

INVESTIGATING THE COMPLEXITY OF TRANSITIONING SEPARATION ASSURANCE TOOLS INTO NEXTGEN AIR TRAFFIC CONTROL

*Ashley N. Gomez, Lynne Martin, Jeffrey Homola, Susan Morey, Christopher Cabrall, Joey Mercer
San Jose State University/NASA Ames Research Center, Moffett Field, California*

Thomas Prevot

NASA Ames Research Center, Moffett Field, California

Abstract

In a study, that introduced ground-based separation assurance automation through a series of envisioned transitional phases of concept maturity, it was found that subjective responses to scales of workload, situation awareness, and acceptability in a post run questionnaire revealed as-predicted results for three of the four study conditions but not for the third, Moderate condition. The trend continued for losses of separation (LOS) where the number of LOS events were far greater than expected in the Moderate condition. To offer an account of why the Moderate condition was perceived to be more difficult to manage than predicted, researchers examined the increase in amount and complexity of traffic, increase in communication load, and increased complexities as a result of the simulation's mix of aircraft equipage. Further analysis compared the tools presented through the phases, finding that controllers took advantage of the informational properties of the tools presented but shied away from using their decision support capabilities. Taking into account similar findings from other studies, it is suggested that the Moderate condition represented the first step into a "shared control" environment, which requires the controller to use the automation as a decision making partner rather than just a provider of information. Viewed in this light, the combination of tools offered in the Moderate condition was reviewed and some tradeoffs that may offset the identified complexities were suggested.

Introduction

Introducing automation into a critical system (that cannot be taken "offline") is a complex process, made more complicated when this system is not software alone but a dynamic and safety-critical human-automation system like Air Traffic

Management (ATM). Therefore, for the transition from current day Air Traffic Control (ATC) to a full Next Generation Air Transportation System (NextGen) environment [1], not only is the end-point of concern but also the transitional phases, where automation is being first introduced. For example, in transitional phases it is likely that aircraft equipage will vary, and while some controller tools are available, not all planned features and functions will be accessible. The human-automation interaction in these phases is of concern. There is a requirement for the tools in transitional phases to assist controllers with managing the predicted increase in air traffic demand [2] without any degradation in the performance of the system relative to the high bar set by current day operations.

Current day Air Traffic Control relies heavily on a human operator possessing the skills and cognitive resources to manage the traffic in their sector. This is true at a broad level but also for specific ATC functions, for example, the controller is responsible for nearly all separation-related tasks. If the separation assurance function is viewed as a human supervisory control system [3], then current day ATC could be deemed to function at an automation level of 1 [4], that is: "the computer offers no assistance: human must take all decisions and actions" (p287, Parasuraman, Sheridan, & Wickens, 2000, [5]). Although the controller has an enhanced radar display, it provides information to enable controller situation awareness rather than decision support (advice about what action to take).

The Concepts and Technology Development Project of the NASA Airspace Systems Program is exploring fundamental changes to the separation assurance process [6]. It will take advantage of advances in automation to allocate many separation assurance tasks to ground-based automation, freeing the controller to manage their traffic and oversee the

aircraft separation function. Fully operational, and after full implementation, such a system would “execute automatically, then necessarily inform humans” (p287, Parasuraman, et al., 2000 [5]) for routine separation tasks, which is an automation level of 7 in Sheridan and Verplank’s [4] distinction of levels of automation.

Endsley and Kaber [7] explored different levels of automation (LOA) based on a taxonomy they developed from the work of Sheridan and Verplank. They found that the LOA significantly impacted task performance in a radar monitoring task in ways that were sometimes positive and sometimes negative, often varying with LOA, e.g., a level 7 implementation produced slightly better performance than a level 8 instantiation [7]. They also suggested participants became distracted under mid-to-low levels of automation. Endsley and Kabers’ findings, and those of others (e.g., [8]), suggest that the impact on task performance of the particular combination of function allocation to both human and automation under different phases of implementation (different LOA) needs to be examined.

While some systems can be implemented in an all-at-once fashion, advancing the level of supervisory control from 1 or 2 to 7 or beyond in one step, larger systems are likely to be implemented in phases for numerous practical reasons. However, if there is a choice, selecting which tools should be introduced in early phases and which should be reserved for more advanced stages is a problem in itself [9], as it is possible that some combinations of tools may work better together than others.

The separation assurance concept, employed in this study, was developed from the Advanced Airspace Concept (AAC) proposed by Erzberger [10]. It comprises three ground-based elements: the Automated Airspace Computer System (AACS), the Tactical Separation Assisted Flight Environment (TSAFE) and a controller interface.

A human-in-the-loop separation assurance study was conducted that focused on human performance when automated tools were combined into ATC operations [6]. Researchers observed whether controllers could accommodate higher levels of demand in the en route airspace (i.e., more traffic) with no reduction in current levels of safety and no increase in Air Traffic Controller load. The study reported the impact on human performance and

safety as it was tested through four potential phases of implementation, beginning with the current day ATC system as a baseline, through two transitional phases, to a full-implementation “Maximum” phase, where ground-based automation performed the separation task from detection to resolution for all aircraft without direct operator involvement [6]. Each of the phases were characterized not only by differences in the level of automation available but also by different levels of aircraft equipage and data communications. This paper will delve into human interactions with the automation system and investigate how certain combinations of tools affected performance of ATC duties, with specific focus on the third phase of implementation where the impact of the level of automation was unexpected.

Operational Environment/ Procedures

Study Conditions

The four study conditions were chosen to represent a baseline Current Day phase and three envisioned future phases of implementation of a ground-based separation assurance concept based on the AAC [10]. Implementing the concept made two significant changes to the ATC environment. The most obvious was the addition of conflict resolution tools on the controller workstations. However, the second, and possibly more significant change, was to the controllers’ responsibilities. That is, the ATC environment developed into an ATM environment where the controller moved from being responsible for safe separation to being responsible for general oversight and system exceptions (Table 1). Additionally, study conditions were varied by the level of aircraft equipage – the number of aircraft with Data Comm capabilities (TFR) – and environmentally – in the amount of traffic in the scenario – so that the environmental complexity was commensurate with the LOA available to control it. Thus, the mixture of aircraft and the tools available in any given condition were designed to reflect transitional environments.

Table 1. Study Conditions and Variables

	Current Day	Minimum	Moderate	Maximum
Traffic level vs. Today's	Same as today	1.2x	1.5x	~2.0x
Data Comm equipage	0%	~25%	~50%	100%
Controller's responsibility	Safe separation	Safe separation	Safe separation	System exceptions
Ground automation's responsibility	Current day	Current day	Critical support role as conflict detection and resolution tool	Conflict detection, strategic resolutions, tactical resolutions, exception alerting
Ground automation tools	Conflict probe, conflict list, pre-probed fly out altitude menu	Conflict probe, conflict list, pre-probed fly out altitude menu, time to go in flight data block (FDB), lateral trial planner. Data Comm, auto top of climb (TOC), auto handoff (HO) for equipped aircraft	Conflict probe, interactive conflict list, pre-probed fly out altitude menu, time to go in FDB, 5 th line conflict advisories, trial planner; Data Comm, auto TOC, auto HO, interactive auto resolver for equipped aircraft	Conflict probe, interactive conflict list with auto resolver status info, FDB pre-probed fly out altitude menu, auto TOC, auto HO, auto uplink auto resolver conflict resolutions, auto uplink of TSAFE advisories, post TSAFE auto back-on-course auto uplink, turn TSAFE auto off & on
Frequency changes	Via voice	Via voice or Data Comm	Via voice or Data Comm	Via Data Comm
Trajectory changes	Via voice	Via voice	Via voice or Data Comm	Via voice or Data Comm
Handoff behavior	Manual initiate, manual accept	Manual initiate & accept (IFR); auto initiate, auto accept (TFR)	Manual initiate & accept (IFR); auto initiate, auto accept (TFR)	Auto initiate, auto accept – all TFR

Airspace

The simulated airspace used for this study consisted of five adjacent, high-altitude, en route test sectors, constructed to represent Cleveland Center airspace (ZOB) (see Figure 1). The floor of the overall test airspace was set to flight level (FL) 330. One participant R-side controller and one supporting D-side controller pair was assigned to each of the five test sectors. Confederate positions maintained the airspace surrounding the test area, both below FL 320 and the surrounding FL 330 and above.

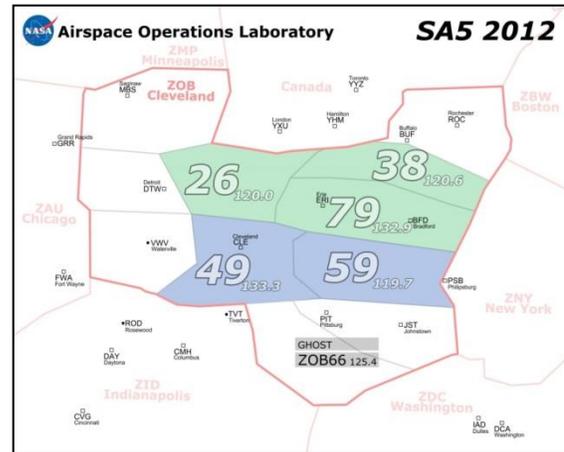


Figure 1. Cleveland Center Test Airspace. The five sectors were divided up into north and south areas, identified here by color.

Traffic

The traffic scenarios used were based on actual traffic from the Cleveland Center area and each scenario was comprised of a mix of ZOB arrivals and departures to and from area airports with level overflights. As shown in Table 1, in the Current Day condition, traffic was set at today's levels and sectors had a Monitor Alert Parameter (MAP) of 18. Traffic was increased from this level by 20% in the Minimum condition, 50% in the Moderate condition, and was nearly doubled in the Maximum condition (see figure 2). As the traffic increased, the percentage of aircraft equipped with air-ground data communications capabilities also increased, from 100% unequipped (IFR) traffic to all Data Comm equipped aircraft (TFR) (see figure 3).

Through scenario design and the controller's interaction with the system, a number of conflicts occurred that required some level of controller involvement. Conflicts were detected and displayed

automatically to the controller; how they were displayed varied depending on the condition.

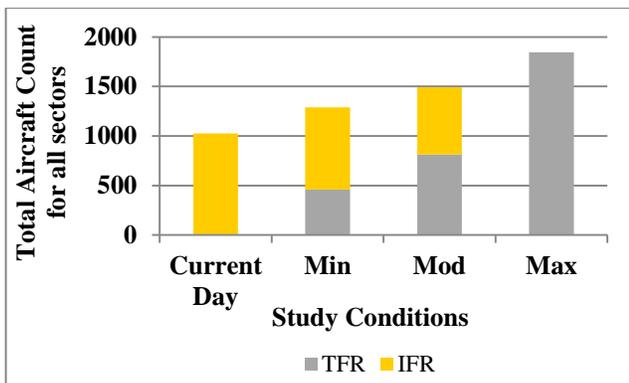


Figure 2. Total Aircraft Count in the Simulation per Condition. Note that these are averages across all runs in a given condition.

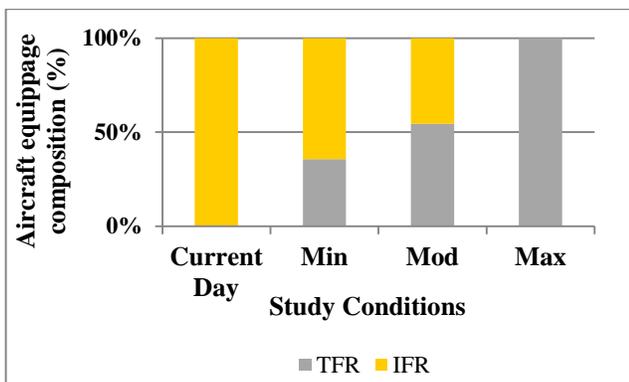


Figure 3. Percentage of Unequipped (IFR) & Equipped (TFR) Aircraft in the Traffic Scenarios for the Four NextGen Study Conditions

Apparatus

The simulation platform used for this study was the Multi Aircraft Control System (MACS) [11], a software platform developed in the Airspace Operations Laboratory (AOL) that has a wide range of simulation and rapid prototyping capabilities. Each controller workstation was equipped with a Barco display and data system replacement (DSR) trackball and keyboard. Voice communications were conducted through a custom, stand-alone voice system with a dedicated server.

Participants

Seven individuals served as test participants for this study. Five were radar controllers (R-side), and two served as supervisors. Test participants were current or recently retired. Additionally, five retired

controllers served as radar associate (D-side) positions in support of the test participants, and three as confederate “ghost” controllers responsible for the traffic outside the test airspace. Ten airline pilots operated eight mid-fidelity, single-aircraft flight simulators (referred to as ASTORs), and ten general aviation/corporate pilots operated multi-aircraft stations.

Procedures

The study was run over a total of eight days, with two consecutive days paired for a given condition. The first day in each pair was spent training for the condition of interest and the second day was spent collecting data by way of six study runs for that condition. The first two days focused on the Current Day condition, the second two days on the Minimum condition, the third two days on the Moderate condition, and the last two days tested the Maximum condition. Thus, a total of 24 data collection runs were completed across the four conditions.

Each of the 24 runs was 40 minutes in length with the traffic building up gradually to be at its peak by the midpoint of the run. The North and South areas (Figure 1) were housed in physically separate rooms, each with an assigned area supervisor who monitored the traffic situation as well as the workload of the participant radar controllers. It was up to the supervisor to judge whether/when D-side support was needed.

For most conditions, the R-side controllers were responsible for the safe separation of traffic, although in the Maximum condition their task was to oversee the automation and manage traffic by exception. Their other tasks, such as traffic hand-offs and managing local arrivals and departures were also affected by the level of automation available (see Table 1). In the Current Day condition they had to complete these tasks manually but by the Maximum condition they oversaw the automation as it completed these tasks.

The automation’s role changed through the four conditions of the study. In the Current Day condition, it detected conflicts and alerted the controllers who then had to solve the problems themselves (Table 1). In the Minimum condition, the automation offered more information about the conflicts and provided resolutions along the lateral

dimension. By the Moderate condition, the automation found resolutions to strategic conflicts, but controllers had to send the resolution to the TFR aircraft, which was their opportunity to review it. In the Maximum condition, the automation solved both tactical and strategic conflicts autonomously, informing the controller through a status panel on the display. The controller could intervene at any point, especially in tactical conflicts, and take over the problem if they desired.

Thus, the eventual endpoint in our four phase transition into NextGen (being the Maximum condition) would allow for a shift in controllers' responsibilities to a more exception-based role. The controller was alerted to conflicts and resolutions that fell outside of the pre-defined thresholds. Once the controller took control of a conflict it became his/her responsibility until s/he "released" the aircraft pair back to the automation using a keyboard command.

Tools

For this paper, there is a focus on the variability in the LOA of tools available to the controllers and how the tools were used in the context of the condition. The tools provided were either informational tools, decision support tools (DSTs), or supervisory control systems, as defined by [7]. The chief difference between these tools is their function, but this is characterized by the way they display information, their part in the decision process of an action to be taken and the implementation of that action, and whether they alter the controller's role in ATM. Our working definition of an information tool is one where a computer-based system provides information about a situation that is likely to require an action but identification, selection and execution of the action is the controller's responsibility. Here, the decision support tools generate a single resolution, usually represented as the best option given the circumstances, which provides a decision-action option and again, it is the controller's decision to select the action. However with DSTs, once an option is selected, it is turned over to the computer/automation to implement. In our "supervisory system", automation plays a larger role by generating options, selecting the option to implement, and carrying out that action. The controller mainly monitors the system and intervenes if necessary.

Depending on the condition, controllers had access to the various tools, but were not required to use them (refer to table 1). In the Current Day condition, controllers had three information tools available to assist them with separation assurance – the conflict probe, conflict list, and pre-probed altitude flyout menu. In the Minimum condition, controllers had the same information tools as in the Current Day condition, with the addition of time to go in the flight data block (FDB), and a lateral trial planner in addition to the vertical flyout menu. And with the addition of equipped aircraft, data comm, auto top of climb (TOC), auto handoff (HO) were also available for equipped aircraft. In the Moderate condition, all of the same informational tools were available with more interactive capabilities; however two DSTs were added with the 5th line advisories for all conflicting aircraft and an interactive strategic auto resolver for equipped aircraft. Again Data Comm, auto TOC, and auto HO were available for equipped aircraft (see figure 4 below).

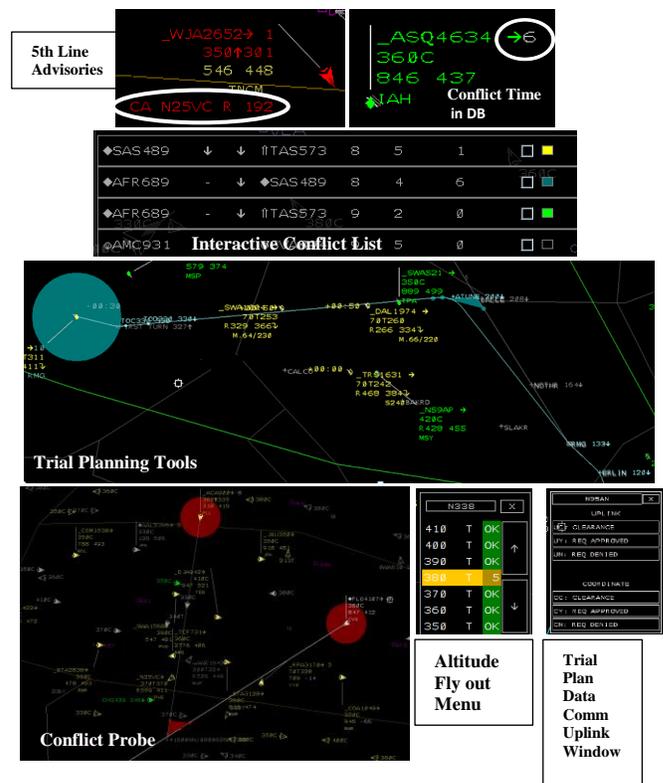


Figure 4. Example of Controller Tools from the Moderate Condition

With the advancement of the conditions explained thus far, the separation assurance tools increased in number and capability, but it was still

the responsibility of the controller to decide on and implement a particular action. In the Maximum condition, was the first instance of a supervisory system, where the automation detected conflicts, computed resolutions, selected a resolution, and implemented the resolution for both short term (tactical/TSAFE) and long term (strategic) conflicts by automatic uplink to the aircraft. Controllers monitored clearances issued by the automation through a status window, and were alerted to conflicts and resolutions that fell outside of the pre-defined thresholds, and intervened to solve the conflicts that the automation could not.

Data Collected

Non-voice data communications as well as both periodic and event driven data were collected via MACS' internal collection capabilities and were supplemented through audio-video screen capture software. At the end of each run, participants completed a questionnaire that asked for their opinions about different aspects of the run, including their perceived workload, situation awareness and the acceptability of the automation configuration for the run.

Results

Human Performance Findings

The results below are presented in two sections; firstly, a set of (predominantly subjective) human and system performance findings, and a second set that explores how and why the participants used the automation available (the "tools"). In both sections, while results are presented for all four study conditions, the focus is on the Moderate condition findings and where they were, or were not, in line with our predictions for these data specifically, and the study aims in general. The study aims were that performance in future conditions would equal or exceed Current Day performance. The predictions for our specific analysis reported below were that perceived operator load would gradually reduce across the four conditions from the current day to the far future. The remainder of this paper will explore this prediction.

Workload Ratings

The NASA Task Load Index (TLX), a measure of perceived workload, was developed by Hart &

Staveland [12] and is comprised of six subscales. In the present study, four of these subscales were asked in the post run questionnaire where participants responded on a 1 to 7 scale ("very low" to "very high") to the level of time pressure, effort, mental demand, and frustration they felt. A mean was calculated for each subscale across the six runs in each condition, and these are depicted in Figure 5.

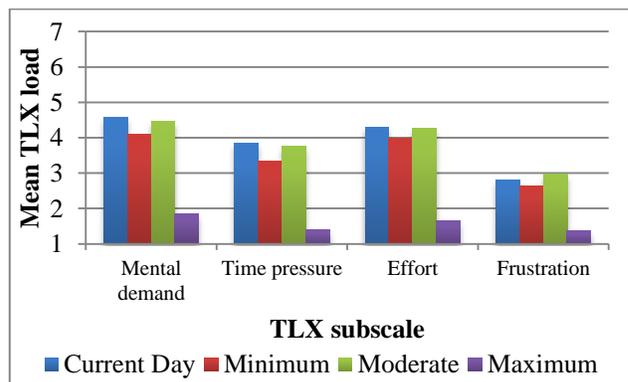


Figure 5. Mean Controller Ratings for NASA TLX Subscales

On all four TLX subscales, participants reported the highest mean load in the Current Day condition. Mental demand, time pressure and effort were all reported to be "moderate" (around 4) and there was "some" frustration ($m=2.81$). For the Minimum condition, mean load is slightly lower on all four subscales, not by much, in general a half-scale point, but there is a trend. Looking to the Maximum condition, mean load is much lower on all four scales – with the mean rating being "low" or "very low" reflecting the change in the controllers' role in this condition.

Our prediction was that workload would decrease from the Current Day across levels of automation, with workload reducing to its lowest in the far future (Maximum) condition. Workload across the study followed this pattern in general but the mean workload for the Moderate condition was higher than for the Minimum condition, contrary to our prediction, and instead, on average, reported loads close to Current Day levels.

Situation Awareness Ratings (SART)

A second subjective measure obtained through self-report was the Situation Awareness Rating Technique (SART), which was developed by Taylor & Selcon [13] and has three subscales. Again, participants responded from 1 to 7 on each sub-scale ("very low" to "very high") to give their estimation

of their situational Understanding, their Capacity to take in more information and the Demand on their attention. The SART score was calculated from these ratings, producing an overall score from -5 to 13 for each participant, and then a mean for each condition was obtained, shown in Figure 6.

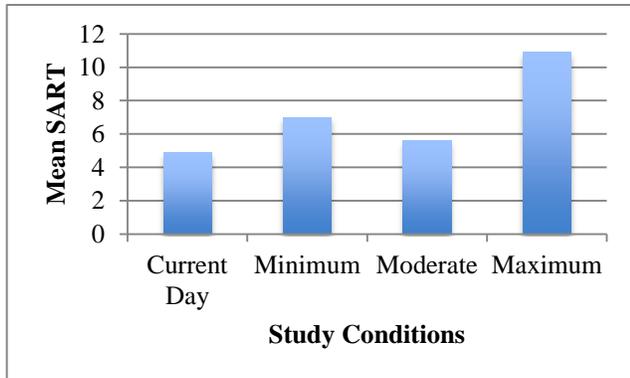


Figure 6. Mean SART Scores Given by Controllers

Overall, mean situation awareness showed an increasing trend from the Current Day (4.87, “moderate awareness”) to the Minimum condition (7.00) and on to the Maximum condition (10.93, “high awareness”). The Minimum awareness scores were significantly greater than those in the Current Day ($Z=2.639$, $n=29$, $p=.008$) when tested using a Wilcoxon Signed Rank Test. The exception to these results was a downward shift in the Moderate condition rating ($m=5.6$, “moderate awareness”), which is still slightly higher than awareness reported in the Current Day but not markedly higher, as predicted. The difference between the Current Day and the Moderate condition was not significant.

Our prediction was that situation awareness would increase from the Current Day to the Maximum condition, with awareness being at its highest in the far future (Maximum) where participants should feel the automation increased their Capacity over the Demand. Situation awareness under three of the conditions followed our predictions, increasing with LOA, but for the Moderate condition it did not.

Acceptability Ratings (CARS)

A third subjective measure gathered through the questionnaire was the Controller Acceptance Rating Scale (CARS) [14] that uses a scale derived from the Cooper-Harper to assess controllers’ comfort level with a system. Participants work through a series of

yes/ no questions to reach a three-point rating scale where they assess their acceptance of the system. These three-point scales can be combined into one ten-point scale to give a CARS score out of ten, see Figure 7.

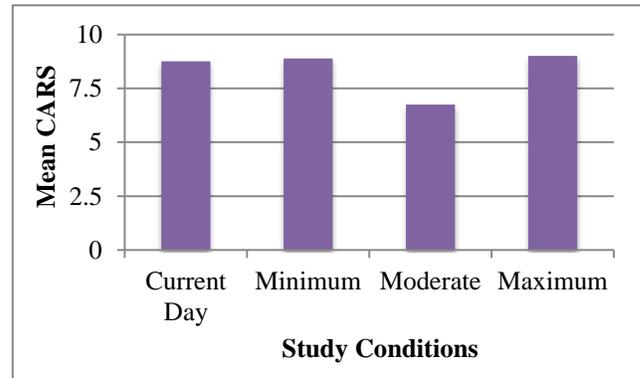


Figure 7. Mean Acceptability Ratings Given by Controllers

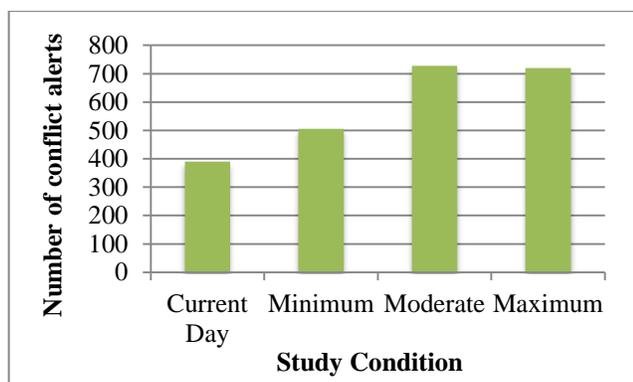
As with the previous two scales, our prediction was that mean ratings on the CARS would increase from the Current Day to the Maximum condition. Although the increase is slight, from a mean rating of 8.75 (Current Day) to 9 (Maximum), the trend is present for three of the four study conditions. Again, the Moderate mean rating, of 6.75, does not follow the trend and is lower than predicted. It is significantly lower than the CARS means for the three other conditions (e.g., the comparison with the Minimum condition values was significant at the $p<.05$ level: $Z=2.375$, $n=10$, $p=.018$), while the other three conditions were not significantly different between themselves.

This trend, where data for the Moderate condition showed unexpected patterns when compared with the other three conditions in the study, is consistent with controller responses to other questionnaire items that the Moderate condition was more challenging than the Minimum or Maximum. However, there is not the space in this paper to cover these. To ascertain whether the observed differences were solely in participant perception or whether there were performance effects, the instances of loss of separation (LOS) were reviewed. LOS events were chosen as an objective measure because the focus of the automation enhancements was on supporting separation assurance.

Losses of Separation

In the Current Day, Minimum and Moderate conditions, controllers were alerted to conflicts when they owned one or both aircraft, or if the loss of separation was predicted to occur within the controller's airspace. Aircraft had to meet specific proximity criteria to alert as a conflict, and these are discussed in detail in [15]. In the Maximum condition, controllers were only alerted to conflicts predicted to occur within their airspace – a different rule that reduced the number of conflicts controllers were alerted to.

Across the four study conditions, there were 2,323 conflicts detected by the automation, and some of these alerted to multiple controllers based on the rules above. Figure 8 shows the 2,323 conflicts distributed across the four study conditions.



CONFLICTS	CD	Min	Mod	Max
Total	390	486	727	720

Figure 8. Number of Detected Conflicts Across the Four Study Conditions

The increase in alerted conflicts in the future conditions was expected, due to the increase in traffic level within the same area of airspace. However, very few of these alerts developed into a loss of separation event. A LOS was recorded any time two aircraft were simultaneously closer than 5 nautical miles (nmi) laterally and less than 800 feet apart vertically. To be included in the analysis, a LOS had to occur within one of the test sectors after the first five minutes of a run and last for more than twelve consecutive seconds (one full, simulated radar position update), see Cabrall [16] for more details. Of the 2,323 conflicts alerted, 48 developed into losses of separation (2%). After further analysis of these 48 incidents, 23 were discounted as simulation artifacts, leaving 25. Of these, 15 (or 60%) occurred in the

Moderate condition (Figure 9). Compared with the Maximum condition, which essentially had the same number of alerts (Figure 8) but far fewer LOS (Figure 9), the disparity in the number of LOS in the Moderate condition is marked. In terms of a percentage based on the alerts, there is an increase in the rate of LOS from 1.2% (Minimum) to 2.1% (Moderate). A doubling of the rate of LOS may not seem a large increase but in terms of safety is considered a marked upswing in losses of separation.

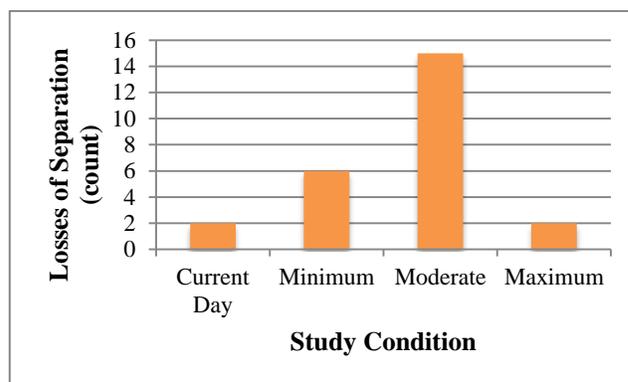


Figure 9. Losses of Separation Across the Four Study Conditions

As for the predictions regarding the subjective data, the study predictions for the LOS were similar – that the Moderate LOS rate would fall somewhere between the Minimum and Maximum rates. Figure 9 shows that this was not the case. It is likely that the subjective results discussed above are linked both directly and indirectly to the LOS observed. That is, a greater number of LOS are likely to directly affect a controller's system acceptability rating and indirectly affect a controller's rating of their situation awareness and workload. The way traffic scenarios were increased across the phases increased the traffic complexity as well as the demand. The controller experience measures suggest that managing the traffic demand in the Moderate condition with DST did not follow the predicted trend that controllers should find the situations easier to manage than the Minimum condition.

However, although it is straightforward to suggest a link between the subjective and objective results, none of the data discussed to this point in the paper offers an account to explain or describe why the Moderate condition was perceived to be harder to manage than the Minimum condition and only just a little easier than the Current Day. The remainder of this paper will delve into other data to identify

potential factors that could have made the Moderate condition more difficult to manage than expected.

Exploring the Moderate Condition Findings

Increased Complexity

There are a number of possible accounts for the irregular pattern of controller experience results in the Moderate condition. Possibly the most simple reason might be the increase in traffic. All categories of traffic were increased: overflights, arrivals and departures. Because of this, each test sector was more densely packed and this increased the chances for conflicts (as seen in Figure 8 above). It can be suggested that the increase in traffic level in these conditions resulted in more events and situations and this increased demand on the controllers. Specifically, the increased traffic became more complex rather than just more numerous, which possibly made individual situations harder to solve. Increasing the traffic levels with each condition meant that there was an increase in the number of aircraft coming in and out of each sector, and some of the hardest conflicts to solve were those at the sector boundary, or with transitioning aircraft (Figure 11). Since the two transitional phases (with automation) still required some level of controller involvement, it is possible that the number and type of conflicts increased demand on operators' cognitive resources, and this was not yet offset by the increased automation functions until the Maximum condition.

A second effect of more traffic was an increase in perceived communication and coordination load. Figure 10 below shows the controller's mean reported level of both voice communication and Data Comm load across the study conditions. These measures were subjective, and point at the number of controller resources that were devoted to managing both types of communication. The controller's level of perceived voice load was fairly consistent between Current Day ($m=3.9$), Minimum ($m=3.8$), and Moderate conditions ($m=3.6$). This itself would not be interesting, but in both the Minimum and Moderate conditions, controllers' perceived Data Comm load ($m=3$) was similar to the voice load.

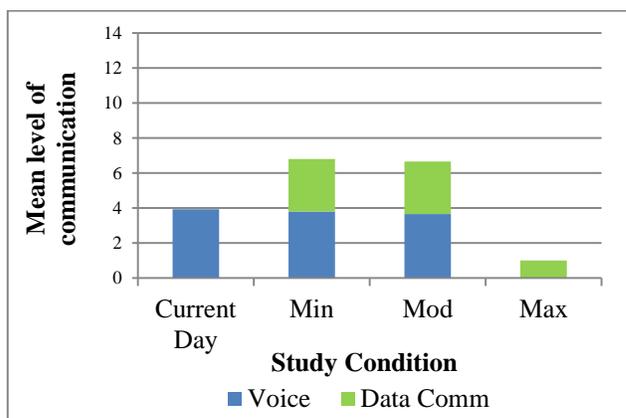


Figure 10. Controllers Perceived Level of Communication/Coordination Load

Taken together, the perceived communication operations in the Minimum and Moderate conditions were greater than those found in the Current Day or Maximum conditions. And although Data Comm was introduced in the Minimum condition as a means to reduce the voice load, it can be observed that its functional implementation still required a large degree of controller involvement, thus Data Comm may have added to, rather than offset, the overall higher communication demand in these future conditions.

Alongside this notion, controllers also identified which operations required the most coordination (Figure 11).

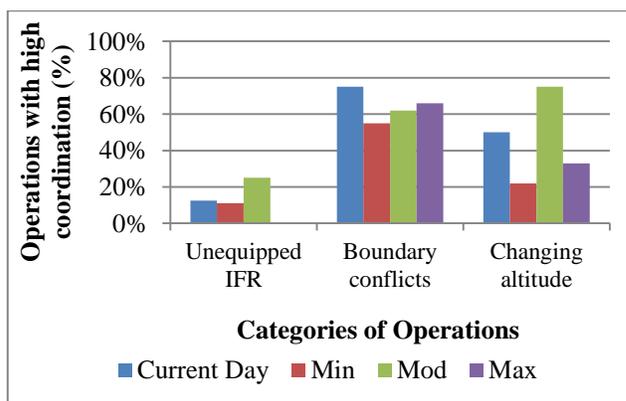


Figure 11. Controllers' Opinions of Operations that Required the Most Coordination, as a Percentage. *Note that controllers could mark multiple categories.*

Figure 11 shows that both boundary conflicts and conflicts caused by aircraft changing altitudes were marked as requiring the high levels of coordination. With the increase in traffic across the study

conditions, the volume of these two types of conflicts also increased and the corresponding coordination load that came with them was not offset by the tools as was intended. Instead, it is likely that these conflicts increased task demands on controllers.

A third impact on demand could have been the mix of aircraft equipage. As the conditions advanced, not only did the overall total number of aircraft increase, but the ratio of TFR (equipped) to IFR (unequipped) aircraft also increased. As seen in Figure 12, the majority of conflicts detected were between IFR-TFR pairs. With the variation in display and procedures between aircraft equipage levels, it could have been difficult for controllers to establish a strategy for resolving mixed equipage conflicts; and with limitations to the time and effort that can be spent on any given task, it could be the case that the added complexity of the transitional phases was not proportional to the aid provided by the tools available.

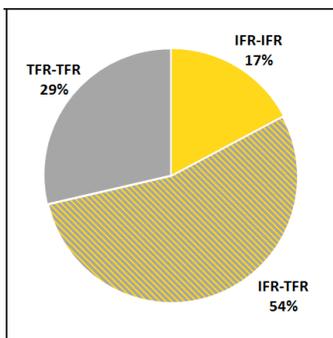


Figure 12. The Composition of Conflict Pairs by Flight Rule in the Moderate Condition

With the intention of providing controllers with tools that were meant to offset the increased demands of a given condition, researchers took a closer look at how controllers used the tools as they developed across the study conditions.

Tool Use/Preference

After each run, controllers indicated which tools they had used from the set of tools available for that condition. Figure 13 shows the way reported tool usage varied across the four study conditions. (Note that only the conflict list and conflict probe were available in all four conditions.) To summarize, most informational tools were reported to be used most of the time. For example, R-side controllers reported using the conflict probe 80% of the time or more. One informational tool used slightly less often was

the conflict list. It provided information, but its use declined from 75% in the Current Day to 60% in the Maximum condition. However, controllers reported using the updated interactive conflict list more in the Moderate condition than any other condition. The DSTs that were introduced from the Moderate condition onward had varying levels of popularity. Controllers reported using the auto-resolver for strategic conflicts 80% of the time or more, and the auto-uplinking T-SAFE resolution 75% of the time in the Maximum condition, although the short-term resolution suggestion in fifth line of the FDB was the least popular DST.

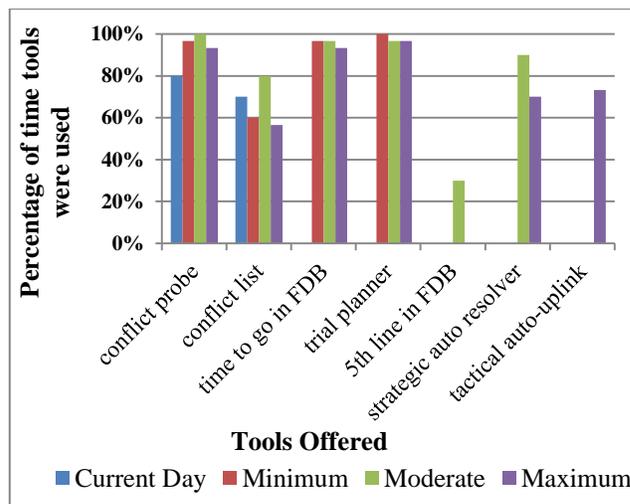


Figure 13. Tools Used by Controllers Across the Four Study Conditions

Controllers tended to use the informational tools – conflict probe, conflict list, time to go in FDB, and trial planning – across all the conditions but there were indications they shied away from using the more automated DSTs where partial controller involvement was required (such as the 5th line in the FDB). Evidence (of the way controllers used the DSTs) suggests they were not used to their full advantage as decision support tools. That is, controllers referred to and used the information provided by the DSTs, but did not relinquish decision making or responsibility of implementation to the automation. As evidence, in the Moderate condition, although all of the conflicts were alerted to the controllers, the auto-resolver was only invoked 22% of the time, and in those instances where it was invoked, controllers rejected the solution 54% of the time [15]. This reiterates that controllers tended to

use the more automated tools as informational resources rather than decision support.

A possible reason for why controllers did not use the tools as intended could have been a mismatch in the developmental stage of the tool and the task demands of the condition. To delve a little deeper into this idea, researchers asked controllers to rank the tool variations by “usefulness in aiding their simulation tasks.” Three groups of tools were selected – short term conflict tools, strategic conflict and planning tools, and conflict information tools – and controllers ranked the list of tools within each category.

For the short-term conflict tools, participants ranked the Maximum condition auto-uplink TSAFE with the back-on-course features as their top choice (the most useful) over the “stripped-down” versions of the tool presented in earlier conditions (Table 2). However, they did not rank the tools’ usefulness in descending order of their functionality, after the Maximum condition short term conflict tool, they ranked the Current Day, when they had no tools, as their second preference. Importantly, it seems participants preferred all-or-nothing for short term conflict assistance.

Table 2. Participants’ Rank Ordering of Usefulness of the Short Term Conflict Tools

Rank Order	NextGen Levels	Short Term Conflict Tools
1	Maximum	Auto-uplink of T-SAFE advisories plus post-T-SAFE auto back-on-course auto-uplink
2	Current Day, Minimum	Solving short term conflicts manually
3	Moderate	Conflict vector advisories in the 5th line of the flight data block

For the strategic conflict and planning tools, controllers ranked the Moderate interactive auto-resolver tool with the manual data link feature as the most useful over the more advanced version of the tool presented in the Maximum condition (Table 3). However, the Maximum tool was their second preference and the Current Day option of no tools was ranked least useful. Participants preferred having tools to assist them with strategic conflicts.

Table 3. Participants’ Rank Ordering of Usefulness of the Strategic Conflict and Planning Tools

Rank Order	NextGen Levels	Strategic Conflict & Planning Tools
1	Moderate	Interactive auto-resolver with manual data comm uplink
2	Maximum	Auto-uplink of within-limits strategic resolutions by an auto conflict resolver
3	Minimum	Manual trial planning tools (vertical and lateral) via voice
4	Current Day	No trial planning tools

For conflict information tools, participants ranked the Minimum conflict list, with the time to go in the FDB feature, as the most useful over the more advanced versions of the list presented in later conditions (Table 4). However, continuation of the ranking shows a preference for the more advanced tools, with controllers ranking the auto-resolver status list as their second preference, over the basic conflict list, which was ranked least useful. Other analyses support these subjective rankings; see Homola, et al. [15].

Table 4. Participants’ Rank Ordering of Usefulness of the Conflict Information Tools

Rank Order	NextGen Levels	Conflict Information Tools
1	Minimum	Conflict list with time-to-go in the flight data block
2	Maximum	Auto-resolver status information included in the interactive conflict list
3	Moderate	Interactive conflict list with auto-resolve on request
4	Current Day	Basic conflict list

Overall, tools in the Maximum condition received the highest usefulness rankings, and those in the Current Day received the lowest rankings. However, Minimum tools received higher rankings than Moderate tools overall. This may shed some light on the subjective performance data discussed above, which suggested that the Moderate condition was more difficult to manage, as these tool ranking reports suggest the controllers found the tools in the Moderate condition less useful – slightly above that

of the Current Day. A reasonable explanation for such results might be that perceptions of tool accuracy and controller confidence in those tools suffered and so are examined in the next section.

Tool Accuracy

In general, questionnaire responses indicate that tools were not perceived to be less accurate in the Moderate condition. Controllers were asked to rate the accuracy of the conflict list, auto resolver, and TSAFE advisories on a scale from 1 to 7 (from “very inaccurate” to “very accurate”). Controllers thought the conflict list was “quite accurate” in the Current Day (m=5.42) and Minimum (m=5.25) conditions. Their opinion did not change in the Moderate condition where they still felt the conflict list was “quite accurate” (m=5.58). They also felt the auto-resolver (a DST) was “quite accurate” (m=5.53). However, the fifth line advisories in the FDB were rated less positively as only “somewhat accurate” (m=4.33). In the Maximum condition, controller opinions improved slightly, as they rated the T-SAFE resolutions as “quite accurate” (m=4.8).

Tool Confidence

Controllers also rated their confidence in the trial planning tools and strategic conflict advisories, again on a scale from 1 to 7 (from “not at all confident” to “very confident”). Controllers’ confidence grew in the trial planning tools as they used them. They were “quite confident” when they used the trial planner in the Minimum condition (m=5.2) but this confidence increased to “very confident” in the Maximum condition (m=6.5). Their confidence grew in a similar way when using the strategic conflict advisories: controllers were “quite confident” when they used the strategic advisories in the Moderate condition (m=5.5) and this confidence increased to “confident” in the Maximum condition (m=6.1).

As controller confidence was reasonably high and the tools had positive perceived accuracy, it suggests that the reasons for low tool use were not that the controllers thought they worked incorrectly.

Discussion

In the Moderate condition, participants were introduced to two tools that provided a level of decision support, in addition to informational tools.

While they felt confident in one of these DSTs and that it was accurate overall (the interactive auto-resolver for equipped aircraft) they did not rank the Moderate toolset as more useful than the Minimum toolset. Other controller self-reports supported that they found the Moderate toolset less useful than the Minimum toolset through generally increased workload and reduced situation awareness, and the recorded rate of LOS events nearly doubled. Taken as a whole, results suggested that controllers may not have drawn on tools’ decision support capabilities as much as they could and continued to use even DSTs as sources of information only. It begs the question of why the more automated tools did not aid the controllers to the same degree as the informational tools.

The Moderate condition tools were advanced logically from the Minimum condition: the conflict list became interactive and two DSTs were introduced. This raised the level of automation of the tools to around a level 4 [4], which Endsley & Kaber [7] describe as “shared control”. Under shared control, the human and the automation share the task of generating solutions in addition to sharing monitoring and implementing tasks. Looking back to Sheridan & Verplank’s levels of automation [4], we can see that controllers tended to prefer tools on either end of the spectrum: those that were more basic and involved little computer assistance, or those that were fully autonomous and required little controller interaction. Those iterations of tools that provided more informational resources – with little to no computer involvement in decision making appeared to be preferred over their more automated counterparts. However, when tools that moved controllers to a supervisory role were introduced, they were able to use these tools – and even reported a preference for some of them (Table 2). This suggests that in the Moderate condition, controllers may have tried to maintain a “current day” approach or strategy to using the tools and did not modify their approach to a “Moderate strategy” that took advantage of the tools’ capabilities.

Using lower “LOA strategies” in a higher LOA environment may have increased load for the controllers in a second way, in that the DSTs added another task for the controllers – to monitor what the automation was doing, so now the controller was still generating a resolution idea (for an increased number

of alerts) *plus* they were monitoring the automation. So, the tools offered a raised LOA and ensured the controllers were informed but they did not support them in terms of reducing their tasks. Kaber and Endsley [17] in a follow-on study to [7] obtained similar findings: that intermediate levels of automation were related to lower performance.

Taking the suggestion that controller strategies did not advance at the same rate as the LOA in the study conditions, there is still the question of why controllers preferred information-only tools when more automated tools were available. One explanation is that informational tools, like the conflict list, were used more frequently because, firstly, the controllers were familiar with these tools having used or seen versions of them at Center facilities. Secondly, they had more time to train with them as the Current Day tools carried over through all four conditions. Thirdly, the variations in the tools' capabilities were not so drastic from condition to condition to have to re-learn how to use them. All of these factors meant controllers were more familiar with the informational tools' functions, thus allowing them to more fully integrate the tools into their accustomed strategies.

The DSTs introduced in the Moderate condition had increased functionality and allowed for controller interaction, but to use them to their full extent there was a learning curve. And because these tools were introduced later in the study, there was both less training with, and less exposure to them. It could be the case that controllers were unable to use the tools as intended (to compensate for the increase in demands for the given condition), which may also explain why the controllers used these tools as informational resources rather than for decision support as they were intended.

Although traffic increased to 1.5 times the baseline (with 50% TFR) in the Moderate condition, which increased demand (and possibly complexity) of the problems, if a different combination of decision support tools had been available in the Moderate condition, participants may have found it easier to manage. Following this line of questioning, if the Moderate tools did not support controllers in the way they needed, what properties could have been changed to make them more useful?

As noted above, in the present study these considerations are speculative, as all the separation

assurance tools that were introduced were new for the participants. It is impossible to know whether additional training alone would have reduced the reports of higher demand in the Moderate condition as DSTs were introduced. Even if this is the case, some improvements to the method and time of introducing this LOA may ease the process. For example, if this level of tools is introduced earlier in time, when traffic demand is at 1.3 or 1.4 times current day levels, this may help the controllers by giving them time to become familiar with the automation before they truly need it. A second facilitator may be to focus on a part of the automation and introduce that more fully. The Moderate condition introduced an interactive resolver for strategic conflicts and a close-in alerting tool, both had value but neither reduced operator load. For the Maximum condition, both of these tools were automated, reducing load for both types of conflict. What if the strategic auto resolver had been introduced in two modes in the Moderate condition, where conflicts meeting some criteria were automatically solved and others could be requested? As a trade-off close-in conflict alerting and resolution could have been reserved until the Maximum condition. This would have provided the same number of tools in the Moderate condition but may have reduced its taskload.

A third possible way to ease the introduction of shared control may be to look outside the focus automation (of the separation assurance tools) to the other sources of increased equipage. In this study Data Comm was also automation that may have been a greater benefit to controllers if it had been advanced further in the Moderate condition. For example, if 75% of the traffic had been Data Comm equipped, rather than 50%, then an additional 25% of the aircraft would have had automatic handoffs and transfers of communication. This could result in a reduction in housekeeping taskload.

Conclusion

The present study considered separation assurance LOA through transitional phases of development, between current day ATC and fully implemented separation assurance ATM. The way traffic scenarios were increased across the phases increased the traffic complexity as well as the demand. Subjective reports from participants indicated they found the transitional phase that

presented “shared control” [7], where some tools offered decision support, harder to manage than earlier and later study phases. The “Moderate tools” did not offset the increased complexity of the condition. The nature of shared control may have served to increase controllers’ monitoring duties, but not reduce their implementation or decision making tasks. Three modifications to our approach to the shared control LOA phase would be to: introduce the technology earlier to give a longer training period; make the LOA of each tool lower or higher on a case-by-case basis to avoid shared control tools; and, ensure an increase in aircraft equipage.

Taking a wider view, our initial findings would suggest that, as automation is introduced, consideration should be given to the taskload that semi-automated tools will add to an operator’s load. If one kind of task is removed (e.g., decision making) but a different kind is added (e.g., monitoring automation) the result could be a reduction in performance. With this in mind, presenting combinations of tools that are automated enough to reduce operator load by at least as much as they add may assist with decisions about which tools should be implemented in interim phases that introduce automation to a system.

References

- [1] Joint Planning and Development Office (JPDO), 2010, *Concept of Operations for the Next Generation Air Transportation System (Version 3.2)*, http://jpe.jpdo.gov/ee/docs/conops/NextGen_ConOps_v3_2.pdf.
- [2] Federal Aviation Administration, 2012, FAA Aerospace Forecasts FY 2012-2032, http://www.faa.gov/about/office_org/headquarters_of_fices/apl/aviation_forecasts/aerospace_forecasts/2012-2032.
- [3] Sheridan, T., 1992, *Telerobotics, Automation, and Human Supervisory Control*, Cambridge, MA: MIT Press.
- [4] Sheridan, T. & W. Verplank, 1978, *Human and Computer Control of Undersea Teleoperators*, Cambridge, MA: Man-Machine Systems Laboratory, Department of Mechanical Engineering, MIT.
- [5] Parasuraman, R., T. Sheridan, & C. Wickens, 2000, A Model for Types and Levels of Human Interaction with Automation, *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30 (3) 286-297.
- [6] Wing, D., et al., 2013, Pilot and Controller Evaluations of Separation Function Allocation in Air Traffic Management, *Tenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2013)*, Chicago, July.
- [7] Endsley, M. & D. Kaber, 1999, Level of automation effects on performance, situation awareness and workload in a dynamic control task, *Ergonomics*, 42, 462-492.
- [8] Moray, N., T. Inagaki, & M. Itoh, 2000, Adaptive automation, trust and self-confidence in fault management of time-critical tasks, *Journal of Experimental Psychology: Applied*, 6, 44-58.
- [9] Prevot, T., J. Homola, L. Martin, J. Mercer, & C. Cabrall, 2012, Toward Automated Air Traffic Control – Investigating a Fundamental Paradigm Shift in Human/Systems Interaction, *International Journal of Human-Computer Interaction*, 28, 77-98, <http://dx.doi.org/10.1080/10447318.2012.634756>.
- [10] Erzberger, H., 2001, The Automated Airspace Concept, *Fourth USA/Europe Air Traffic Management Research and Development Seminar (ATM2001)*, Santa Fe, NM.
- [11] Prevot, T., 2002, Exploring the Many Perspectives of Distributed Air Traffic Management: The Multi-Aircraft Control System MACS, In S. Chatty, J. Hansman, & G. Boy. (Eds). *HCI-Aero 2002*, AIAA Press, Menlo Park, CA. pp 149-154.
- [12] Hart, S. & L. Staveland, 1988, Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.) *Human Mental Workload*, Amsterdam, North Holland Press.
- [13] Selcon, S., R. Taylor, & E. Koritsas, 1991, Workload or situational awareness?: TLX vs. SART for aerospace systems design evaluation, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 35 (2), 62-66.
- [14] Lee, K. & K. Kerns, R. Bone, M. Nickelson, 2001, Development and Validation of the Controller Acceptance Rating Scale (CARS): Results of Empirical Research, *Fourth USA/Europe Air Traffic*

Management Research and Development Seminar (ATM2001), Santa Fe, NM.

[15] Homola, J., S. Morey, C. Cabrall, L. Martin, T. Prevot, 2013, Analysis of Interactive Conflict Resolution Tool Usage in a Mixed Equipage Environment. *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Boston, Massachusetts.

[16] Cabrall, C., A. Gomez, J. Homola, S. Hunt, L. Martin, J. Mercer, & T. Prevot, 2013, Transitioning Resolution Responsibility between the Controller and Automation Team in Simulated NextGen Separation Assurance. *Proceedings of the Electronic Navigation Research Institute International Workshop on ATM/CNS (EIWAC 2013)*, Tokyo, Japan.

[17] Kaber, D. & M. Endsley, 2004, The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task, *Theoretical Issues in Ergonomic Science*, 5 (2), 113-153.

Acknowledgements

The study from which this paper was derived, was sponsored and fully funded by the NASA Airspace System Program, Concept and Technology Development Project. Many thanks to the dedicated researchers, developers, pilot trainers, and support personnel in the Airspace Operations Laboratory who helped make it happen.

*32nd Digital Avionics Systems Conference
October 6-10, 2013*