Experimental Evaluation of a Joint Cognitive System for 4D Trajectory Management

Rolf Klomp†, Clark Borst†, Max Mulder†,
Gesa Praetorius‡

†Control and Simulation, Aerospace Engineering
Delft University of Technology
Delft, The Netherlands
r.e.klomp@tudelft.nl

‡Department of Shipping and Marine Technology
Chalmers University of Technology
Gothenburg, Sweden
gesa.praetorius@chalmers.se

Martijn Mooij§, Dennis Nieuwenhuisen‡

§Human Factors and Cognition Team
Thales Research and Technology Netherlands
Delft, The Netherlands
martijn.mooij@d-cis.nl

‡Department of ATM and Airports
National Aerospace Laboratory NLR
Amsterdam, The Netherlands
dennis.nieuwenhuisen@nlr.nl

***Abstract*— Effective joint human-automation coordination is essential in order to support the central role of the human operator in foreseen future trajectory-based air traffic operations. The SESAR WP-E project C-SHARE aims to achieve this by taking a Cognitive Systems Engineering approach, based upon accomplishing joint human and automation cognition through a shared representation of 4D-trajectory management. In foregoing research, a work domain model and a joint human-machine interface has been developed to support the human operator in the task of en-route 4D trajectory re-planning. This paper presents the findings of two experiments that aimed to determine the effect of both the initial level of *traffic orderliness* (i.e., structured versus unstructured traffic) and the *scale of perturbations* acting upon the airspace (e.g., number of conflicts and restricted areas) on the overall effectiveness of such a system. The findings of the experimental evaluation show that the C-SHARE approach to joint human-automation coordination in perturbation management is promising. Further, the experiment subjects accepted the tool and found it supportive for the task at hand, resulting in a manageable degree of workload during all experiment scenarios.**

Index Terms- Human factors, joint cognitive system, air traffic management

I. Introduction

THE Air Traffic Management (ATM) domain is foreseen to undergo a paradigm shift in the way in which air traffic is controlled [1]. Rather than the current form of hands-on tactical control, (4D-)Trajectory Based Operations (TBO) will enable the human controller to become a strategic airspace manager. This new form of control leans heavily upon the introduction of new decision support tools with higher levels of automation to support the human controller in performing new tasks [2].

Although the introduction of higher levels of automation is not good or bad in itself, in other complex socio-technical domains this has often shown to create new problems (e.g., coordination breakdowns, skill degradation, overreliance, transient workload peaks, etc.) [3]. In order to mitigate the risk for these so-called automation surprises by design, the C-SHARE consortium has taken an approach based upon the Cognitive Systems Engineering (CSE) and Ecological Interface Design (EID) paradigms [4], [5]. Instead of focusing on replacing human operators with automated systems, CSE and EID put the focus on supporting humans to conduct work in a work environment that is governed by constraints and opportunities. The goal is to develop a functional model of the work domain that represents the complete space of possibilities to do work. By making this representational model explicit in the human-machine system, the system operator, either human and/or automated agent, can jointly and robustly respond to system perturbations.

Previous work in aviation has demonstrated promising *ecological* concepts in the fields of airborne self-separation [6] and in-flight 4D trajectory management [7]. However, these interfaces were developed for pilots and thus required an egocentric perspective of the control problem. As such, these concepts propagated a distributed form of control in line with the EID and CSE approaches advocated by amongst others Rasmussen and Vicente. However, an air traffic controller seldom needs such an egocentric perspective. The nature of a controller's task is much more centralized and therefore would require a more exocentric perspective.

Within the C-SHARE project an exocentric representational model, called the *Travel Space*, for 4D trajectory management has been developed for short-term perturbation management in an airspace environment that has been de-conflicted a priori. The Travel Space representation enables human and automated agents to correct small scale system perturbations that take place on individual flights. This includes rerouting aircraft to compensate for small delays and prevent predicted separation violations. In a way, the Travel Space still supports a distributed form of control in a centralized environment. Therefore, it remains questionable how effective this representation would be for larger scale perturbations, such as large weather cells, that require multiple aircraft to be significantly rerouted. Such extreme control actions could potentially do more harm than good if the organization of the airspace itself is less structured.

The aim of this study has been to determine the effectiveness of the Travel Space representation under varying airspace and traffic conditions. By means of two human-in-the-loop experimental evaluations it is investigated how the initial level of traffic orderliness (i.e., structured versus unstructured traffic) and the initial scale of perturbation acting upon the airspace affect safety, performance, operator workload, and automation usage as well as controller acceptance. The first experiment featured a manual control task with a relatively low level of automation (i.e., only information integration by means of the Travel Space), whereas the second experiment featured the addition of an automated agent which would provide automatic trajectory resolutions for selected flights on controller request.

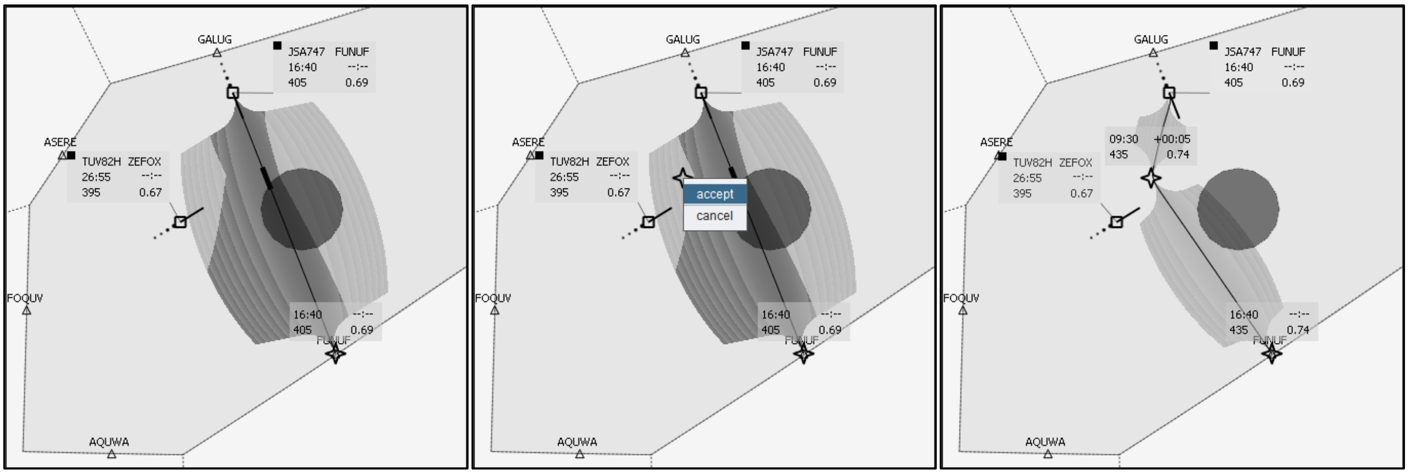
This paper presents the findings of the final human-in-the-loop evaluation performed within the framework of SESAR WP-E project C-SHARE. First, a brief description of the Travel Space representation, and how it can be used by the human operator, is given in Section II**.** In Section III the experiment set-up and methodologies are described in more detail. Section IV summarizes the most significant results which followed from the experiments. Finally, Section V provides a discussion of the results, followed by conclusions and recommendations in Section VI.

II. TRAVEL SPACE REPRESENTATION

The concept of the *Travel Space*, and its visual representation on the Air Traffic Control (ATC) plan-view display, forms the basis for shared human-automation cognition in the C-SHARE Joint Cognitive System (JCS) [8]. The task of manipulating and revising a 4D trajectory is supported by presenting the boundaries for safe control actions rather than single and optimized trajectory advisories. Inspired by the theoretical reasoning of Gibson on fields of travel for automobile-driving [9], the Travel Space visualizes the complete safe and restricted fields of travel in which a rerouting command, either human or computer generated, ensures adherence to aircraft performance (e.g., speed envelope, turn characteristics) and timing constraints as well as the overall airspace safety (e.g., separation assurance, restricted area avoidance).

Figure 1 shows three subsequent screenshots of the implementation of the Travel Space representation, and shows how this representation can be used to support the human operator in the manual control task. The task consists of de-conflicting the selected flight, rerouting it around a restricted area (RA), and maintaining timeliness at its sector exit point. Figure 1(a) shows a section of controlled airspace with two conflicting flights; the selected flight (JSA747) and a second en-route flight at the same flight level (TUV82H). Furthermore, the dark circular area inside the sector indicates a RA which is to be avoided by both aircraft. The Travel Space is visualized for the selected flight and consists of a *safe field of travel* (the lighter shaded area) and a *restricted field of travel* (the darker shaded area). The safe field of travel indicates the (2D) area in which the placement of an intermediate waypoint will be feasible (e.g., adhere to aircraft performance and timing constraints), and not lead to a new conflict with traffic. Figure 1(b) shows how the human operator can select and accept a position within the safe field of travel (indicated by the star symbol) to introduce an intermediate waypoint into the trajectory of the selected aircraft. Note that restricted areas are not explicitly taken into account in the restricted field of travel, thus waypoint placement requires some additional caution by the controller. Figure 1(c) shows the resulting valid trajectory for the selected flight. The original straight trajectory is divided into two equal-speed segments, both for which the new Travel Space is visualized.

The visualization of the work domain constraints and their relationships by the Travel Space allows the human controller to reason about, and directly act upon the airspace environment. However, the Travel Space visualization in itself does not indicate any (set of) *discrete optimal* solutions (i.e., shortest added path length, following the rules of the air, optimized for fuel burn, etc.) within the safe field of travel. Automated agents, on the other hand, are able to process large quantities of data and can calculate optimized discrete solutions. By combining the strengths of both humans (i.e., context aware creative problem solver) and automation (i.e., information fusion and computational power), effective joint problem solving can be facilitated. An important prerequisite for this is that the rationale which guides the automation



(a) Travel space representation for an aircraft in conflict and flying through a restricted area

(b) Placement of intermediate waypoint to ensure separation and restricted area avoidance

(c) Resulting valid trajectory for the observed aircraft

Figure 1. Travel Space support for the task of manual trajectory revision

should be based upon the *same* common ground which supports operator cognition (i.e., the concept of Travel Space). As such, this would allow for a shift back and forth across various levels of automation, i.e., from fully manual control to fully automatic control.

The current implementation of the Travel Space features two levels of automation: fully manual control, supported by a high level of information integration (i.e., the Travel Space visualization), and management-by-consent, where the human operator can request a computer-generated advisory for a selected flight that adheres to the work domain constraints and that is plotted inside the Travel Space visualization. After inspecting the validity of the advisory, the human can either choose to accept or reject it.

III. Experimental Design

A. *Goal*

The evaluation of the Travel Space has been conducted through an experimental study consisting of two separate experiments. Experiment I focused on the effectiveness of the manual control task with the Travel Space representation (without any automated advisories), whereas experiment II included the option to request an automated trajectory advisory for resolving perturbations. The goal of the study has been two-fold: firstly, to investigate how well the Travel Space supports the task of perturbation management in various traffic and airspace settings, and secondly, how often controllers would request advisories and actually accept and implement them.

B. *Subjects and instructions to the subjects*

Each of the two experiments was performed with a total of twelve subjects divided into three groups. Group A consisted of four air traffic controllers (area controllers, both certified and in training), Group B consisted of four domain experts, who are currently working in the ATM domain, and Group C consisted of four PhD students who perform similar flight-deck and/or ATM-related research. All subjects in Groups B and C, and two subjects in Group A participated in both experiments. All in all, 14 subjects contributed to the evaluation.

In the evaluation, the subjects were asked to manage traffic within an artificial two-dimensional airspace throughout six scenarios representing six experimental conditions. During each run the overall goal was to plan and guide the traffic through the controlled sector safely (e.g., without Loss of Separation (LoS) or restricted area intrusions) and efficiently (e.g., adhere to timing constraints at the sector exit points).

After the initialization of a scenario the subjects were free to resolve them by issuing changes to the 4D-trajectories (i.e., manipulating waypoints) of each individual aircraft. The resulting updated trajectories were generated and executed automatically by the aircraft.

C. *Apparatus*

The evaluation was performed on a dedicated software-based ATM-platform, running on a single computer. The same set-up (e.g., scenario design, display lay-out and input methods) has been used for both experiments.

Both the JCS Human Machine Interface (HMI) and the automated advisory software were run from the same computer. The shared representations within the JCS were integrated in a *traditional* plan-view display (PVD), providing a top-down view of airspace and air traffic (Figure 1). The PVD was presented on a 30-inch screen (60Hz LED, 2560 x 1600 pixels) in front of the participant. Input was given by a standard mouse and control options could be selected by on-screen drop-down menus.

D. *Independent Variables*

Both experiments featured a within-subjects design with two independent variables, which were:

- **Orderliness,** the initial traffic orderliness, with two levels: *structured traffic* (TS) and *unstructured traffic* (TU);
- **Perturbation**, the scale of an introduced airspace perturbation, with three levels: *small perturbation* (PS), *medium perturbation* (PM), and *large perturbation* (PL).

In total, the independent variables defined six traffic conditions. The orderliness variable defined the initial traffic set-up of the scenario. In structured traffic (TS), all aircraft would traverse the sector in structured (e.g., predictable) streams. This implied that aircraft initially traversed the sector in-trail on a limited set of fixed routes. In unstructured traffic (TU) all aircraft would enter and exit the sector by a unique combination (entry/exit point) of the eight fixed waypoints.

The *perturbation* variable was defined by the number of control actions which the subject initially had to perform during a scenario. In the small perturbation (PS) condition three aircraft pairs were initially in conflict and had to be re-routed. Note that the TS-PS and TU-PS conditions were the two baseline traffic conditions for all six scenarios. In the medium perturbation (PM) condition a restricted area (circular area with a radius of 10NM) was introduced in the sector at a location which required the additional re-routing of five aircraft. In the large perturbation (PL) condition the same restricted area was placed at a location which required a total of seven additional aircraft to be re-routed. An overview of the six experiment conditions is given in Table 1.

Table 1. Definition of the six experiment conditions

Condition		Structure	Perturbation
TS-PS	(scenario 1)	Structured	Small
TS-PM	(scenario 2)	Structured	Medium
TS-PL	(scenario 3)	Structured	Large
TU-PS	(scenario 4)	Unstructured	Small
TU-PM	(scenario 5)	Unstructured	Medium
TU-PL	(scenario 6)	Unstructured	Large

E. Dependent Measures

The following dependent measures were used to investigate the effect of traffic orderliness and perturbation scale (and their interactions) on the effectiveness of the system in both experiments:

- **LoS and RA intrusions:** Number of losses of separation and restricted area intrusions per condition.
- **Added path length:** Cumulative additional path length with respect to the original -shortest- trajectory for all aircraft per condition.
- **Acceptance:** Participants were asked to rate the use of the system after each given scenario using the Controller Acceptance Rating Scale (CARS). The subjective CARS rating was given on a modified version of the Cooper-Harper Scale , a decision-tree like structure developed by NASA to assess the acceptance of novel assistance systems in the ATC domain [10].
- **Workload:** A subjective Instantaneous Self-Assessment (ISA) workload rating was used to measure experienced workload. The digital ISA rating (continuous scale from low workload (0) to high workload (100)) popped up on the left-hand side of the PVD at three key points in each scenario.

In addition to the above, Advisory Usage was used as a dependent measure in experiment II.

- **Advisory Usage:** the number of requested, accepted and rejected advisories during each run.

F. Scenarios

The subjects were asked to manage traffic a hypothetical en-route sector (~40.000 KM^2) under six different control conditions. The rotation and orientation of the sector varied uniquely between scenarios consisting of the same (baseline) traffic structure to avoid a control bias due to scenario recognition. Each scenario presented approximately 15 aircraft and eight sector entry/exit points and lasted 24 minutes in scenario-time. The simulator ran at four times the normal speed, such that each scenario lasted six minutes.

The average traffic density was set to 8 aircraft under control at a given point in time, with the exemption of the first and last minute of the scenarios, in which the traffic either built up or reduced to compensate for the absence of handovers in between sectors, and the lack of verbal communication.

All aircraft entered the controlled sector at FL300 through one of eight fixed waypoints and were given an initial (straight) 4D trajectory leading towards one of the other waypoints. Aircraft could only be controlled laterally (i.e., vertical manipulation of the 4DTs was not possible), and only if they were physically inside the sector. Nevertheless, the aircraft entering the sector in the future were indicated by a grey representation when approaching the sector, such that the subjects had a certain amount of time to prepare for future traffic situations.

Furthermore, the performance of all aircraft was simulated using a single generic aircraft type. The initial conditions of each scenario were set such that the controller had to resolve a fixed set of perturbations (i.e., de-conflict aircraft pairs and avoid restricted areas) by manipulating the trajectories of individual aircraft. However, the control actions themselves could introduce new perturbations further ahead in time.

G. Hypotheses

It was hypothesized that the Travel Space representation supported the manual control task in small and medium perturbations, but that it was less effective for managing traffic in a condition with a large perturbation and unstructured traffic, leading to an increased number of LoS and intrusions into the restricted area.

It was also hypothesized that the decentralized representation of the travel space might impact the coordination between human and automation, leading to an increased number of advisory requests and implementations in less ordered traffic with large perturbations.

Furthermore, workload was hypothesized to be highest for the large perturbation conditions, and higher in unstructured than in structured traffic.

Finally, controller acceptance was hypothesized to decrease with increased degree of perturbation and in less structured traffic conditions.

IV. RESULTS

A. LoS and RA Intrusions

Out of the 2232 controlled flights in both experiments, two losses of separation events occurred in experiment I (one student in the TU-PL condition and one domain expert in the TS-PM condition), and none in experiment II. In experiment I there was one RA intrusion (one domain expert in the TU-PM condition), and one in experiment II (one domain expert in the TU-PL condition). The data did neither indicate any significant influence of traffic orderliness nor of perturbation scale on LoS and RA intrusions.

B. Added Path Length

Figure 2 shows a boxplot of the cumulative added path length (with respect to the original shortest aircraft trajectories) per condition for both experiments. As is to be expected, the figure shows an increase in added path length with respect to a larger perturbation scale due to the initial number of aircraft that required re-routing. The anomaly in added path length between the TU-PM and TU-PL conditions is likely caused by an artefact in the scenario design.

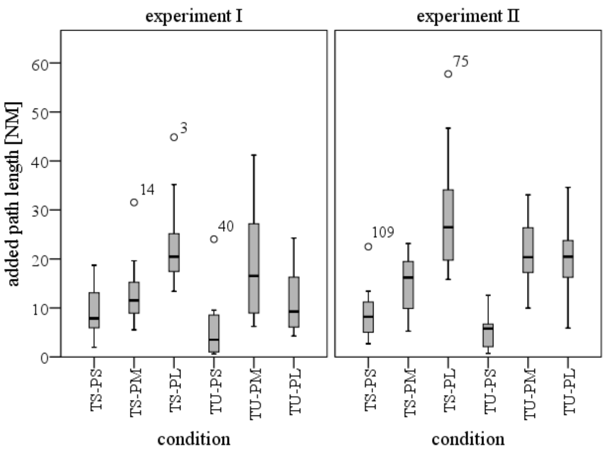


Figure 2. Boxplot of added path length per experiment, per condition

In both experiments, no between subject group effect was found on added path length. Furthermore, by comparing the cumulative added path length for all scenarios between both experiments, the possibility to use an automated advisory showed no significant effect on added path length.

C. *Acceptance*

Figure 3 shows a histogram of the subjective CARS ratings per experiment and per subject group. By using a non-parametric Friedman ANOVA the acceptance score was found to differ significantly between groups for both experiment I ($\chi^2(6)=11.57$, $p < 0.01$) and experiment II ($\chi^2(6)=12.00$, $p < 0.01$). In both experiments acceptance was found to be significantly higher for students compared to the domain experts, and significantly higher for domain experts compared to the ATCos. Furthermore, by comparing the difference in cumulative acceptance scores for each group by a Wilcoxon signed ranks test, acceptance was found to be significantly higher for each condition in experiment I compared to experiment II ($W(18)=-2.218$, $p < 0.05$). Traffic orderliness and perturbation scale both did not show a significant effect on the subjective acceptance rating.

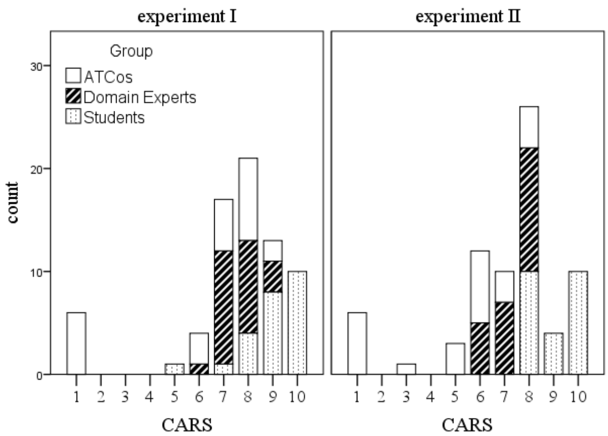


Figure 3. Histogram of subjective CARS rating per experiment, and per subject group

From Figure 3, it is also interesting to see that overall the ATCos used lower CARS ratings than both the domain experts and the students. This effect seems even more pronounced in Experiment II that featured a higher level of automation. This result suggests that professional air traffic controllers are more reluctant to accept newer and higher levels of automation in their work.

D. *Workload*

Firstly, by using a Kolmogorov-Smirnoff test, the average ISA data was found to be normally distributed for both experiment I ($D(72)=0.076$, $p > 0.05$) and experiment II ($D(72)=0.095$, $p > 0.05$). Figure 4 shows a boxplot of the within-subject z-scored average ISA values per condition for both experiments. By using a repeated-measures ANOVA a significant increase of workload was found as result of an increasing perturbation scale (i.e., small to large) for both experiment I ($F(1.90, 20.80)=12.26$, $p < 0.01$) and experiment II ($F(1.47, 16.20)=11.27$, $p < 0.01$). Furthermore, a significant decrease of workload was found for unstructured traffic compared to structured traffic in experiment II ($F(1,11)=8.47$, $p < 0.05$).

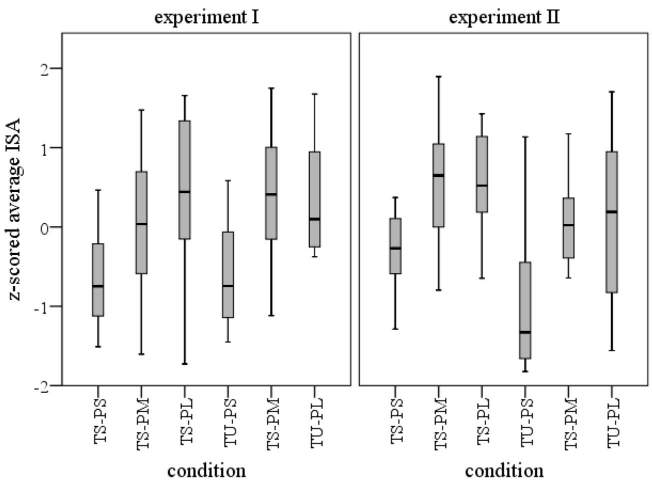


Figure 4. Boxplot of z-scored average ISA per experiment, per condition

E. *Advisory Usage*

Figure 5 shows a bar chart of the average number of accepted and rejected advisories per scenario and per subject group. By using a Wilcoxon signed ranks test, the ATCo group was found to request and accept significantly less advisories than both domain experts ($W(6)=-2.201$, $p < 0.05$) and students ($W(6)=-2.201$, $p < 0.05$). In total, the requested advisory count over all scenarios was 33 (accepted 52%) for the ATCo group, 86 (accepted 69%) for the domain expert group and 120 (accepted 56%) for the student group. Neither traffic orderliness, nor perturbation scale were found to have any effect on the amount of accepted and rejected advisories.

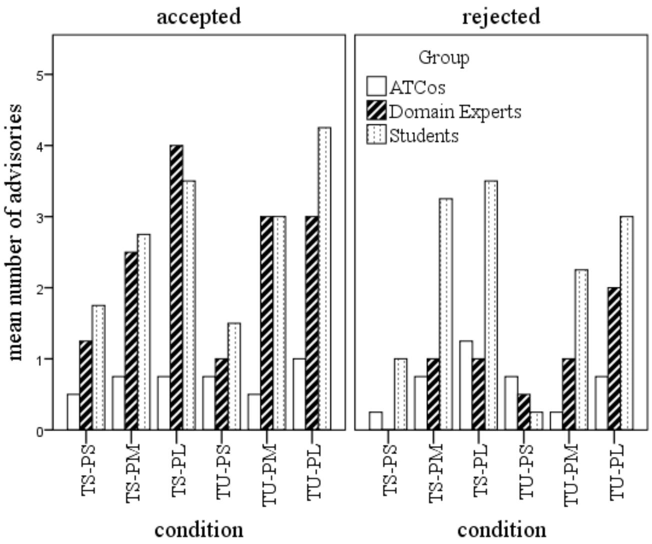


Figure 5. Average accepted and rejected advisory count per group, per condition

V. DISCUSSION

The aim of the presented research has been to study whether the C-SHARE approach supports human-automation coordination in perturbation management under varying airspace and traffic conditions. The following section will discuss the findings of the two experiments used to evaluate the approach.

It was anticipated that safety and efficiency would decrease in conditions with a higher level of perturbation and unstructured initial traffic conditions. However, the results have shown that there were only two loss of separation events and two RA intrusions throughout a total of 2232 controlled flights. This indicates that the human-automation ensemble –as used in the experiment– can effectively support the task of in-flight trajectory manipulation by ATC in various levels of perturbation.

However, the experiment system itself presents a simplification of the real-world work domain. Firstly, the vertical dimension which is frequently used by en-route ATCos in current operations has been omitted; in this study the human operators were restricted to manage traffic solely by lateral separation. Secondly, the mode of control in the system was implemented in a deterministic manner meaning that no uncertainty in the execution of trajectories by the aircraft was taken into account. All flights would immediately comply to any control actions issued by the controller. Through a visual analysis of the playback of the experimental runs it was apparent that the subjects were frequently operating at the boundaries of the presented constraints (e.g., actively controlling close to the 5 NM lateral separation constraint and RA boundaries to minimize path deviation), and thereby in a sense reducing the flexibility and robustness of the overall system to additional perturbations. In future research an attempt will be made to quantify these measures and to re-evaluate the sensitivity of the control actions to varying levels of additional perturbations.

The human factors measurements which have been applied in this study were the Instantaneous Self-Assessment Technique (ISA) to assess the operators' mental workload and the Controller Acceptance Rating Scale (CARS) to obtain data on the overall acceptance of the system by the controllers.

The results of the workload assessment show a significant increase with increasing degree of perturbation, while traffic orderliness did not show an effect for the first experiment. In experiment II, on the other hand, the workload, against what was hypothesized, is on average lower for unstructured traffic in comparison to structured traffic. While the increase in workload due the level of perturbation can be explained by the increased number of operator interaction needed to resolve all perturbations within a scenario, it remains unclear if ISA and similar techniques to assess mental workload subjectively are actually meaningful to apply for the evaluation of new systems following the initiative of NextGen and SESAR for TBO. In an ATM system in which TBO is exercised, the control task itself is significantly different from the tactical control that governs the work of ATCos today. TBO will focus on a strategic level of traffic management, meaning that tasks such as real-time traffic monitoring, which constitutes a large part of the workload an ATCo experiences today, will shift towards monitoring of the future trajectories, and will pose different cognitive demands towards the human-machine ensemble. Within the frame of this evaluation this indicates that the workload score measured every second minute by a subjective rating might actually not reflect the experienced workload of the specific traffic situation at that specific time, but might rather be related to the future state of the system in terms of trajectories which are and need to be addressed by the operator. Further, as also emphasized by Loft et al. [11], workload as a phenomenon is rather emergent based on the context within which an operator works and which an operator impacts on. This means that it is not a distinct point in time that can be analysed in isolation. Instead mental workload is influenced and therefore heavily dependent on the strategy a human operator develops to deal with the complexity of the tasks at hand. This can possibly explain why traffic structure did not seem to impact on the experienced workload, and also, why the increase of the score was most salient for the student participants in comparison to the domain experts and ATCos.

A crucial part of designing systems for future ATM is the overall controller acceptance and their usage for a defined task. In the evaluation presented within this paper it was hypothesized that acceptance would decrease in conditions with increased degree of perturbation and unstructured traffic. In contrary to the expectations prior to the evaluation, acceptance showed not to be affected significantly by perturbation scale or traffic orderliness. This shows that the tool is overall supporting the operator task in various settings of perturbation and traffic orderliness, and can facilitate the management of 4D trajectories in en-route ATM even for more complex traffic scenarios.

Further, the ATCos showed the least acceptance for the tool, meaning that they judged it as being less acceptable in supporting the trajectory-based control task than the students and the domain experts. A possible explanation might be that the ATCos judged the tool based on their current operational perspective. The ATM domain is a safety-critical domain in which the human operator is often identified as the last line of defence in unanticipated events. Transferring control and sharing decision making with an automated agent requires that the advisory heavily portrays what is considered a human operator's strategy of perturbation management (e.g., *strategic conformance)*, so that the operator can understand and feel that he/she is in the loop although parts of the problem-solving process might be externalized and executed by the automation. The significant decrease of acceptance for each condition in experiment II compared to experiment I could also be attributed to a possible mismatch in strategic conformance between the automated advisories and human strategy. Conformance has shown to be important for automation acceptance and the degree of acceptance with automation advisories [12].

Finally, as new systems with increased degrees of automation are introduced into the ATM domain, the human factors measurements applied in the assessment of today's systems might not be able to account fully for the impact of changed roles and functional distributions between human operator and automated agent [13]. As methods today heavily focus on the distinction between the "human" system parts, and those that are technical, they do not assess the overall performance of the joint human-automation system. While CSE offers methods to design system parts and to increase the understanding of the inner system workings, but a clear toolset to evaluate its overall performance does not yet exist.

VI. Conclusion

The overall objective of the C-SHARE experimental evaluation has been to see whether the Travel Space representation and the automated advisories support the task of en-route ATM in various traffic and perturbation conditions. At no point in time the system suffered from a breakdown, and only four safety-critical events occurred in the total of 2232, which shows that the approach itself is promising. In addition, the results also showed that the experiment subjects accepted the tool and found it supportive for the task at hand. The workload, although increased, was judged manageable while using a novel system within fairly advanced settings.

Acknowledgements

The authors acknowledge the inspiration from EUROCONTROL and the SESAR Joint Undertaking. This work is co-financed by EUROCONTROL acting on behalf of the SESAR Joint Undertaking (the SJU) and the EUROPEAN UNION as part of Work Package E (project C-SHARE: Joint ATM Cognition through Shared Representations) in the SESAR Programme. Opinions expressed in this work reflect the authors' views only and EUROCONTROL and/or the SJU shall not be considered liable for them or for any use that may be made of the information contained herein. Further, we would like to thank the Fru Mary von Sydow, född Wijk, foundation and the Region of Västra Götland for their financial support. Finally, we would like to thank the students, experts and the Air Traffic Controllers for their participation within the experiment.

References

[1] Eurocontrol, "SESAR Concept of Operations," Brussels, Belgium, 2007.

[2] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation.," *IEEE Transactions on Systems, Man, and Cybernetics. Part A, Systems and Humans*, vol. 30, no. 3, pp. 286–297, 2000.

[3] L. Bainbridge, "Ironies of automation," *Automatica*, vol. 19, no. 6, pp. 775–779, 1983.

[4] J. Rasmussen, A. M. Pejtersen, and L. P. Goodstein, *Cognitive systems engineering*, no. October. Wiley-Interscience, 1994, p. 396.

[5] K. J. Vicente and J. Rasmussen, "Ecological Interface Design: Theoretical Foundations.," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 4, pp. 589–606, 1992.

[6] J. Ellerbroek, M. Visser, S. B. J. Van Dam, M. Mulder, and M. M. Van Paassen, *Design of an airborne three-dimensional separation assistance display*, vol. 41, no. 5. IEEE, 2010, pp. 863–875.

[7] B. J. A. Van Marwijk, C. Borst, M. Mulder, M. Mulder, and M. M. Van Paassen, "Supporting 4D Trajectory Revisions on the Flight Deck : Design of a Human-Machine Interface," *The International Journal of Aviation Psychology*, vol. 21, no. 1, pp. 35–61, 2011.

[8] R. Klomp, M. M. R. Van Paassen, C. Borst, and M. Mulder, "Joint Human-Automation Cognition through a Shared Representation of 4D Trajectory Management," in *Proceedings of the SESAR Innovation Days (2012)*, 2012, pp. 1–7.

[9] J. J. Gibson and L. E. Crooks, "A Theoretical Field-Analysis of Automobile-Driving," *Journal of Psychology*, vol. 51, no. 3, pp. 453–471, 1938.

[10] K. K. Lee and T. J. Davis, "The development of the Final Approach Spacing Tool (FAST): A cooperative controller-engineer design approach," *Nasa Technical Memorandum*, no. August, 1995.

[11] S. Loft, P. Sanderson, A. Neal, and M. Mooij, "Modeling and predicting mental workload in en route air traffic control: Critical review and broader implications," *Human Factors*, vol. 49, no. 3, pp. 376–399, 2007.

[12] C. Borst, C. A. Westin, and B. H. Hilburn, "Mismatches Between Automation and Human Strategies: an Investigation Into Future Air Traffic Management Decision Aiding," in *Proceedings of 17th International Symposium of Aviation Psychology*, 2013, pp. 530–535.

[13] J. Langan-Fox, M. J. Sankey, and J. M. Canty, "Human factors measurement for future air traffic control systems," *Human Factors*, vol. 51, no. 5, pp. 595–637, 2009.