E.02.08-D14-MUFASA Extension-D14-Final Project Report

Document information	
Project Title	**Mu**ltidimensional **F**ramework for **A**dvanced **S**ESAR **A**utomation
Project Number	E.02.08
Project Manager	CHPR
Deliverable Name	**Final Project Report**
Deliverable ID	D14 (DE.3 in contract)
Edition	01.00.00
Template Version	03.00.00
Task contributors	
CHPR BV, and Technical University of Delft	

Abstract

This is the third and final deliverable in the MUFASA extension phase (M2) research. The MUFASA 1 (M1) project showed benefits to controller acceptance and performance of a conflict detection and resolution (CD&R) decision-aiding system, when suggested resolution advisories were conformal with the controller's own way of solving the conflict. M2, building on these results, set out to investigate why controllers sometimes rejected their own previous solutions, when they (mistakenly) believed these came from automation. The empirical results and conceptual developments associated with the MUFASA project have been disseminated over the last few years through journal articles, conference presentations, and technical reports. This report summarises this body of work.

Authoring & Approval

Prepared By - *Authors of the document.*		
Name &Company	Position & Title	Date
[redacted] CHPR	[redacted]	02/09/2015

Reviewed By - *Reviewers internal to the project.*		
Name & Company	Position & Title	Date
[redacted] TUD	[redacted]	04/09/2015

Reviewed By - *Other SESAR projects, Airspace Users, staff association, military, Industrial Support, other organisations.*		
Name & Company	Position & Title	Date

Approved for submission to the SJU By - *Representatives of the company involved in the project.*		
Name & Company	Position & Title	Date
[redacted] CHPR	[redacted]	04/09/2015

Rejected By - *Representatives of the company involved in the project.*		
Name & Company	Position & Title	Date

Rational for rejection

Document History

Edition	Date	Status	Author	Justification
00.00.01	02/09/2015		[redacted]	New Document
00.00.02	22/09/2015			Edited based on review

IPR (foreground)

This deliverable consists of SJU foreground.

Table of Contents

PUBLISHABLE SUMMARY 4
1 INTRODUCTION 8
1.1 PURPOSE OF THE DOCUMENT 8
1.2 INTENDED READERSHIP 8
1.3 DOCUMENT STRUCTURE 8
1.4 INPUTS FROM OTHER PROJECTS 8
1.5 ACRONYMS AND TERMINOLOGY 8
2 MUFASA 1 (M1) 10
2.1 BACKGROUND, M1 10
2.2 METHODS, M1 10
2.3 RESULTS, M1 10
2.3.1 Acceptance 10
2.3.2 Agreement with automated advisories 11
2.3.3 Response time to advisories 12
2.3.4 Difficulty ratings 12
2.4 CONCLUSIONS AND RECOMMENDATIONS, M1 12
3 MUFASA 2 (M2) 14
3.1 BACKGROUND, M2 14
3.2 METHODS, M2 14
3.3 RESULTS, M2 14
3.3.1 Conformance and acceptance 15
3.3.2 Source and acceptance 15
3.3.3 Consistency 15
4 TOWARD A REFINED CONCEPTUAL VIEW OF STRATEGIC CONFORMANCE 16
5 TECHNICAL PROJECT DELIVERABLES 18
6 DISSEMINATION ACTIVITIES 19
7 TOTAL ELIGIBLE COSTS 20
8 PROJECT LESSONS LEARNT 21
9 REFERENCES 22

Publishable Summary

It was demonstrated over 25 years ago that humans might be more likely to find fault in automation than in their own performance, even when “automation“ is in fact an unrecognisable replay of their own previous performance [1]. More recently, EUROCONTROL’s own CORA project struggled with how to present controllers strategic decision aiding advice [2]. The question centred on whether controllers would accept optimised or algorithmic solutions, which might not fit with their own preferred strategies. A more recent survey of conflict detection and resolution (CD&R) modelling methods noted a mismatch between such automation and controllers’ preferred working methods [3].

The MUFASA project started from the assumption that the inevitable trend toward automation in ATM will involve greater use of strategic decision aiding-- that is, automation that will assume control of more strategic and cognitive functions, functions that used to be exclusively the domain of the air traffic controller. If this is so, controller acceptance of automation would likely be even more critical in the future. Automation would not only have to perform “correctly,” but also perform in a way that conforms to the controller’s own preferred strategies. The MUFASA project introduced the concept of *strategic conformance*, which we defined as follows:

> *The degree to which automation’s behaviour and apparent underlying operations match those of the human.*

Starting from this view, MUFASA laid out an initial predictive framework for automation usage and acceptance, and set out to explore how these would be impacted by the possibly interactive effects of three factors: traffic complexity, level of automation, and strategic conformance. Specifically, M1 posed the following research question: Would controllers be more likely to accept automated CD&R advisories that matched their own previous solutions? Secondarily, would such strategic conformance demonstrate other performance effects, in terms of workload, response time, and / or agreement?

Methods, M1

M1 conducted a series of three real time simulations, of increasingly higher fidelity. The first two (SIMBA and PUMBA) were used to develop a simulated automation capability and interface, confirm heterogeneity of controller responses and solution strategies, and to refine experimental procedures (e.g. display and timing issues). This effort culminated in a large-scale real time simulation (dubbed NALA), in which three factors were varied within subject: level of automation (LOA, defined as Management-by-Consent, and Management-by-Exception [1]); Complexity (Low versus High) and Conformance of provided advisories (Conformal versus Non-conformal), Dependent measures included acceptance (the binary choice of whether to accept or reject a given advisory), response time (to advisory), difficulty rating, agreement rating (notice that this can differ from acceptance), and various survey responses. Details of the specific methods, including participants, equipment, traffic test scenarios, experimental design, and procedures, can be found in [4].

Results and Conclusions, M1

Conformal advisories were accepted more often, rated higher, and responded to faster than were non-conformal advisories (see Table 1). Similarly, under high (as opposed to low) complexity traffic, advisories were accepted more often, rated higher, and responded to faster.

These effects, in particular those associated with Conformance, were consistent with our original research hypothesis: that strategic conformance can benefit acceptance and performance with strategic advisory automation. In the end, though, one M1 result stood out in particular: of 256 conformal solutions (i.e. replays of controllers’ very own previous performance), 61 (or 23.8%) were rejected by controllers. How could it have been that controllers would disagree with themselves nearly one quarter of the time? Could it have been that controllers were demonstrating “dispositional bias” against automation—that is, that they were simply disinclined to use automation? Or could it have been that, as some have speculated, controllers are simply inconsistent over time in the solutions and strategies they might use?

Conformance	↑ Acceptance ↑ Agreement ↓ Response time --- Difficulty*

[1] For theoretical and practical reasons, LOA factor levels were ultimately collapsed for analysis.

	*differed by acceptance status
Complexity	↑ Acceptance ↑ Agreement ↓ Response time ↑ Difficulty

Table 1 Main effects, for Conformance and Complexity (M1).

The M1 revised framework report [5] laid out a number of recommendations that arose from the M1 project. Some of the main ones included

- Acceptance of, and agreement with, automation can be very different, depending on (for instance) how driven one feels by the automation. Automation development should consider the potential for the two to dissociate;
- Controller strategies do differ. There was some question going in that we would not be able to establish heterogeneity of controller responses (notice that the theoretical importance of strategic conformance rests on the assumption that controllers can differ in their preferred strategies. This heterogeneity should be kept in mind by system developers [6]; and

Strategic conformance is probably best considered early in the training / implementation cycle, when trust in automation is first being developed.

MUFASA 2 - Background

On the basis of M1 results, the M2 extension phase set out to address the following three issues:

1. **Consistency--** If controllers, when repeatedly presented the same air traffic problem, are inconsistent over time in the strategy they use (and thus the solution they choose), this would undermine the concept of strategic conformance. It could have been that the roughly 25% conformal rejection rate seen in M1 had been simply a reflection of such inconsistency;
2. **Source bias--** Alternatively, the M1 conformal rejection effect might have reflected a source bias. That is, had M1 controllers been biased against automation, or against receiving any sort of advisory? This led us to speculate whether the same conformal rejection effect would be seen if controllers had been instructed that advisories had come, not from automation, but from a colleague.
3. Representation effect-- "Representation" refers here to properties of the interface that impact the transparency / opacity of that interface, including the choice of display parameters and richness of information.

Methods, M2

M2 conducted a pair of real time simulations, with two samples. The Source effect was studied using a sample of five licensed ATCOs, and the Representation effect was studied using a sample of nine trainee ATCOs. Because we were able to collect data for the Consistency analysis through repeated manual (prequel) simulation runs, we were able to base the Consistency analysis on the entire combined sample of 14. Figure 1 shows the overlap between the three research questions and two experiments (each experiment involved a real-time simulation). Further details of the M2 simulation methods can be found in [7].

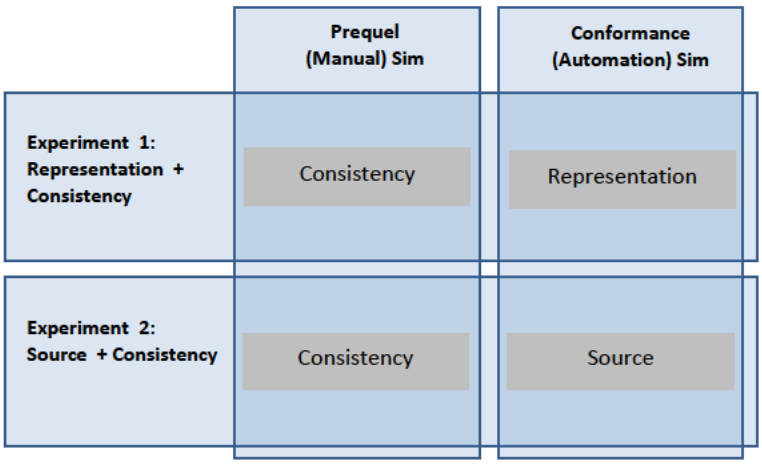

Figure 1. Experimental design matrix (M2).

Results, M2

The following touches only on selected highlights of the M2 analysis and results. Fuller coverage can be found in [7].

Conformance and acceptance: The M1 conformal acceptance effect was replicated in M2, albeit smaller overall and differentially by sample: overall, acceptance rate was 95% (19 of 20) for ATCOs, and 72.2% (26 of 36) for trainees. Acceptance rate in the trainee sample was indeed higher for conformal than non-conformal advisories (77.8% vs 66.7%), but the difference was much smaller in M2 than in M1. For ATCOs (n=5), acceptance was nearly complete (again, 95%), so there was no conformance effect.

Source and acceptance: The Source analysis was complicated by a ceiling effect on acceptance (19 of 20 advisories were accepted overall). However, Visual Analogue Scale (VAS) responses indicated that (and even though solutions were identical between the Source conditions), participants felt that "human-based" solutions were safer, less risky, and similar to the way they themselves would have solved a given conflict.

Consistency: Overall, four different consistency patterns were identified, all of which can be explained in a dichotomous choice of doing either action A or the opposite action B. Participants varied considerably in how consistently they solved the designed conflict. Four different groups of consistency patterns were identified based on: Solution parameters hierarchy analysis; Number of required interactions; Solution geometry; and control problem analysis. ATCOs were more homogenous than trainees. For all ATCOs, the common consistency pattern was that described as a control problem solution whereby the controlled aircraft was vectored either behind or in front of the intruder aircraft. All except one participant consistently vectored the controlled aircraft behind the intruder aircraft. The other participant consistently did the opposite: this participant vectored the controlled aircraft in front of the intruder aircraft. Finally, participants' self-rated consistency correlated moderately well with their calculated consistency.

Discussion and Conclusions

The MUFASA project originally put forth a conceptual model of controller acceptance, built around the notion of strategic conformance (SC). Over the course of the project, the team refined its view of the SC construct, and the role that it can play in fostering operator acceptance of advanced automation systems, particularly at the initial stages of implementation.

In situations, like ATM, that can lack "gold standard" criteria for optimal solution (minimisation of total flight path trajectory length is not always the "best" solution), the human can be the best judge. Moreover, there is a critical potential paradox: given that automation is becoming increasingly capable of assuming strategic decision making control, offered as advisories, automation is becoming more of an advisor and colleague. However, as with a human colleague, advice can be ignored or misused. The potential paradox lies in the implementation and familiarisation phase, when trust must develop. A controller might not develop trust until he / she has adequate experience using the machine; But he / she might not use the machine under it is trusted. Although the concept of SC says nothing about the quality of decisions (indeed, some rightly argue that it will sometimes just reproduce human errors), it suggests that benefits can accrue in terms of acceptance and trust.

Figure 2 shows a hypothetical relationship that illustrates the research team's evolved thinking on the potential benefit of the SC concept in automation development. At the initial roll-out of advanced decision-aiding automation, SC remains high for some period of time. As trust, usage and acceptance develop, automation can adaptively reduce SC (as appropriate), so as to increase the gap between the operator's baseline and current strategies. In this way, automation can begin to function not merely as an advisor but as a higher level trainer, ideally optimising solutions (which might differ from the human's previous ones) while maintaining acceptance. A much fuller discussion of this concept, and relevant support literature, can be found in [6].

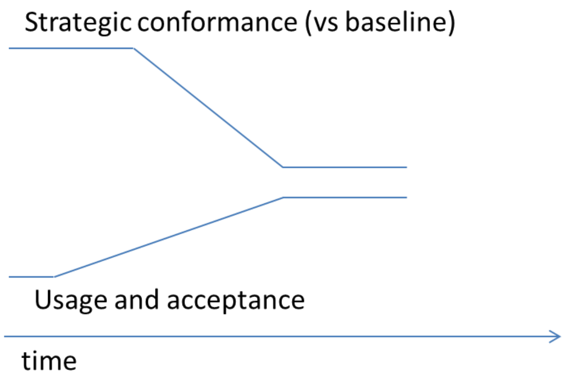

Figure 2. Hypothetical role of SC in training for automation.

1 Introduction

1.1 Purpose of the document

This document summarises the results of the Multidimensional Framework for Advanced SESAR Automation (MUFASA) project, including both the initial (M1) and extension phase (M2) work. It will review results of both M1 and M2 real time simulations, present a refined conceptual view of strategic conformance (SC), and summarise the dissemination activities associated with the entire MUFASA project.

1.2 Intended readership

The intended audience for this report consists primarily of representatives of the policymaking and R&D communities, who might have a stake in future ATM design. Those involved directly with aeronautical research, HMI design and human factors might also welcome the topics covered in this report. The report is written from an engineering and experimental perspective. Those from these and related domains will probably be most conversant with the terminology, but every attempt has been made to keep the document readable and approachable, so as to invite a wider audience.

1.3 Document structure

The document is structured as follows: This current Section 1 provides a high-level orientation of the document in regards to its purpose and target audience. Sections 2 and 3 review the background, methods and results of the M1 and M2 simulations, respectively. Section 4 presents a refined conceptual view of strategic conformance, and the role it can play in human-automation interaction. Section 5 summarises the dissemination activities that grew out of the MUFASA project.

1.4 Inputs from other projects

This research is a direct continuation of the MUFASA project, which was conducted under SESAR's WP-E research umbrella. The MUFASA project builds on a broad body of theoretical and empirical research into ATM automation and HMI design. The vast majority of this research is publicly accessible from academia and research organisations. The test platform for human-in-the-loop simulations was built around the Technical University of Delft's Solution Space Diagram (SSD), which has been under development for some years.

1.5 Acronyms and Terminology

Term	Definition
ATCO	Air Traffic Controller
ATM	Air Traffic Management
CD&R	Conflict Detection and Resolution
CHPR	Center for Human Performance Research
HMI	Human-Machine Interface
LOA	Level of Automation
M1	MUFASA 1
M2	MUFASA 2
MUFASA	Multidimensional Framework for Advanced SESAR Automation
NALA	Nominal Advisory Level Automation
PUMBA	Preliminary Update / Modified Baseline Automation
SESAR	Single European Sky ATM Research Programme

Term	Definition
SIMBA	Simulated Baseline Automation
SC	Strategic Conformance
SSD	Solution Space Diagram
TUD	Technical University of Delft
VAS	Visual Analogue Scale

2 MUFASA 1 (M1)

2.1 Background, M1

It was demonstrated over 25 years ago that humans might be more likely to find fault in automation than in their own performance, even when "automation" is in fact an unrecognisable replay of their own previous performance [1]. More recently, EUROCONTROL's own CORA project struggled with how to present controllers strategic decision aiding advice [2]. The question centred on whether controllers would accept optimised or algorithmic solutions, which might not fit with their own preferred strategies. A more recent survey of conflict detection and resolution (CD&R) modelling methods noted a mismatch between such automation and controllers' preferred working methods [3].

The MUFASA project started from the assumption that the inevitable trend toward automation in ATM will involve greater use of strategic decision aiding-- that is, automation that will assume control of more strategic and cognitive functions, functions that used to be exclusively the domain of the air traffic controller. If this is so, controller acceptance of automation would likely be even more critical in the future. Automation would not only have to perform "correctly," but also perform in a way that conforms to the controller's own preferred strategies. The MUFASA project introduced the concept of *strategic conformance*, which we defined as follows:

> *The degree to which automation's behaviour and apparent underlying operations match those of the human.*

Starting from this view, MUFASA laid out an initial predictive framework for automation usage and acceptance, and set out to explore how these would be impacted by the possibly interactive effects of three factors: traffic complexity, level of automation, and strategic conformance. Specifically, M1 posed the following research question: Would controllers be more likely to accept automated CD&R advisories that matched their own previous solutions? Secondarily, would such strategic conformance demonstrate other performance effects, in terms of workload, response time, and / or agreement?

2.2 Methods, M1

M1 conducted a series of three real time simulations, of increasingly higher fidelity. The first two (SIMBA and PUMBA) were used to develop a simulated automation capability and interface, confirm heterogeneity of controller responses and solution strategies, and to refine experimental procedures (e.g. display and timing issues). This effort culminated in a large-scale real time simulation (dubbed NALA), in which three factors were varied within subject: level of automation (LOA, defined as Management-by-Consent, and Management-by-Exception [2]); Complexity (Low versus High) and Conformance of provided advisories (Conformal versus Non-conformal), Dependent measures included acceptance (the binary choice of whether to accept or reject a given advisory), response time (to advisory), difficulty rating, agreement rating (notice that this can differ from acceptance), and various survey responses. Details of the specific methods, including participants, equipment, traffic test scenarios, experimental design, and procedures, can be found in [4].

2.3 Results, M1

The following is a brief summary of the M1 results. For details, and additional results, see [4].

2.3.1 Acceptance

Acceptance was defined as a controller's binary acceptance or rejection of a given resolution under automated conditions. Overall, controllers accepted 340 of 512, or 66%, of all advisories. Figure 1 shows the interaction between complexity and conformance on acceptance rate (defined as the number of acceptances as a percentage of total advisories).

[2] For theoretical and practical reasons, LOA factor levels were ultimately collapsed for analysis.

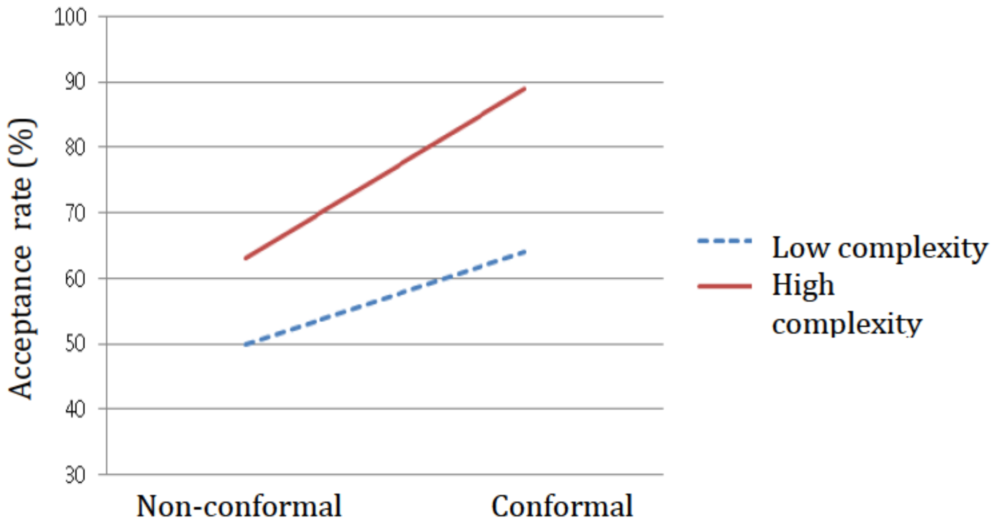

Figure 1. Acceptance rate, by Complexity and Conformance (M1).

Both complexity and conformance showed a significant main effect on acceptance rate. As shown in figure 1, controllers accepted roughly 75% and 56% of advisories under high and low complexity conditions, respectively. Conformance showed a nearly identical effect: controllers accepted 76% and 57% of advisories under conformal and non-conformal conditions, respectively. There was also a very weak interaction trend ($p=.37$) on acceptance: controllers tended to show higher acceptance for conformal solutions, and this effect was more pronounced under high complexity conditions.

2.3.2 Agreement with automated advisories

Regardless of whether a given advisory was accepted or rejected, controllers were instructed to indicate (on a scale of 1-100) their agreement with the advisory, immediately after each resolution scenario. Because of individual differences in ratings among controllers, absolute ratings were transformed into standardised *z* scores within subject.

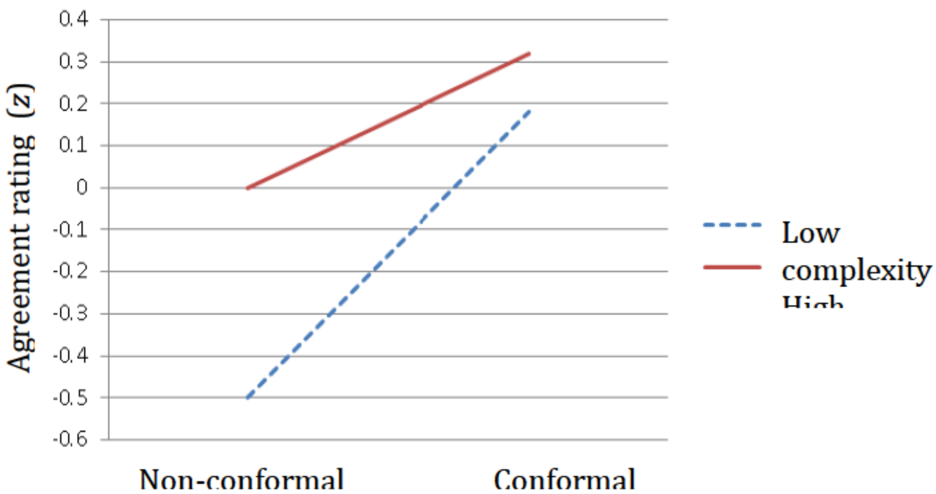

Figure 2. Average standardised agreement rating (z), by Complexity and Conformance (M1).

As with acceptance rate, standardised agreement ratings showed a significant main effect of both complexity and conformance. Acceptance ratings were significantly higher for conformal than non-conformal solutions (with an average *z* score of +.11 and -.11, respectively). Agreement was higher under complex conditions (average *z* score of +.05 and -.05 under high and low complexity). As shown in figure 2, agreement rating showed a borderline-significant interaction trend ($p<.1$) between complexity and conformance. Whereas controllers tended to show higher agreement with highly conformal solutions, this effect was <u>less</u> pronounced under high complexity conditions.

2.3.3 Response time to advisories

Response time was defined as elapsed time from onset of a given advisory to the controller's acceptance or rejection of that advisory. Response time ranged from approximately 0.6 sec to 14.6 sec (note that the system timed out at 15 sec). Response time showed a significant main effect of conformance, but not of complexity. Overall, response time was significantly higher for non-conformal advisories (5.9 sec vs 4.9 sec)—that is, controllers responded faster to conformal advisories. Response time was also lower high complexity conditions (5.7 sec vs 5.1 sec)—that is, controllers responded faster under complex traffic (though this difference was not statistically significant).

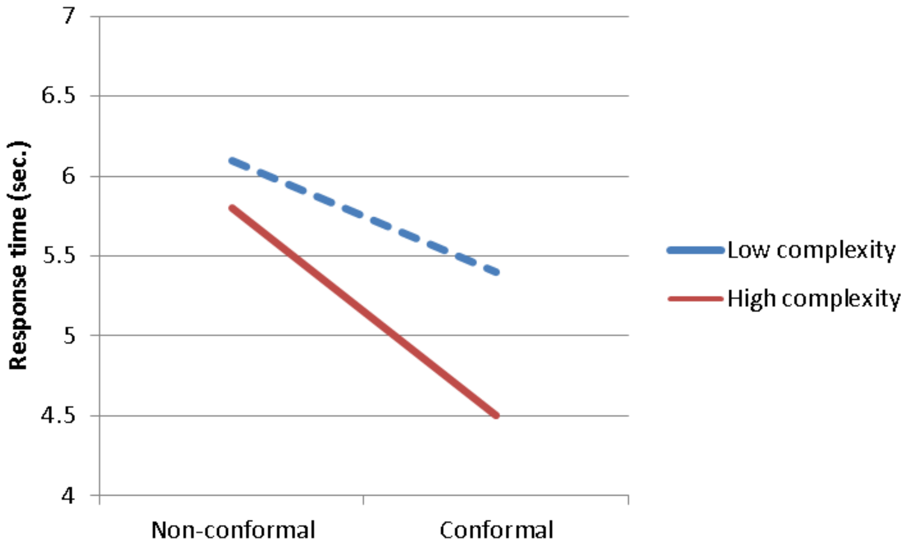

Figure 3. Response time (sec), by Complexity and Conformance (M1).

2.3.4 Difficulty ratings

	Conformal	**Non-Conformal**	*row avg*
Complexity – H	.19	.19	+ .19
Complexity – L	-.19	-.19	- .19
column avg	0	0	

Table 1 Average difficulty rating (*z*), by Complexity and Conformance (M1).

Difficulty ratings were obtained after each session, on a scale of 0-100. Notice that these ratings referred to the entire scenario, not just the advisory. Because of individual differences, difficulty ratings were standardised as *z* scores within subject. As shown in table 1, difficulty ratings varied by complexity, but not by conformance.

2.4 Conclusions and Recommendations, M1

Conformal advisories were accepted more often, rated higher, and responded to faster than were non-conformal advisories (see Table 2). Similarly, under high (as opposed to low) complexity traffic, advisories were accepted more often, rated higher, and responded to faster.

Conformance	↑ Acceptance ↑ Agreement ↓ Response time --- Difficulty* *differed by acceptance status
Complexity	↑ Acceptance ↑ Agreement ↓ Response time ↑ Difficulty

Table 2 Main effects, for Conformance and Complexity (M1).

These effects, in particular those associated with Conformance, were consistent with our original research hypothesis: that strategic conformance can benefit acceptance and performance with strategic advisory automation. In the end, though, one M1 result stood out in particular: of 256 conformal solutions (i.e. replays of controllers' very own previous performance), 61 (or 23.8%) were rejected by controllers. How could it have been that controllers would disagree with themselves nearly one quarter of the time? Could it have been that controllers were demonstrating "dispositional bias" against automation—that is, that they were simply disinclined to use automation? Or could it have been that, as some have speculated, controllers are simply inconsistent over time in the solutions and strategies they might use?

This final result led us to pose questions for further research, which ultimately led to the simulations of M2. First, why did controllers reject a fair number of conformal advisories? Was it that controllers are, as some have speculated, simply inconsistent over time? Further, was the rejection of advisories that we saw a reflection of controllers rejecting (what they believed to be) automation, or was it a rejection of advisories per se? Would controllers be similarly biased (i.e., inclined to reject a fair number of conformal advisories) if they had believed these had come not from automation, but from a trusted colleague? These research questions were addressed in M2, as outlined in section 3.

The M1 revised framework report [5] laid out a number of recommendations that arose from the M1 project. Some of the main ones include

- Acceptance of, and agreement with, automation can be very different, depending on (for instance) how driven one feels by the automation. Automation development should consider the potential for the two to dissociate;
- Controller strategies do differ. There was some question going in that we would not be able to establish heterogeneity of controller responses (notice that the theoretical importance of strategic conformance rests on the assumption that controllers can differ in their preferred strategies. Again, this heterogeneity should be kept in mind by system developers [7]; and
- Strategic conformance is probably best considered early in the training / implementation cycle, when trust in automation is first being developed.

3 MUFASA 2 (M2)

3.1 Background, M2

Again, M2 set out to investigate why controllers sometimes rejected their own previous solutions, when they (mistakenly) believed these came from automation. Three research questions were identified for further investigation. These were:

4. **Are controllers consistent?** If controllers, when repeatedly presented the same air traffic problem, are inconsistent over time in the strategy they use (and thus the solution they choose), this would undermine the concept of strategic conformance. It could have been that the roughly 25% conformal rejection rate seen in M1 had been simply a reflection of such inconsistency;
5. **Do controllers show a source bias?** Alternatively, the M1 conformal rejection effect might have reflected a source bias. That is, had M1 controllers been biased against automation, or against receiving any sort of advisory? This led us to speculate whether the same conformal rejection effect would be seen if controllers had been instructed that advisories had come, not from automation, but from a colleague.
6. **Is there a representation effect?** "Representation" refer to properties of the interface that impact the transparency / opacity of that interface, including the choice of display parameters and richness of information.

3.2 Methods, M2

M2 conducted a pair of real time simulations, with two samples. The Source effect was studied using a sample of five licensed ATCOs, and the Representation effect was studied using a sample of nine trainee ATCOs. Because we were able to collect data for the Consistency analysis through repeated manual (prequel) simulation runs, we were able to base the Consistency analysis on the entire combined sample of 14. Figure 4 shows the overlap between the three research questions and two experiments (each experiment involved a real-time simulation). Further details of the M2 simulation methods can be found in [6].

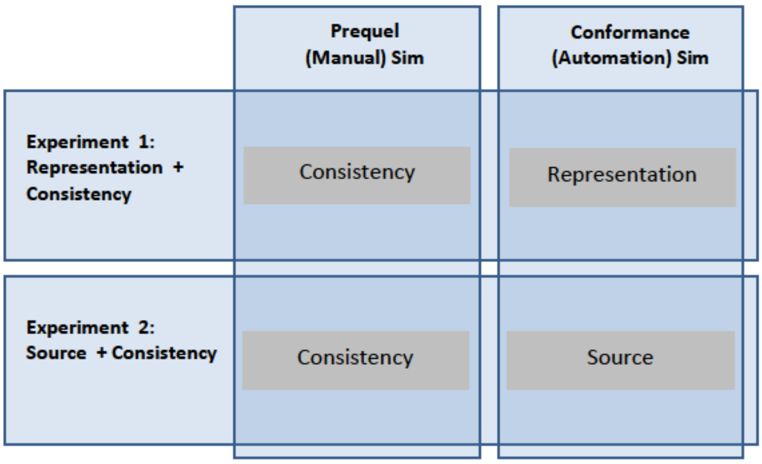

Figure 4. Experimental design matrix (M2).

3.3 Results, M2

The following touches only on selected highlights of the M2 analysis and results. Fuller coverage can be found in [6].

3.3.1 Conformance and acceptance

The M1 conformal acceptance effect was replicated in M2, albeit smaller overall and differentially by sample: overall, acceptance rate was 95% (19 of 20) for ATCOs, and 72.2% (26 of 36) for trainees. Acceptance rate in the trainee sample was indeed higher for conformal than non-conformal advisories (77.8% vs 66.7%), but the difference was much smaller in M2 than in M1. For ATCOs (n=5), acceptance was nearly complete (again, 95%), so there was no conformance effect.

3.3.2 Source and acceptance

The Source analysis was complicated by a ceiling effect on acceptance (19 of 20 advisories were accepted overall). However, Visual Analogue Scale (VAS) responses indicated that (and even though solutions were identical between the Source conditions), participants felt that "human-based" solutions were safer, less risky, and similar to the way they themselves would have solved a given conflict.

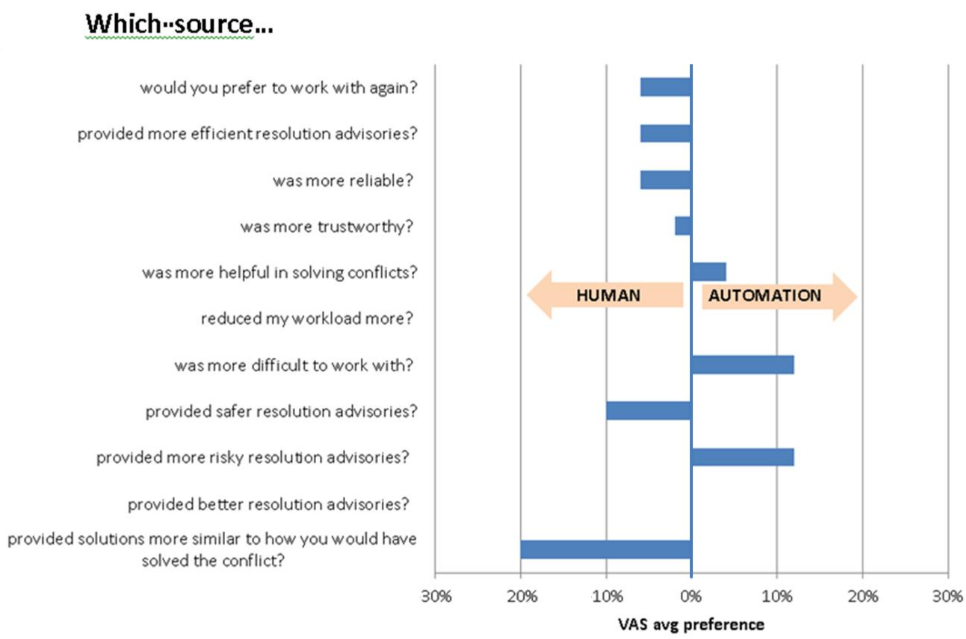

Figure 5. Average VAS response, Source effect (M2).

3.3.3 Consistency

Overall, four different consistency patterns were identified, all of which can be explained in a dichotomous choice of doing either action A or the opposite action B. Participants varied considerably in how consistently they solved the designed conflict. Four different groups of consistency patterns were identified, based on: Solution parameters hierarchy analysis; Number of required interactions; Solution geometry; and control problem analysis. ATCOs were more homogenous than trainees. For all ATCOs, the common consistency pattern was that described as a control problem solution whereby the controlled aircraft was vectored either behind or in front of the intruder aircraft. All except one participant consistently vectored the controlled aircraft behind the intruder aircraft. The other participant consistently did the opposite: this participant vectored the controlled aircraft in front of the intruder aircraft. Finally, participants' self-rated consistency correlated moderately well with their calculated consistency.

4 Toward a refined conceptual view of Strategic Conformance

The MUFASA project originally put forth a conceptual model of controller acceptance, built around the notion of strategic conformance (SC). Over the course of the project, the team has refined its view of the SC construct, and the role that it can play in fostering operator acceptance of advanced automation systems, particularly at the initial stages of implementation. This current section draws upon the article recently accepted for publication in IEEE Transactions on Human-Machine Systems [7]. Since this article captures the simulation results and conceptual evolution over the entire MUFASA project, it represents a collection of lessons learnt from the project.

Technology resistance is a real concern across various fields. In domains such as healthcare and ATC, resistance by eventual users has meant that advanced automation (such as robotic surgery in the former case, or strategic advisory tools in the latter) has sometimes gone un-, under- or misused. One potential source of this resistance can be an incompatibility between the heuristic and deterministic (optimised) processes that often underlie human vs automated processes. ATM initiatives such as AERA, ARC2000 and PHARE have all faced this potential issue.

Alternative approaches have been taken to designing automation in such as way that it better embraces the psychological and behavioural variables that can colour human decision making. One possibility is to model algorithms after human control strategies, so as to make automation act in a manner consistent with human decision-making strategies, thereby enhancing compatibility.

Human-machine compatibility can be seen within a hierarchy of compatibility levels [7]. Figure 6 shows how we conceptualise these levels, in comparison to respective cognitive engineering constructs. At the lowest level is response compatibility, the mapping between response and control device (such as an up/down lever for gear retraction/extension). At the other extreme is decision making compatibility, a relatively recent consideration as automation has evolved. If the machine arrives at the same solution as the human, the human is likely to infer the same underlying rationale at work. This offers the potential workload, response time and acceptance benefits.

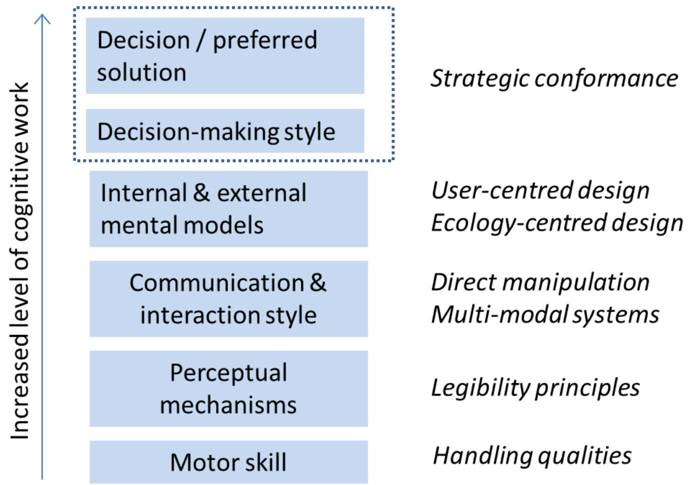

Figure 6. Levels of human-machine compatibility [after [6]).

Notice that SC can be framed in terms of either underlying strategy (i.e., process), and the outcome or solution (product). A conformal solution does not guarantee a conformal process (chess master and computer might arrive at the same solution, though the former generated a solution based on quick recognition of the board layout, whereas the computer might have performed an exhaustive lookup of all possible moves and counter-moves). Although we cannot infer process from product, in fact we all do this daily when interacting with others. We make inferences about underlying strategies and motivations based on repeated exposure to observable behaviours. This forms the basis for trust. Research suggests a similar process is involved when we interact with automation.

In situations, like ATM, that can lack “gold standard” criteria for optimal solution (minimisation of total flight path trajectory length is not always the “best” solution), the human can be the best judge.

Moreover, there is a critical potential paradox: given that automation is becoming increasingly capable of assuming strategic decision making control, offered as advisories, automation is becoming more of an advisor and colleague. However, as with a human colleague, advice can be ignored or misused. The potential paradox lies in the implementation and familiarisation phase, when trust must develop. A controller might not develop trust until he / she has adequate experience using the machine; But he / she might not use the machine under it is trusted. Although the concept of SC says nothing about the quality of decisions (indeed, some rightly argue that it will sometimes just reproduce human errors), it suggests that benefits can accrue in terms of acceptance and trust.

Figure 7 shows a hypothetical relationship that illustrates our thinking on the potential benefit of the SC concept in automation development. At the initial roll-out of advanced decision-aiding automation, SC remains high for some period of time. As trust, usage and acceptance develop, automation can adaptively reduce SC (as appropriate), so as to increase the gap between the operator's baseline and current strategies. In this way, automation can begin to function not merely as an advisor but as a higher level trainer, ideally optimising solutions (which might differ from the human's previous ones) while maintaining acceptance. A much fuller discussion of this concept, and relevant support literature, can be found in [7].

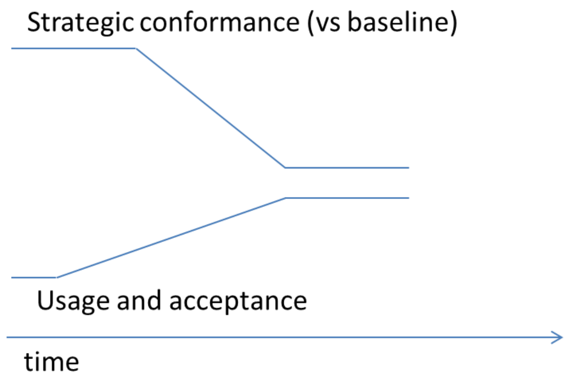

Figure 7. Hypothetical role of SC in training for automation.

5 Technical Project Deliverables

Management deliverables such as progress reports, gate report or this final report need not be included.

Number	Title	Short Description	Approval status
DE.1	Experimental Design	This report presents the experimental design of the planned human-in-the-loop simulations scheduled to be performed in Sweden early 2015. The experimental design associated with the three main research themes (consistency, source bias, and interface representation) is described in detail. Further, changes and improvements to the HMI and simulator used in the MUFASA research are explained. In particular, this report outlines the methods for manipulating experimental factors, as well as the dependent measures and analysis protocols we intend to use. Special focus is given to the development of the main stimuli consisting of designed conflict and traffic scenarios.	Approved
DE.2	Real-time Simulation Report	This report describes the methods and results of two real-time simulations recently conducted as part of the MUFASA extension phase (M2) research. On the basis of earlier MUFASA results, this work explored three research questions: [1] Are controllers consistent in their solution strategies? [2] Are controllers biased if they believe advisories come from a human colleague vs automation?, and [3] Can varying the interface representation benefit interaction with an automated advisory system? Two real-time simulations were conducted in winter-spring 2015. Participants included fully-licensed controllers (n=5) at Norrköping TMC, Sweden, and controller trainees (n=9) at Malmö ATCC, Sweden. The conformance effect first seen in M1 (i.e., that controllers were more likely to accept advisories when these matched their own strategy) was replicated, though the effect size was smaller here, especially among full ATCOs. Although most of the other performance measures were statistically insignificant, some interesting patterns emerged, especially among the self-report data, which favoured a “triangle” representation. Moreover (and even though they differed only by the participant’s belief in the source), “human” advisories were seen as safer, less risky, and better matching the controller’s own strategy, than were “automated” advisories.	Approved
DE.3	Final Project Report	This is the third and final deliverable in the MUFASA extension phase (M2) research. The MUFASA 1 (M1) project showed benefits to controller acceptance and performance of a conflict detection and resolution (CD&R) decision-aiding system, when suggested resolution advisories were conformal with the controller’s own way of solving the conflict. M2, building on these results, set out to investigate why controllers sometimes rejected their own previous solutions, when they (mistakenly) believed these came from automation. The empirical results and conceptual developments associated with the MUFASA project have been disseminated over the last few years through journal articles, conference presentations, and technical reports. This report summarises this body of work.	This document

Table 1 - List of Project Deliverables

6 Dissemination Activities

Following is a list of the publications and presentations to date, that have arisen from the MUFASA project, including both M1 and M2 phases.

<u>2011</u>

- **SID 2011**-- C. Westin, B. Hilburn, and C. Borst. Mismatches between automation and human strategies: An investigation into future air traffic management (ATM) decision aiding. Presented at the 1st SID. Toulouse, France.

<u>2012</u>

- **SID 2012**-- C. Borst, C. Westin, and B. Hilburn. An investigation into conflict detection and resolution strategies in air traffic management. Presented at the 2nd SID. Braunschweig, Germany.

<u>2013</u>

- **SID 2013**-- C. Westin, B. Hilburn, and C. Borst. The effect of strategic conformance on acceptance of automated advice: Concluding the MUFASA project. Presented at the 3rd SID. Stockholm, Sweden.
- **ISAP 2013**-- C. Westin, C. Borst, and B. Hilburn. Mismatches between automation and human strategies: An investigation into future air traffic management decision aiding. Presented at the 17th International Symposium on Aviation Psychology (ISAP), Dayton, Ohio, USA.

<u>2014</u>

- **Air Traffic Control Quarterly**-- B. Hilburn, C. Westin, and C. Borst, Will controllers accept a machine that thinks like they think? The role of strategic conformance in decision aiding automation. vol. 22, no. 2, pp. 115–136.
- **ICRAT**-- B. Hilburn, C. Westin, and C. Borst. Strategic conformance: An important concept in future automation design? Presented at the 6th International Conference on Research in Air Transportation (ICRAT), Istanbul, Turkey.
- **SESAR magazine**— article in issue 12, October, 2014.

<u>2015</u>

- **IEEE Transactions on Human-Machine Systems**--C. Westin, C. Borst, and B. Hilburn. Strategic conformance: Overcoming acceptance issues of decision aiding automation. In press.
- **CEAS**-- C. Westin, B. Hilburn, and C. Borst, Air traffic controller decision-making consistency and consensus in conflict solution performance, Presented at the 5th Challenges in European Aerospace (CEAS), Delft, The Netherlands.
- **LFV magazine**— article in summer issue, 2015.
- **TUD magazine**—article in winter issue, 2015.

7 Total Eligible Costs

This section is based on the Project Costs Breakdown Forms of the eligible costs incurred by project participants.

Date	Deliverables on Bill	Contribution for Effort	Contribution for Other Costs (specify)	Status
1/7/2015	DE.1	€ 49,826.07	*	paid
6/10/2015	DE.2, DE.3	€ 95,631.41	*	submitted
GRAND TOTAL		€ 145,457.48		

*Costs of travel included in daily rate

8 Project Lessons Learnt

What worked well?
Streamlined administrative procedures allowed us to focus most of our resources on technical work.
Network mechanism facilitated beneficial interaction between projects
Technical mentors helped refine our thinking, and improved quality control.
What should be improved?
It is hoped that administrative procedures could be streamlined even further in the future

Table 2 - Project Lessons Learnt

9 References

1. Fuld, R. B., Liu, Y., & Wickens, C.D. (1987). The impact of automation on error detection: some results from a visual discrimination task. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 31(2):156-160.

2. Kirwan, B. (2001). Towards a cognitive tool for conflict resolution assistance: A literature review (No. ASA.01.CORA.2.DEL04-A.LIT).

3. Kuchar, J.K., & Yang, L.C. (2000). A review of conflict detection and resolution modelling methods. IEEE Transactions on Intelligent Transportation Systems, 1(4), 179-189.

4. MUFASA (2013a). D3.2 Report on NALA 3D Validation.

5. MUFASA (2015). D2 Real-time Simulation Report.

6. MUFASA (2013b). D4 Revised Framework Report.

7. Westin, C., Borst, C. & Hilburn, B. (2015). Strategic conformance: Overcoming acceptance issues of decision aiding automation. IEEE Transactions on Human-Machine Systems, in press.

-END OF DOCUMENT-