

An Investigation into the Use of Novel Conflict Detection and Resolution Automation in Air Traffic Management

Clark Borst

Control and Simulation Division
Delft University of Technology
Delft, The Netherlands
c.borst@tudelft.nl

Carl Westin and Brian Hilburn

Center for Human Performance Research, CHPR BV
Voorburg, The Netherlands
carl@chpr.nl, brian@chpr.nl

Abstract—The Multidimensional Framework for Advanced SESAR Automation (MUFASA) project is exploring issues concerning the acceptance and usage of advanced decision aiding automation. Through a series of human-in-the-loop simulations, it ultimately aims to examine the interactive effects of automation level, air traffic complexity, and strategic conformance (i.e. the fit between human and machine strategies) on automation usage and acceptance. This paper, however, only presents the design and results of an exploratory experiment that aimed to investigate conflict detection and resolution (CD&R) automation usage among professional air traffic controllers and novices (students) utilising a novel decision-support tool: the Solution Space Diagram (SSD). This preliminary study featured a manual air traffic control task with a fixed level of automation under high and low traffic complexity. The results indicate that students, in contrast to the controllers, reacted more immediately and promptly to conflict warnings. With no separation losses, students outperformed controllers in keeping aircraft safely separated. Controllers, on the other hand, had multiple separation losses. Observations and debriefing of controllers revealed a general scepticism towards the SSD display, its accuracy and usefulness. This allows us to speculate that controllers, compared to students, had less trust in the SSD as a CD&R tool, and rather used their own judgment in conflict management.

Foreword - This paper describes a project that is part of SESAR Work Package E, which is addressing long-term and innovative research.

Keywords- automation, air traffic, acceptance, strategy, decision aiding, SESAR, complexity

I. INTRODUCTION

Future Air Traffic Management (ATM) will have to rely on more, and more sophisticated, automation to accommodate predicted air traffic. This belief is captured in the SESAR programme's definition of five operational Service Levels [1], which are intended to guide the evolution from current to far-term European ATM operations. These five Service Levels assume increasingly greater performance requirements and greater information sharing between all stakeholders. At the heart of SESAR's five Service Levels is the expectation that

automation will become more advanced in terms of the types of tasks it can perform, and the level of authority and autonomy it can assume. However, the general consensus is that in complex domains, there will always be a potential for problems that cannot be anticipated in the design of automated systems. Thus, the creative human expert will remain an important resource for dealing with this unanticipated variability.

Studies across various domains have shown that user acceptance of automation decreases when the authority of decision-making automation increases [2-6]. Research has also shown that the predominantly algorithmic approaches used in automation seldom fit well with the more heuristic methodology employed by humans. In their exhaustive survey of conflict detection and resolution (CD&R) modelling methods, Kuchar and Yang [7] concluded that CD&R automation correlate poorly to how controllers prefer to work. Bekier, Molesworth, and Williamson [8] suggested that there is a "tipping point" for any automated tool, above which controller acceptance of that tool quickly drops.

Designing human-centred automation with which humans retain ultimate responsibility and control authority will certainly present technical challenges. However, some of the greatest challenges will be those having to do with human factors - for example, how do we build automation as a team player that keeps the human in the loop, while assuring acceptance, and yet that assumes unprecedented levels of authority and autonomy?

The MUFASA project aims to explore the interaction of strategic conformance, complexity, and higher levels of automation, and ultimately present a framework for guiding future research and development of ATM automation. The experimental design relies on a series of three human-in-the-loop simulations of increasing fidelity. The experimental design is novel in the sense that we are simulating automation capable of providing conflict resolutions of the same calibre as a human controller would be able. This is achieved by presenting an automated solution that in fact is a replay of the controllers own solution to the conflict scenario extracted

This paper describes a project that is part of SESAR Workpackage E, which is addressing long-term and innovative research. The project was started early 2011 so this description is limited to an outline of the project objectives augmented by some early findings.

from an earlier run. To date, two of three simulations have been conducted. The final simulation sessions (Nominal Advisory Level Automation, NALA) is planned to take place in early 2013. The preceding simulations, SIMBA (Simulated Baseline Automation) including novices, and PUMBA (Preliminary Update / Modified Baseline Automation), including experienced air traffic controllers, have been completed.

This paper presents a preliminary investigation of ATM CD&R automation usage associated with a sophisticated CD&R support tool: the Solution Space Diagram (SSD) under development by the Technical University Delft (TUD). In its most succinct form, the SSD is a tactical decision-support tool that helps controllers' vector aircraft in conflict-free speeds and/or headings. The experiment presented in this paper aims to compare resolution strategies with the SSD among (and between) controllers and novices. As the SSD can be modified to represent different levels of automation, the ultimate goal will be to present conformal and non-conformal machine advisories within the SSD in order to study automation acceptance under the three main factors of air traffic complexity, level of automation, and strategic conformance.

In order to measure conformance, however, it is important that we will first identify and extract different strategies used by controllers when issuing conflict resolutions. If there would be no difference in strategies, and ultimately resolution action taken, there would be no conformance. This extraction stage is the referred to as the *Prequel*. Note that this paper does not address the subsequent stage where we measure the level of conformance (i.e. replaying the participant's own solution). As such, this paper *only* contains the results of the manual "Prequel" runs performed in the SIMBA and PUMBA trials and featured a manual air traffic control task with a fixed level of automation under high and low traffic complexity.

This paper is organised as follows. First, the SSD display will be briefly explained. Then, a detailed description of the experimental design will be provided, followed by the results of the experiment. Finally, the results will be discussed followed by conclusions and recommendations.

II. SOLUTION SPACE DIAGRAM

The TUD has for some time carried out research and development into innovative display concepts for CD&R. One such display is the SSD that aims to represent an aircraft's control space in terms of speed and heading [9]. In the robotics domain, this concept is also known as the "velocity obstacle theory".

A velocity obstacle is a collection of relative velocities between moving vehicles that will result in a collision. When considering a given aircraft's "protected zone" as a 5nm radius circle, the "velocity obstacle" represents the collection of relative velocities that will result in a separation violation (Fig. 1). In the horizontal plane, a velocity obstacle forms a visually attractive geometrical shape – a triangle – that can be directly visualized on a display to support separation assurance tasks.

When also showing the minimum and maximum velocity boundaries of the aircraft, the SSD reveals the conflict space as well as the solution space of an aircraft in terms of heading and speed.

A controller can use the SSD to vector aircraft into conflict-free areas so as to maintain safe separations of 5nm between aircraft. She can do this by ensuring that the velocity of the controlled aircraft lies outside the velocity obstacle (formed by the observed aircraft) and inside the speed envelope of the controlled aircraft. In Fig. 2 a screen capture is shown of two SSDs surrounding the aircraft blips on a plan view display. Note that in order to be useful for controllers, the velocity obstacles need to be shifted toward the absolute space such that controllers can issue absolute speed and heading clearances an aircraft can easily work with.

In the current state of development, the SSD only supports separation assurance in the horizontal plane. Developments to expand the SSD representation to include the altitude domain are still ongoing. However, these efforts will not be part of the MUFASA project.

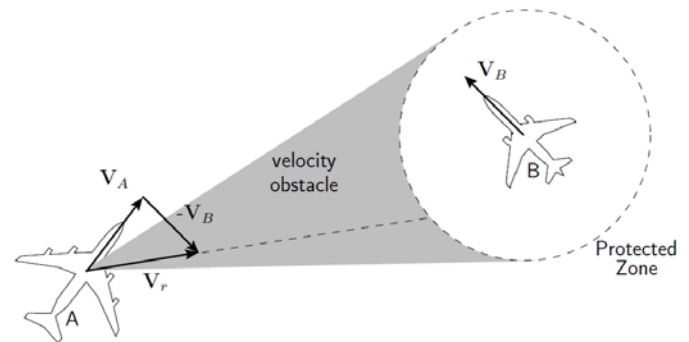


Figure 1. The geometry of a velocity obstacle between two aircraft in the relative speed and heading domain.

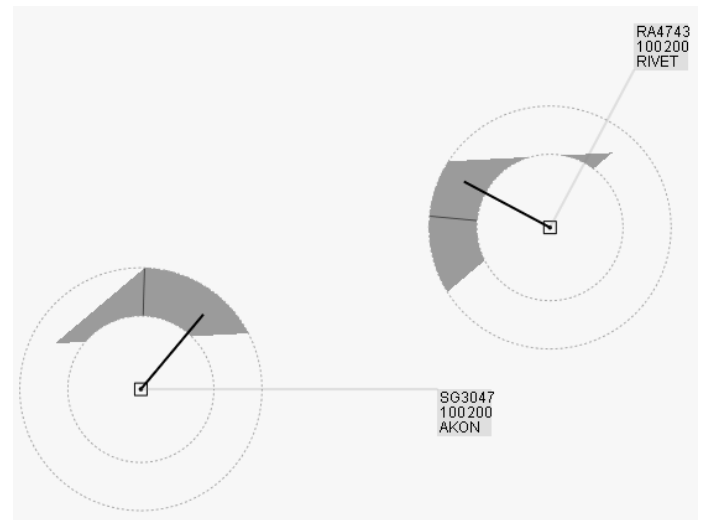


Figure 2. The SSD on a controller's plan view display showing the velocity obstacles in the absolute speed and heading domain.

III. EXPERIMENT

An exploratory experiment has been designed to investigate the usage of the SSD under various traffic loads in a manual air traffic control task. The goal of the experiment was to investigate controller CD&R strategies and resolutions, while using the SSD as a decision-support tool, and see if there would be a difference in strategies, and ultimately resolution action taken, that could serve as simulated machine advisories in the conformance study planned to take place in early 2013.

A. Apparatus

The MUFASA simulator is a Java-based application that allows air traffic controllers to control short traffic scenarios within a square volume of airspace (Fig. 3). All aircraft fly on the same flight level and each aircraft has its own designated exit point out of the sector. The designated exit point is shown in an aircraft's flight label. The SSD used in the simulator is a simplified version of the one shown in Fig. 2. That is, the SSD in the simulator only shows the conflict areas on a heading band at the aircraft's current velocity. Further, the conflict areas have been color-coded to indicate on what time frame a loss of separation would occur when an aircraft is vectored into a conflict area. A yellow colour indicates that a loss of separation would occur within 5 to 20 minutes, whereas the red colour indicates a time frame between now (0 minutes) and 5 minutes.

To vector an aircraft, a controller can use a mouse pointer device to click on an aircraft of interest, drag the velocity trend vector to a new conflict-free area on the heading ring, and press the ENTER key on a keyboard to implement the vector. Speed clearances (and combined speed and heading clearances) can also be given by using the mouse scroll wheel to either increase or decrease the speed by 10 knots. This also increases the radius of the heading band and shows the corrected conflict zones for the new speed settings. This allows a controller to

quickly browse through different speed settings and preview the conflict and conflict-free heading areas for different speeds. Further, no wind conditions were taken into account and all aircraft velocities (and speed clearances) are given in knots Indicated Airspeed (IAS).

The traffic motion in the simulator has been simulated by simple, linear kinematic equations:

$$\begin{aligned} x(t) &= x_0 + \int_{t=0}^T V \cos \psi \cdot dt \\ y(t) &= y_0 + \int_{t=0}^T V \sin \psi \cdot dt \end{aligned} \quad (1)$$

with x and y the lateral position coordinates, V the aircraft speed, ψ the heading angle, and dt the simulator time step. Whenever a controller proposes a new velocity vector command and executes/activates it, first order transfer functions are used to simulate the aircraft turn dynamics and the transient in changing the magnitude of the aircraft velocity.

To keep the traffic scenarios short, repeatable, and interesting it was decided to run the simulator faster than real time. Speeding up traffic scenarios in ATC simulators is a common technique to serve this purpose.

In the experiment, a four-times-faster-than-real-time update frequency was implemented for the aircraft motion. That is, a ten-minute scenario would last 2.5 minutes in our simulator. As a result, the forecasting time of the aircraft velocity trend line is 15 seconds in our simulator rather than 60 seconds. Further, the aircraft plots on the display are updated every second to simulate a 1 Hz ADS-B update frequency.

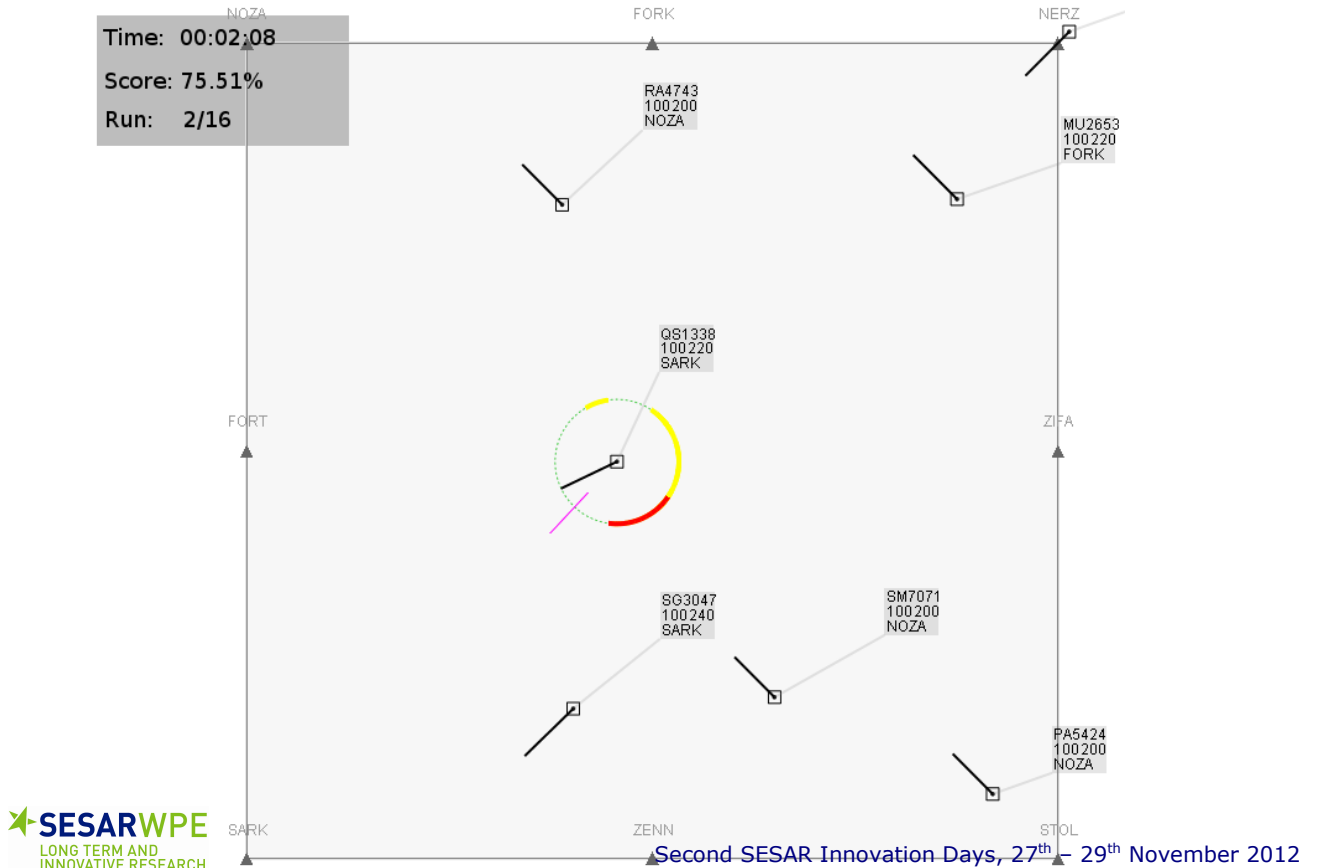


Figure 3. The MUFASA simulator.

B. Independent variables

The experimental design is a within-subjects design with only one independent variable, traffic complexity, which featured two levels (high and low). For more details concerning the scenarios and their complexity levels, see section III.D.

C. Subjects and tasks

The participants in the experiment were three retired professional air traffic controllers, and two aerospace students from Delft. The professionals were all experienced en-route controllers. The students were familiar with the SSD concept, but did not have any experience as air traffic controllers. We aimed to compare CD&R strategies between controllers and novices to see how differently or similarly they would use the SSD.

Participants were all given two main tasks, and they were instructed to attend to each of them. These two main tasks were conflict resolution, and clearing aircraft to their intended exit points. How they prioritised between the two tasks was up to them. In some occasions, clearing an aircraft to its exit point was not possible without introducing a conflict.

a) *Conflict resolution task.* Participants were instructed that at some point during each scenario, a flight-path conflict could occur. They were further instructed that the underlying conflict detection algorithm employed a limited look-ahead time to minimise false alarms. This was intended to prevent situations in which controllers might anticipate trajectory conflicts. Conflicts were highlighted in a typical red alert fashion when a loss of separation would occur within 5 minutes for each of the associated plots. Participants were not specifically instructed which conflict resolution strategies to use, and were therefore free to evaluate the SSDs of each of the involved aircraft, and to issue clearances to either one or both of the involved aircraft. Participants were also instructed on how to use the SSD, and instructed that they were able to use a combination of heading and/or speed clearances.

b) *Exit clearance task.* The data label of each aircraft includes an indication of the “cleared sector exit point,” or COPX (Fig. 3). In some cases aircraft are already on appropriate headings to reach their COPX. For other aircraft, however, heading clearances must be issued to direct the aircraft to its COPX. Relative bearing to the COPX varies across aircraft. For “uncleared” aircraft the trajectory vector does not aim toward the assigned COPX. In most cases, early detection of this discrepancy requires opening the SSD command display for a given aircraft, and issuing an appropriate heading clearance. Participants were instructed to be as accurate as possible in this heading clearance.

To support the exit clearance task, the SSD showed a magenta line on the heading band to indicate the desired heading of the aircraft toward its exit point. As such, participants needed to minimize the heading deviation between the current flight direction and the desired flight direction whenever possible. To provide feedback on the controller’s performance, a dynamic performance score was

shown on the display that was updated every second (Fig. 3). The performance score is calculated every by the average heading deviation of all aircraft inside the sector. All aircraft aligned with their designated exit point would then result in a 100% score. When conflicts were triggered, a score reduction of 10% per involved aircraft was given. A mid-air collision was severely punished by providing a 0% score and immediately stopping the scenario. After each scenario, an averaged total score was provided.

D. Traffic scenarios

We aimed to create scenarios that were repeatable yet unrecognisable. This was done to reduce potential confounding factors and assure that complexity was the same across scenarios, facilitating comparison between low and high complexity scenarios. For this reason, we decided to make use of scenario rotations in which the relative trajectories and closure angles were kept constant but the entire sector was rotated, and sector entry points renamed. This technique has been used in the past to create scenario cognates for investigating ASAS concepts [10]. Moreover, we had to create scenarios that were somewhat realistic, and sufficiently complex and challenging. Otherwise, there would be no motivation to use a higher level of automation in later trials.

There were two baseline scenarios for each of two (high and low) complexity levels. We then created three variants of each baseline scenario. Each of these variants was identical in terms of number of aircraft within the sector, aircraft entry rate, number of designed conflicts, routing structure, and sector volume. Where they differed was in view orientation (imagine a rotation in which a scenario was a square in one session, and a diamond shape the next), and name of exit points. This was done to ensure that the scenario variants had identical geometrical properties.

Scenarios were created on the basis of primary sector flows with slight alterations of individual flight paths to achieve some distribution in the traffic pattern with respect to sector rotation angle, presumed complexity, number of total aircraft, maximum number of aircraft, minimum number of aircraft, and the type of designated conflict. Conflicts were defined as closure angles of either 50, 85, or 90° crossing.

A fair amount of pre-testing and tryout was necessary in the development of traffic scenarios. The main challenge here was that we had to adjust complexity to the point that solutions were not trivial, and therefore that there was some variability in the choice of solutions. Although we attempted to create scenarios of “low “ and “high” complexity, we recognised early on that attempts to experimentally manipulate complexity or difficulty might have unpredictable results. For this reason, we intended to use our binary initial classifications (high versus low) but also to ultimately use participants’ own complexity ratings as a covariate.

E. Dependent Measures

Dependent measures focused on performance, safety, and subjective complexity ratings. Performance was measured in terms of the number of SSD inspections, number of speed, heading, or combined commands, and the performance score. Safety was measured by the minimum separation distance (in nautical miles) between aircraft in each scenario and by the number of mid-air collisions. The subjective complexity ratings for each scenario were given on a 5-point Likert scale. A single rating was requested after completing a scenario (or experiment run). After the experiment the participants were asked to complete a questionnaire that contained general questions about their opinions of the simulator and the SSD interface.

IV. RESULTS

It should be stressed that comparing SIMBA and PUMBA results is difficult since there has been continuous iterative development of both the simulator and the HMI between the two experiments. However, considering that both the experimental design and test scenarios have remained unchanged, some interesting comparisons can be drawn between the two experiment groups. It should also be underlined that since sample size has thus far been very small, analysis and conclusions should be approached with caution. As such, no statistical analysis has been done of the results due to limited test subjects.

A. Complexity

Data from students in the SIMBA simulation argued for the effectiveness of our complexity manipulation on workload ratings. This conclusion was, however, not supported by data derived from controllers in the PUMBA simulation. As can be seen in Fig. 4, students provided higher workload ratings for the high complexity scenarios than the controllers. For the low complexity scenarios the opposite is true.

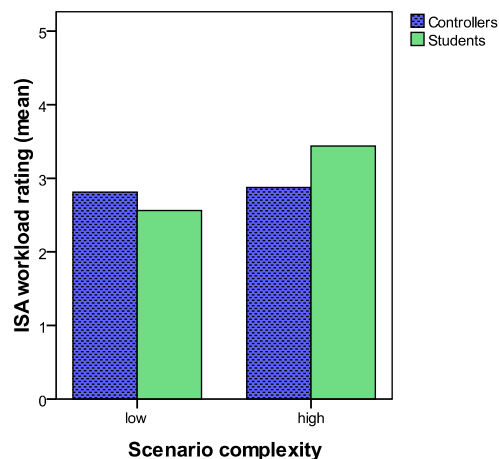


Figure 4. Comparison of workload between controllers and students

Mean performance scores for low and high complexity scenarios show that students outperformed controllers for both variables. This result was surprising as it was expected that

controllers would outperform students. However, controller performance score might have been negatively affected by two aspects not applicable to the students. First, screen resolution in the PUMBA simulation (i.e., with controllers) was lower than that used in the SIMBA simulation (with novices). With lower resolution (i.e. fewer pixels), the screen is smaller which affects the scale making it difficult to accurately judge distances (and separation) and implement commands. (i.e. vector traffic exactly to an exit point). In half of all scenarios, the controller group had separation losses whereas the students had no separation losses. Second, the controllers tended to vector traffic parallel to exit points if that traffic was close in proximity or in a catch-up situation. This has a negative effect on the performance score pertaining to the accuracy of the exit clearance task performance (measured through average closest-point-of-approach to exit point). Again, the performance was in fact mainly intended as a secondary task.

There might, however, be another aspect at play here. In contrast to students, controllers disregarded the performance score. In fact, two of the controllers stated that they did not notice it at all. While the reasons underlying this omission are unclear (e.g. it could be due to preoccupation with the task of controlling traffic, or lack of trust in the performance score), it signals a failure of monitoring the performance score as a secondary task. However, it does not reveal whether controllers considered, or neglected, the performance aspects and tasks stipulated in the experimental briefing as important for attaining a high performance score.

The purpose of the performance score was twofold:

1. to keep participants motivated and involved in the simulator session, and
2. to obscure the reiteration of scenarios and design conflict within.

Data from both students and controllers indicate the achievement of these goals. Participants from both groups were highly motivated and involved in the simulator session. One controller found the SSD and simulation to be “*a new and intuitive approach to CD&R*”, and it reminded the controller about the “*joy of working as an air traffic controller*”.

There were no indications that controllers recognised scenarios or the associated scripted conflict. This is in line with our observations from SIMBA trials with students. By rotating the sector and changing exit waypoint names, we have been able to create geometrically and mathematically, but unrecognisable scenario variants.

B. Safety

In the SIMBA simulation both students successfully completed all scenarios without any separation losses or mid-air-collisions. A similar result, and in any case not worse, was expected from the controllers used in the PUMBA simulation. Therefore, the extent of separation losses among the controllers came as a surprise (Figure 5). However, the probable cause was attributed to display resolution. In SIMBA, the resolution was higher than that used in PUMBA.

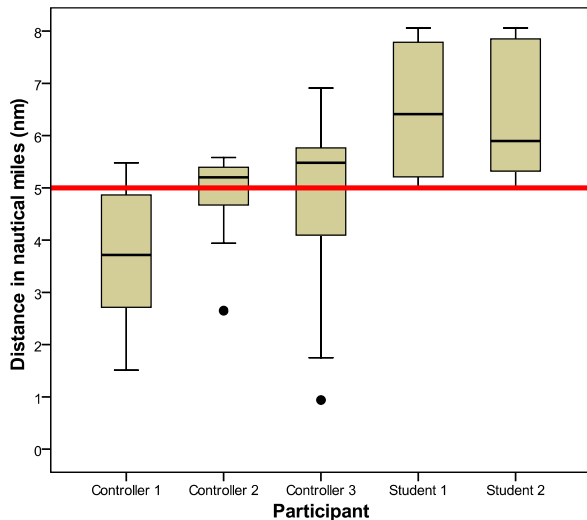


Figure 5. Minimum separation in nm between aircraft.

C. Strategies

Controller and student strategy and SSD usage is deduced on a general basis and predominately based on observations during the trials. Students appeared to interact less with traffic than controllers on all parameters (Figure 6). This shows that students had higher control efficiency than controllers. Both students, however, had previous experience of the SSD which might have affected their use of the SSD.

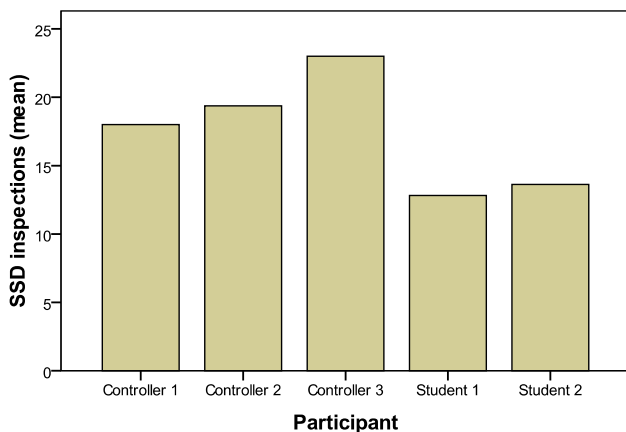


Figure 6. SSD inspections

Both controllers and students preferred heading commands. In contrast to controllers, students also made extensive use of combined commands (Fig. 7). This can possibly be attributed to the combined command capability of the SSD interface, which cannot be found in real ATC working stations. In real ATC, heading and speed commands are rarely combined in a clearance. Students not familiar with ATC are not restricted by such working routines, and were therefore expected to use the combined command option more frequently.

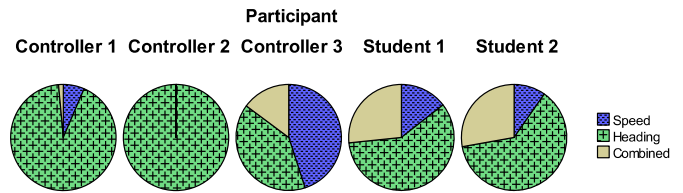


Figure 7. Distribution of commands used by participants.

Both students and the majority of the controllers indicated that they prioritised the “Exit to COPX” task. One controller stated that conflict avoidance was prioritized, with exit to COPX task as secondary.

Controllers and students appear to deal with the red conflict warning differently. Students reacted earlier than controllers to conflict alerts. The tendency was to immediately solve a conflict. Controllers seemed to disregard the indications of the SSD. It was observed that controllers repeatedly controlled traffic within red SSD zones. It was also observed that controllers disregarded the conflict warnings. Note that aircraft in red only highlight a potential loss of separation within 5 minutes, and is thus not a confirmed separation loss. One controller defended the vectoring in red zones with the argument that the controller was certain that separation was maintained through an assessment of display scale and distances between aircraft. Interesting was that this controller created multiple extra conflict, beyond the only designed conflict in each scenario. This difference in strategy is reflected in the performance score in that the longer time an aircraft is in conflict with another, the more penalty will be applied. Students were more aware of the performance score, whereas controllers gave it little attention. This leads us to speculate that controllers had little confidence in the SSD. Although they were forced to use the SSD (by means of issuing resolutions commands), their actions show that they repeatedly disregarded conflict alerts, as well as red / yellow zones. We can unfortunately not address this suspicion with the data available, but will make sure to better probe this aspect in future trials.

Another clear difference between students and controllers was the vectoring of aircraft in close proximity to sector border and exit point. Controllers stated that they preferred to turn aircraft parallel if close to sector border and going to the same exit point. This was especially true for catch-up situations. In real life, they would coordinate a parallel handover with the adjacent sector. Students did not apply this strategy in similar situations.

TABLE I. CONTROLLER RESOLUTIONS FOR THE DESIGNED CONFLICT PER SCENARIO

Scenario	Controller 1			Controller 2			Controller 3		
	aircraft ID	heading	speed	aircraft ID	heading	speed	aircraft ID	heading	speed
1	OM3185	right	-	OM3185	right	-	-	-	-
1B	OM3185	right	-	OM3185	right	-	-	-	-
1C	OM3185	right	-	OM3185	right	-	OM3185	-	increase
1D	OM3185	right	-	OM3185	right	-	OM3185	-	increase
3	RG3628	left	-	RG3628	left	-	PA5424	-	increase
3B	RG3628	left	-	RG3628	left	-	RG3628	left	increase
3C	RG3628	left	-	RG3628	left	-	PA5424	left	increase
3D	RG3628	left	-	PA5424	right	-	PA5424	-	increase
4	SM7071	-	decrease	SM7071	left	-	RG3628	left	-
4B	SM7071	-	decrease	SM7071	left	-	SM7071	-	decrease
4C	SM7071	right	-	SM7071	left	-	SM7071	-	increase
4D	RG3628	left	-	RG3628	right	-	SM7071	-	increase
5	PA5424	right	-	QS1338	right	-	QS1338	right	increase
5B	PA5424	right	-	PA5424	right	-	PA5424	right	-
5C	PA5424	right	-	PA5424	right	-	QS1338	right	increase
5D	QS1338	-	increase	PA5424	right	-	QS1338	-	increase

When analysing the conflict resolutions of the controllers for the designed conflict in each scenario (Tab. 1), it can be observed that the controllers were quite consistent among themselves. This is especially true for the second controller who almost always picked the same aircraft to control and implemented the same resolution. The third controller appears to be less consistent compared to the other two controllers. Although the pool of controllers for the PUMBA study was quite limited, there is reason to believe (based on Tab.1) that with more controllers there would be a sufficient large pool of resolution command variations to conduct the final NALA conformance study.

V. DISCUSSION

SIMBA and PUMBA have been critical steps in the development of the MUFASA simulator and shaping the experimental design. Considering the ambitious research scope and exploratory nature of the MUFASA project, there have been, and still are, many questions to be answered. First and foremost, it has to be mentioned again that due to limited sample size, no statistical analysis could be performed of the results and therefore no hard conclusions can be drawn.

A. Complexity

Complexity levels may need to be adjusted based on data derived from PUMBA. Workload/complexity ratings from students in the SIMBA trials revealed a difference in complexity between scenarios. Controllers did, however, not indicate a difference in complexity between scenarios. This would encourage the need for a finer measurement scale to

increase the sensitivity of the subjective complexity ratings. The difference between high and low scenario complexity can be questioned. Neither workload ratings, nor the number of SSD inspection or commands were affected by complexity. However, a higher scenario complexity did have a negative effect on performance score. But this result should be taken with caution as the performance score is related to specific objectives (see Section 3.1. Experimental design) whereas complexity is a much broader concept.

B. Strategy

Strategy is deduced on a general basis, but not for the designed conflict. Conformance is only evaluated based on the strategies used in regards to the designed conflict. Even though we have been able to identify different resolutions between participants, we have been less successful with identifying the strategies underlying these resolutions. In the initial experimental design, verbal protocols were suggested as a method to probe participant strategies. During early simulator development it was decided to abandon the use of verbal protocols due to the disruptive effect, and difficulty to explain one owns action in real time. Instead, the decision was made to rely on strategy extraction through debriefings and performance metrics. However, the data from SIMBA and PUMBA does not provide adequate information on the strategies used.

On the other hand, we may not need to actually identify the underlying resolution strategies since the main objective in order to diversify conformal and non-conformal resolutions have been achieved. If the resolution is conformal, does it

matter if the underlying strategy is non-conformal? In this experiment, the underlying strategy used by the automation for the resolution is not readily explained to the operator. Will an operator not infer that the underlying strategy of the automation is conformal to his own strategy in resolution selection? Even though this may be the case, we would like to know more about the underlying strategies used by participants. In order to probe this the most attractive solution is perhaps running a few scenarios, possibly one of each scenario versions (in total four scenarios), either as replays and have participants talk through their own actions, or new ones where participant explain their actions in real time through verbal protocols. Both options have its flaws, with post-hoc debrief subjective to memory-bias, and verbal protocols disruptive.

Rotation and exit point names could be seen as confounding factors. Mathematically speaking the scenarios of each scenario group are exactly the same, but they might be perceived differently by controllers. For example, one controller may be used to having traffic only going east-/westbound. Scenario versions with traffic flows going north-/southbound may therefore be perceived as more difficult even though they are mathematically identical to east-/westbound scenario versions. Future debriefings will address this issue.

VI. CONCLUSION AND FUTURE WORK

This paper presented the design and results of an exploratory experiment that aimed to investigate conflict detection and resolution (CD&R) automation usage among professional air traffic controllers and novices (students) utilising a novel decision-support tool: the Solution Space Diagram (SSD). From the experiment results we observed that participants interacted differently with traffic in each scenario and that students, in contrast to the controllers, outperformed controllers by reacting more immediately and promptly to red conflict warnings. Although we can see a difference in work methods and SSD usage, the data does not tell us much about the underlying strategies used by controllers when solving conflicts. However, observations and debriefing of controllers revealed a general scepticism towards the SSD display, its accuracy and usefulness. This allows us to speculate that controllers, compared to students, had less trust in the SSD as a CD&R tool, and rather used their own judgment in conflict management.

Despite the lack of statistical analysis of the experiment results due to the small sample size, this exploratory experiment did indicate that the controllers used the SSD differently and this resulted in a variety of conflict resolutions to serve as conformal/non-conformal automated advisories under a higher level of automation. This puts us in a good position to launch the final NALA simulations scheduled early 2013. The NALA trials will be a complete study that investigates the interactive effects of automation level, air traffic complexity, and, above all, strategic conformance. Additionally, the NALA trials will also feature a large pool of

air traffic controllers in order to increase the sample size and the statistical power.

ACKNOWLEDGEMENT

We would like to thank the students at Delft and the controllers at Brussels for their participation in the experiments.

ACRONYMS

CD&R	Conflict Detection and Resolution
COPX	Cleared sector exit point
ISA	Instantaneous Self Assessment
LOA	Level of Automation
NALA	Nominal Advisory Level Automation
PUMBA	Preliminary Update / Modified Baseline Automation
SIMBA	Simulated Baseline Automation
SSD	Solution Space Diagram

REFERENCES

- [1] SESAR (2007). Deliverable D3: the ATM target concept, DLM 0612 001 0200a, September 2007.
- [2] Endsley, M.R. and E.O. Kiris, *The out of the loop performance problem: Impact of level of automation and situation awareness*, in *Human Performance In The Automated Systems: Current Research And Trends*, M. Mouloua and R. Parasuraman, Editors. 1994, Lawrence Erlbaum Associates: Hillsdale, NJ. p. 50-56.
- [3] Kaber, D.B. and M.R. Endsley, *The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task*. *Theoretical Issues in Ergonomics Science*, 2004. 5(2): p. 113-153.
- [4] Parasuraman, R. and V. Riley, *Humans and automation: Use, misuse, disuse, abuse*. *Human Factors*, 1997. 39(2): p. 230-253.
- [5] Ruff, H.A., S. Narayanan, and M.H. Draper, *Human interaction with levels of automation and decision-aid fidelity in the supervisory control of multiple simulated unmanned air vehicles*. Presence: Teleoperators and Virtual Environments, 2002. 11(4): p. 335-351.
- [6] Wickens, C.D., A.S. Mavor, and J.P. McGee, eds. *Flight to the future, human factors in air traffic control*. 1997, National Academies Press: Washington, DC.
- [7] Kuchar, J.K. and L.C. Yang, *A review of conflict detection and resolution modeling methods*. *IEEE Transactions on Intelligent Transportation Systems*, 2000. 1(4): p. 179-189.
- [8] Bekier, M., B.R.C. Molesworth, and A.M. Williamson, *Tipping point: The narrow path between automation acceptance and rejection in air traffic management*. *Safety Science*, 2012. 50(2): p. 259-265.
- [9] Mercado Velasco, G.A., Mulder, M., & van Paassen, M.M. (2010). *Analysis of Air Traffic Controller Workload Reduction Based on the Solution Space for the Merging Task*. Proceedings of the AIAA Guidance, Navigation, and Control Conference, Toronto, Canada.
- [10] Ellerbroek, J., Mulder, M., & van Paassen, M.M. (2011). *Evaluation of a Separation Assistance Display in a Multi-Actor Environment*. Proceedings of the 16th International Symposium on Aviation Psychology, Dayton, OH.