Ecological Approach to Train Air Traffic Control Novices in Conflict Detection and Resolution

Clark Borst, Roeland Visser, M.M. Van Paassen, and Max Mulder

Control & Simulation, Faculty of Aerospace Engineering
TU Delft, 2629 HS, Delft, Netherlands
contact: c.borst@tudelft.nl

Abstract*—In the near future, air traffic controllers will need to work with more sophisticated automation to meet higher demands in flight technical performance and operational safety. However, the aviation community fears that more intelligent support systems could diminish the 'hands-on' skills and cognitive expertise of controllers. To overcome this, training will become a critical issue to help them understand the rationale underlying the automation as well as to develop and preserve their expertise. But what if we could design automation in such a way that the emphasis lies on *cooperation* by making the technical systems *share* their deep knowledge? Can such 'transparent' systems actually help to *develop* and *maintain* the cognitive expertise required for the job? In this paper, an ecological interface for conflict detection and resolution, developed in a previous study, was used as a 'transparent' system that provides a deeper insight into the causal constraints governing traffic situations. It was hypothesized that this interface would expedite novice learning and convergence to 'best practices' when compared to an instructional training method. Results from the experiment, in which 16 students participated, revealed that the overall control performance in the final measurement session, featuring a conventional radar display, was not significantly different between the ecological and the instructional group. However, observations and trends suggested that the ecological group revealed more critical reflection behavior (leading to delayed actions), which occasionally allowed participants to solve 'novel' conflict scenarios. Further research is needed to analyze and map the student's development of cognitive expertise.

***Keywords*—Ecological Interface Design, Training, Human-Machine Interface, Air Traffic Control.**

I. Introduction

Predicted increases in air traffic complexity will force a fundamental shift in air traffic control (ATC), requiring humans to supervise more complex and more intelligent automation [1]–[3]. This, however, has also given rise to a growing concern within the ATC community: will controllers remain competent and skilled enough to safely assume control in case the automation fails? [4], [5] Similar to how flight deck automation and autopilots have been reported to play a significant role in skill erosion of commercial airline pilots [6]–[8], the fear is that smarter automation will create a dependency that will ultimately diminish the controller's cognitive expertise. As such, properly training the next generation air traffic controllers will become a critical issue to keep their hands-on skills and expertise alive [9], [10].

Besides new training procedures, another way to ensure efficient training for the future air traffic management system would be to change the way in which ATC systems (and automation in general) are designed. The majority of automation design approaches put the emphasis on supplanting human involvement by technical systems [6], [11]. Consequently, essential knowledge required to control a process remains hidden inside the black box and such systems would require additional training to fill knowledge gaps necessary to safely intervene in case the automation reaches its boundaries. But what if we could design automation in such a way that the emphasis lies on *cooperation* by making the technical systems *share* their deep knowledge [11]? Can such 'transparent' systems actually help to *develop* and *maintain* the cognitive expertise required for the job?

In this paper the Ecological Interface Design (EID) paradigm [12] will be empirically explored as a way to design 'transparent' support systems with unique training potential. Ecological interfaces typically reveal the deep structure underlying a control problem [13] and are therefore hypothesized to lead to a deeper knowledge into the control problem through insightful learning [14]. Here, the scope will be training air traffic control novices (who are unbiased by previously developed strategies) in a conflict detection and resolution task in the horizontal plane. An ecological interface developed in a previous study, called the Solution Space Diagram [15], [16], will be used to achieve this goal. The general approach will be to analyze the decision-making strategies and the control performance gained after two concurrent programs, where a simplified instruction-based training program (representing the current state of the art in ATC training) will be compared to a training program featuring the SSD in addition to the instructions. Note that preliminary insights into the training potential of ecological interfaces have been gathered before in the process control domain, featuring slow system dynamics, over a period of six months [14], [17]. In this paper, however, training time is reduced to only two days and the control task features fast system dynamics As such, new empirical insights obtained from a different work domain (i.e., aviation) may provide a deeper understanding of the merits and versatility of the EID approach as a training method.

II. Theoretical Motivation

A. EID for ATC training

In ATC, trainees are taught to expedite air traffic safely and as efficiently as possible using a combination of learned strategies and procedures in response to recognized patterns and conflict geometries [10], [18]. Expert controllers can actively switch between strategies and procedures 'on the go' to adapt their task performance warranted by the demands of the operational situation. Before trainees reach this level of expertise, it will take up to four years of intensive training, featuring a steep learning-curve [18].

In their first year of basic training, trainees learn ATC rules and regulations in classroom sessions, practice ATC basics in simple simulators and have to master radio communication. In their second and third year, trainees practice traffic scenarios in high fidelity simulators and perfect their skill set to manage all sorts of scenarios. Self discovery of what works and what does not work is typically how they learn the required skills. Unexpected disturbances challenge earlier proven solutions, so as to discourage a 'solve-all' strategy and to encourage knowledge-based problem-solving rather than memorizing 'tricks'. ATC The Netherlands (LVNL) has developed the ATC Cognitive Process & Operational Situation (ACoPos) model to make the criteria underlying controller expertise more salient (see Figure 1) [10]. The white boxes are all categories used in proficiency assessment of the trainee and describe how well the controller interacts with the operational situation (left side) as well as how his or her own competences develop over time (right side).

Given the description of the ATC training process and its objectives, similarities can be discovered with the ideas underlying the EID paradigm. First, similar to how ATC trainees are taught 'robust' control strategies instead of fixed procedures, the EID framework was founded on the basic principle of providing support for unanticipated events for which no procedure exists. Ecological interfaces typically do so by portraying the space of possibilities (e.g., governed by the laws of physics) instead of single optimized solutions that may fall short in situations that violate their specific assumptions [13].

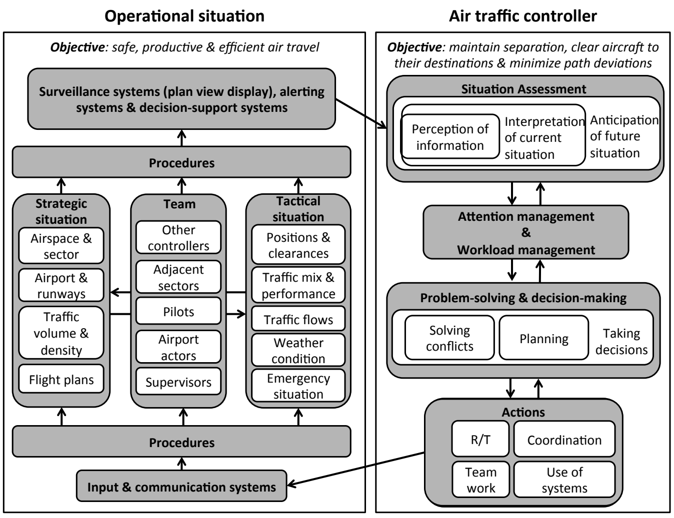


Figure 1. The ACoPos model (adapted from [10]).

Second, as can be observed from the ACoPos model in Figure 1, the air traffic controller needs to develop a mental model of the operational situation and how aspects of the tactical traffic situation (e.g., aircraft positions, their flight directions, atmospheric conditions, etc.) relate to a higher-level strategic situation that contain information about the current and future state of the overall airspace environment. In EID, Rasmussen's Abstraction Hierarchy and/or the Abstraction-Decomposition Space serve a similar purpose by grouping domain-relevant constraints at different levels of abstraction, ranging from lower-level states and whereabouts of objects to their relationships with higher-level functional goals in the operational environment. To goal of EID is to portray this work domain structure on a display to serve as an externalized mental model of the system under control. Preliminary evidence indicates that under certain conditions people can use this information to update their internal mental model [14].

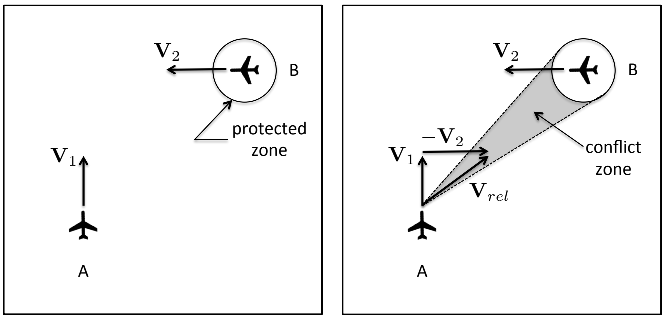
Finally, from Figure 1 one can readily see that controller expertise is influenced by a large number of perceptual factors. This is not surprising, considering that a controller needs to gain knowledge about the state of the airspace entirely from a Plan View Display (PVD), or, radar display. An ideal ATC training tool should thus support the trainee to become familiar with intricacies of the operational context by actively supporting the action-perception cycle. In this view, ecological interfaces typically aim to transform a cognitive task into a perceptual task [12], enabling their users to expedite situation recognition and formulate solutions to problems.

In sum, the general objectives of ATC training are largely consistent with the theory and ideas behind the EID paradigm. In particular, encouraging the development of knowledge-based problem-solving activities and robust control strategies (by targeting the action-perception cycle) are the most striking similarities that should make ecological interfaces suitable for ATC training.

B. Supporting solution strategies for workload mitigation

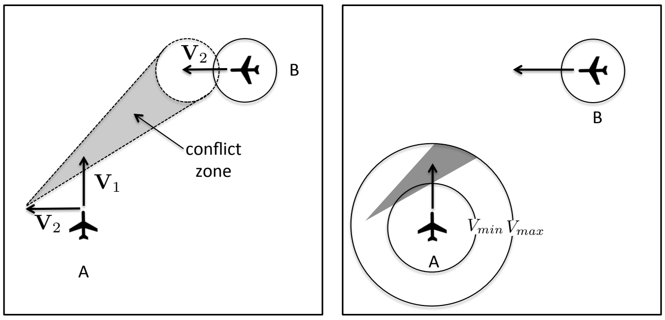
An important trait that separates an expert controller from a novice is that experts apply solution strategies that will minimize their monitoring time. Management of attention and workload is therefore also explicitly represented in the AcoPos model (see Figure 1). Although ATC instructors do not teach specific solution strategies, literature indicates that expert controllers tend to converge to a range of 'best practices' with workload-mitigating properties [19]–[24]. Also here, ecological interfaces are expected to support the development of such practices.

To illustrate this, consider the Solution Space Diagram (SSD) shown in Figure 2, an ecological interface developed in a previous study [15], [16]. In its most succinct form, the SSD portrays velocity obstacles (or, conflict zones) in speed and heading within the maneuvering envelope of the aircraft under



(a) traffic geometry

(b) conflict zone in relative space



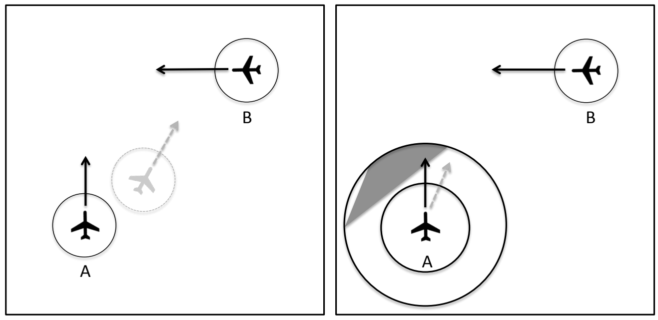
(c) conflict zone in absolute space

(d) resulting solution space diagram for aircraft A

Figure 2. The Solution Space Diagram (SSD), showing the triangular velocity obstacle (i.e., conflict zone) formed by aircraft B within the speed envelope of the controlled aircraft A.

control. The velocity obstacles constrain the maneuvering opportunities of the controlled aircraft in terms of potential loss of separation events. That is, if the velocity vector of the controlled aircraft lies within a triangular conflict zone, a loss of separation will occur in the near future. Vectoring the controlled aircraft outside such a conflict zone would thus resolve the conflict.

An example ATC 'best practice' to resolve a crossing conflict in the horizontal plane, featuring two aircraft flying at different speeds, is to vector the slow aircraft behind the faster aircraft. This is a typical 'set-and-forget' strategy that minimizes the required monitoring time [23], [24]. The merit of this strategy is justified when considering it within the SSD. From Figure 3(b) it can easily be seen why the best practice is a robust solution that requires less monitoring time than vectoring the slow aircraft in front of the fast aircraft. That is, the available solution space on the right side of aircraft A's maneuvering envelope is much richer than on the left side. Additionally, placing the speed vector of the aircraft A outside the velocity obstacle involves a small heading change to the right, making this a quick solution to resolve the conflict. As such, the SSD has an explanatory value that amplifies the best practice, potentially expediting the development of control expertise.



(a) ATC 'best practice': put slow aircraft A behind faster aircraft B.

(b) SSD: put slow aircraft A behind faster aircraft B.

Figure 3. The SSD explains and amplifies the ATC 'best practice', potentially expediting the development of control expertise.

III. Experiment Design

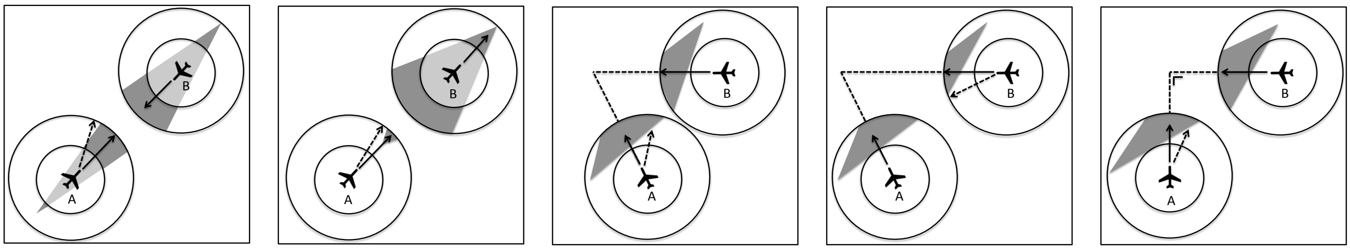
A custom-made, simplified ATC training course was created for conflict detection and resolution (CD&R) in the horizontal plane to test the influence of the SSD on learning and control performance. Participants in the focus group (trained with the SSD and best-practice instructions) and the control group (trained with best-practice instructions, but without the SSD) were both taught to recognize specific conflict geometries along with their corresponding solutions. Participants were asked to learn and train as best as they could, and to put this knowledge into practice during a measurement phase where both groups only had access to a conventional PVD.

A. Participants

Sixteen aerospace engineering students (TU Delft) with an interest in air traffic control and a mean age of twenty-six years (standard deviation of 1.9) were asked to participate voluntarily. They were first asked to complete a questionnaire with general questions about their familiarity with ATC systems and tools (e.g., having played ATC computer/smartphone games). After that, a short test was administered requiring them to observe traffic stills on paper and indicate whether they were in conflict. From this test, their entry skill level was derived. In this manner, two balanced groups of eight participants with on average similar skill levels were formed. This pre-test also acted as an elementary selection test, excluding participants with too much knowledge, or too little knowledge in ATC.

B. Instructions

All participants were given a mini lecture (slide show) about aspects of the ATC experiment. From this mini lecture, five learning goals needed to be mastered, which were five aircraft conflict types (see Figure 4), each with a 'best practice' solution (see Table I). The 'best solutions' dictated in a specific conflict geometry of an aircraft pair what the best and most efficient action was to solve the conflict. The solutions to the conflicts were distilled from the Rules of the Air, general rules of thumb that were adapted from research about the strategies of controllers [22], [23] and feedback from external experts on ATC training programs. These 'best solutions' provide a straightforward and quick fix for a conflicting pair of aircraft. Note that in these solutions, the conflict and corresponding solution was limited to two aircraft. During subsequent training exercises, also scenarios with three or more aircraft could



(a) Head on (HON) (b) Overtake (OVR) (c) Crossing (CRO) (d) Crossing + distance bias (CRB) (e) Perpendicular (PER)

Figure 4. Conflict types and their visualizations within the SSD. The length of the speed vectors indicate the speed magnitude and the dashed speed vectors indicate the best practice solution to the conflict. The other dashed lines indicate the distance toward the crossing point of the aircraft pairs.

TABLE I
CONFLICT TYPES AND THEIR 'BEST PRACTICES'.

Conflict type	Heading difference [deg]	'Best practice' with speed difference	'Best practice' with equal speeds
Head on (HON)	170 - 180	Faster aircraft evades conflict	Either aircraft, depending on surrounding aircraft
Overtake (OVR)	0 - 10	Overtaking aircraft evades conflict	–
Crossing (CRO)	10 - 170	Slower aircraft evades conflict	Either aircraft, depending on surrounding aircraft
Crossing + bias (CRB)	10 - 170	Aircraft arriving later evades conflict	Aircraft arriving later evades conflict
Perpendicular (PER)	80 - 100	Slower aircraft evades conflict	Either aircraft, depending on surrounding aircraft

be encountered. Two-aircraft scenarios always had one 'best solution'. In the three-aircraft scenarios, the third aircraft could either support (strengthen) that 'best-solution' *or* cause the original 'best solution' to create a conflict with the third aircraft, thus compel the controller to deviate from the learned best practice (see Figure 5).

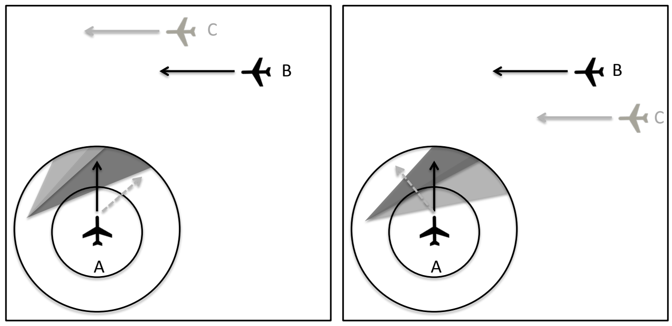
When practicing in the simulation environment, the task of the participant was to first guarantee safe separation of aircraft at all times by solving or preventing conflicts, and secondly to vector aircraft as efficiently as possible toward their respective exit waypoint. More difficult tasks such as the merging of conflict streams or creating flight patterns for airport arrivals were not considered, due to the short training time in this experiment (about three hours).

Participants were also instructed to think out loud during the simulator sessions to gain insight in their decision-making strategy. Feedback after each exercise could then be given on their decision-making strategy.

C. Independent variables

The two independent variables in this experiment were: (a) **training**, with or without the SSD and (b) the **traffic scenarios** presented to the participants. Training with or without the SSD varied between-participants and only applied to a training phase. During a measurement phase, both groups only had access to a legacy PVD to control traffic. Further, training participants in the five conflict types and their corresponding solutions (Figure 4 and Table I) featured aircraft pairs without any other traffic. Traffic scenarios were varied within-participants, and this was realized by changing the order of appearance of the rehearsal exercises. Hence, a mixed design was used.

The traffic scenarios have been developed with the help of two external experts on ATC training from the LVNL,



(a) amplifying best practice. (b) deviating from best practice.

Figure 5. Example of a three-aircraft scenario where the third aircraft (C) either amplifies, or requires deviation from, the best practice.

with the specific focus on a single horizontal plane and vectoring of aircraft by heading clearances. These restrictions (or simplifications) of the scenarios aimed to prevent confounds caused by the increased number of conflict solution possibilities. Scenarios were classified by the type of conflict (one of the five learning goals), whether the aircraft pair in conflict flew at different speeds, and the number of aircraft surrounding the conflicting pair. The geometrical orientations of the conflicts were rotated or mirrored in order to keep scenarios unrecognizable despite their similar geometries. All traffic scenarios had predefined solutions, such that 1) near-unbiased performance comparisons could be made, as well as 2) the same feedback could be given afterwards to participants.

In total, three types of training elements were designed: stills of conflict scenarios (pictures), short dynamic scenarios (90 seconds) and long dynamic scenarios (900 seconds). In the dynamic scenarios participants could interact with the traffic and provide heading clearances. The short scenarios were variations of the five learning goals with each a single conflict between two or three aircraft. The longer scenarios were a compilation of at least seven consecutive conflicts, all with different geometries. In these scenarios, multiple 'noise' aircraft were present in the sector to distract the participant, increase complexity and force the participant to consider multiple solutions when a conflict situation presented itself. A new conflict would present itself (turn amber) at least 120 seconds after its previous conflict.

D. Procedure

The experiment entailed two separate sessions, with exactly one day in between and at different times during the day. The introductory mini lecture on day 1 was followed by practice exercises, which increased in difficulty. Day 2 started with a short recap exercise, several training runs with a legacy PVD, and concluded with a measurement session and a questionnaire. The group that trained with the SSD did the last ten exercises without the SSD in order to get accustomed to the legacy interface. All participants were asked to complete a final retrospective questionnaire to gain insights in their strategies, what they found hard and/or easy and concluded with a few general questions on how they had experienced the experiment.

E. Control variables

In order to boost the signal in the experiment data and to limit the number of potential confounding factors, the following elements were kept constant during the entire experiment:

- **Flight level**: All traffic was limited to the two-dimensional horizontal plane on flight level 290.
- **Aircraft type**: all aircraft were of the same type featuring a speed envelope ranging from 150 kts to 290 kts and a fixed turning performance.
- **Instructions**: the amount and type of instructions was kept equal for both participant groups, all except for a short explanation on how the visualizations on the SSD were constructed. Both groups were given the same explanations on why each of the five 'best solutions' was preferred. The cues directly visualized in the SSD were explained as 'extra tips' to participants who trained without the SSD.
- **Sector layout**: sector size and shape (squared 50x50 nautical miles area) and waypoint names were constant throughout the experiment.
- **Procedure**: exactly one day was in between the two training days. The timing of the training sessions was varied over the two days, to counterbalance the effect of 'time of day' on their control performance.
- **Feedback**: the feedback for each scenario consisted out of a set of pre-prompt options, depending on how the conflict had been resolved. This was done such that participants would not get significantly unequal explanatory feedback. Feedback would for instance consist of the 'best solution', and why this solution was best even if this was the same as the choice of the participant.

Only in the measurement phase, the following measures were also kept constant:

- **Number of aircraft**: the number of aircraft in the traffic scenarios in the measurement phase was always three or more. In the short scenarios, three aircraft were always present in the airspace and during the long scenarios even more aircraft were on screen, although not all aircraft influenced the conflicting aircraft. The conflict to be solved in the traffic scenarios was always a conflict between two aircraft, and the aircraft in the neighborhood would either affect the solution, or not. The idea behind including a third aircraft was to encourage participants to scan the airspace to double check a solution before executing it.
- **Scenario ordering**: the short dynamic scenarios presented the five conflict types in a random order, and the long scenario was a compilation of eight conflicts in a constant order.

Note that these control variables have come at the cost of a less realistic simulation environment, but as the participants were novices with hardly any ATC experience, this difference was believed not likely to negatively affect their performance or motivation.

F. Dependent Variables

The experiment collected dichotomous performance data during each scenario for three choices made: (1) correct/incorrect conflict recognition, (2) correct/incorrect choice of aircraft and (3) correct/incorrect choice of direction of the solution for the conflict. As the solutions to each conflict problem were pre-defined, simple yes/no answers were noted and cumulative error percentages could be calculated for each participant group. Also, the response times of these three decisions were recorded in seconds. In case the recognition of a scenario or the choice for a solution was altered, these would be recorded separately as well. The response time after an initial action was then noted such that in this way the 'penalty' time of the first incorrect choice was included. Other control performance variables included the number of heading clearances (before and after solving the conflict) and how often a loss of separation occurred.

G. Apparatus

The software used for the ATC simulations featured a Java application. Aircraft were simulated by linear kinematic equations and described by their position coordinates, velocities and heading angles. In order to simulate aircraft turn dynamics, first order transfer functions were used assuming a fixed bank angle of 30 degrees. Simulations ran on a desktop computer with a 30-inch HD display with a resolution of 2560x1600 pixels and a refresh rate of 60 Hz. Interaction with aircraft was done by direct manipulation using a regular computer mouse and keyboard.

H. Hypotheses

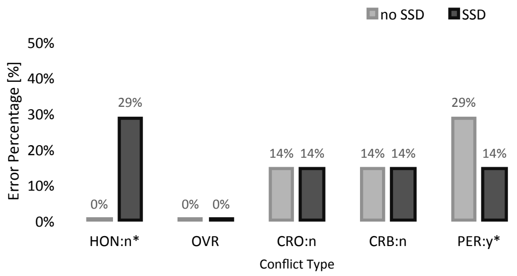
It was hypothesized that training with the SSD would expedite novice learning, eventually resulting in less errors in conflict type recognition, choosing the correct aircraft to solve the conflict and implementing the correct solution. It was also expected that participants in the SSD group would show a reduction in the number of heading clearances before (i.e., evasive maneuver) and after the conflict (i.e., re-align aircraft with exit waypoint), thus improving their decisiveness in terms of control actions.

IV. Results

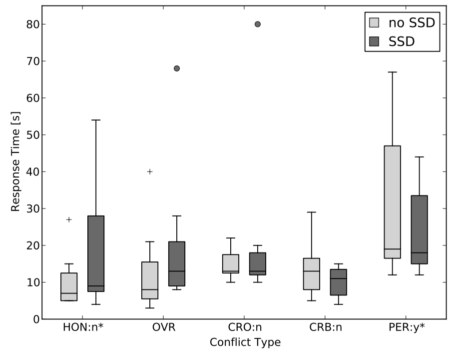
In this section, only the results of the short dynamic scenarios at the end of the second training day are provided and discussed. This session was considered a benchmark for the efficacy of the two training programs and involved control performances with only the legacy PVD. The scenarios in this session featured a subset of the five conflict types and their possible speed settings (see Table I). It was decided to omit the data analysis from the long scenarios, because the large variability in the evolution of traffic situations (due to various different control actions) made the results very difficult to compare between participants and groups.

Additionally, the data of two participants have been removed from the analysis, because they caused significant outliers and showed deviating behavior. Fortunately, these two participants belonged each to a different group, resulting in two balanced groups of each seven participants.

Finally, given the relatively low sample size for each experimental condition, non-parametric tests were used to compare the control performance between the two participant groups. Kruskal-Wallis and Friedman tests were applied to analyze between- and within-group effects, respectively.



(a) Incorrect conflict recognition



(b) Correct conflict recognition response time

Figure 6. Cumulative error percentages and response times of conflict type recognition. The * symbol indicates the conflict type where the third aircraft required deviation from the best practice and 'y' or 'n' designates the presence of a speed difference 'yes' or 'no.'

A. Conflict recognition

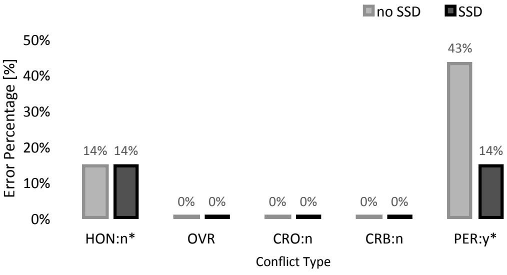
Figure 6 shows the error percentages of conflict type recognition and their corresponding response times. From this figure it is clear that the head on (HON) and perpendicular (PER) scenarios, that both required a departure from the best practice due to the location of the third aircraft, appeared to be relatively difficult to identify. The most remakable result between the two participant groups is that the SSD group had more difficulty to identify the HON conflict than the non-SSD group (higher median and large spread), but for the PER scenario the opposite result was found. This suggests that the SSD had a positive effect on identifying the PER conflict, but a negative effect on identifying the HON conflict. This difference, however, was not marked significant by a Kruskal-Wallis test.

The response times to identify the correct conflict type paints a similar picture. That is, participants in the SSD group took slightly longer to identify the correct conflict type (because more error were made), but the opposite result seems to hold for the PER scenario. Again, Kruskal-Wallis did not report this difference between the two groups to be significant. Friedman test indicated a significant result for conflict type ($\chi^2(4) = 19.476$, $p < 0.01$) across the two groups. Pair-wise comparisons with a Bonferoni correction revealed that the PER scenario had significant longer response times than the HON scenario. The remaining scenarios (OVR, CRO and CRB) were not significantly different in both conflict identification and response times and were relatively easy to spot for both groups.

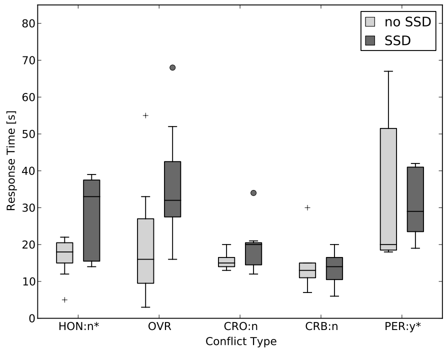
B. Aircraft choice

In Figure 7 the error percentages and response times in aircraft choice are shown. This step occurred after conflict type identification and thus some similarity with Figure 6 is to be expected. The error percentages reveal that HON and PER again resulted in the highest errors in choosing the correct aircraft to solve the conflict. Friedman test did not find this result significant, however. A Kruskal-Wallis test also did not find a significant effect of the participant group on aircraft choice, despite the relatively high error percentage for the non-SSD group in the PER scenario, which suggests that the SSD had the most beneficial effect for the PER conflict.

For the cumulative response times (i.e., absolute time it took to choose the correct aircraft after the conflict was correctly identified), a significant effect was found for conflict type ($\chi^2(4) = 15.928$, $p = 0.03$). Pair-wise comparisons revealed that CRB was significantly different from OVR and PER. Despite that no group effect was found, the response time plot indicates that in the OVR conflict the SSD group shows a trend

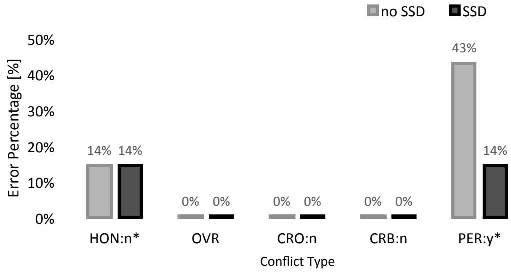


(a) Incorrect aircraft choice

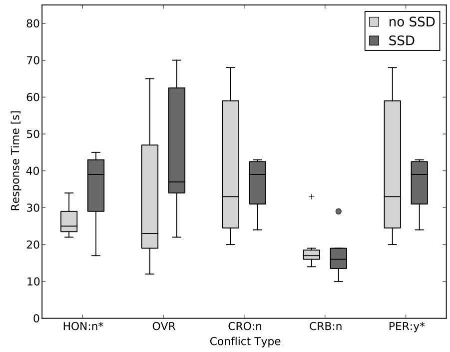


(b) Correct aircraft choice response time

Figure 7. Cumulative error percentages and response times of aircraft choice. The * symbol indicates the conflict type where the third aircraft required deviation from the best practice and 'y' or 'n' designates the presence of a speed difference 'yes' or 'no.'



(a) Incorrect solution choice



(b) Correct solution choice response time

Figure 8. Cumulative error percentages and response times of solution choice. The * symbol indicates the conflict type where the third aircraft required deviation from the best practice and 'y' or 'n' designates the presence of a speed difference 'yes' or 'no.'

toward higher response times. This suggests that participants that trained with the SSD found it slightly more difficult to choose the correct aircraft to solve the conflict.

C. *Solution choice*

The results for the solution choice, the step that occurred after choosing the correct aircraft, are shown in Figure 8. The error percentage is equal to the error percentages in aircraft choice, likewise the lack of non-significant effects for both group and conflict type. More interestingly, the cumulative response times (that now show the total absolute time it took from the start of the scenario toward implementing a conflict resolution) now also show larger spread patterns in the OVR and CRO scenarios, especially for the non-SSD group. However, no significant effect was found for between-group effects.

Friedman test revealed a significant effect for conflict type ($\chi^2(4) = 29.789$, $p < 0.01$), where pair-wise comparisons showed a significant difference between CRO and CRB, OVR and CRB, and PER and CRB. As such, it appears that the CRB scenario was overall the easiest and fasted conflict to detect and resolve, irrespective of participant group.

D. *Heading clearances*

Figure 9 shows the total number of heading clearances given by the participants in the two groups, categorized by conflict type and split by clearances before (i.e., commanding

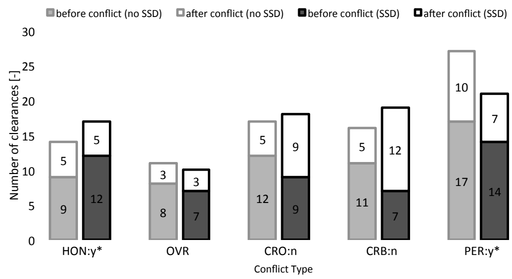


Figure 9. Total number of heading clearances. The * symbol indicates the conflict type where the third aircraft required deviation from the best practice and 'y' or 'n' designates the presence of a speed difference 'yes' or 'no.'

an evasive maneuver) and after (i.e., commanding aircraft back on their target waypoint) the conflict. The minimum total number of heading clearances was 140 (i.e., 14 participants $\times$ 5 conflict types $\times$ 2 vectors (evade and recover)). In total, 170 heading clearances were given, 85 for the SSD group (15 more than minimally required) and 55 for the non-SSD group. The reason for less-than-minimally-required vectors in the non-SSD group is that they sometimes waited longer than the scenario runtime (90 seconds) before steering the aircraft back to its designated waypoint. This also happened for the SSD group in the HON and OVR scenarios.

Statistically, however, the difference between the two participant groups was not significant. Kruskal-Wallis tests reported only a significant group effect in the CRB scenario

($H(1) = 4.396$, $p = 0.036$) regarding heading clearances after solving the conflict. Within the participant groups, Friedman tests indicated a significant effect of conflict type on heading clearances before ($\chi^2(4) = 11.716$, $p = 0.020$) and after ($\chi^2(4) = 10.049$, $p = 0.040$) the conflict. Although Figure 9 suggests that this significance is caused between the OVR and PER conflicts, pair-wise comparisons did not support this observation after taking into account a Bonferoni correction factor.

In general, the overall result runs counter to what was hypothesized. That is, it was expected that the SSD group would show a reduction in heading clearances, but in some cases an increase is observed, especially after the conflict was solved. In the training phase, when the SSD was accessible, participants used the SSD to very precisely vector aircraft just outside the boundaries of conflict zones and could precisely determine the moment to put the aircraft back on its desired course. This 'optimization' strategy was also observed in the measurement phase, but because the SSD was not accessible anymore, participants in the SSD group had a tendency to put the aircraft on their designated course too early, requiring many corrective actions to stay clear of the conflict that was initially solved. The group that trained without the SSD showed less corrective heading adjustments, especially after the conflict was solved. They waited longer before clearing aircraft to their exit waypoints, or not provide a recovery vector at all, resulting in less control actions.

E. *Questionnaire*

After the experiment, all participants completed a questionnaire, containing questions about their general opinion of the experiment and several reflective questions on how they solved conflicts and what strategy they used if any.

Nearly all participants explained to have a very clear strategy that they tried to follow when solving a conflict. Although these differed slightly among participants, the general idea was starting with an analysis of the situation by scanning the airspace, focussing on the exit waypoints of aircraft and aircraft speeds. This was followed by predicting upcoming conflicts (i.e., estimating their crossing points) and identifying the conflict type. By considering the learned 'best solution' first and checking if this would create new conflicts or not, a plan could be made to execute the best practice solution, or formulate a new solution. Before execution, several participants from the SSD group noted that they would double check whether their plan was safe or whether a better solution existed. But overall, the majority of the participants devised fairly similar strategies.

V. Discussion

The work described in this article set out to empirically investigate the training capabilities of an ecological interface for conflict detection and resolution (i.e., the SSD) within an air traffic control context. A comparison study has been conducted between two alternative training programs, one with and one without the SSD as additional support, after which all participants fell back to a concentional radar display to test their gained 'deeper' knowledge and skills. Although the majority of the results are inconclusive due to the lack of statistical significance, interesting trends indicate that the addition of the SSD in training did have some effect.

In general, it seems that the SSD was most helpful for recognizing and correctly solving particular conflicts (i.e., perpendicular conflict geometry that required deviation from the best practice). For other conflict types, such as the head-on conflict, it actually seemed to diminish control performance. There is, however, no clear explanation for this result. Further, it was also observed that the decision-making behavior of the participants in the SSD group involved more critical reflections on proposed solutions, thereby increasing their response times.

But beyond discussions on statistical significance, perhaps a more fundamental question is: does training with an ecological interface lead to desirable operator behavior? In ATC, an important trait of an expert controller is decisiveness, but the experiment results revealed that participants in the SSD group were less decisive, indicated by delayed response times and more reflective behavior. On the one hand, this could be interpreted as the SSD being unfit for training novices in becoming the type of control experts demanded for ATC. On the other hand, however, the emphasis of the experiment was put on *training* and *learning*, in which critical reflection is believed to play a crucial role in bridging the gap between abstract theory and practice [25]. As such, it can be argued that participants in the SSD group were put on a sound learning path from the beginning, but that the duration of the experiment per participant (i.e., just two days) was simply too short to notice significant effects in control performance.

VI. Conclusion

This paper investigated the effects of training a group of novices (in air traffic conflict detection and resolution) with an ecological interface (i.e., the Solution Space Diagram (SSD)) and compare that with a group that only received instructions. A human-in-the-loop experiment was conducted wherein two participant groups underwent a specially designed two-day training program. In the final measurement scenarios, both groups reverted back to a conventional interface. Results from the experiment revealed that the overall control performance between the ecological and the instructional group was not significantly different. In terms of decision-making behavior, the ecological group exhibited more reflective behavior, leading to delayed actions and decreased decisiveness, but occasionally allowed participants to solve 'novel' conflict scenarios. Further research, featuring a larger sample size and longer training session, is needed to further investigate the viability and effectiveness of training controllers with ecological decision-support tools.

References

[1] T. Prevot, J. R. Homola, L. H. Martin, J. S. Mercer, and C. D. Cabrall, "Toward Automated Air Traffic Control – Investigating a Fundamental Paradigm Shift in Human / Systems Interaction," *International Journal of Human-Computer Interaction*, vol. 28, no. 2, pp. 77–98, 2012.

[2] Federal Aviation Administration, "NextGen Implementation Plan," Federal Aviation Administration, Washington, DC, Tech. Rep., 2013.

[3] SESAR Consortium, "European ATM Master Plan," *The Roadmap for Sustainable Air Traffic Management*, no. October, pp. 1–100, 2012.

[4] K.-P. L. Vu, H. Silva, J. Ziccardi, C. A. Morgan, G. Morales, S. Lee, A. Kiken, T. Z. Strybel, and V. Battiste, "How Does Reliance on Automated Tools During Learning Influence Students' Air Traffic Management Skills When the Tools Fail?" *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting*, vol. 56, no. 1, pp. 16–20, 2012.

[5] SESAR Joint Undertaking, "SESAR 2020 Exploratory Research: First Call for Research Projects," SESAR, March 2015. [Online]. Available: http://www.sesarju.eu/sites/default/files/documents/procurements/h2020-call-doc-er-sesar-ju_en.pdf

[6] C. E. Billings, *Aviation Automation: The Search for A Human-centered Approach*, 1st ed. CRC Press, 1996.

[7] R. N. Charette, "Automated to Death," IEEE Spectrum, December 2009. [Online]. Available: http://spectrum.ieee.org/

[8] N. Carr, *The Glass Cage: Automation and Us*. W.W. Norton & Company, 2014.

[9] C. Dow and J. Histon, "Review of Controller Training Requirements for Advanced Air Traffic Control Environments," in *Canadian Aeronautics and Space Institute (CASI) Annual General Meeting*, Toronto, Ontario Canada, 2012, pp. 1–9.

[10] M. Schuver-van Blanken, "Clarifying Cognitive Complexity and Controller Strategies in Disturbed Inbound Peak ATC Operations," in *Advances in Aviation Psychology*, 1st ed., M. A. Vidulich, P. S. Tsang, and J. M. Flach, Eds. Burlington: Ashgate Publishing Ltd, 2014, ch. 6, pp. 85–102.

[11] K. Christoffersen and D. D. Woods, "How to make automated systems team players," in *Advances in Human Performance and Cognitive Engineering Research*, 2nd ed., E. Salas, Ed. Bradford, England: Emerald Group Publishing Limited, 2002, vol. 2, ch. 1, pp. 1–12.

[12] 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.

[13] C. Borst, J. M. Flach, and J. Ellerbroek, "Beyond Ecological Interface Design: Lessons From Concerns and Misconceptions," *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 2, pp. 164–175, 2015.

[14] "A Longitudinal Study of the Effects of Ecological Interface Design on Deep Knowledge," *International Journal of Human-Computer Studies*, vol. 48, pp. 729–762, 1998.

[15] S. B. J. van Dam, M. Mulder, and M. M. van Paassen, "Ecological interface design of a tactical airborne separation assistance tool," *IEEE Transactions on Systems, Man and Cybernetics, Part A*, vol. 38, no. 6, pp. 1221–1233, 2008.

[16] J. Ellerbroek, K. C. R. Brantegem, M. M. van Paassen, and M. Mulder, "Design of a Co-Planar Airborne Separation Display," *IEEE Transactions on Human-Machine Systems*, vol. 43, no. 3, pp. 277–289, 2013.

[17] K. Christoffersen, C. N. Hunter, and K. J. Vicente, "A Longitudinal Study of the Effects of Ecological Interface Design on Skill Acquisition," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 38, no. 3, pp. 523–541, sep 1996.

[18] Federal Aviation Administration, "Review and Evaluation of Air Traffic Controller Training at the FAA Academy," U.S. Department of Transportation, Tech. Rep., 2013.

[19] J. D'Arcy and P. Della Rocco, "Air Traffic Control Specialist Decision Making and Strategic Planning-A Field Survey," Federal Aviation Administration, DOT/FAA/CT-TN01/05, Washington D.C., Tech. Rep., 2001.

[20] S. Loft, S. Bolland, M. S. Humphreys, and A. Neal, "A theory and model of conflict detection in air traffic control: incorporating environmental constraints." *Journal of Experimental Psychology: Applied*, vol. 15, no. 2, pp. 106–124, jun 2009.

[21] E. M. Rantanen and A. Nunes, "Hierarchical Conflict Detection in Air Traffic Control Hierarchical Conflict Detection in Air Traffic Control," *The International Journal of Aviation Psychology*, vol. 15, no. 4, pp. 339–362, 2005.

[22] E. M. Rantanen and C. D. Wickens, "Conflict Resolution Maneuvers in Air Traffic Control: Investigation of Operational Data," *The International Journal of Aviation Psychology*, vol. 22, no. 3, pp. 266–281, jul 2012.

[23] B. Kirwan and M. Flynn, "Investigating Air Traffic Controller Conflict Resolution Strategies," EUROCONTROL, Brussels, Belgium, Rep. ASA.01.CORA.2.DEL04-B.RS, March 2002.

[24] S. Fothergill and A. Neal, "Conflict-Resolution Heuristics for En Route Air Traffic Management," in *Proceedings of the Human Factors and Ergonomics Society 57th Annual Meeting*, vol. 57, no. 1. SAGE Publications, September 2013, pp. 71–75.

[25] A. S. Helsdingen, K. van den Bosch, T. van Gog, and J. J. G. van Merriënboer, "The Effects of Critical Thinking Instruction on Training Complex Decision Making," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 52, no. 4, pp. 537–545, 2010.