their initial training for this study, the pilots were told they would be flying at the altitude assigned to them by their flight dispatchers in all the scenarios. Therefore, they may have inferred that an altitude change was suboptimal for fuel. This overriding concern with fuel also appeared in the pilots comments: once they had satisfied safety concerns (avoiding conflicts), they chose resolutions that would not negatively impact their fuel consumption. The pilots were prepared to execute any resolution that satisfied these goals: *“It depends on the situation. Sometimes a vector is better, sometimes altitude is better and sometimes a speed change is better. It depends on terrain, weather, turbulence, fuel, traffic flow. Each situation is independent of the other.”*

3.4 Time Pressure

In a three-times normal traffic density environment, time is of the essence. The pilots were given 90 seconds to review and accept, or review and modify/create, an acceptable conflict resolution. If they were unsuccessful in completing the process in that time frame, the resolution became null and void and the scenario ended. Nine out of the 12 pilots stated that they wanted more than 90 seconds to resolve some conflicts, especially multiple conflicts – *“Yes. Oh, yes. It was a 90 second crunch.”* However, the pilots felt the allotted time was sufficient in certain cases, usually when they only needed to resolve a single conflict. A majority of the pilots felt that if they were given more time to resolve the conflicts, they would have considered more alternative resolution strategies. *“Absolutely. Absolutely! I just went with the first one that worked.”*

3.5 Flightdeck Display

Pilots were asked a series of questions regarding the 3D CSD display and its suite of tools, in particular: the usefulness of the information displayed, display clutter management, the display's impact on decision-making, and what additional information might be of assistance to them. Each pilot reported that they used a combination of features and tools that worked best for them: the assigned colors of the aircraft (blue = above, green = below, white = co-altitude), adjusting the range and pulse predictor, aircraft identification tags, distance to airport or top of descent, and others. All the pilots felt the display was cluttered; this was most likely due to the traffic density being set to three-times current day traffic levels. Pilots provided a number of ideas for clutter mitigation in addition to those currently provided by the 3D CSD. The pilots also provided useful ideas for information they would like to see on future 3D CSD applications, which are available, but not used in this study (e.g., terrain, weather). Of particular note, most pilots (7 out of 12) suggested that non-conflicting aircraft or aircraft a specified distance from Ownship should be removed from the display – *“Remove the aircraft that are nowhere near a possible conflict. Some that could potentially be conflicts, yeah, leave them there.”*

With regards to decision-making information, the pilots were queried on both the provided information that supported their evaluation and additional information that might have been useful. As noted, each pilot modified the tools to best support their own strategy, but most felt they had sufficient information to make an informed resolution decision. When asked about the minimum information needed to make a safe resolution decision, responses varied by pilot. The most common requests were altitude, heading, conflict color change, and predicted future position.

3.6 Trust in Automation

When asked whether pilots would trust automated resolutions compared to the manual resolutions in the field, and what could be provided to increase their level of trust in the automation, all of the pilots involved stated they would trust the automation. The majority of pilots (8 out of 12) would like the automation to be always activated in their flights, but the pilots would make the final decision regarding the suggested resolutions. *“Well, the suggested resolution always on doesn't bother me because I can always fly the airplane myself without using the resolution.”* All of the pilots felt that their conflict resolutions were either *“very similar”* or *“somewhat similar”* to the automated resolutions. Also, all the pilots felt that with sufficient training and proper certification that they would have an increased trust in the presented automation. *“Yeah, I think I would. If it were certificated by the airline or by Boeing or whoever, Airbus, and the FAA. If everybody blessed it, why shouldn't I?”*

4 Conclusion

The majority of the pilots were naïve to the distributed traffic management concept before their participation in this study (10 out of 12 pilots); therefore, hesitations

about utilizing and performing with this new task were expected. Despite this, we found that the pilots' trust and confidence levels in conflict resolution automation were high, provided that there is a human-in-the-loop resolution review process. This can be expected for several reasons. Before pilots can be expected to trust automation without human review, thorough familiarity and support in application is necessary: pilots need to be confident that all groundside and airside operators are familiar with, and are optimally efficient, in utilizing this system. Furthermore, ensuring pilots have a basic understanding of the primary factors that are considered by the automation when creating a conflict resolution - such as weight, winds, speed and altitude - may support an increase in their overall level of confidence in the system.

This system is designed to alleviate groundside workload within a high traffic density environment. However, this traffic management concept will not be optimally achieved if pilots do not trust automation without a pilot or controller review process in every conflict situation. Pilots stated that, while initial development of these automated systems should focus on ways to effectively enable this review, trust and confidence will increase with first-hand experience of using the system. They cited their experience with TCAS as an example of a system where such trust evolved over time. Overall, the pilots were optimistic that similar confidence and trust would be achieved with this new automation-assisted system.

5 References

1. Erzberger, H.: Automated Conflict Resolution for Air Traffic Control, Proceedings of the 25th International Congress of the Aeronautical Sciences (ICAS), Germany (2006)
2. Battiste, V., Johnson, W.W., Dao, A.Q., Brandt, S.L., Johnson, N., Granada, S.: Assessment of Flight Crew Acceptance of Automated Resolutions Suggestions and Manual Resolution Tools. Proceedings of ICAS 2008, Anchorage, Alaska (2008)
3. Johnson, W.W., Battiste, V., Dao, A.Q., Brandt, S.L., Johnson, N., Granada-Vigil, S.: Pilot Acceptance of Automated Resolution Advisories: Preliminary Evaluation. Presentation at Human Factors and NextGen: The Future of Aviation, Arlington, TX (2008)
4. Homola, J.: Analysis of Human and Automated Separation Assurance at Varying Traffic Levels. Master's Thesis, San Jose State University, San Jose, CA (2008)
5. 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. Proceedings of the 12th International Symposium on Aviation Psychology, Oklahoma City, OK (2005)
6. Dao, A.Q., Brandt, S.L., Battiste, V., Vu, K.-P.L., Strybel, T., Johnson, W.W.: The Impact of Automation Assisted Aircraft Separation on Situation Awareness. Paper presented at the 13th annual International Conference on Human-Computer Interaction, San Diego, CA (2009)
7. Erzberger, H. Transforming the NAS: The Next Generation Air Traffic Control. Proceedings of the 24th International Congress of the Aeronautical Sciences, Yokohama, Japan (2004)