Faculty of Aerospace Engineering

Supporting Discovery Learning in Air Traffic Control through Ecological Interfaces

Master of Science Thesis

J. Stoof

25 February 2020

TU Delft Delft University of Technology

Challenge the future

Supporting Discovery Learning in Air Traffic Control through Ecological Interfaces

Master of Science Thesis

For obtaining the degree of Master of Science in Aerospace Engineering
at Delft University of Technology

J. Stoof

25 February 2020

Faculty of Aerospace Engineering · Delft University of Technology

TU Delft

Delft University of Technology

Delft University Of Technology
Department Of
Control and Simulation

The undersigned hereby certify that they have read and recommend to the Faculty of Aerospace Engineering for acceptance a thesis entitled **"Supporting Discovery Learning in Air Traffic Control through Ecological Interfaces"** by **J. Stoof** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: 25 February 2020

Professor:

prof.dr.ir. M. Mulder

Supervisor:

dr.ir. C. Borst

Readers:

dr. M.A. Mitici

Preface

Before the reader lies the product of my Master Thesis research project that concludes seven and a half years of studying at Delft University of Technology. This thesis was written in order to graduate at the Control & Simulation section of the Faculty of Aerospace Engineering and presents the most important elements and findings of my research project. The report starts with the scientific paper that presents the research project and the results. Additionally, the preliminary research has been included, followed by the book of appendices for further background information, support and clarification.

The past months (and years) have been an interesting, incredible and unforgettable journey and I can think back to the many memories made with a smile on my face. Along this journey, many friends, family and especially, my supervisors have supported and guided me, and thereby helped me to get the most out of this research project and to obtain the results that are presented in this report. I would therefore first like to acknowledge my supervisors, as I am very grateful for all their advice, knowledge and guidance. First, thanks go out to my daily supervisor Clark Borst, for always being available for questions and discussions (and for being such an easy target at archery tag). The weekly meetings were always very insightful and each time resulted in many more new ideas. Although the never-ending stream of ideas and feedback sometimes drove me crazy while trying to reduce the scope, it more often drove my research to greater excellence. Secondly, I would like to thank Max Mulder for taking time for me and for students in general despite a busy schedule and for the sharp questions and ideas that further strengthened the foundation of my research. I would furthermore like to express my gratitude to the training specialists and coaches of the LVNL, who provided me with ample amounts of information, and to everyone that participated in my experiment. The results presented in this thesis could not have been established without them and I would therefore like to express my gratitude for their time, dedication and enthusiasm.

My graduation and time at the faculty would not have been the same without the support and love of my friends and family. I am extremely grateful to my parents for their love and support and for making it possible for me to study at the TU Delft. Additionally, thanks go out to my brother and sister, who both actually contributed to this thesis despite the topic being completely out of their field of study. I would like to thank my fellow students from room SIM 0.08 for the support and fun times during and after research hours and my friends from Stabilo, as together with everyone from SIM 0.08 you made my time at C&S feel like being part of the C&S family and being at the faculty feel like home. It was a great time!

Jennifer Stoof
Delft University of Technology
February 25th, 2020

Contents

Acronyms

ACC	Area Control
ACoPOS	ATC Cognitive Process & Operational Situation
ACS	Area Control Surveillance
AH	Abstraction Hierarchy
APP	Approach Departure Control
ATC	Air Traffic Control
ATCO	Air Traffic Control Officer
ATM	Air Traffic Management
BPS	Basic Practical Skill
CD&R	Conflict Detection & Resolution
COPX	Cleared Sector Exit Point
CPA	Closest Point of Approach
CTA	Control Area
CTR	Control Zone
CWA	Cognitive Work Analysis
DCT	Direct To
DL	Decision Ladder
ECOM	Extended Control Model
EID	Ecological Interface Design
FAA	Federal Aviation Administration
FAB	Functional Airspace Block
FBZ	Forbidden Beam Zone
FIR	Flight Information Region
FL	Flight Level
HB	Heading Band
HDG	Heading
IAS	Indicated Airspeed
ICAO	International Civil Aviation Organization
KBB	Knowledge-Based Behavior
LoS	Loss of Separation
LVNL	Luchtverkeersleiding Nederland
MTCD	Medium-Term Conflict Detection

NextGen	Next Generation Air Transportation System
OJT	On-the-Job Training
PVD	Plan View Display
PZ	Protected Zone
RBB	Rule-Based Behavior
RMS	Root Mean Square
SA	Situation Awareness
SBB	Skill-Based Behavior
SESAR	Single European Sky ATM Research
SID	Standard Instrument Departure
SPD	Speed
SRK	Skills, Rules and Knowledge
SSD	Solution Space Diagram
STAR	Standard Terminal Arrival Route
STCA	Short-Term Conflict Alert
SUA	Special Use Airspace
TMA	Terminal Control Area
TOC	Transfer of Control
TWR	Tower
UAC	Upper Airspace Control
UTA	Upper Control Area
ZPD	Zone of Proximal Development

List of Figures

List of Tables

Part I

Master of Science Thesis Paper

Supporting Discovery Learning in Air Traffic Control through Ecological Interfaces

Jennifer Stoof*
Supervisors: Clark Borst† and Max Mulder‡

Abstract—**In an effort to further evaluate the effects of Ecological Interface Design on acquiring expertise, this paper presents the results of an investigation into how Ecological Interface Design could possibly promote discovery learning processes in the Air Traffic Control domain through the application of an instructional design model, and how it could thereby support novices in picking up an implicit rule of thumb. A between-participants experiment (N = 28) was performed through which participants, divided over two groups, were guided by means of a training script developed according to the Four-Component Instructional Design model. Participants from both groups had to perform a simplified Conflict Detection & Resolution task in which the rule of thumb was implicitly integrated. In addition, one group received additional support from an ecologically-designed cognitive tool that was unexpectedly removed after a transfer manipulation. Results show that applying an instructional design method to the experiment design indeed leads to a more structured learning process and participants making significantly more use of optimal strategies, corresponding to the execution of the rule of thumb. Additionally, being trained with the cognitive tool leads to an increased awareness of the implicit rule of thumb but does not lead to an increased use of this rule. It was furthermore found that participants, who were trained with the tool, showed increased dependency on this tool. It is therefore recommended to further extend the experiment design and investigate whether including additional instructional design elements, such as phased visual elimination of the elements of the cognitive tool, could reduce participants' dependency.**

Index Terms—**Air Traffic Control, Ecological Interface Design, Solution Space Diagram, Air Traffic Control Training, Conflict Detection & Resolution, Instructional Design, Discovery Learning, Cognitive Tools, Air Traffic Control Strategies, Competencies, Learning Tasks, Scaffolding, Transfer of Learning.**

I. Introduction

THE mission statement, to ensure a safe, orderly and expeditious flow of air traffic, and the means with which Air Traffic Control Officers (ATCOs) ensure the safe separation of aircraft, have remained largely unchanged over the past decades [1]. However, with today's rapidly advancing technologies, the annually increasing number of flights as well as the pressure from environmental organizations, a shift in Air Traffic Control (ATC) and a modernization of the systems used by ATCOs is inevitable, requiring controllers to work in a more complex environment [2]. In order to facilitate this shift, the capacity of the current system will have to be increased by moving toward higher levels of automation, meaning that changes in the tools and procedures nowadays used in ATC are unavoidable [3], [4].

Several challenges arise with the introduction of increasing levels of automation in a system or a domain, also described by the ironies of automation [5]. Although automation generally improves safety during normal operations, over-reliance on automation as well as deteriorated knowledge and skills of controllers might actually result in an unsafe situation during abnormal operating conditions [6]. The human controller should furthermore be kept actively aware of system performance and is likely to take on a more supervising role as the level of autonomy and authority of the systems increases with increased levels of automation: controllers will only be intervening in case of unexpected situations [7]. As humans are known to be very creative and flexible in unanticipated or unexpected situations (e.g., automation failures), they should in such a case thus be provided with 'right-time' access to the appropriate information in a format that supports the controller in quickly and accurately assessing the situation and effectively stepping in when necessary.

To make optimal use of this creativity and flexibility, the design of the human-machine interface and the presentation of information is of critical importance. Ecological Interface Design (EID) is a design philosophy that focuses on making constraints and relationships in the work domain visible to the controller and thereby improve the controller's deeper understanding of the work domain [8], [9]. Additionally, by offering support in adapting to change and novelty through these visual constraints and relationships, EID allows controllers to limit their core activities to higher-order problem-solving and decision-making [8]. While EID thus supports expert controllers in assessing (unexpected) situations and gaining a deeper understanding of the work domain, it can be argued that these same aspects of EID can aid in familiarizing novice controllers with a new work domain [10]. However, when using technology to familiarize novice controllers with a new work domain or with certain aspects of this domain, the risk exists of their knowledge becoming dependent on the availability of this technology [7], [11]. The question that could thus be raised is whether EID could be used during training and could contribute to building expertise of novice controllers or whether this expertise becomes dependent on the technology or interface?

Within process control, it has previously been shown that EID can indeed lead to better knowledge development, a functionally-organized knowledge base and increased performance after an exposure of six months [11]. However, little

* Faculty of Aerospace Engineering, Delft University of Technology, Email: jenniferstoof@hotmail.com,
† Faculty of Aerospace Engineering, Delft University of Technology,
‡ Faculty of Aerospace Engineering, Delft University of Technology.

research has yet been conducted on training with EID tools in the ATC work domain [12]. Research on short-term effects of EID on knowledge development within this domain suggests that EID encourages Knowledge-Based Behavior (KBB) and generates goal-oriented thoughts and can therefore play an important role in the early stages of knowledge development [10]. However, this study also showed several limitations.

First, all participants of both study groups received an initial training to which no instructional design model was applied. This caused four different learning curves to be intertwined during the study. Participants were (1) still familiarizing themselves with the simulation environment and (2) learning to work with the newly introduced ecologically-designed decision-support tool, the Solution Space Diagram (SSD) [13], while (3) also already having to perform the control task and (4) simultaneously (try to) think out loud, which altogether may have negatively influenced the results. Second, none of these learning curves or processes were monitored or tracked throughout the experiment. This made it very difficult to determine each participant's level of knowledge prior to the measurement phase and whether the participants had already reached a learning plateau or if they were still in the process of learning. Third, although participants were required to engage in discovery learning and to learn to detect and resolve conflicts by means of so-called 'best practices', they were already explained these 'best practices' or rules of thumb for solving different conflict situations during an initial information session. While this was supposed to give participants of both study groups a head-start in the knowledge development, the effect of the interface on the knowledge development became less salient and thus the added value of the interface became less evident. The research presented in this paper will be an effort to further evaluate the short-term impact of ecological interfaces on performance and knowledge development of novices being engaged in discovery learning in the ATC domain and to overcome the above-mentioned limitations.

Although the SSD is argued to be useful for shaping the operator's mental model and hence supporting a deeper understanding of the system and developing expert-like behavior [10], [13], a risk of this interface is that surface learning occurs which leads to the development of shallow knowledge in case the interface is used as a Rule-Based tool only [11]. Furthermore, a larger dependency on the interface might be developed when the interface is used in such a way. Designing the training task according to an instructional design method might, however, help to structure the mental model and manage the cognitive load during the training [14]. First, it is expected that this could reduce the occurrence of surface learning and could therefore contribute to a deeper knowledge of the system as well as more knowledge-based problem-solving. Second, it is expected that applying such a method to the learning process will result in the sequential occurrence of the above-mentioned learning curves rather than a mixed or simultaneous occurrence.

This paper presents a human-in-the-loop experiment, designed according to the Four-Component Instructional Design method [15], that represents a small-scale Conflict Detection & Resolution (CD&R) course in which novice participants are trained with the SSD. The setup of the experiment is similar to the study of Borst et al. [10]: the knowledge and performance of the two groups of participants is tested and compared after a transfer manipulation where the SSD support will unexpectedly be removed. One group, the focus or SSD group, will thus be trained with the SSD, whereas the control group will be trained without this tool. Several adaptations are, however, made to the experiment setup.

It was previously found that discovery learning is most successful when students have the right prerequisite knowledge and undergo structured experiences [16]. First, our participants will thus be guided through the experiment by means of a step-by-step script, containing background information, exercises or tasks as well as questions. This script and the overall experiment set-up are designed according to the Four-Component Instructional Design model and are meant to not only provide participants with all the required background knowledge, but also to structure and separate the learning curves mentioned before. The set-up of the script can additionally be used to monitor participants' learning curves and test whether these learning curves have actually been separated. Second, the experiment will focus on only one conflict geometry with one 'best practice' or rule of thumb that is implicitly integrated into this conflict geometry by means of three levels of salience. Participants will not receive any information about this rule of thumb prior to or during the experiment but are required to obtain this knowledge by means of discovery learning in order to increase the salience of the effect of the SSD on the performance and knowledge development. As discovery learning and EID both promote problem-solving and facilitate lateral thinking, it is expected that structuring the learning process in this way and combining this with (implicit) structured information from the SSD about the rule of thumb will lead to more participants developing the right solution strategies. Finally, next to the performance development, strategy development will thus also be evaluated, as this provides information on whether the transfer of learning (i.e., obtaining the required knowledge about the rule of thumb) has indeed been successful.

By investigating how EID could possibly promote or support discovery learning processes through the application of an instructional design method, this research can provide new empirical insights into the benefits of ecological interfaces. It furthermore provides insight into whether such interfaces can be used to teach students rules of thumb or 'best practices' and how they facilitate further knowledge and performance development.

This paper is organized as follows. First, background information about ATC, EID and learning methodologies is given in Section II. Section III elaborates on how this information has been integrated in the step-by-step script that guided participants through the experiment. The experiment design is then presented in Section IV. Section V presents the results of the experiment. The paper ends with a discussion and conclusions.

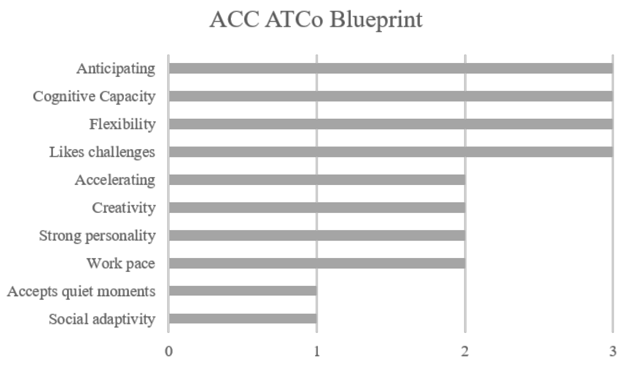


Fig. 1. Blueprint displaying the most important competencies for Air Traffic Control Officers within Area Control. [1]

II. Background

A. Structure of Air Traffic Control Training

Since ATCs is subject to very strict safety regulations, there is little to no room for (human) errors or incompetence and thus performance standards for the ATC task are high. The task is considered to be both dynamic and highly complex as it requires processing of large amounts of constantly changing information [17], [18]. Several competencies have been defined to determine whether someone is capable of conducting the ATC task [18]. When a competency is successfully obtained, it allows people to come up with solutions for new and unexpected events or situations.

Figure 1 shows to what extent a controller within Area Control (ACC) should possess certain competencies. Cognitive capacity, flexibility and being able to anticipate are especially important as controllers need to learn to deal with unexpected situations in an effective manner while ensuring safety. They need to be able to anticipate these situations, come up with plans to solve them and be flexible in which solution they choose and how it affects their other plans and strategies. While some competencies shown here might seem less important for ACC ATCOs, they can be quite important for ATCOs of other units. An example is the competency where ATCOs need to be able to accelerate or change the pace of their work. Within ACC, the airspace contains relatively similar amounts of air traffic most of the time, whereas ATCOs stationed in, for example, the tower at Groningen Airport Eelde generally experience large differences in the amount of traffic throughout the day. They thus need to be able to quickly accelerate for traffic peaks after (long) periods with relatively little air traffic.

Because of the complex nature of the task, there is only a minority of people that is able to acquire the competencies within the predetermined training period (approximately 3 years, depending on the student and the unit). The workload during ATC training is considered to be high and the learning curve that is expected of students is steep. This leads to many students ending their training prematurely (often already quite far in the program) as they are not able to meet the high standards set for the training [19], [20].

Nowadays, the ATC training program consists of several phases. All students start by learning basic practical skills, such as how to use the ATC equipment and how to handle flight strips in the right way. Furthermore, students take several theoretical courses about subjects such as meteorology, aircraft mechanics and performance, radio communication, aircraft recognition, air traffic law, equipment, human factors and navigation. After the basic training, students are admitted to a unit, based on their own preferences and how well their skills/competencies match the competencies that are required for that particular unit, such as the competencies that were defined for ACC in Figure 1.

After being admitted to a unit, students start training in a basic simulator and follow unit-specific exercises during the so-called Initial Training. The Initial Training consists of sets of exercises related to inbound flights, outbound flights, neighboring fields and a consolidation phase. There is no specific set of exercises dedicated to only CD&R, but CD&R is instead integrated in each of the sets mentioned above. For every set, complexity is increased throughout the exercises, that generally take about 25-45 minutes. Although each exercise might focus on a different subject or part of ATC, the goal of all exercises is to maintain a safe, orderly and expeditious flow of air traffic. It is furthermore expected that the student indicates the competencies or skills he/she likes to improve or work on during an exercise, as discovery learning and self-reflection are considered important aspects during the training.

Once the student successfully finishes the Initial Training, the student obtains his/her Student Controller License and is allowed to move on to the next phase of training: Unit Training. During this phase, students are taught subjects such as air structure, classification, aircraft recognition and routes specific for that unit and are further familiarized with all rules, regulations, procedures and protocols specific to the sectors controlled by this unit. The final stage of the ATC training is the On-the-Job Training. During the On-the-Job Training (OJT), the student will follow and watch professional ATCOs perform their job but will also get the chance to work as an ATCO in that specific sector while still being guided by professional ATCOs.

B. Control Strategies in Air Traffic Control Training

As mentioned before, the main focus during the ATC training is put on expediting traffic in the safest and most efficient way. During the training, students are taught several strategies to accomplish this goal, such as perception, interpretation, anticipation, workload management and planning strategies, but also strategies to detect and resolve conflicts. Which strategy is used by a controller usually depends on the characteristics of a situation as well as operational constraints. To develop more robust knowledge of the system, students are discouraged from trying to develop a solve-all strategy as having a range of strategies available generally reduces the risk of performance being compromised during disturbed operational situations [21]. By introducing unexpected and unfamiliar events in the training scenarios and assignments, the students' versatility and resourcefulness are trained and knowledge-based problem-solving is encouraged.

Although strategies are crucial in ATC performance and appear to be a key element to success during the ATC training [21], they are not specifically taught to students during training, as was confirmed by anecdotal evidence from a training coach during a visit to the Luchtverkeersleiding Nederland (LVNL) on April 24th, 2019. ATCOs usually obtain the required knowledge by means of discovery learning. Additionally, coaches might give ad-hoc input and steer the students towards the use of certain strategies, depending on the scenarios and their complexity. This does, however, cause the development of these strategies and the competencies that are associated with these strategies, to be influenced by the personal preferences of instructors.

Little is known about the ways ATCOs are taught to detect and resolve conflicts. Guidelines are provided by the International Civil Aviation Organization (ICAO) (Doc 10056, Annex 2, Doc 4444 [22]–[24]), but no set lesson plan exists on how to solve conflicts in an effective and expeditious manner. Research has shown, however, that there are three factors that are of influence to the decision-making strategies of ATCOs [25]. These three factors are expediency, preservation of airspace structure and visualization: maneuvers that resolve conflicts more rapidly, are least disruptive to the overall traffic flow and can be perceived more rapidly on the radar screen, are preferred as such maneuvers generally require less monitoring time [25]. For expediency as well as for preservation of airspace structure, vertical separation is preferred as this requires the least amount of attention. For visualizations and conflict resolutions, however, lateral movements are preferred as the effect is immediately visible to the controller on the radar screen [25].

C. Assessment in Air Traffic Control Training

The time-critical, complex and dynamic nature of the ATC task, as well as the fact that cognitive skills (e.g., information processing) cannot be measured or observed directly and can be influenced by external factors, make it difficult to develop a reliable objective assessment system for the skills and competencies that are important for ATCOs. Several models have been developed to map the cognitive processes taking place inside an ATCO's head. The ATC Performance Model, shown in Figure 2, is a model that describes the competencies that are of importance to ATCOs and serves as a framework for the identification and design of performance criteria, which are used in the competence-based assessment of ATCOs [18].

The ATC Performance Model roughly consists of four parts: information processing, actions, influencing factors and outcome. The competencies related to information processing are split into different categories or skills and thereby show the dominant role of the information-processing component within the ATC task [18]. Important for the cognitive process, or the processing of information, are competencies such as

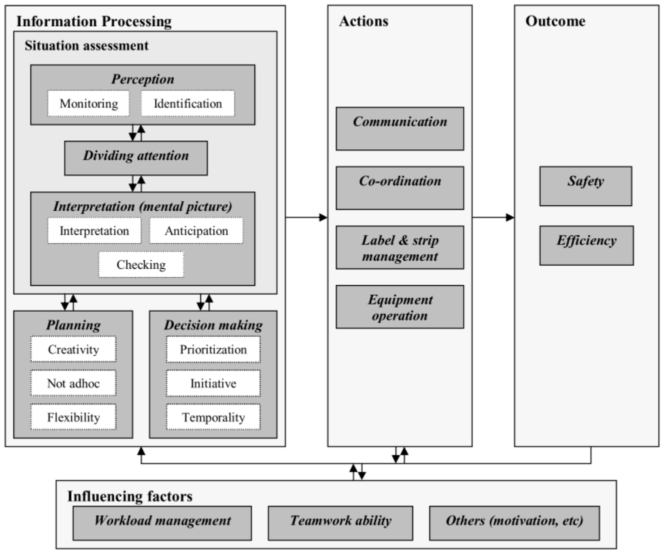


Fig. 2. The ATC Performance Model showing the competencies that were identified as important for ATCOs and are used as a basis for the competence-based assessment for ATCOs [18].

situation assessment, planning and decision-making. The situation assessment component can be divided further into the elements of perception, dividing attention and interpretation, similar to the three levels of the Situation Awareness theory [26]. Information processing subsequently forms the basis for the actions. Both information processing and executing tasks can be influenced by external (e.g., personal) factors, such as dealing with the workload, teamwork and attitude, whereas actions can in turn also influence the influencing factors; for example, when label and strip management is not performed well, it could influence the experienced workload. These three components then lead to the outcome, which in turn corresponds to the mission statement and highest goal of ATC: safely and efficiently organizing and expediting the flow of air traffic.

Next to providing information about the competencies that are most important to ATCOs, the model provides information on how the competencies can be assessed [18]. It forms the basis for the performance criteria in the competence-based assessment of ATCOs during the training and separates criteria that can be measured subjectively and objectively. The objective criteria are the competencies and criteria that are part of the outcome and action blocks, whereas the competencies and criteria belonging to invisible information processing activities can mostly be measured subjectively. Assessment of ATCOs during training is currently done by performing so-called over-the-shoulder observations in either simulations or during the OJT and rating controllers' performance on a 6-point rating scale. In general, it holds that the better the performance of a student for a specific competency is, the more often the student shows the required type of behavior and the more often the student recognizes and corrects any present errors. Because the competencies can for a large part be observed through behavior, the assessment or performance criteria are usually formulated as behavior descriptions or behavioral markers.

Observations and performance data alone are, however, usually not enough as the information-processing activities are not visible to the assessors and hence should be complemented with additional information about the thinking patterns and strategies of the students. Methods often applied to obtain this information are think-aloud protocols, critical incident analysis, interviews and re-runs of training scenarios [18]. Combining performance data with information about the thinking patterns of students, obtained from thinking-out-loud protocols as well as over-the-shoulder observations, leads to a more or less complete picture of the skills and competencies of the student during the execution of a certain task [18]. Assessment should furthermore take place at a higher level, as skills and certain pieces of knowledge might be mastered in different orders or in different amounts of time and social, emotional and environmental factors should additionally be taken into account during the assessment [18]. Assessment at higher levels allows for distinguishing learning curves of these skills and parts of knowledge. The over-time-increasing scenario complexity is what furthermore allows coaches to monitor the students' progression and learning curves as the required competencies remain the same during the different training phases. Learning curves are important in monitoring the students' progress: they can serve as an indicator for whether students are still in the process of learning or whether they have already reached a learning plateau [18].

When insight is gained into the students' learning curves of different skills/competencies, appropriate measures can be taken if it turns out a student lacks certain skills or parts of knowledge. Next to the competence-based assessment, continuous assessment is therefore applied during the training phase of ATCOs. The assessors continuously interact with the students during and after the training exercises. During the training, the students are asked questions about their decisions or to probe their situation awareness in order for the assessors to gain insight into the cognitive processes and strategies applied by the students and hence give appropriate feedback on strategies that were well- or ill-chosen. Feedback and self-reflection are considered very important parts of the learning process. Different types of feedback or evaluation forms are used that are usually filled out by the coach and student together, which leads to a better insight into the student's points for improvement and learning progress. A student has multiple coaches during a training phase. By having multiple coaches assess a student, the assessment becomes less subjective and hence more reliable and fair.

D. Theories and Practices in Complex Learning

When students are learning complex tasks, such as the ATC task, they can sometimes be overwhelmed by the amount of information and the complexity of the task. Hence, it is of importance to manage the cognitive load during the learning process by providing the right amount and type of support and guidance that is fully integrated in the task [27]. To increase the chances at a successful transfer of learning of a complex and dynamic task such as the ATC task, a holistic design approach to the learning process is necessary according to instructional design, that does not lose sight of the separate elements and their interrelations, but deals with the system or learning domain as a whole [28]. One such an approach, of which the separate elements can also be observed in the current ATC training, is the Four-Component Instructional Design (4C-ID) model which assumes that blueprints for complex learning can always be described by four components:

- *Learning tasks* should consist of easy-to-difficult task classes, a high variability and a decreasing level of support and guidance when moving through a task sequence in order to increase the chances at a successful transfer of learning [15], [28]. Examples during ATC training are the simulation exercises that increase in complexity throughout the training.
- *Supportive information* explains how a learning domain is organized and is usually presented when students start working on a new task class. It should always be available during that task class in order for students to go back and forth between the task and the information [15], [28]. An example of supportive information that is provided to ATC students during the training is a pre-simulation or pre-OJT briefing.
- *Procedural information* allows students to perform routine aspects of a task [15], [28]. It specifies how to perform these routine aspects during a task and is preferably presented just in time. During the ATC training, students will, for example, learn about new protocols for adjacent or new sectors at some point. These protocols can be provided in the form of procedural information such that students can make just-in-time use of this information.
- When a certain aspect is required to be performed on a highly-automated cognitive level, additional *part-task practice* should be provided (e.g., practicing Radio/Telephony (RT), getting speech therapy or practicing working with the equipment). Part-task practice involves repetition (strengthening) and should only be provided after the aspect has been introduced in the context of the whole task [15], [28].

Initially, the support in a learning task allows a student to perform a task or achieve a goal that would not be achievable without the support. As the student's expertise increases, support is gradually decreased until the student no longer needs the support and is able to perform the task independently [29], [30]. This is also referred to as scaffolding and is based on Vygotsky's concept of the Zone of Proximal Development (ZPD) [27], [29]. Scaffolding support can be provided in many ways and is traditionally used to refer to the process in which a teacher or more knowledgeable person assists the learner in accomplishing a task that would otherwise be out of reach. [31], [32]. Central to this definition is that a second person intervenes the learner at appropriate times and what the learner can actually accomplish by means of these interventions [27], [32], [33]. Within the ATC training, this scaffolding support is thus provided by coaches when they give ad-hoc input or feedback during exercises.

Recent instructional design research, focused on current-day learning environments, has been aimed at applying scaffolding to software [30], [33]. Rather than teachers or peers supporting a learner, computers can support learners by explicitly supporting or representing cognitive processes and changing the task in such a way that learners can accomplish tasks that would otherwise be out of reach [33], [34]. In supporting the cognitive processes required to perform a task, computers or more specifically, cognitive tools, can thus serve as an extension of the mind and extend the limits of the human cognitive capacities [34].

Scaffolding approaches to cognitive tools include scaffolding by computer and a human tutor combined or by means of a fully-embedded cognitive tool in a computer-based learning environment [30]. Cognitive tools that are used in this way can lead to a deeper understanding of the system by actively helping to organize a controller's knowledge and by helping learners to reflect on their own problem-solving processes and skills [35]. Cognitive tools can furthermore bridge the difference between open learning environments, such as the discovery learning environment, and more traditional (expository) learning environments [34].

With discovery learning, learners construct their own knowledge by generating hypotheses and experimenting within a domain and hence, by inferring rules from these experiments [34]. Because the learners are actively constructing their own knowledge, it is assumed that the domain is understood at a higher level than when the required information is considered a mere transfer when it is for example presented by a teacher in a traditional (expository) learning environment [34]. The active involvement of the learner in constructing his/her own knowledge is said to result in a better and more structured knowledge base and thus discovery learning is perceived as a promising way of learning. It should, however, be noted that learners, in general, often require additional information next to solely domain information and the extent to which learners are found to be successful in discovery learning generally depends on a number of discovery skills such as hypothesis generation, experiment design, prediction, data analysis and planning [34]. Because these skills are often required in a complex information society or domain, they are also considered learning goals in itself [34].

Next to providing information about the domain to learners, learners often need assistance in selecting and interpreting the domain information to construct and test hypotheses that contribute to their knowledge base. In order to assist in this process and thereby increase the chances at successful discovery learning, the discovery learning process can be supported. Cognitive tools can in this case serve as a means of support in discovering a domain as they can serve as an extension of the mind and can thereby add to the required discovery skills. When using a cognitive tool for such a purpose, it should be kept in mind that the learner should have sufficient freedom in selecting and interpreting information during the learning process, as a limitation in this freedom goes against the very nature of discovery learning.

E. Solution Space Diagram as Cognitive Tool

The SSD, shown in Figure 3b, is an example of a decision-support tool designed according to EID principles that could be

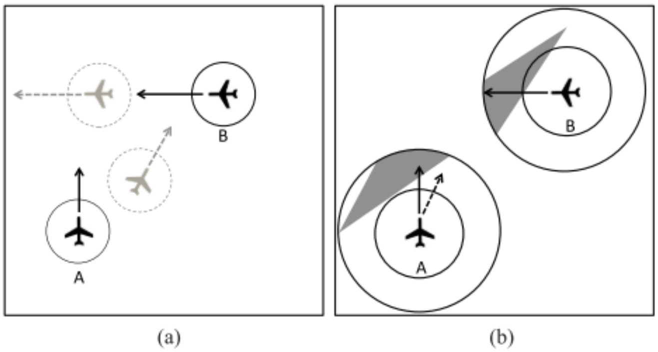


Fig. 3. Schematic of CD&R rule of thumb or "best practice". (a) Conflict situation involving slow aircraft A and fast aircraft B, along with the "best practice" to solve this situation: slower aircraft should be vectored behind faster aircraft in case of conflict. (b) Conflict situation along with the aircraft's SSDs [13], illustrating that the shown solution is indeed a robust solution and that the SSD promotes the 'best practice' [10].

helpful in developing expert-like behavior [10], [13]. The EID principles work especially well with the open and dynamic nature of the ATC environment. The figure shows a conflict involving slow aircraft A and fast aircraft B along with their SSDs. The gray area in Figure 3b represents a conflict in the near future. It can be seen that for both aircraft, the solution space on the aircraft's right-hand side is larger (more white and less gray area). In resolving this conflict, it would thus be most efficient to give either one of the aircraft a small heading deviation to the right such that the speed vector of either of the aircraft is directed out of the conflict zone.

As the SSD accounts for unanticipated and unfamiliar events or situations by providing the controller with the complete range of solutions, an overlap is found between EID and ATC training. Next to this, in order to successfully accomplish the mission statement of ATC, it is of importance that ATCOs are able to perform a correct situation assessment and are aware of the complete situation. As EID contributes to situation awareness, another overlap is found [26], [36]. Referring back to the ATCO blueprint and the skills that are especially important within ACC, the SSD thus matches especially well with supporting these skills. By showing the controller an instant overview of the solution possibilities in terms of heading and speed in the 2D plane, it helps controllers anticipate, increases their flexibility and reduces cognitive load by increasing situation awareness.

An example of an ATC strategy or so-called rule of thumb in resolving conflicts between an aircraft pair in the horizontal plane, is to vector the slower aircraft behind the faster aircraft. Figure 3 shows that this is indeed a robust solution that leads to the smallest track deviation and the least amount of additional monitoring time. The SSD thereby makes the rule of thumb visually salient and allows the controller to evaluate the information that is presented by the SSD about potential conflicts as well as about the rule of thumb. The SSD is thus argued to be useful for shaping the internal mental model and hence, for gaining a deeper understanding of the system [10]. As it helps organizing the controller's knowledge, it can be considered a cognitive tool. An overlap is then found

TABLE I
FOUR-COMPONENT INSTRUCTIONAL DESIGN MODEL TRANSLATED TO THE EXPERIMENT DESIGN

4C-ID Model Component	Experiment Component
Learning Task	Experiment Scenarios
Supportive Information	Training Script
Procedural Information	Experiment Information ('cheat') sheet
Part-task Practice	RT/Think-aloud Protocol

between cognitive tools as a means of support and EID, as both concepts help to organize the controller's knowledge and thereby decrease the controller's cognitive load. Using the SSD as a cognitive tool in an ATC learning task could help to further organize and structure the controller's knowledge and increase his/her performance during training, without negatively influencing the development of deep knowledge.

The means with which ATCOs fulfill their tasks, have remained largely unchanged over the past decades, just as the training of ATCOs has seen little change over the same time span. Although significant effort has already been put into objectifying the training and especially the assessment of the students by developing cognitive models and visualizing thinking patterns, parts of the training and assessment are still based on subjective expert opinions. To objectify the training even more, an ecologically-designed decision-support tool, such as the SSD, could be a promising tool as coaches are then able to use the tool as an objective basis for their feedback and instructions, thereby reducing the influence of personal preferences on the students' strategies and performance.

III. DESIGN OF THE CUSTOM CONFLICT DETECTION AND RESOLUTION TRAINING

As explained in Section II-D, a holistic design approach is necessary to increase the chances at a successful transfer of learning of a complex and dynamic task such as the ATC or CD&R task. In designing the custom CD&R training for the experiment, the instructional design methodologies discussed in Section II-D were taken into consideration as well as the current ATC training to increase the fidelity of the task, and thereby contribute to the holistic view of the participants. The scope of the experiment task was thus designed to resemble the real ATC task as much as possible. As explained before, the Four-Component Instructional Design model assumes that blueprints for complex learning can always be described by four components. The custom-made training has therefore been set up according to these four elements. Table I shows the four components and how each of these components has been translated to the experiment design. The rest of this section further elaborates on the design of the custom CD&R training and how each of the components has been taken into consideration in the design.

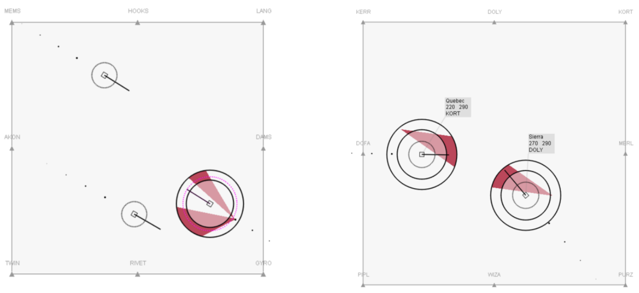
A. Training Script

As stated before, it is important to manage the cognitive load during the learning process by providing fully-integrated support and guidance. The SSD can be considered such fully-integrated support in relation to obtaining the required knowledge about the rule of thumb. However, as that is only one learning process and participants are also required to obtain knowledge about the domain before they can successfully execute their task, it was decided to translate the *supportive information* component, that explains how a learning domain is organized, to a step-by-step script guiding the participants through the experiment. The script, found in Appendix M, was available to the participants throughout the entire experiment, allowing them to go back and forth between the task and the information. It was meant to inform the participants on several elements of the experiment and thereby separate the different learning curves that had been found to be intertwined in previous research, but still take the interrelations into account. The separate elements or scaffolds that the script is built up from, that each have a different learning objective, are described below.

1) Simulation Environment: After globally explaining the goal of the experiment in the introduction of the script, the script first elaborated on the simulation environment. The goal of this chapter was to familiarize participants with the simulation environment, the interaction with aircraft and the standard tools and information available to them, such as the flight labels, the color coding of the aircraft, the number of the current scenario, the total number of scenarios and a 10NM-scale, similar to how ATC trainees start by learning basic practical skills and thereby learn how to use the available equipment.

The first chapter was accompanied by one dynamic scenario. At the start, this scenario was paused allowing participants to go back and forth between the information in the script and on the screen, such that participants were able to observe and interpret the information without time pressure. First, the Plan View Display (PVD) was described to familiarize participants with the simulation environment and the airspace sector. Additionally, the command display was explained in order for participants to be able to interact with aircraft in the simulation environment. After the explanation, the scenario could be started and participants were instructed to perform several commands in order to experience the dynamics of the aircraft (e.g., how long it takes to complete a turn) and the simulation environment in general (e.g., update rate, color coding).

2) Conflict Detection & Resolution and the Solution Space Diagram: The second chapter of the script elaborated on CD&R. Again, the chapter was accompanied by a single scenario that was paused at the start of the chapter in order for participants to first take a careful look at all the elements on the screen that were being described to them. The goal of this chapter was to familiarize the participants with CD&R in general: how is a conflict defined, what is the minimum separation distance and what is a Loss of Separation (LoS). It furthermore explained what a Short-Term Conflict Alert (STCA) is and after starting the scenario, participants got to experience an STCA as they were explicitly told to wait until both the orange and red STCAs appeared, indicating a time to LoS of 40 and 20 seconds, respectively. After they experienced the alerts, participants were instructed to give a heading clearance to one or both aircraft to solve the situation, whichever solution they thought would be best.



(a) (b)

Fig. 4. Excerpt figures from the script illustrating two types of information that can be derived from the SSD. (a) Distance between the controlled and observed aircraft influences the width of the conflict triangle. (b) Larger white area on one side indicates a larger (and more efficient) solution area; tip of speed vector pointing *inside* the red area indicates a conflict.

For the group that would be trained with the SSD, the CD&R chapter contained additional information about the SSD. The construction of the SSD was explained as well as important information that could be derived from the conflict triangles in the SSD about the traffic situation and the solution space. Figure 4a, an excerpt figure from the script, shows the effect of the proximity of neighboring aircraft on the conflict triangles in the SSD. Figure 4b was additionally shown to demonstrate that the solution space on the right of both aircraft was larger (less red and more white area) and hence it would be most efficient to vector either one or both aircraft to the right. The rule of thumb (i.e., to vector slower aircraft behind faster aircraft in case of conflict) was not explicitly mentioned as participants were required to obtain this knowledge by means of discovery learning, nor was the speed difference between the aircraft discussed as this could hint at the correct rule of thumb.

After explaining all information, the participants were instructed to answer several questions to probe whether they fully understood the information about the SSD that was just presented to them. One set of questions showed simple conflict situations such as in Figure 4b where participants had to state whether the two aircraft were in conflict or not. The goal of this set of questions was to test whether the participants understood that they had to look at the tip of the aircraft's speed vector to determine whether an aircraft was in conflict or not. If the vector was directed through the conflict zone but the tip was positioned outside the conflict zone, the aircraft would not be in conflict, which was considered to be potentially confusing to participants in case the information was not properly read. The second set of questions contained two similar questions of which one is presented in Figure 5. The participants had to match a traffic situation to an SSD. The goal of these questions was to test whether the participants correctly understood the information that could be derived from the conflict angles.

The SSD group thus received additional training compared to the control group, but related to the basic understanding of the SSD only. As the extra amount of training time could potentially be a confound, it was decided to provide

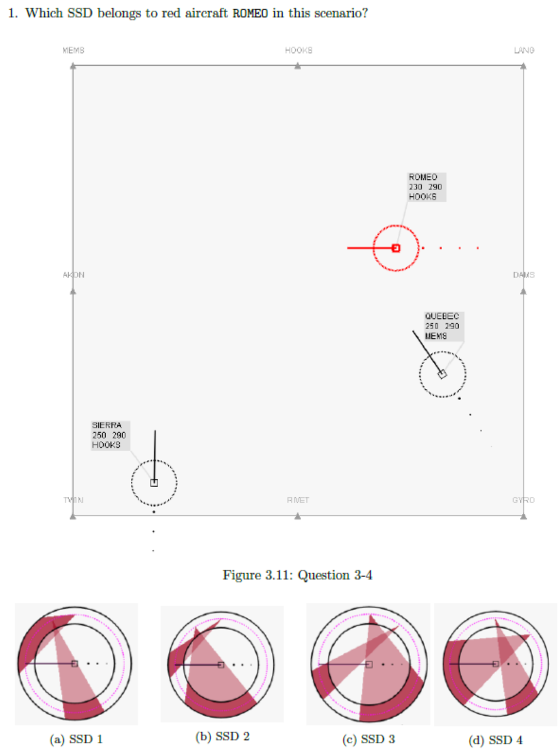


Fig. 5. Excerpt question from the script to test participants' understanding of the SSD: closer proximity results in a larger conflict triangle width and the direction of the observed aircraft's speed vector shows the position of the origin of the conflict triangle in the SSD; hence, B is the correct answer.

this additional training in a written form only and without additional dynamic scenarios, such that the simulated training time would be equal for both groups.

3) Think-aloud Protocol: The third chapter in the script elaborated on the think-aloud protocol and can be considered one of the elements from the Four-Component Instructional Design model, namely the *part-task practice*. Similar to the ATC training, observations and performance data are not enough to be able to assess the participants' performance and thus additional information about their thinking patterns was required. In order to be able to classify participants' strategies, participants were required to think out loud, a method also often applied by assessors during the ATC training. As processing the think-aloud data proved to be very laborious in the study performed by Borst et al. [10], it was decided to structure the think-aloud process by asking participants to follow a predefined protocol, that was a simplified but representative version for RT in current ATC environments.

As this task might (negatively) influence participant performance by taking up too much cognitive capacity, it had to be performed on an automated cognitive level and therefore required additional practice. Participants were first shown what the protocol entailed and were afterwards able to practice the thinking-out-loud during three scenarios. The scenarios were all still paused and had to be started by the participant, allowing them to first inspect the traffic situation. At this point, the control task was not yet explained and participants could thus freely vector the aircraft and practice the thinking-out-loud. After this chapter, additional scenarios were presented to the participants along with the control task so participants could further strengthen their thinking-out-loud 'skill'. As the Four-Component Instructional Design model furthermore states that the part-task practice should only be introduced in the context of the whole task and domain, it was decided to introduce the thinking-out-loud in a separate third chapter, after the simulation environment and CD&R concept had been explained, similar to how RT is given as a separate course during the ATC training. Participants were furthermore explained the basic goal of RT so that they could link this to the ATC training and understand why it was part of the experiment.

For the think-aloud protocol, participants were given several examples on how to mention certain thinking steps out loud. More specifically, they were asked to:

- Name the aircraft they wanted to give a command to by its call sign: e.g., `ALPHA`;
- Name the aircraft involved in the conflict (in case they identified a conflict): e.g., `Conflict ALPHA TANGO`;
- Name the chosen solution as well as the aircraft used to resolve the conflict: e.g., `TANGO Heading 240 (two-forty)`;
- Name the reason for giving a certain command: e.g., `TANGO Heading 240 to resolve conflict`; `Direct To to send ALPHA to waypoint COZA because conflict has been resolved`;
- Name any other information/observations regarding the chosen solutions and scenarios that came to mind during the scenarios:
  - `More room to send TANGO left/right to resolve conflict, thus Heading 240`;
  - `More room to send TANGO in front/behind to resolve conflict, thus Heading 240`;
  - `Heading 240 to minimize additional track miles / results in less additional track miles`;
  - `Heading 240 to minimize monitoring time / results in less monitoring time`;
  - Etc.

4) Control Task: After explaining the simulation environment and how to interact with aircraft, the necessary background about CD&R, the SSD in case of the SSD group and the think-aloud protocol, the control task was explained. This was again done in several steps, accompanied by both still and dynamic scenarios. As the name suggests, participants could interact with aircraft in the dynamic scenarios whereas the still scenarios were pictures of conflict situations where participants had to indicate whether a conflict was present, which aircraft they would choose to resolve the conflict in case this was present and whether they would send this aircraft left or right. The chapter started by presenting the participants with three still scenarios. An example of such a still scenario is shown in Figure 6, where the number in the second row on the right of the flight label represents the aircraft's speed. In this case, the correct answers would be *Yes, a conflict is present* that can be resolved by either sending *Victor to the right* or *Bravo to the right* as this would result in the slower aircraft passing behind the faster aircraft. Participants had to write down only one solution.

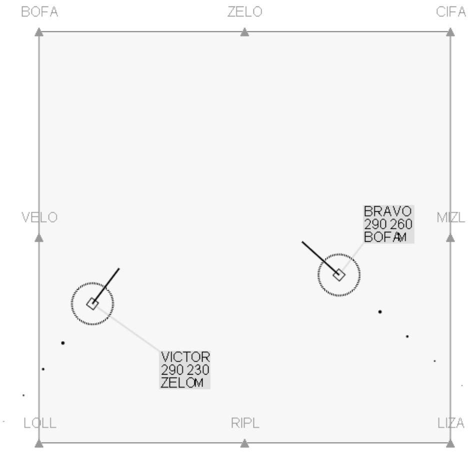


Fig. 6. Excerpt question or still scenario from the script.

After answering these three questions, the participants were instructed to perform 4 dynamic scenarios by executing commands that were stated in the script and by observing what happened to the aircraft (and their respective SSDs in the case of the SSD group). Again, each scenario was paused at the start, allowing participants to first carefully observe the traffic situation and the SSDs in case of the SSD group. The dynamic scenarios consisted of three scenarios with two aircraft and one scenario with three aircraft. For the three-aircraft scenario as well as two of the two-aircraft scenarios, the correct (safe and efficient) solution was shown: the slow aircraft would just pass at the rear of the faster aircraft. For one of the two-aircraft scenarios, an incorrect solution was demonstrated: the slow aircraft would pass in front of the faster aircraft, resulting in a large detour and a longer monitoring time. As participants also had to think out loud during these scenarios, it encouraged them to evaluate their observations and possibly already form some hypotheses for the control task and preferred strategies.

After the 'introduction' to the control task, the control task was explicitly stated. As will further be explained in Section IV, the control task was focused on expediting and maintaining a safe and orderly flow of air traffic, as is also the case during the exercises in the actual ATC training. Participants were furthermore asked to write down what their strategies were for detecting and for resolving conflicts and to perform three more still scenarios, similar to the first set of still scenarios. Once they had done so, the next phase of the experiment started: the *training phase*. During this phase, participants were presented with 11 dynamic scenarios that were not paused at the start, but immediately started running. After performing these scenarios, participants were again asked to perform three still scenarios and to write down their strategies for detecting and resolving conflicts, to probe whether their strategies might have changed after having been able to properly practice the control task.

The last phase of the experiment was the *test phase* (see Section IV for more details). The test phase was similar to the training phase: it contained 11 dynamic scenarios that were again not paused at the start, but immediately started running after pressing start. Whereas participants from the SSD group were able to perform the scenarios during the practice and training phase with the SSD, at the start of the test phase they were told that the SSD would no longer be available during the rest of the experiment, to resemble the introduction of an unexpected event as is also done during the ATC training. In case learners are able to successfully come up with solutions for new and unexpected situations, they are said to have successfully obtained a competency and the transfer of learning can, subsequently, be considered successful.

The dynamic scenarios were followed by 9 still scenarios, similar to the still scenarios from the practice and training phase, and several debriefing questions where participants were again required to write down their strategies, but were also asked about their general experience with the experiment. The participants were asked to perform the still scenarios and answer questions about their strategies at different points in the experiment to allow for continuous assessment.

The scenarios can be interpreted as the *learning task* component of the Four-Component Instructional Design model. As will be explained in Section IV, they consisted of different levels of support (for the rule of thumb) and thereby different 'difficulty classes'. The learning task additionally consisted of a high variability as the scenarios were being rotated throughout the experiment, and was tried to be kept as realistic as possible in order to add to the holistic view of the participant. Similar to the exercises that ATC trainees perform during their training, the length of both the training and test phase was between 25 and 45 minutes.

B. Additional Documents

Next to the training script or manual, the participants were provided with several other documents before and during the experiment. These are described below.

1) Experiment Briefing: Before the experiment, each participant was sent a two-page document explaining the goal and motivation for the experiment: looking into short-term knowledge development while learning an ATC task by means of discovery learning. It was explained that during the current ATC training, trainees obtain the required knowledge about the rules of the air, strategies and rules of thumb by means of discovery learning. In order to gain new empirical insights in this subject, the participants would learn to perform an ATC task in a similar fashion. Additionally, the participants were given information about the procedure of the experiment.

2) Experiment Answer Sheet: As the participants were required to answer questions about the still scenarios and state their strategies for detecting and resolving conflicts at different moments throughout the experiment, they were given an answer sheet. This sheet could be used to write down their answers and it could additionally be used to draw any helplines in the still scenarios that might help them converge to an answer. The document also contained the retrospective debriefing questions.

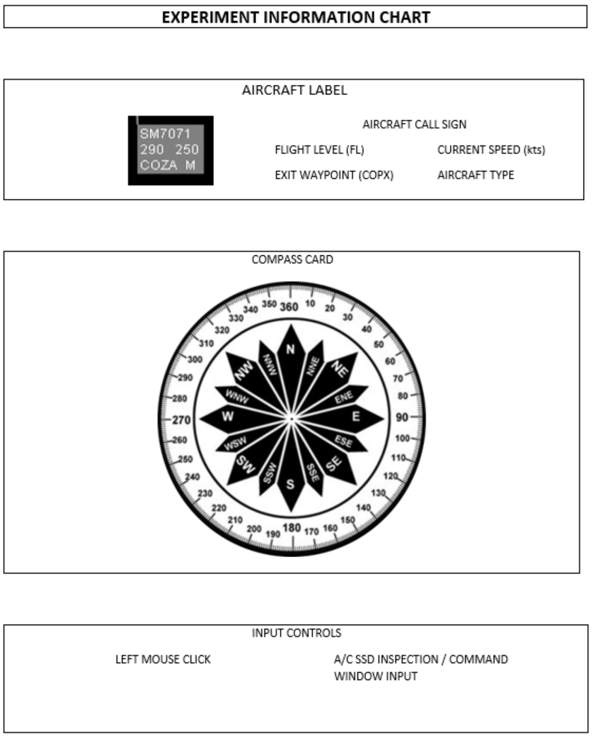


Fig. 7. Illustration of the Experiment Information Sheet or "cheat sheet" that all participants were allowed to use throughout the entire experiment.

3) Experiment Information Sheet: Finally, the last document that was provided to the participants, that also represents the final component of the Four-Component Instructional Design model (*procedural information*), was the Experiment Information Sheet or the 'cheat sheet', shown in Figure 7. By serving as a 'cheat sheet', the procedural information could be presented just in time, as is also desirable according to the instructional design model. As explained in Section II-D, procedural information generally allows students to perform routine aspects of a task. One such an aspect is vectoring aircraft. Although vectoring aircraft might seem easy and straightforward, the different heading directions are easy to confuse when a participant's stress level increases during a scenario. A heading compass was therefore provided in order to allow participants to successfully vector aircraft. Additionally, the input controls and the aircraft label with its corresponding information were stated on this document.

IV. Experiment Design

The custom-made CD&R training was developed to test whether an ecological interface could contribute to novices picking up an industry rule of thumb by means of discovery learning. A between-participants experiment, with 28 participants divided over two groups, was performed in the ATM lab of the Faculty of Aerospace Engineering at TU Delft. The experiment had been approved by the Human Research Ethics Committee of the TU Delft prior to the experiment.

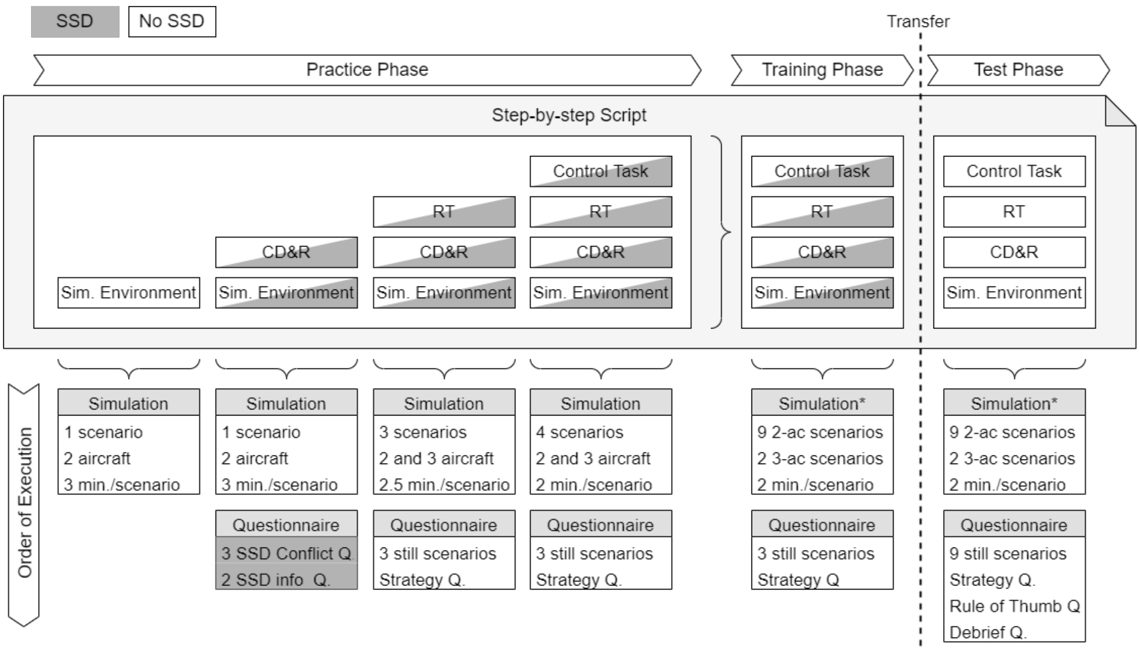


Fig. 8. Schematic of the experiment procedure. Moving from left to right indicates the sequence of the experiment. Per phase the tasks (simulation and questions) that had to be performed are shown from top to bottom. *See Table IV for the scenario order and characteristics.

A. *Participants*

The experiment was performed by 28 participants (27 MSc students and 1 assistant professor, all with a background in Aerospace Engineering) with a mean age of 24.8 years (standard deviation of 2.1 years) of the TU Delft. As the research was aimed at novice discovery learning in ATC training, mostly task-naive participants were asked to voluntarily participate. Each participant was asked to fill in a questionnaire prior to the experiment, to probe their knowledge of and familiarity with ATC systems, goals, strategies and practices as well as experience with computer or cellphone (ATC-related) games or applications. The results of this questionnaire were used to create two balanced groups of 14 participants, with on average similar (self-indicated) knowledge and (expected) skill levels. Additionally, each participant first signed a consent form, approved by the Human Research Ethics Committee of the TU Delft.

B. *Procedure*

Participants were guided through the experiment by means of an interactive step-by-step script that contained all steps that had to be performed as well as all the necessary background information. The experiment took approximately 2.5 to 3 hours, depending on the group and the reading pace of the participant, and can be divided in roughly three parts: a practice phase, a training phase and a test phase. An overview of the set up of the experiment is shown in Figure 8. The schematic shows that the complete experiment task consisted of four main items, namely: the simulation environment, CD&R, RT and the actual control task.

Moving from left to right, participants started the experiment with the practice phase, which was built up from several items or scaffolds that each had a separate learning objective. First, participants had to be familiarized with the simulation environment and all the tools and features available to them, as well as how to interact with aircraft. Secondly, the basic concept of CD&R was explained (i.e., the definition of a conflict). Additionally the participants were familiarized with conflicts in the simulation environment and the SSD (in case of the SSD group). The third item that was added to their knowledge was RT, where participants received instructions about the think-aloud protocol that had to be followed during the rest of the experiment. Finally, the information about the control task was presented, which concluded the practice phase. Each step, in which an extra item (scaffold) was added, was accompanied by one or multiple dynamic scenarios and some steps were accompanied by questions and still scenarios, as was also explained in Section III. Both groups had to perform the still scenarios without the SSD, such that the results could serve as baseline measurements throughout the entire experiment. Moving from top to bottom in the lower half of Figure 8 indicates the task sequence for a specific step.

After a short break, participants continued with the training phase of the experiment. During 11 scenarios of two minutes each, participants had the chance to practice the complete task. Finally, during the test phase, participants had to perform 11 scenarios that were similar to the scenarios from the training phase to test their newly-obtained knowledge. After each dynamic scenario (in every phase of the experiment), participants were asked to rate how they perceived the difficulty of that scenario by using a sliding bar. For the SSD group, the SSD was available during the dynamic scenarios of both the practice and the training phase, indicated by the half gray blocks in Figure 8. Both groups performed the test phase on the baseline

TABLE II
EXPERIMENT PROCEDURE

Experiment Phase	Elements	Duration
Briefing	Procedural information	$\sim$ 5 min
Practice	Background information, 9 dynamic scenarios (2-3 minutes each), 6 still scenarios	$\sim$ 60-90 min
Training	11 dynamic scenarios (2 minutes each), 3 still scenarios	$\sim$ 35 min
Test	11 dynamic scenarios (2 minutes each), 9 still scenarios	$\sim$ 35 min
Debrief	Retrospective questionnaire	$\sim$ 10 min

PVD, where the SSD was not available.

Prior to the experiment, the procedure was explained by going over the experiment briefing and the participants were given the opportunity to ask questions. No information was given on the content of the experiment or any further background information, other than stated in the briefing that was sent in advance, as all content-related information was stated in the script. By doing so, the explanation about the experiment was kept equal for both groups and the risk at a confound was minimized.

The experiment ended with a debrief where participants were asked to answer some retrospective questions on the answer sheet to gain an insight into their strategies, elements of the experiment they found easy/difficult/boring/fun and their general experience with the experiment. An overview of the procedure and the duration of each phase is found in Table II, where the highlighted rows indicate phases that were included in the training script.

C. Instructions

Prior to the experiment, all participants received an experiment briefing which globally explained the goal and procedure of the experiment. At the start of the experiment, participants were given more procedural information about the experiment. Each participant received three documents: a training guide, an answer sheet and a 'cheat sheet', shown in Figure 7. The training guide, discussed in Section III, contained all the steps participants had to perform during the experiment, all questions participants had to answer throughout the experiment, as well as all necessary background information about ATC, the PVD and CD&R, that was required to perform the experiment. The guide did not elaborate on ATC strategies related to CD&R or ATC strategies in general, as the participants were expected to obtain these by means of discovery learning.

In line with the mission statement of ATC, to safely and efficiently organize and expedite the flow of air traffic from origin to destination [1], the participants were instructed to first guarantee safe separation of aircraft and secondly, to clear aircraft to their respective Cleared Sector Exit Point (COPX) in the most efficient way possible. Participants were thus instructed to ensure the minimum horizontal separation distance of 5 NM between neighboring aircraft at all times, to avoid STCAs (orange and red) and to resolve these as soon as possible in case they did occur, before efficiently clearing aircraft to their respective COPX. Additionally, participants were instructed to communicate their thoughts and strategies during all scenarios according to the "think-aloud" protocol, discussed in Section III, in order to gain an insight into their strategies and thinking patterns.

D. Apparatus

The experiment was performed on a desktop computer in the ATM Lab of the Faculty of Aerospace Engineering at TU Delft. The software that was used to construct the training scenarios and to conduct the measurements is the Java application MUFASA [37]. The traffic motion in the simulator is simulated by simple, linear kinematic equations, no wind conditions are taken into account and all aircraft velocities are given in knots Indicated Airspeed (IAS) [37]. Figure 9 shows the experiment setup.

Fig. 9. Experiment setup

A screenshot of the simulation environment is shown in Figure 10. Note that the colors have been adapted for better visibility. The scenarios contained a 50 NM by 50 NM square airspace for all scenarios in order to be able to rotate the scenarios and thereby increase variability of practice. The airspace contained eight waypoints, at each of the corners of the airspace and in between, of which the names were randomized for each scenario to increase variability and prevent recognition of the scenarios. On the top left of the simulation environment, the time that the current experiment run was running, the number of the current experiment run as well as the pause/play button that was available during the practice phase were displayed. On the bottom right, a scale of 10 NM was displayed. In its most basic form, each aircraft on the radar screen was accompanied by a flight label, speed vector, history dots and the half-Protected Zone (PZ) circle.

To assure separation, which was achieved by making sure the half-PZ circles with a 2.5 NM radius around the aircraft did not overlap, and to clear aircraft to their respective COPX, participants could give the aircraft only heading changes in the 2D plane. The bottom line of the flight label of each aircraft

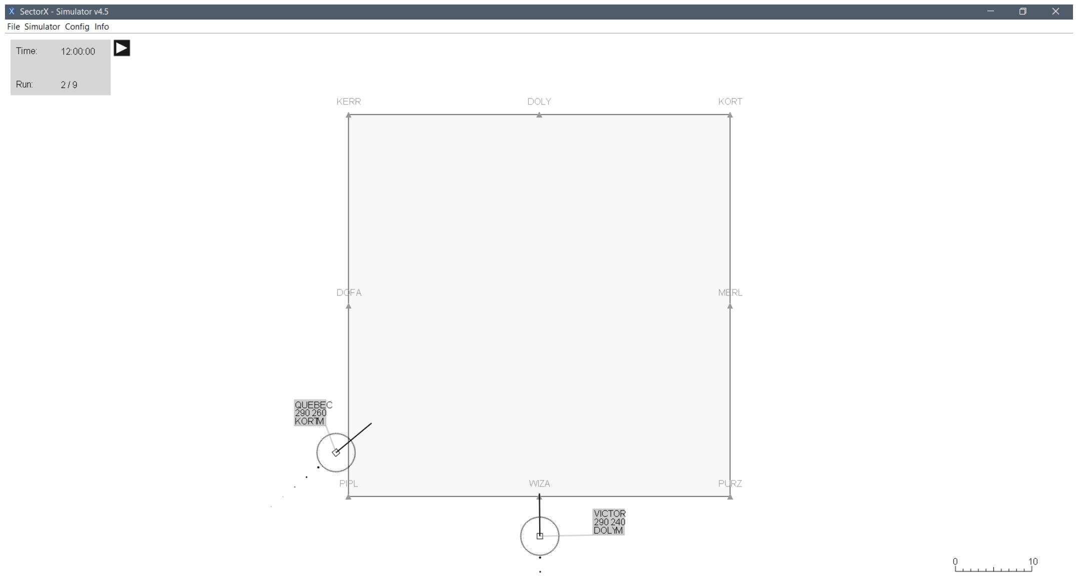


Fig. 10. Inverted screenshot of the simulation environment

showed its designated COPX. Aircraft that were on course to their respective COPX were displayed in green, the other aircraft were shown in gray. Aircraft could furthermore be orange or red, which indicated an STCA where a LoS was imminent within 40 or 20 seconds, respectively.

The route of an aircraft could be modified by making use of the Command Display, shown in Figure 11. The two types of commands that could be given to modify aircraft routes during this experiment were:

- Heading (HDG): a heading command was given by selecting an aircraft in the PVD, either by clicking on the aircraft or clicking on its label, clicking on the HDG button in the command display, clicking on the individual numeric buttons that together form the new heading value, and finally, clicking on Execute (EXQ). After issuing a heading clearance, this became clear by the changing direction of the speed vector and the changing track of the history dots.
- Direct To (DCT): a DCT command was used to send an aircraft immediately in the direction of its assigned COPX and was given by selecting an aircraft in the PVD, either by clicking on the aircraft or clicking on its label, clicking on the DCT button in the command display, and finally, clicking on EXQ. After issuing a DCT command, this became visible by the aircraft turning green once it was on track to its assigned COPX.

Additionally, participants could make use of the Clear (CLR) button. In case of a wrongly-entered command, this could be corrected for by clicking on CLR. Participants were furthermore told to ignore the TOC and PRV buttons on the command display as these were not required in the experiment. The participants did thus only need the mouse to give clearances to aircraft and did not need to make use of the keyboard. In addition, no voice communication was required to command heading changes to the aircraft, as this would interfere with the "think-aloud" task of the participants.

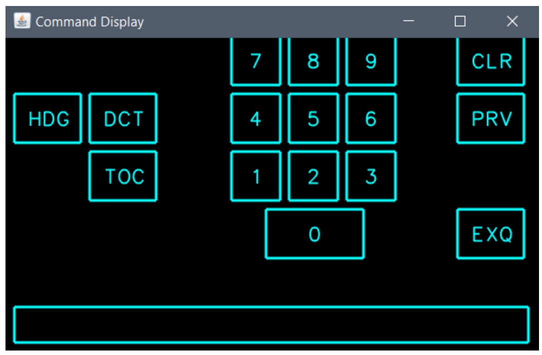


Fig. 11. Screenshot of the command display

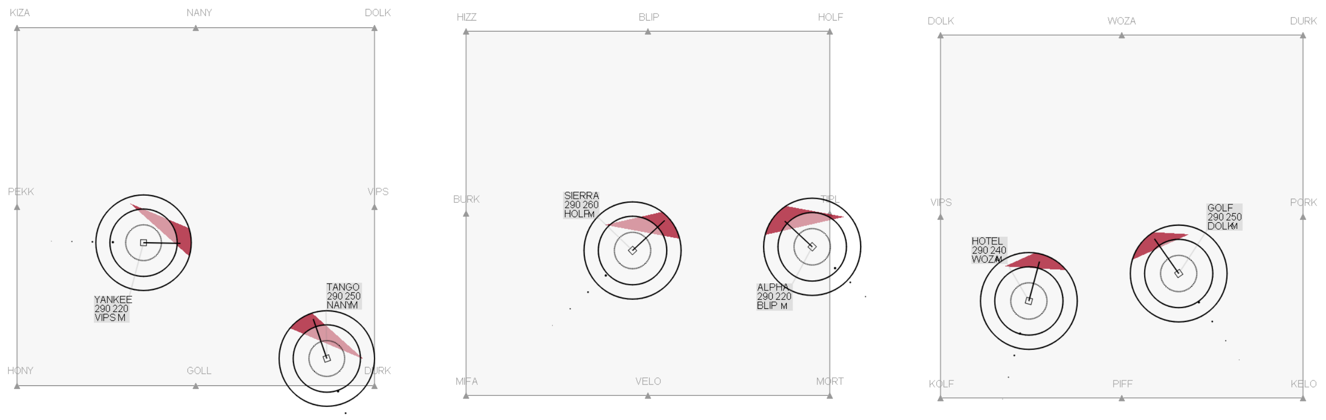
E. Independent Variables

The independent variable in this between-participants experiment design was the display configuration. Training with or without the SSD meant that not only the display was different, but the training guide for both groups was also different. While both training guides contained a chapter on CD&R, the chapter for the SSD group also elaborated on the SSD and contained information on how the SSD is constructed and what kind of information could be derived from it. Participants furthermore had to answer five questions about the SSD to make sure that they understood the SSD, the information they could derive from it and how to derive this information.

F. Traffic Scenarios

The merging or converging conflict type was chosen as main conflict type for this experiment, as this is generally considered to be one of the more difficult conflict types[2] and thus, a decision-support tool might be beneficial when learning this

[2]Based on previous research and anecdotal evidence from ATC experts during a visit to the LVNL on April 24th, 2019.



(a) (b) (c)

Fig. 12. Examples of the three conflict geometries used in the experiment to illustrate the salience of the rule of thumb. (a) Easy scenario where the rule of thumb is most obvious. (b) Medium scenario where the rule of thumb is less obvious but where one direction is still more clearly preferable. (c) Difficult scenario where the rule of thumb is least obvious.

TABLE III
CONFLICT VARIABLES PER SCENARIO CATEGORY

Category	CPA [NM]	CA [deg]	Speed difference [kts]
1: Easy	4.5	100-115	50-60
2: Medium	2.5	70-85	30-40
3: Difficult	1	45-60	10-20

task [38]. It was decided to only focus on a 2D horizontal plane in which aircraft could only be given heading commands to limit the number of strategies and solutions that could be applied by participants. Two types of training elements were developed, namely still and dynamic scenarios.

As the implicit rule of thumb, that students were required to learn by means of discovery learning, was to vector slower aircraft behind faster aircraft in case of crossing and/or conflicting traffic, it was decided to create dynamic and still scenarios in which this rule was always applicable. However, the extent to which this rule was evident from the conflict geometry and thus from the SSD, was varied. Table III shows the three types of conflict geometries that were defined based on the Closest Point of Approach (CPA), Conflict Angle (CA) and the speed difference between the involved aircraft. The first level indicates a scenario for which sending the slower aircraft behind the faster aircraft was the best and easiest or most obvious solution while the third level indicates a difficult scenario for which this was least obvious. The three resulting conflict geometries are also shown in Figure 12.

All scenarios furthermore contained a conflict, for which an orange STCA was scheduled to occur after 30 seconds into the simulation. Next to this, aircraft were not aligned to their assigned exit waypoints at the start of the scenario and giving a DCT command as first command in the scenario would always make the conflict worse, as it decreased the CPA. This way, participants always had to actively think about the aircraft choice and solution direction and scenarios could not be solved by performing a ‘trick’ (i.e., simply giving both aircraft a DCT command).

For each scenario category, three dynamic and three still scenarios of two aircraft without any other traffic were created while for the third and most difficult category, two additional dynamic scenarios with three aircraft and no other traffic were created. Scenarios with two aircraft always had one “best solution”, whereas for scenarios with three aircraft, the third aircraft would always strengthen this solution and make the implicit rule of thumb more evident, despite being of the third and most difficult category. The three-aircraft scenarios were thus additionally created to direct participants more towards using the correct strategy or rule of thumb, but also to prevent participants from becoming bored and to increase the variety of the scenarios.

TABLE IV
DYNAMIC SCENARIO ORDER AND CHARACTERISTICS FOR THE TRAINING AND TEST PHASE

Scenario	No. Aircraft	Category
1	2	Easy
2	2	Medium
3	2	Difficult
4	2	Easy
5	2	Medium
6	2	Difficult
7	2	Easy
8	2	Medium
9	2	Difficult
10	3	Difficult
11	3	Difficult

The order of dynamic scenarios was kept constant for both groups during both the training and the test phase of the experiment in order to be able to compare learning curves. This order can be found in Table IV. It can be seen that the scenarios containing two aircraft can be divided in three sets, each increasing in difficulty or complexity, as this is also considered an important factor of the exercises during the ATC training. By ordering the scenarios in such a way, the learning task contained easy-to-difficult task classes as well as a decreasing level of support: the salience of the rule of thumb in the SSD decreased by increasing the level of difficulty of the scenarios.

The order of the still scenarios was also kept constant for both groups before and after the transfer, as every participant had to work through the same training script. Similar to the dynamic scenarios, the still scenarios had also been divided into three sets, each increasing in difficulty or complexity. The first set of three still scenarios before the transfer was furthermore identical to the final set of three still scenarios after the transfer, in order to be able to compare performance at the very start and very end of the experiment. In order for participants to develop more robust knowledge and increase their resourcefulness as well as to prevent recognition of the conflicts, the geometrical orientations of the conflicts were rotated for the dynamic as well as the still scenarios.

G. Control Variables

Several measures were kept constant throughout the experiment to minimize the risk at confounds. Although these control variables came at the cost of less realistic scenarios and a less realistic simulation environment, it was believed that participant performance would not or barely be affected as the participants were mostly novices. The control variables for this experiment were:

- *Flight Level:* All traffic was limited to the horizontal 2D plane at flight level 290;
- *Aircraft Type:* all aircraft were of the same type with a speed envelope ranging from 200 to 290 kts;
- *Display Layout:* Conform industry standards. Sector size and shape of 50 NM by 50 NM, display colors and waypoint locations were kept constant throughout the experiment;
- *Conflict Type:* Converging traffic of three predefined levels;
- *Type of Commands:* To limit the variance in strategies and to put more implicit focus on the learning objective, aircraft could only be given heading clearances. It was not possible to give altitude or speed commands.
- *Update Frequency:* The update frequency was 0.33 HZ.
- *Scenario Length and Speed:* All scenarios during the training and test phase were two minutes and were played at three times the actual speed.
- *Scenario Configuration and Order:* The constructed scenarios were the same for both groups and were presented in the same order to both groups to be able to compare their learning curves;
- *Experiment Briefing:* Participants were given information about the goal and procedure of the experiment prior to the experiment;
- *Instructions:* Apart from the chapter on CD&R where the SSD was introduced to the SSD group, the training script was equal for both groups;
- *Feedback:* No feedback was given as discovery learning was encouraged as much as possible.

H. Dependent Measures

To quantify the performance development, aircraft positions and states as well as the given commands were logged every 3 seconds. These data also contained information about the strategies and choices made by the participants. Furthermore, as only the data from the training and test phase would be used for data analysis, these phases were recorded with a camera to obtain the think-aloud data that would be used to gain an insight in participants' thinking patterns. Finally, all participants had to answer questions during the experiment on an answer sheet. From the above, roughly four categories were derived for the data analysis.

1) Questionnaire: The answers from the questionnaire were used to get an impression of the effectiveness of the manual. The number of still scenarios that were solved correctly before and after the transfer manipulation could be compared. Furthermore, participants were asked about their strategy for detecting and resolving conflicts in the dynamic and still scenarios at multiple times throughout the training phase as well as the test phase. Finally, after the experiment, participants were asked what they thought the rule of thumb for resolving conflicts was that they had learned during the experiment.

2) Participant Strategies: As all aircraft positions, states and given commands were logged during the experiment, dichotomous data about the choices made by controllers could be derived from the logged data. Two choices that could lead to the development of a wide range of strategies were especially important: (1) aircraft selection (slow/fast) and (2) solution direction (correct/incorrect). The strategies that were developed by the participants could be categorized based on the number of aircraft involved in the solution, aircraft choice, solution type (optimal or sub-optimal) and number of corrections, or they could be categorized based on being an optimal solution (i.e., slow aircraft is vectored behind fast aircraft) or a sub-optimal solution and the aircraft choice. The strategies were first derived from the logged data and in case a strategy was not evident from these data, the video recordings were used to complete the analysis. It should, however, be noted that this analysis was based on the interpretation of the researcher. Additionally, the response times of the first command given to an aircraft were recorded in the logged data as well as in the video recordings.

3) Performance: Performance was first and foremost measured in terms of safety and could additionally be measured in terms of how efficiently the task was performed. In terms of safety, performance was characterized by the number of PZ violations as well as the minimum separation distance between aircraft. Other control performance measures included the track deviation from the initial track, participant response times and the number and type of commands.

Since measurement runs were conducted throughout the entire experiment, the dependent measures above could be used to identify learning curves as the performance, chosen strategies or specific cognitive processes might change and/or improve during the training. To analyze the learning curves, learning gradients as well as deltas (jumps) between the training and the test phase were derived for each participant for each of the scenario categories. A schematic of the definitions of the learning gradients and delta is shown in Figure 13. The learning gradient was determined by linear regression as only three data points were available per participant per experiment phase. The delta was determined per participant

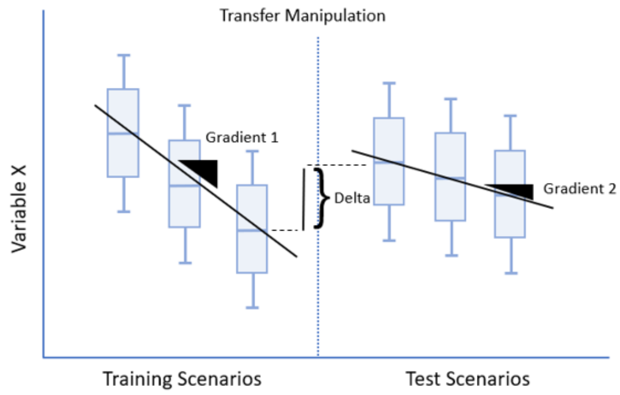


Fig. 13. Schematic of the expected learning curves along with the definition of the learning gradient and the delta.

and was defined as the difference between the first scenario of the test phase and the last scenario of the training phase.

I. Hypotheses

The experiment was structured according to the Four-Component Instructional Design model and all participants were guided through the experiment by means of the training script (with or without the SSD). It was expected that applying such a method to the learning process would result in the sequential occurrence of the above-mentioned learning curves rather than a mixed or simultaneous occurrence. It was thus hypothesized that (H1) the training script would be effective and would result in participants successfully obtaining the knowledge about the predefined rule of thumb. It was furthermore hypothesized that (H2) structuring the learning process like this would result in more salient learning curves throughout the experiment.

The SSD was embedded in the experiment as a cognitive tool. As the SSD aids participants in observing the situation and evaluating different solution options, it was hypothesized that training with the SSD would be beneficial and would (H3) lead to more participants being/becoming aware of and using the rule of thumb when being trained with the SSD as cognitive tool. It was furthermore expected that (H4) the three scenario categories (level of salience of the rule of thumb) would have an effect on the correct execution of the rule of thumb: the largest number of correct strategies was expected to be found in the easy scenarios and the smallest number of correct strategies was expected to be found in the difficult scenarios.

Additionally, although a risk exists that surface learning occurs when the SSD is used as a rule-based tool only and participants might start to become dependent on the interface, it was expected that designing the training task according to an instructional design method would help to structure the mental model and manage the cognitive load during the experiment, and thereby reduce the occurrence of surface learning. It was thus hypothesized that (H5) participants from the SSD group would experience a 'setback' in performance but would quickly recover and (H6) would experience steeper learning curves after the transfer manipulation compared to the control group, and (H7) would continue using the same strategies after this transfer manipulation due to their increased understanding compared to the control group.

V. Results

Using the data that were logged every three seconds by the simulation software, as well as the audio and video recordings, the data analysis could be performed in several steps. First, the training script effectiveness is discussed. After this, participant strategies are classified and analyzed and finally, the performance data and the learning curves that follow from these data are analyzed.

A. Training Script Effectiveness

1) Simulation Environment: Participants started the experiment by being familiarized with the simulation environment and by learning how to interact with aircraft. In order to test whether participants successfully learned how to interact with the aircraft, the input errors were recorded during the training and test phase. An input error was in this case defined as an incomplete command. An example is 'ALPHA180; EXQ' instead of 'ALPHA; HDG180; EXQ'. The HDG command button was not selected in the first case, which was then counted as an input error. Participants could also clear an incomplete command in the command window by using the button CLR. However, as participants noticed their own mistake(s) in this case, before executing the command, the number of CLR commands was not taken into account when looking at the input errors.

Ideally, participants would have zero to one input errors. However, it was found that a number of participants made several input errors during the training phase of the experiment as well as during the test phase. As the majority of the data were obtained from the two-aircraft scenarios and as the number of input errors could have an effect on the results of the learning curves observed for the different participants, only the results for the two-aircraft scenarios are shown in Figure 14.

Figure 14 shows two peaks in the number of errors for participant 10 of the control group. It was found that this participant often clicked on Execute (EXQ) a number of times within a few seconds after finding out the initial command did not come through (due to a yet unknown error in the software), without further specifying a new command. All these 'empty' Execute commands were counted as input errors, while, in fact, the actual number of incorrect commands for this participant was zero. The 'empty' Execute commands are therefore depicted in light blue.

Figure 14 furthermore shows that especially within the SSD group more participants made errors after the transfer. This could be a result of the unexpected removal of the SSD, raising participants' stress levels. The results for the control group, however, show that less participants made mistakes after the transfer. Except for participant 10 from the control group, the participants of this group also made less errors in general after the transfer which indicates a positive learning process.

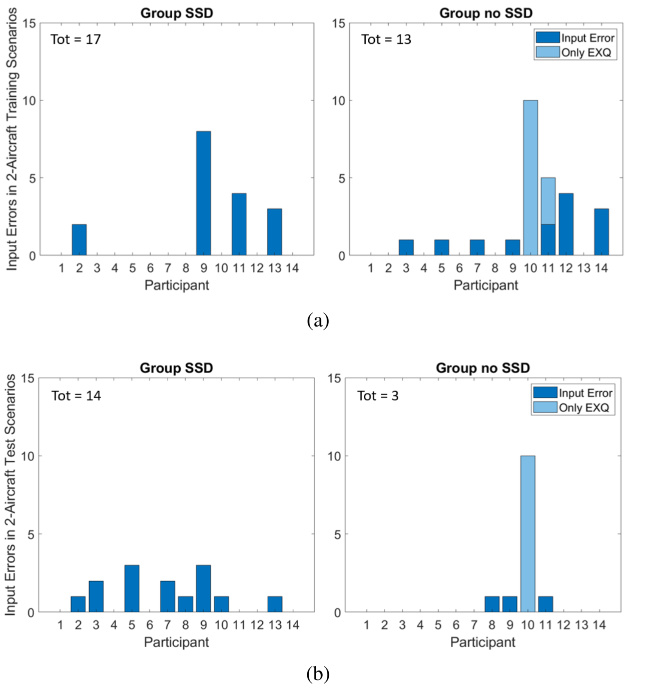


Fig. 14. Number of input errors or incorrect commands during the two-aircraft scenarios (a) per participant during the training phase and (b) per participant during the test phase.

Not taking the correction for the 'empty' Execute commands into account, Wilcoxon Signed Rank Tests confirmed that the control group experienced a significant difference in input errors after the transfer (*No SSD: z = -1.983, p = 0.047*), whereas the SSD group did not experience a significant difference when comparing the experiment phases (SSD: z = -0.179, p = 0.858). Additionally, the effect of the SSD on the number of errors was not found to be significant during the different phases of the experiment (Training: H(1) = 1.378, p = 0.240; Test: H(1) = 2.381, p = 0.123). After excluding 'empty' Execute commands from the analysis, Wilcoxon Signed Rank Tests showed that the difference in number of input errors between the training and test phase was insignificant for both groups (No SSD: z = -1.930, p = 0.054; SSD: z = -0.179, p = 0.858). When looking at the effect of the SSD on the number of input errors after applying the correction, it was found that the difference between the groups during the test phase was significant (Training: H(1) = 0.361, p = 0.548; *Test: H(1) = 4.665, p = 0.031*).

2) Solution Space Diagram: After being familiarized with the simulation environment, participants were introduced to CD&R and, more specifically, the SSD group was introduced to the SSD. During this phase, the participants of this group had to answer several questions about the SSD and tell their answers to the researcher. This allowed the researcher to check whether the participants correctly understood the information that could be derived from the SSD, before they continued with the rest of the experiment, where it was crucial to correctly understand this information.

The questions were divided into two sets. First, three conflict solutions were shown where participants had to indicate whether a conflict was present or not, based on the information of the SSD. All participants answered all three questions correct. Second, two questions were asked about additional conflict geometry information that could be derived from the SSD. 11 of the 14 participants answered these questions correctly. Again, the answers were discussed with the participants to make sure that all participants understood the SSD at the same level before continuing to the next phase.

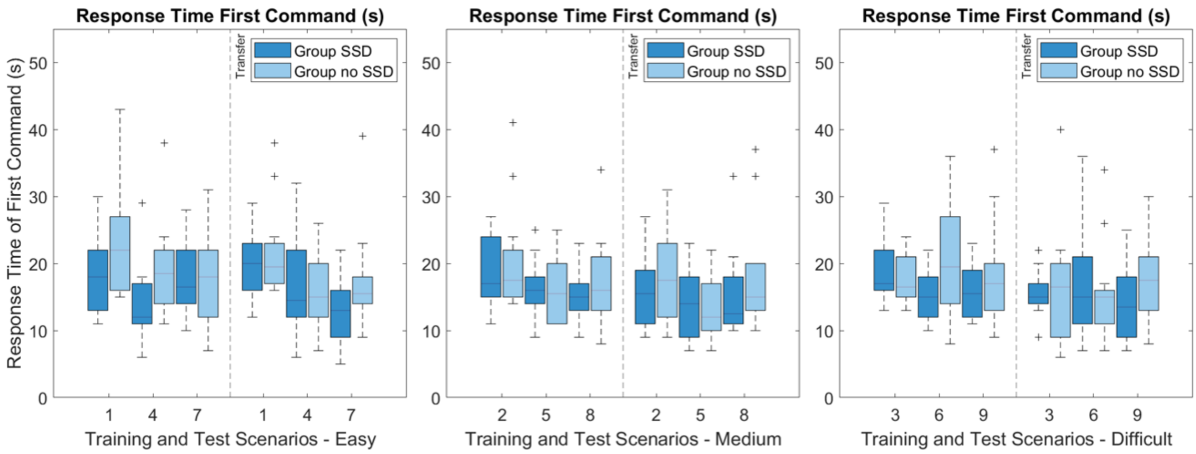
3) Think-aloud Protocol: The third item in the practice phase was RT or the think-aloud protocol. Participants were asked to think out loud according to the protocol in order for the researcher to use these data in the assessment and gain an insight in the participant's thinking patterns and cognitive processes. It was found that some participants strictly adhered to the protocol, whereas other participants elaborated on their choices beyond the protocol. Additionally, several participants seemed to require more time and effort than others to put their thoughts to words (i.e., many 'hm's and 'eh's).

The think-aloud protocol specified an order for naming the different thoughts. It did, however, not specify where in this order participants were required to give commands to aircraft. It turned out that many participants first stated and reasoned their approach before actually executing the commands. The difference in amount of elaboration or hesitation in combination with performing commands after stating the approach thus resulted in some participants executing commands later than they would have probably done without the think-aloud protocol, indicating that they might have required additional repetition (part-task practice) for this element.

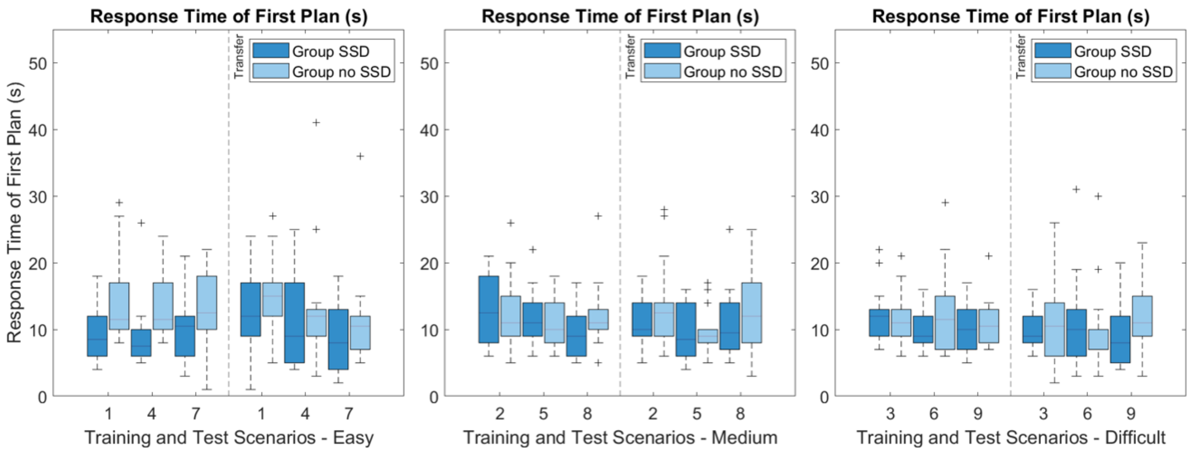
This effect can also be seen when looking at participants' response times. Response times have been determined in two ways. First, the timing of the first command given to the first aircraft was extracted from the logged data. It was furthermore decided to also analyze the audio files and extract the times at which participants first started mentioning their approach or plan of execution (e.g., 'aircraft X has more room on the right so I'm sending aircraft X in this direction by giving a HDG command of 90 degrees'). The time that was noted was thus not the time a conflict was detected but the time participants first mentioned the plan for the first command. The results are shown in Figures 15a and 15b.

Figure 15 shows indeed that the response times extracted from the audio files are lower, as was expected, since participants usually first mentioned their plan before executing it rather than mentioning their plan after executing a command. Figure 15b furthermore shows that for the easier scenarios a large gap is present between the groups in response times of the first plan during the training phase. This could be a result of the use of the SSD as it increased participants' situation awareness and allowed them to formulate an approach earlier than the control group. Additionally, mostly negative trends are present in both figures during the training and test phases of each scenario category, indicating decreasing response times.

4) Rule of Thumb in Still Scenarios: Participants were furthermore asked to answer questions about still scenarios before and after the transfer manipulation. Recognizing a conflict and choosing to vector the slow aircraft behind the fast aircraft, or to vector the fast aircraft in front of the slow



(a)



(b)

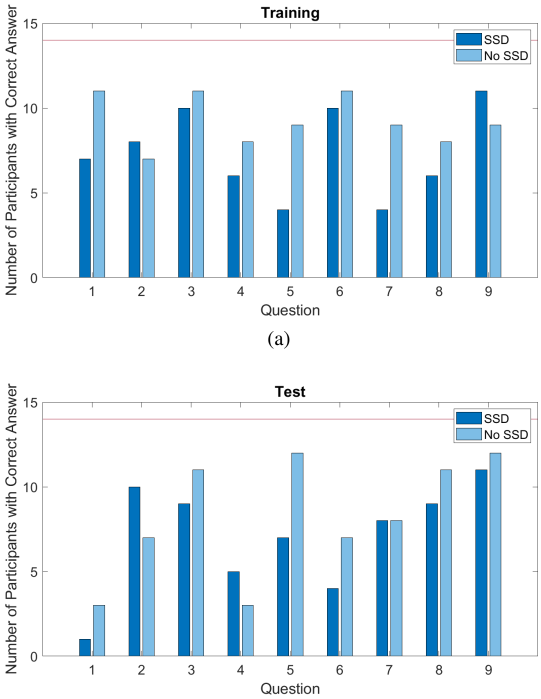
Fig. 15. Conflict resolution response times per scenario category. (a) Response times of first action, derived from the logged data. (b) Response times of mentioning the first conflict resolution plan, derived from the audio data.

aircraft, was both counted as a correct answer. Dichotomous yes/no data were thus collected from these questions. Figure 16 and Table V show the number of participants that answered a question correctly for each of the still scenario questions before and after the transfer. The red line in the figures displays the maximum number of participants per group (i.e., 14). The highlighted numbers in the table show the largest number of correctly answered questions. Before the transfer manipulation, the first three questions had to be answered before the control task was explained and before the participants had to perform and observe several conflict scenarios with optimal and sub-optimal solutions. The second set of three questions had to be answered after these observations and the last set of three questions had to be answered after the control task was explained and after the participants got to practice the complete task during 11 training scenarios, as was also shown in Figure 8.

It was expected that participants (especially from the SSD group) would perform better after each set of three questions during the training phase and would thus answer more questions correctly. However, the data revealed that this was not the case. Furthermore, when comparing the first set of three questions from the training phase to the last set of three questions from the test phase to which they were identical, a slight improvement in performance is seen for both groups. Additionally, during both the training and the test phase, the participants from the control group performed better overall. A larger number of questions was answered correctly by a larger number of participants from this group. Kruskal-Wallis tests, however, revealed that the difference between the groups was

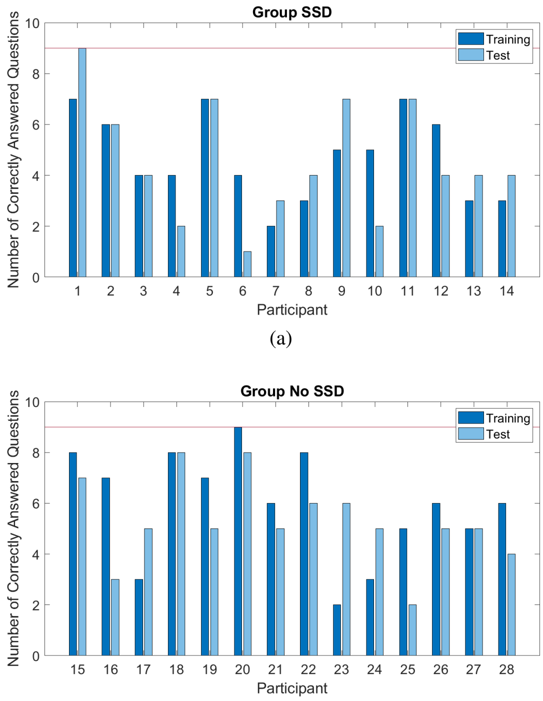
TABLE V
NUMBER OF PARTICIPANTS PER STILL SCENARIO QUESTION WITH A CORRECT ANSWER

	Before Transfer		After Transfer	
Question	SSD	No SSD	SSD	No SSD
1	7	11	1	3
2	8	7	10	7
3	10	11	9	11
4	6	8	5	3
5	4	9	7	12
6	10	11	4	7
7	4	9	8	8
8	6	8	9	11
9	11	9	11	12
Sum	66	83	64	74



(a) (b)

Fig. 16. Number of participants from each group that answered the still scenario questions correctly (a) before the transfer and (b) after the transfer.



(a) (b)

Fig. 17. Number of questions that were answered correctly before the transfer and after the transfer by (a) the participants from the SSD group and (b) the participants from the control group.

not significant (Training: H(1) = 2.568, p = 0.109; Test: H(1) = 1.299, p = 0.254). It should be noted that both groups had to answer the questions without the SSD. The fact that the SSD group was performing the dynamic training scenarios *with* the availability of the SSD but had to answer the still training questions *without* the SSD could have led to the control group performing better and more constantly overall.

Figure 17 and Table VI show the number of questions that were answered correctly per participant during both the training and the test phase. Again, the largest numbers in the table are highlighted and the red line in the figures indicates the maximum number of questions that could be answered correctly (i.e., 9). Comparing the number of questions that were answered correctly before the transfer and after the transfer shows that 6 out of the 14 participants of the SSD group performed better after the test, while only 3 performed better for the control group. The number of participants that answered an equal number of questions correctly before and after the transfer was 4 and 2, respectively. Although a trend is visible, Wilcoxon Signed Rank Tests confirmed that performance of participants did not change significantly when comparing the results from before and after the transfer (SSD: z = -0.456, p = 0.642; No SSD: z = -1.072, p = 0.284). As the questions had to be answered on the answer sheet and without the SSD, some participants of the SSD group tried to draw the SSD themselves on the answer sheet for the still scenarios, which might have contributed to their understanding of the rule of thumb.

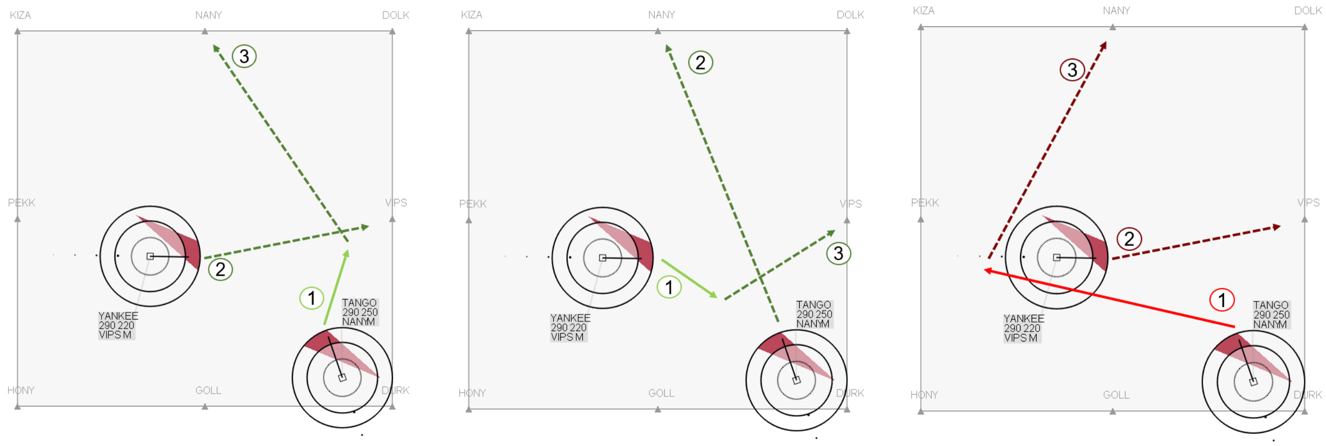
TABLE VI
NUMBER OF CORRECT ANSWERS PER PARTICIPANT FOR THE STILL SCENARIOS

	SSD		No SSD	
Participant	Before Transfer	After Transfer	Before Transfer	After Transfer
1, 15	7	9	8	7
2, 16	6	6	7	3
3, 17	4	4	3	5
4, 18	4	2	8	8
5, 19	7	7	7	5
6, 20	4	1	9	8
7, 21	2	3	6	5
8, 22	3	4	8	6
9, 23	5	7	2	6
10, 24	5	2	3	5
11, 25	7	7	5	2
12, 26	6	4	6	5
13, 27	3	4	5	5
14, 28	3	4	6	4
Sum	66	64	83	74

B. Participant Strategies

Dichotomous data about the choices that participants made in solving the scenarios were collected. As mentioned before, especially the aircraft choice and solution direction were important in identifying the strategies. Below, the results are presented for the two and three-aircraft scenarios, respectively.

1) Participant Strategies in Two-aircraft Scenarios: Two different distinctions were made in the identification of strategies. First of all, strategies could either lead to an optimal solution (i.e., a solution where the slower aircraft would pass behind the faster aircraft) or a sub-optimal solution. Figure



(a) (b) (c)

Fig. 18. Schematics of the definition of optimal (slow aircraft is vectored behind fast aircraft) and sub-optimal strategies. (a) Optimal strategy with faster aircraft as first aircraft. (b) Optimal strategy with slow aircraft as first aircraft. (c) Sub-optimal strategy with fast aircraft as first aircraft.

18 shows two examples of an optimal strategy compared to one example of a sub-optimal strategy for the easiest scenario category. The figures show slower aircraft YANKEE and faster aircraft TANGO along with their SSDs. The SSDs show a larger white solution space on the right of both aircraft (as seen from the perspective of the respective aircraft) and thus that either aircraft should be given a heading command to the right to resolve the conflict in the most efficient way.

Figures 18a and 18b show the strategies where a heading command (continuous arrow) is first given to the faster and slower aircraft (1), respectively, after which both aircraft in the scenarios are cleared to their respective exit waypoints by means of a DCT command (dotted arrow) (2 and 3). These two strategies are considered the 'most optimal' strategies as they are efficient in both additional track miles as well as in controller workload (least amount of additional track miles, commands and monitoring time). Figure 18c shows a strategy where the fast aircraft is first vectored to the left, behind the slower aircraft (1), before the aircraft are both cleared to their respective exit waypoints (2 and 3). It can be seen that by sending the fast aircraft left, the aircraft has to travel through a larger red area in the SSD and thus this solution is less efficient and hence sub-optimal (i.e., a larger number of additional track miles as well as a larger monitoring time). After categorizing the strategies based on being optimal or sub-optimal as described above, it was found that during the two-aircraft scenarios, both participant groups chose an optimal strategy over a sub-optimal strategy significantly more often (*SSD Training:* z *= -3.329,* p *= 0.001; No SSD Training:* z *= -3.223,* p *= 0.001; SSD Test:* z *= -3.322,* p *= 0.001; No SSD Test:* z *= -2.999,* p *= 0.003*).

After categorizing strategies based on being optimal or sub-optimal, the strategies were further categorized based on the aircraft that was chosen to give the first command to (i.e., slow or fast). The results of this strategy categorization for the two-aircraft scenarios can be found in Figures 19 to 21, where the white blocks represent the group with SSD and the gray blocks represent the control group. As the analysis was done for each of the three scenario categories, the data that are shown represent the percentage of scenarios of that category (i.e., 14 participants each performing 3 scenarios per category means 42 scenarios per category) where the respective strategy was observed. Next to the preference for optimal strategies during each category, another trend can be seen in the figures. For all three categories it can be seen that, in case an optimal strategy was chosen for a scenario, this was more often accompanied by choosing the slower aircraft as the first aircraft to give a command to. Similarly, when a sub-optimal strategy was chosen, the faster aircraft was more often chosen as first aircraft than the slower aircraft.

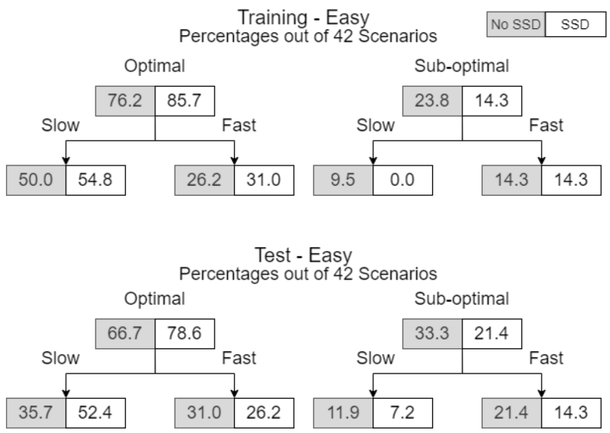


Fig. 19. Strategy trees for the easy scenario category, showing the percentages of scenarios (out of 42 scenarios) where the respective strategy was observed.

Figure 19 shows that for the easy scenarios, the SSD group more often selected an optimal solution compared to the control group. For both groups, the number of times participants chose a sub-optimal solution increased after the transfer. For the medium scenarios, Figure 20 shows that the SSD group performed slightly better than the control group during the training phase in that they more often chose an

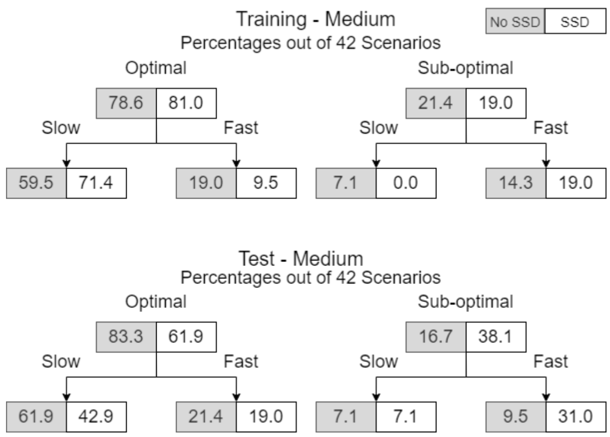


Fig. 20. Strategy trees for the medium scenario category, showing the percentages of scenarios (out of 42 scenarios) where the respective strategy was observed.

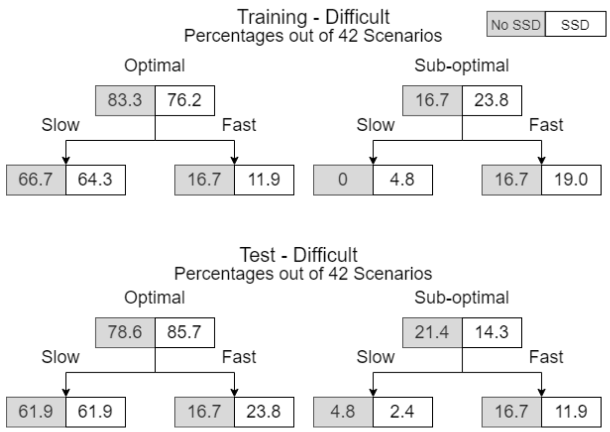


Fig. 21. Strategy trees for the difficult scenario category, showing the percentages of scenarios (out of 42 scenarios) where the respective strategy was observed.

optimal strategy. For the test scenarios, this was the other way around: the SSD group chose more sub-optimal strategies after the transfer while the control group chose more optimal strategies. For the most difficult scenarios, the SSD group started out slightly worse than the control group, as can be seen in Figure 21. However, after the transfer, the SSD group chose the optimal strategies more often than the control group. The trends discussed above can also be seen in Tables VII and VIII, where the largest numbers are highlighted. Table VII shows that both groups only performed better for one scenario category after the transfer. Table VIII shows that when comparing the two groups per experiment phase, the SSD group performed better during two of the scenario categories.

By assigning a score of 1 to every scenario where an optimal strategy was chosen and a score of 0 to every scenario where a sub-optimal strategy was chosen and summing these scores, the differences between the groups, experiment phases and scenario categories could be tested. Kruskal-Wallis tests revealed no significant differences between the groups (Training, Easy: H(1) = 2.250, p = 0.134; Training, Medium: H(1) = 0.135, p = 0.713; Training, Difficult: H(1) = 0.113, p = 0.737;

TABLE VII
NUMBER OF SCENARIOS (OUT OF 42) DURING THE TRAINING AND TEST PHASE WHERE AN OPTIMAL STRATEGY WAS CHOSEN

	SSD		No SSD	
	Training	Test	Training	Test
Easy	36	33	32	28
Medium	34	26	33	35
Difficult	32	36	35	33
Sum	102	95	100	96

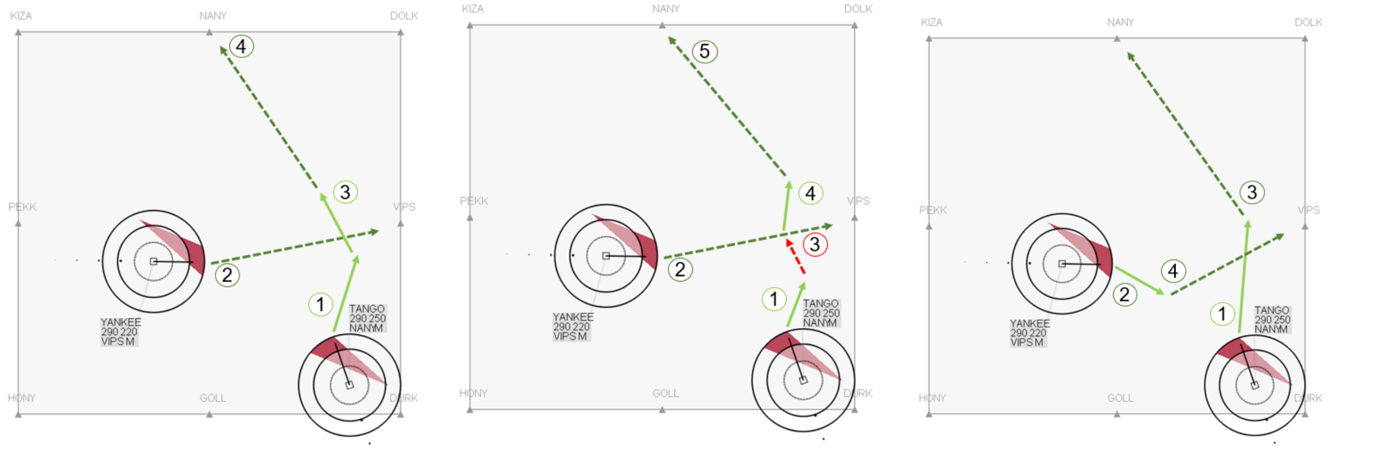
TABLE VIII
NUMBER OF SCENARIOS (OUT OF 42) PER PARTICIPANT GROUP WHERE AN OPTIMAL STRATEGY WAS CHOSEN

	Training		Test	
	SSD	No SSD	SSD	No SSD
Easy	36	32	33	28
Medium	34	33	26	35
Difficult	32	35	36	33
Sum	102	100	95	96

Test, Easy: H(1) = 2.525, p = 0.112; Test, Medium: H(1) = 3.122, p = 0.077; Test, Difficult: H(1) = 0.275; p = 0.600). When comparing the training and test phases per group for each of the categories, Wilcoxon Signed Rank Tests confirmed a significant difference for the control group between the training and test phase results, but only for the easy scenario category (*No SSD, Easy: z = -2.000, p = 0.046,* SSD, Easy: z = -1.342, p = 0.180; No SSD, Medium: z = 0.707, p = 0.480; SSD, Medium: z = -1.469, p = 0.142; No SSD, Difficult: z = -0.816, p = 0.414; SSD, Difficult: z = 1.265, p = 0.206). Finally, when looking at whether the scenario categories had an effect on the strategy scores, Friedman tests revealed that this was the case during the test phase for the control group (SSD, Training: $\chi^2(2)$ = 0.452, p = 0.798; No SSD, Training: $\chi^2(2)$ = 0.941, p = 0.625; SSD, Test: $\chi^2(2)$ = 3.556, p = 0.169; *No SSD, Test:* $\chi^2(2)$ *= 7.316, p = 0.026*). After applying a correction during a post-hoc test, the adjusted significant values, however, showed that no significant difference was present between the scenario categories.

Next to analyzing strategies based on being optimal or sub-optimal, the strategies could be further analyzed by dividing them into categories based on the number of aircraft in a solution and the number of corrections. Figure 22 shows three different optimal strategies that each contain one correction. The number of corrections was in this case defined as the number of heading commands that was given after the first command. While corrections only consisted of heading commands, the first command given to the first aircraft could be a heading or a DCT command.

Both Figures 22a and 22b show an optimal strategy where only one aircraft is used in the solution, that is additionally given one correction. The DCT commands at the end of the scenario are not counted as corrections, as each scenario needed to be completed by giving both aircraft a DCT command. The figures illustrate that although the strategies are still considered optimal because the faster aircraft passes in front of the slower aircraft, the corrections can be of a different nature. For example, the correction given in Figure 22a is a correction given to increase the efficiency of the solution: The controller



(a) (b) (c)

Fig. 22. Schematics of the optimal (slow aircraft is vectored behind fast aircraft) strategies with one correction. (a) Optimal strategy with one efficiency-related correction. (b) Optimal strategy with one safety-related correction. (c) Optimal strategy with two aircraft and one correction.

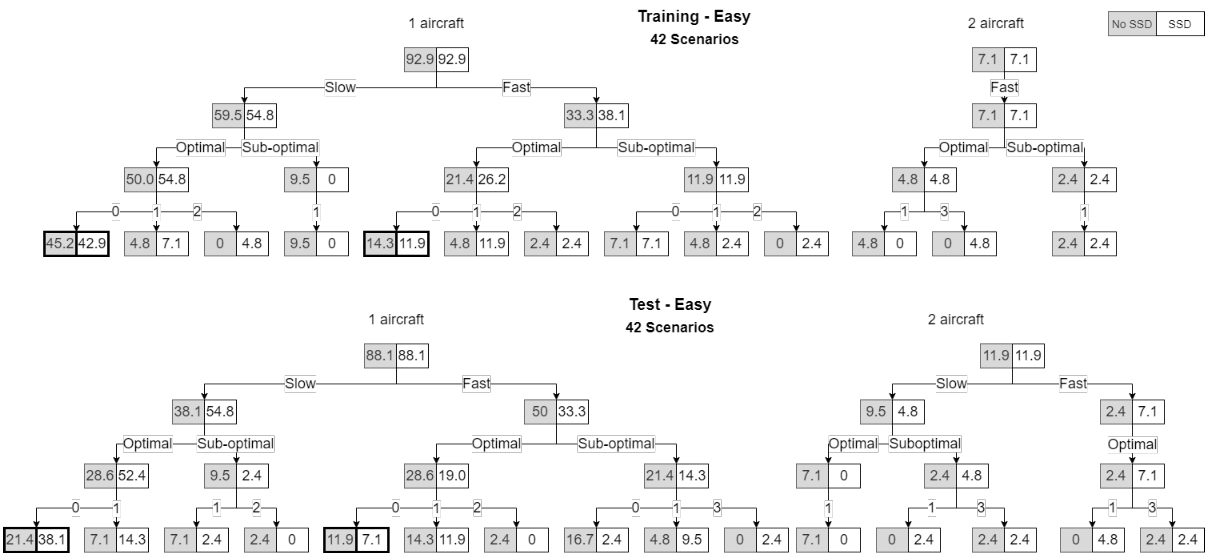


Fig. 23. Strategy trees for the easy scenario category, showing the percentages of scenarios (out of 42 scenarios) where a certain strategy was observed, based on the number of aircraft involved in a scenario, aircraft choice, solution and number of corrections.

realized that the first heading command (1) was too extreme and noticed that there is room to vector the aircraft a little bit more in the direction of the COPX already (3). The correction given in Figure 22b is a result of misjudging the situation and is thus a safety-related correction: After resolving the conflict by giving the faster aircraft its initial (1) heading command, the controller gave this same aircraft a DCT command (3) too soon which introduced a new conflict, which then had to be resolved or corrected for by giving an extra heading command (4). Figure 22c shows an example of a strategy that uses two aircraft to resolve the conflict situation. Again, every heading command given after the first command to the first aircraft is considered a correction. The first command to the second aircraft is in this case thus also considered a correction.

Grouping the strategies based on the number of aircraft involved in the solution, the aircraft choice, the number of corrections and being an optimal or a sub-optimal solution, leads to the results shown in Figures 23 to 25 for the two-aircraft scenarios. Similar to the trees shown before, the white blocks represent the SSD group and the gray blocks represent the control group. The lowest level of the tree represents the number of corrections. The figures show that two sets of blocks in the lower level have been highlighted in bold. These two sets represent the 'most optimal solutions' (i.e., a strategy where 1 aircraft is used to resolve the situation and no corrections were required, such as in Figures 18a and 18b.

Looking at the percentage of scenarios that were solved with one aircraft and the number of optimal strategies with zero corrections in Figure 23, the strategy tree for the easy scenarios, it can be seen that both groups performed similarly during the training. After the transfer, both groups chose sub-optimal solutions more often, as was also seen before.

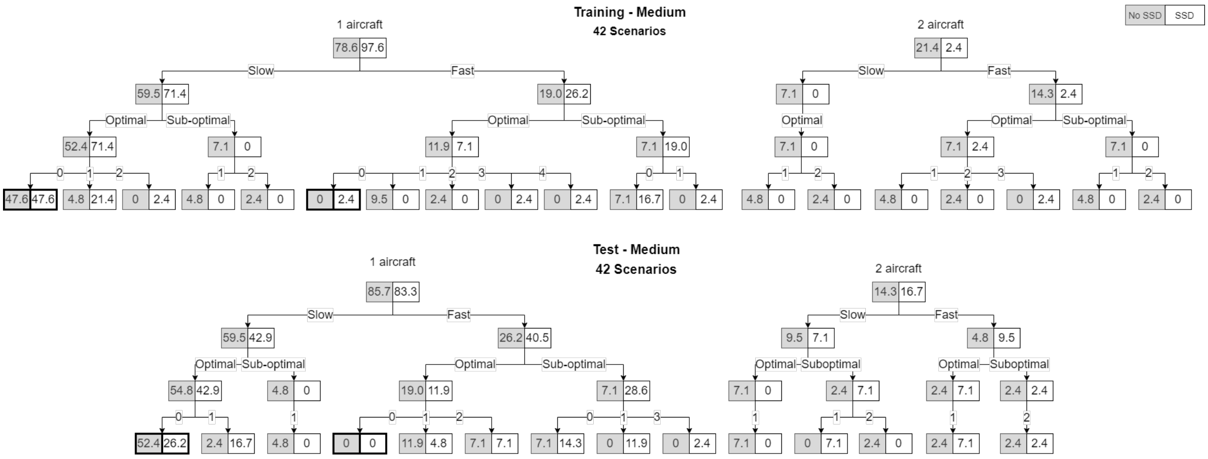


Fig. 24. Strategy trees for the medium scenario category, showing the percentages of scenarios (out of 42 scenarios) where the respective strategy was observed, based on the number of aircraft involved in a scenario, aircraft choice, solution and number of corrections.

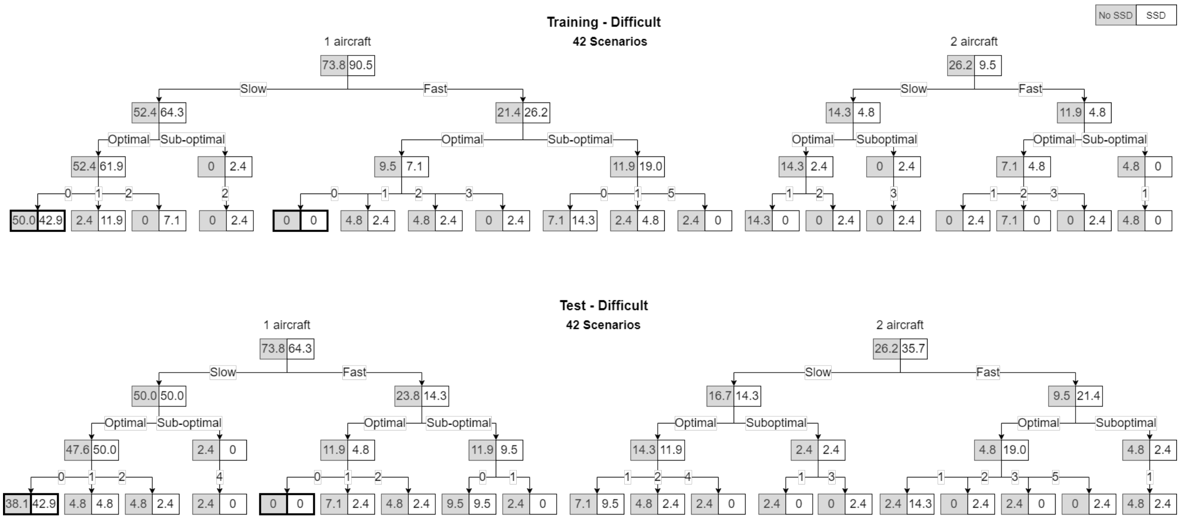


Fig. 25. Strategy trees for the difficult scenario category, showing the percentages of scenarios (out of 42 scenarios) where the respective strategy was observed, based on the number of aircraft involved in a scenario, aircraft choice, solution and number of corrections.

Additionally, the control group made less use of the most optimal strategy (21% compared to 45% during the training) and required more corrections during optimal strategies. The SSD group also made slightly less use of the most optimal strategy but did make a lot more use of optimal strategies involving one correction. A large increase can furthermore be seen in the number of times a fast aircraft was chosen first by the control group. In the one-aircraft strategy tree, this number increased from 33% to 50% after the transfer.

Figure 24 shows that during the medium-level scenarios, participants from the SSD group used only one aircraft for their strategy or solution in 97.6% of the training scenarios, meaning that in only 1 out of 42 scenarios a strategy was used which involved two aircraft. Although the SSD group more often chose an optimal strategy containing one (slow) aircraft in the solution, it can also be seen that they required a correction more often. This could be caused by the SSD as it shows when the aircraft is conflict free and can already be given an extra heading command to reduce additional track miles. During the test, a similar trend is seen. Although the SSD group chose an optimal strategy less often, they still applied one correction during their strategy more often than the control group, even though the SSD was not available anymore. Additionally, the SSD group also chose the fast aircraft for their strategy more often after the transfer. This in turn led to more sub-optimal strategies where the fast aircraft was vectored behind the slower aircraft. Finally, Figure 25 shows the strategy tree for the difficult scenarios. Again, during the training scenarios, the participants from the SSD group used only one aircraft in 90% of the scenarios. Additionally, these participants again used one correction in their optimal strategy more often and made more corrections in general when compared to the control group. Furthermore, choosing the faster aircraft led to a sub-optimal solution more often. For both groups, choosing the faster aircraft never led to an optimal solution with no corrections, as could also already be seen for the medium scenarios. This was the case during both the training and the test phase. During the test phase,

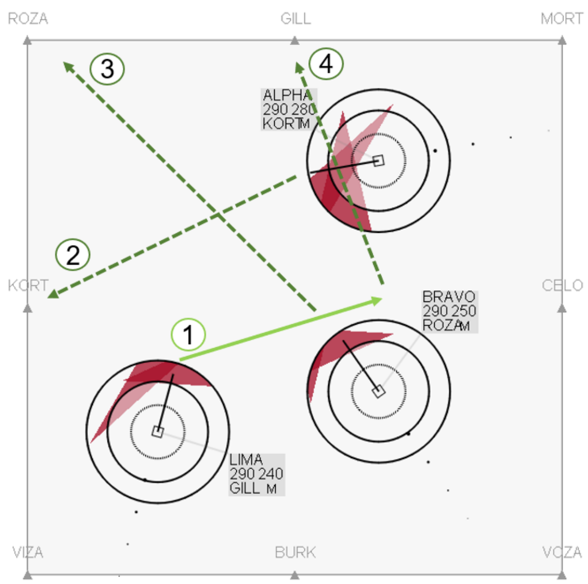


Fig. 26. Schematic of an optimal strategy in a three-aircraft scenario: the slowest aircraft is vectored behind the faster aircraft.

it was found that the participants from the SSD group again used 2 aircraft in their solutions more often while the number remained the same for the control group.

Comparing the three strategy trees shows that while choosing the faster aircraft first more often led to an optimal solution for both groups in the easy scenarios, this was reversed for the difficult scenarios, as choosing a faster aircraft in this case was more often followed by a sub-optimal strategy. Easier scenarios could furthermore more often be solved with less aircraft involved and less corrections compared to the difficult scenarios.

2) Participant Strategies in Three-aircraft Scenarios: A similar analysis was done for the scenarios containing three aircraft. Figure 26 shows an example of a scenario containing three aircraft with three different speeds along with one of the optimal strategies for this scenario. It can be seen that BRAVO and LIMA are in conflict, whereas ALPHA is not in conflict with the other two aircraft as it was placed at this location to only strengthen the optimal solution. The optimal solution for this type of scenario was defined similarly to the two-aircraft scenarios: the slower aircraft should pass the faster aircraft at the rear. In this case, this means that LIMA will pass BRAVO at the rear, which will in turn pass ALPHA at the rear. Compared to the two-aircraft scenarios discussed before, introducing a third aircraft in the scenario introduces many new strategies in approaching a scenario. The example here shows a strategy where the slowest aircraft is chosen for the first command, but it was also possible to solve this scenario by giving the second slowest aircraft (BRAVO) a heading command first. All three-aircraft scenarios could be solved by giving one command to either the slowest or second-slowest aircraft. These strategies were thus counted as the 'most optimal strategies' (in terms of efficiency) in the three-aircraft scenarios.

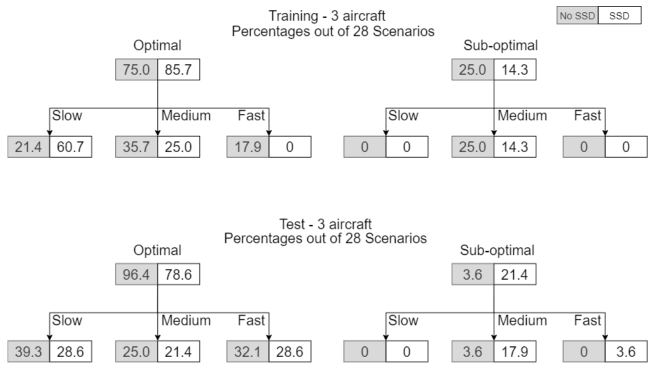


Fig. 27. Strategy tree for the three-aircraft scenarios, showing the percentages of scenarios (out of 28 scenarios) where the respective strategy was observed.

The three-aircraft scenarios did not contain different levels of conflict difficulty and hence only one strategy tree was created that looks slightly different due to the extra aircraft. The results can be seen in Figure 27. Similar to the two-aircraft scenarios, both groups chose an optimal strategy more often compared to sub-optimal strategies, which was also confirmed by the Wilcoxon Signed Rank Test (*SSD Training:* $z = -3.162$, $p = 0.002$; *No SSD Training:* $z = -2.111$, $p = 0.035$; *SSD Test:* $z = -2.530$, $p = 0.11$; *No SSD Test:* $z = -3.606$, $p = 0.000$).

Two interesting trends can furthermore be seen. First, during the training and test phase, participants from both groups showed to mostly use the medium aircraft in case of a sub-optimal solution. This was expected as sending the slower (LIMA) and the faster (ALPHA) aircraft left (sub-optimal strategy) resulted in a large conflict, which can also be seen when looking at the aircraft SSDs. Second, while the SSD group performed worse during the test compared to the training, the control group performed much better. In fact, for only 1 out of 28 scenarios, a sub-optimal solution was chosen. The difference between training and test performance, was however not found to be significant (SSD: $z = -0.707$, $p = 0.480$; No SSD: $z = 1.857$, $p = 0.063$). Furthermore, for the optimal solutions, participants from both groups did not seem to have a preferred aircraft to give a first command to.

Next to analyzing the three-aircraft strategies based on being optimal or sub-optimal, these strategies could also be further analyzed by dividing them into categories based on the number of aircraft in a solution and the number of corrections, similar to the two-aircraft scenarios. An extra aircraft in the scenarios naturally meant that an extra aircraft could be involved in the strategy. The results of the strategy breakdown are shown in Figure 28. It can immediately be noticed that the SSD group never chose the faster aircraft as the first aircraft during the training scenarios, nor did they use three aircraft in their solutions. While this group only used 1 aircraft in the majority of their solutions, the control group used two or three aircraft in the majority of their solutions. With respect to the aircraft choice, it can be seen that while the SSD group preferred the slower aircraft in their solutions, the control group actually preferred to give the first command to the medium-speed aircraft.

During the training phase, the SSD group chose the most

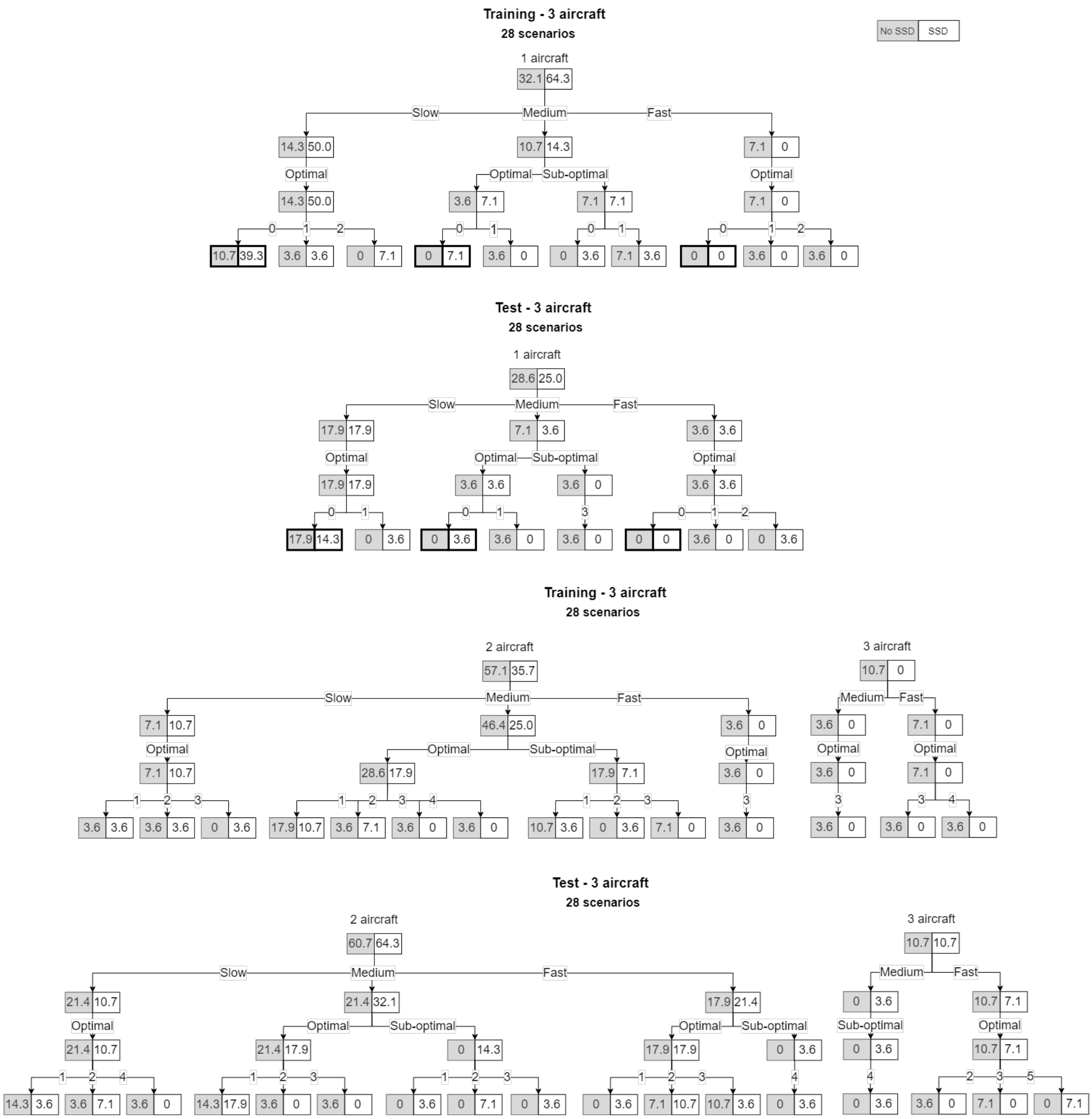
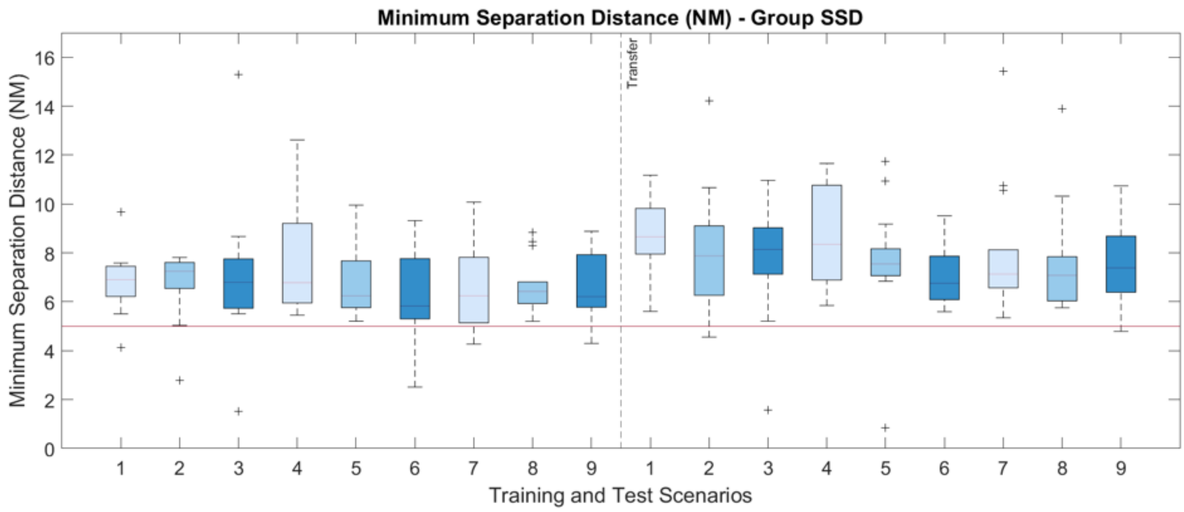


Fig. 28. Strategy trees for the three-aircraft scenarios, showing the percentages of scenarios (out of 28 scenarios) where the respective strategy was observed, based on the number of aircraft involved in a scenario, aircraft choice, solution and number of corrections.

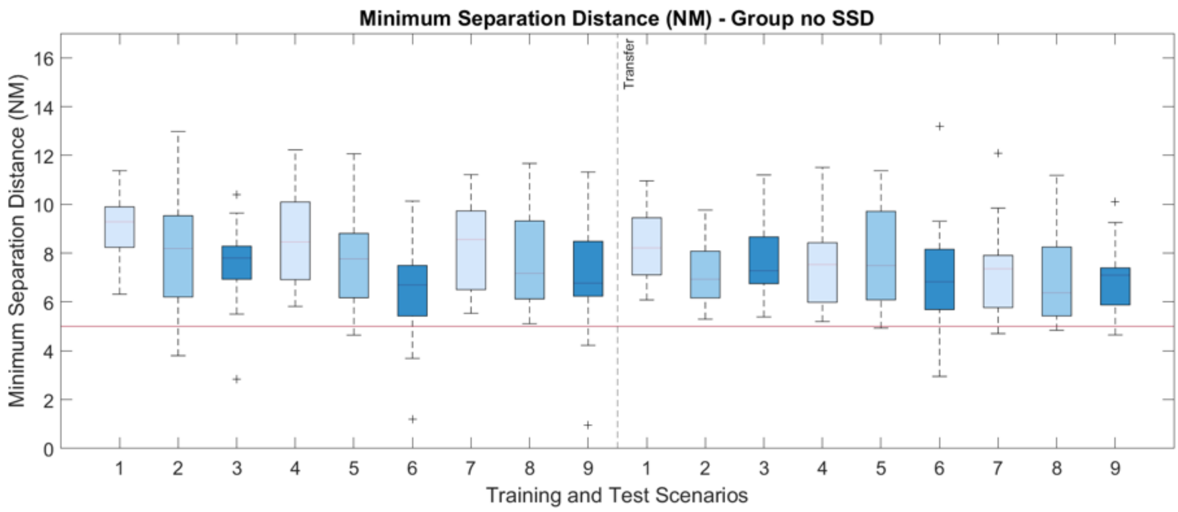
optimal solution, containing 1 aircraft and requiring 0 corrections, in 46.4% of the scenarios compared to only 10.7% for the control group. During the test phase, the control group showed improved use of strategies in that they more often chose an optimal strategy as was also discussed before. The SSD group, however, showed a very different use of strategies after the removal of the SSD. While this group was able to solve the majority of the scenarios with one aircraft during the training, the group now required two or three aircraft in the majority of their strategies. It can also be seen that the same group chose the fastest aircraft several times during the test phase while they never chose to use it during the training phase. Additionally, the SSD group made a larger number of corrections during the test phase. While the SSD group had a seemingly larger preference for the medium-speed aircraft for their strategies during the training phase, this preference seemed to have shifted away from this aircraft towards both the slower and faster aircraft in the test phase.

C. Control Task Performance

During the experiment, participants were instructed to first focus on safety during all scenarios by minimizing the occurrence and duration of STCA alerts before focusing on



(a)



(b)

Fig. 29. Minimum Separation Distance (NM) results displayed in the original scenario order (a) for the SSD group and (b) for the control group.

efficiency. In terms of traffic safety, performance was measured by the number of experienced LoS's as well as the minimum separation distance. In terms of efficiency, performance was measured by means of the additional track miles from the initial track as well as the participant response times and type of commands.

1) Minimum Separation Distance: Figure 29 shows the minimum separation distance for the two-aircraft scenarios for both participant groups in the original experiment order. The horizontal line that is drawn in each of the plots represents the minimum required separation distance of 5 NM. Any data below this line indicate a violation of an aircraft's PZ and hence a LoS. It can be seen that the SSD group operated more efficiently by maintaining smaller safety margins during the training while the control group maintained larger safety margins overall. This was expected as the SSD was available during the training phase and the SSD group was thus better able to operate on the edges of the solution space as all constraints were visualized by the SSD.

This is probably also what caused some of the LoS's for the SSD group during the training phase: participants could have become more focused on efficiency, the second objective, and thereby tried to operate on the edges of the free solution space. While doing so, a heading change might have been slightly too extreme and thereby vectored the aircraft just inside the conflict area, resulting in a 'minor' LoS. Participants from both groups could additionally use the half-PZ circles to determine whether a LoS would take place or not. Minor LoSs could thus also have been the result of an incorrect estimate, in which participants thought the circles would not cross, whereas they actually did. At the time of the crossing, it was then already too late to correct for the LoS.

The LoS's with a minimum separation distance of 3 NM or less were considered severe LoS's and were investigated separately. It was found that apart from the LoS's depicted by the outliers at Scenarios 2 and 3 during the training in Figure 29a, the LoS's experienced by the SSD group were the result of input errors. For example, participants forgot to select the

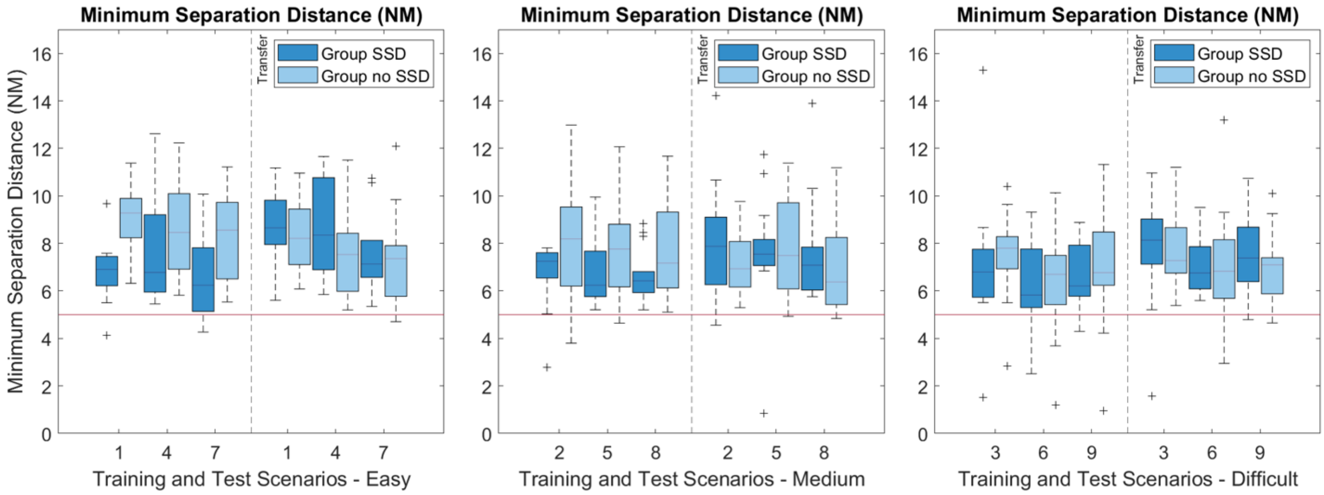


Fig. 30. Minimum Separation Distance (NM) results per scenario category.

HDG button or confused DCT with 'direction' and clicked this command button to give a heading command rather than clicking the HDG button. The outliers at Scenarios 2 and 3 were the result of participants waiting too long with executing a command, even though their strategy for these scenarios was the optimal strategy. Waiting too long with giving a command resulted in having to operate too close to the sector border which left very little space to operate. A similar observation was made for the third training scenario of the control group in Figure 29b. This outlier was the result of waiting too long with giving a command, whereas the outliers at Scenarios 6 and 9 of the training and Scenario 6 of the test were the result of input errors. Although far from a violation of the PZ, the four extreme outliers at 13 NM or more for the SSD group were also looked into. Three of these outliers were caused by the same participant that chose a sub-optimal strategy, resulting in one aircraft having to make a very large detour (e.g., fly in a circle) before it was cleared to its COPX. The fourth outlier was additionally caused by choosing a sub-optimal strategy, resulting in one of the two aircraft having to fly in a circle.

As explained before and as shown in Tables III and IV, the dynamic scenarios consisted of three levels of difficulty based on the conflict geometry and had to be performed by participants in three sets of scenarios ranging from easy to difficult. The three different colors in Figure 29 thus represent the three levels of difficulty. Figure 29b shows a clear 'sawtooth pattern' on the left-hand side of this figure, which corresponds to the three sets of scenarios ranging from easy to difficult. This same pattern can also be seen during the test phase of both groups, although less evident. Comparing the training phase of both groups shows that almost no sawtooth pattern is present for the SSD group. This is considered to be a direct result of having the SSD available as this reduces the effect that the scenario difficulty has on the participant by showing the complete range of solutions. Participants can thus choose a solution close to the 5 NM boundary, resulting in a relatively constant performance when compared to having no SSD available (control group). For the rest of the analysis, the categories are evaluated separately rather than in their original sequence, meaning that, for example, for the easy scenarios, Scenarios 1, 4 and 7 are grouped together, which can also be seen in the following plots.

Figure 30 shows three plots for the minimum separation distance between aircraft, corresponding to each of the three scenario categories. The plots each display the performance of both groups during the training and during the test phase. The first plot, for the scenarios of the first and easiest category where the rule of thumb was most obvious, shows two very evident learning curves. This figure again shows that the SSD group operated more efficiently by maintaining smaller safety margins during the training, while the control group maintained larger safety margins due to the (un)availability of the SSD. It can furthermore be seen that during the test phase, when the SSD is no longer available, the SSD group is maintaining larger safety margins than the control group. They experienced a large 'setback' or 'delta' in performance level. After the transfer, both groups show similar learning curves. It can be seen that the SSD group ends at a similar performance as the control group, but slightly worse than their initial performance during the training, while the control group shows a (relatively) continuous learning curve through the transfer.

For the medium-level scenarios, similar learning curves and deltas can be observed but with a smaller difference between the groups. It can furthermore be seen that more violations of the minimum separation distance occurred in the control group. The two LoS outliers observed in the SSD group were the result of waiting too long to execute a command and an input error, respectively, as was also explained before.

The last plot shows the results for the most difficult category of scenarios. What can first be noticed is the considerable number of LoS's by both groups during the training phase. During the test phase, the SSD group performed better when looking at the number of LoS's. It can furthermore be seen that both groups experience a delta in performance after the transfer rather than only the SSD group and that both groups operate

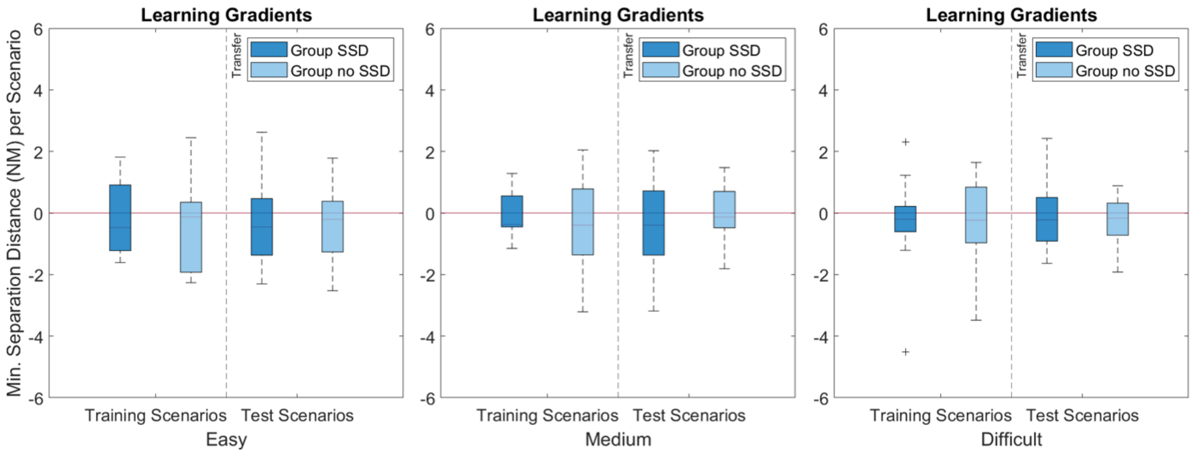


Fig. 31. Minimum Separation Distance learning gradients per scenario category.

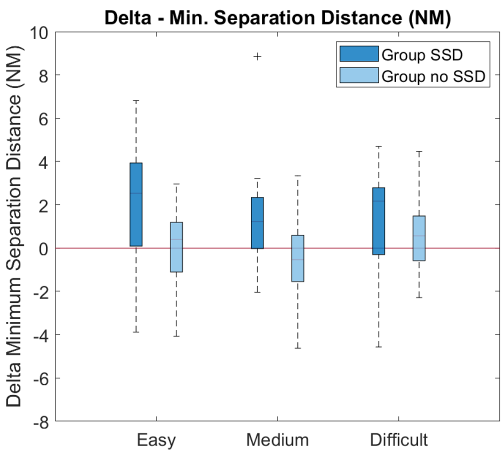


Fig. 32. Delta values for the minimum separation distance per group per scenario category, where the delta is defined as the difference between the first test scenario value and the last training scenario value.

close to the required minimum separation. The difference with the separation distance that participants operated at during the easier scenarios as well as the difference in number of LoS's can be attributed to the conflict geometry and the initial positions of the aircraft. For the difficult scenarios, aircraft are more likely to cross each other at a smaller distance due to the smaller initial conflict angle.

Figure 32 shows the delta values for both groups for each of the scenario categories. The figure shows that the SSD group overall experienced larger deltas. Kruskal-Wallis tests confirm that participants from the SSD group experienced a significantly larger delta during the easy (*H(1) = 4.664, p = 0.031*) and medium scenarios (*H(1) = 4.864, p = 0.027*) (Difficult: H(1) = 1.022, p = 0.312). Additionally, the scenario categories did not significantly influence the magnitude of the delta (SSD: $\chi^2(2)$ = 1.857, p = 0.395, No SSD: $\chi^2(2)$ = 1.000, p = 0.607).

Looking at Figure 30, different learning curves and hence different learning gradients can be observed. For each participant, the learning gradients during the training and test phase have been determined by applying linear regression, as there were only three data points per phase per participant. The resulting learning gradient values are shown in Figure 31, where more negative values indicate steeper (positive) learning curves. Kruskal-Wallis tests revealed that there was no significant difference present between the groups when looking at the different experiment phases for each scenario category (Training, Easy: H(1) = 0.135, p = 0.713; Test, Easy: H(1) = 0.019, p = 0.890; Training, Medium: H(1) = 0.357, p = 0.550; Test, Medium = 0.103, p = 0.748; Training, Difficult: H(1) = 0.019, p = 0.008; Test, Difficult: H(1) = 0.008, p = 0.927). Additionally, when looking at the learning gradients per experiment phase and the effect of the scenario categories on the gradients, no significant difference was found either (SSD, Training: $\chi^2(2)$ = 0.429, p = 0.807, No SSD Training: $\chi^2(2)$ = 1.857, p = 0.395, SSD, Test: $\chi^2(2)$ = 0.143, p = 0.931, No SSD Test: $\chi^2(2)$ = 0.571, p = 0.751). Finally, no significant difference was found between the training and test phases per group per scenario category (SSD, Easy: z = -0.282, p = 0.778; No SSD, Easy: z = 0.094, p = 0.925; SSD, Medium: z = -0.031, p = 0.975; No SSD, Medium: z = 0.596, p = 0.551; SSD, Difficult: z = -0.471, p = 0.638; No SSD, Difficult: z = 0.094, p = 0.925).

2) *Loss of Separation:* Figure 30 already showed that more LoS's occured at the most difficult level and that the LoS's in the medium category were mostly caused by participants from the control group. Figure 33 shows the number of LoS's per two-aircraft scenario throughout the course of the experiment. What stands out is the high number of LoS's in the second half of the test by the control group. A possible explanation is that this is the point where the control group participants became more focused on efficiency and tried to discover the boundaries of the solution space in which they could safely operate, the same point that the SSD group had already reached during the training. Table IX additionally shows the number of LoS's per participant per group during both the training and test phase. The values that are highlighted in gray indicate the smallest number of LoS's when comparing the training and the test phases within a group. While the control group mostly

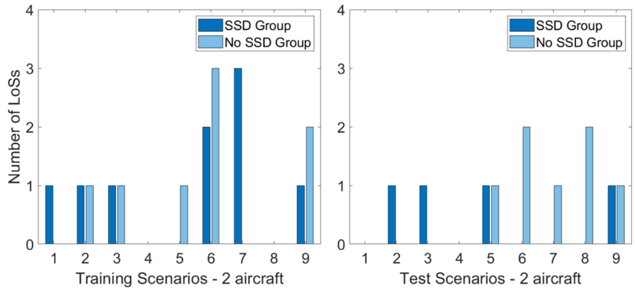


Fig. 33. Number of LoS's per Two-aircraft Scenario for the Training and Test phase.

TABLE IX
Number of LoS's per Participant

	SSD		No SSD	
Participant	Training	Test	Training	Test
1, 15	0	0	0	0
2, 16	1	0	2	1
3, 17	0	0	0	0
4, 18	0	1	0	0
5, 19	1	0	0	0
6, 20	0	1	1	0
7, 21	0	0	0	0
8, 22	1	0	0	1
9, 23	0	0	0	0
10, 24	2	1	0	0
11, 25	1	1	1	2
12, 26	0	0	1	1
13, 27	1	0	0	0
14, 28	2	0	3	2
Sum	9	4	8	7

performed similarly during the training and test phase, the SSD showed a strong improvement in performance as less participants caused less LoS after the transfer when being compared to before the transfer. The differences between the groups (Training: H(1) = 1.808, 0.179; Test: H(1) = 0.003, p = 0.957) and the differences between the two experiment phases (SSD: z = 0.000, p = 1.000; No SSD: z = 1.100, p = 0.271) were, however, not significant.

3) Initial Track Deviation: For each scenario performed by the participants, the track deviation from the initial track was measured for all aircraft after which it was averaged over the number of aircraft present in that scenario. As the aircraft were never aligned with their exit waypoints at the start of the scenarios, they would thus always travel additional miles, compared to their initial tracks. Deviating earlier from this initial track, thus resulted in more additional track miles.

The results for the training and test phase scenarios are shown per category in Figure 34. A positive trend or learning curve can be observed for both groups in the plots for the easy and difficult scenarios. Additionally, there is a large difference in values noticeable between the three scenario categories. This difference can be explained by the conflict geometry. Smaller conflict angles in the difficult scenarios require more extreme heading changes to prevent conflicts, resulting in larger deviations from the original track, compared to large conflict angles in the easy scenarios where only a small heading deviation was required to prevent the conflict from happening. The medium-level scenarios show a different trend with a reasonably large dip in track deviation for the second scenario of this category, meaning that aircraft in this scenario remained on their initial tracks for a longer time and/or less extreme heading changes were given to the aircraft, resulting in smaller track deviations. In an effort to explain this difference, the scenario configurations of the three scenarios of the medium category were investigated as well as the response times and heading values. However, no explanation was found.

Similar to the data for the minimum separation distance, learning curves can be observed as well as deltas in performance. As stated before, the trends that can be observed in the figures are positive trends, which can be caused by multiple factors. First, the positive trend could be caused by the response time. When participants let aircraft fly on their initial track for a longer period of time, this results in a smaller track deviation. A positive trend could thus indicate a difference in response times, indicating that the response time decreased for the later scenarios. Figure 15a shows indeed that mostly negative trends are present during the training and test phases of each scenario category, indicating a decreasing response time, as was expected when looking at the track deviation data.

Second, the positive trend could be caused by the extremity of heading commands. Larger or more extreme heading commands result in larger track deviations from the initial track. While it is expected that participants use smaller heading deviations as their experience increases, a longer response time could result in more extreme (required) heading changes to solve possible conflicts. To accurately estimate the effects of the heading commands and response times, a more in-depth analysis is required per participant per scenario.

Another trend that stands out is the difference in range of track deviation values when comparing the different scenario categories. The track deviation was smallest and increasing a lot for the easy scenarios and largest and most constant for the difficult scenarios, as a result from the conflict geometries. Smaller conflict angles leave less (efficient) operating room and hence result in more extreme heading commands leading to a larger track deviation.

Figure 35 shows the delta values for the track deviation for each of the scenario categories. It can be seen that the majority of the participants of the SSD group experienced positive deltas during the medium and difficult scenarios, indicating a continuous performance curve rather than a curve interrupted by a setback in performance.

Both groups experienced larger deltas in performance during the easy scenarios and smaller deltas (or more positive) as the scenarios got more difficult, indicating a continuous learning curve. The SSD did not significantly influence the delta (Easy: H(1) = 1.319, p = 0.251; Medium: H(1) = 3.549, p = 0.060; Difficult: H(1) = 3.049, p = 0.081). However, looking at the effect of the scenario category on the delta, a significant difference was found for the control group (SSD: $\chi^2(2)$ = 5.286, p = 0.071, *No SSD:* $\chi^2(2)$ *= 7.000, p = 0.030*). Adjusted significant values from the post-hoc tests subsequently showed that only the difference between the easy and difficult scenarios was significant.

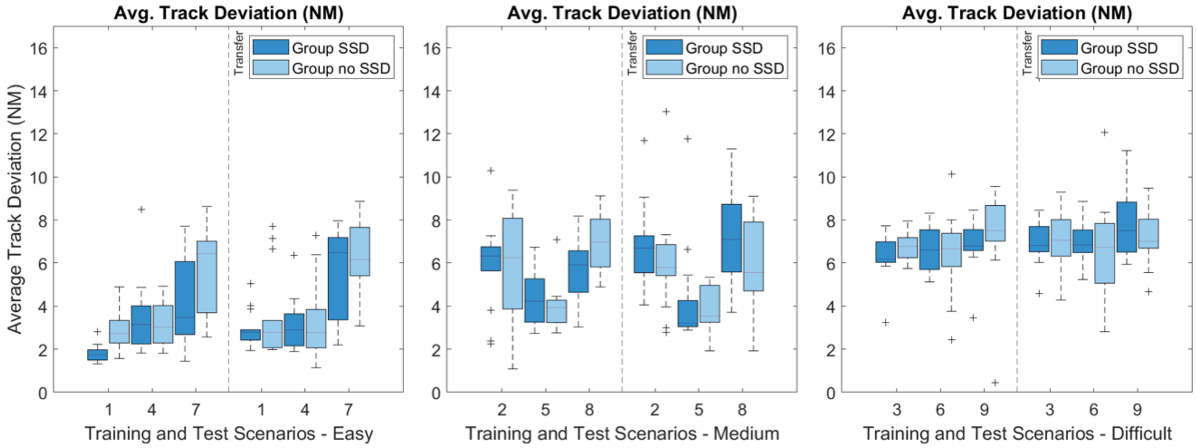


Fig. 34. Average Track Deviation (NM) results per scenario category.

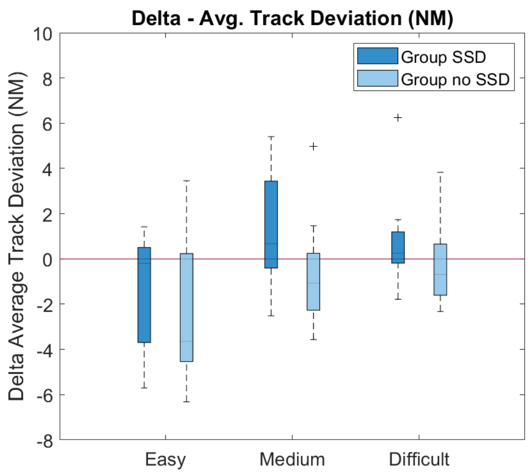


Fig. 35. Delta values for the track deviation per group per scenario category, where the delta is defined as the difference between the first test scenario value and the last training scenario value.

The learning gradients for the track deviation were determined similarly to the gradients for the minimum separation distance. Figure 36 shows the learning gradients for the track deviation for both the training as well as the test scenarios. The figure shows that steeper learning curves are generally observed for easier scenarios compared to the more difficult scenarios where the values lie closer to zero and show a smaller spread, again related to the scenario conflict geometries. Looking at the differences between the groups during the training and test phases for each of the scenario categories, Kruskal Wallis Tests revealed no significant differences (Training, Easy: H(1) = 0.887, p = 0.346; Test, Easy: H(1) = 0.076, p = 0.783; Training, Medium: H(1) = 0.684, p = 0.408; Test, Medium = 0.211, p = 0.646; Training, Difficult: H(1) = 0.540, p = 0.462; Test, Difficult: H(1) = 0.304, p = 0.581). No significant difference was found between the learning gradients in the training and test phases per group per scenario category (SSD, Easy: z = 0.408, p = 0.683; No SSD, Easy: z = 0.220, p = 0.826; SSD, Medium: z = 0.157, p = 0.857; No SSD, Medium: z = -1.287, p = 0.198; SSD, Difficult: z = -0.659, p = 0.510; No SSD, Difficult: z = -1.601, p = 0.109).

Friedman Tests furthermore confirmed that the scenario categories did indeed have a significant effect during the training and test phases for both groups (*SSD Training:* $\chi^2(2) = 8.143$, *p = 0.017, SSD Test:* $\chi^2(2) = 10.429$, *p = 0.005, No SSD Training:* $\chi^2(2) = 10.714$, *p = 0.005, No SSD Test:* $\chi^2(2) = 9.571$, *p = 0.008*). This was followed up by post-hoc tests. For the SSD group, significant differences were found during the training as well as during the test phase. During the training phase, a significant difference was found when comparing the easy scenarios to both the medium scenarios. For the test phase, a significant difference was found between the easy and difficult scenarios. The control group showed slightly different results. During the training phase, significant differences were found when comparing the easy scenarios to the medium as well as the difficult scenarios. For the test phase, a significant difference was found between the easy and medium scenarios. The difference between the medium and difficult scenarios was never significant.

4) Heading Commands: Figures 37a to 37c show the number and type of first chosen heading solutions for the different experiment phases of each of the scenario categories (as displayed in Figure 12). It should be noted that the figures shown here contain all solutions per scenario category, meaning that each figure contains the solutions for three scenarios (42 in total) with slightly different characteristics as well as the solutions of both slow and fast aircraft.

For all figures, one or two strong preferences for a certain heading can be noticed. These correspond to the chosen solutions for the slow and fast aircraft, respectively. In order to accurately determine the effect of the chosen heading solutions on the track deviation, the scenarios should each be analyzed separately for the different aircraft choices rather than performing an analysis per category. The figures shown here are thus used for detecting any trends in the data that might have influenced the track deviation.

Next to the chosen heading values, the figures contain the DCT direction for both aircraft (green continuous lines) and the approximate correct solution directions (green areas),

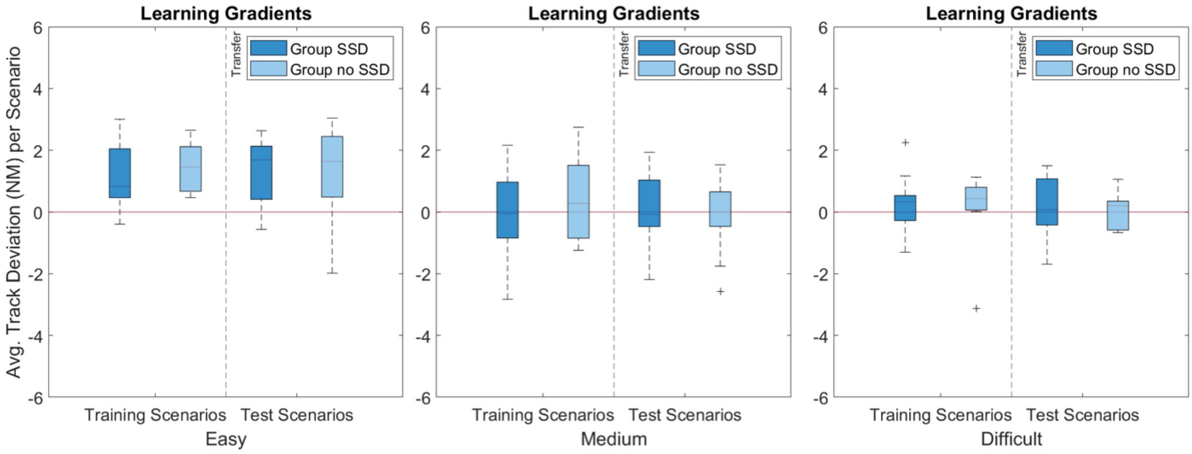
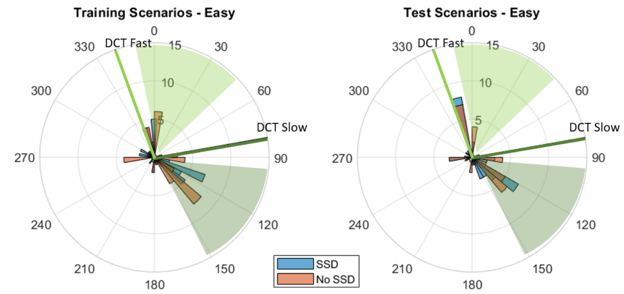
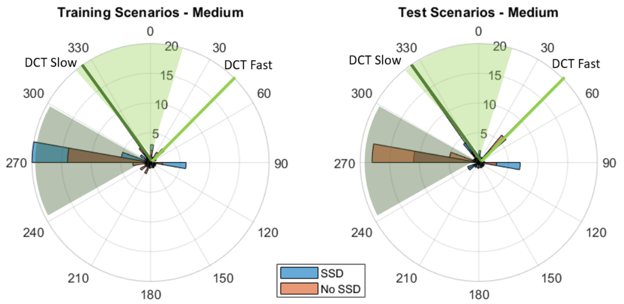


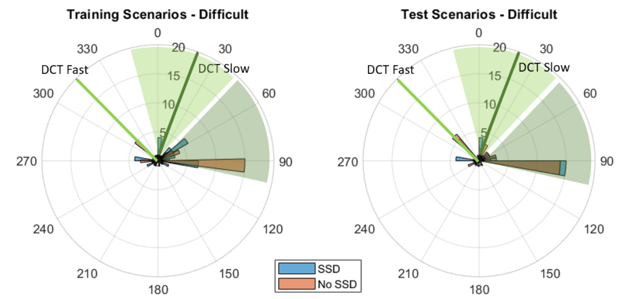
Fig. 36. Track Deviation learning gradients per scenario category.



(a)



(b)



(c)

Fig. 37. Compass roses indicating the chosen heading solutions per experiment phase for (a) the easy scenarios, (b) the medium scenarios, and (c) the difficult scenarios.

where light green corresponds to the faster aircraft and dark green corresponds to the slower aircraft. The correct solution areas approximately indicate the heading range in which the first command should lie in order for it to be counted as a correct heading leading toward an optimal strategy. When comparing the distance between the DCT lines and their corresponding correct solution areas, the difference in conflict geometry can be seen. The smaller gap in the easier scenarios hints at the larger CPA and larger conflict angle when compared to the difficult scenarios where this gap is larger. This larger gap thus also indicates that larger heading changes are required in order to resolve the conflict, explaining the larger track deviation values for the difficult scenarios when compared to the easy scenarios.

Comparing Figure 37a for the easy scenarios to Figures 37b and 37c for the medium and difficult scenarios, respectively, shows that participants chose less extreme headings (i.e., heading directions close to the DCT direction) during the easier scenarios. For the medium and difficult scenarios, the chosen heading solutions are found to be further away from the DCT direction and are thus considered more extreme, as was also expected when keeping the conflict geometries and track deviation results in mind. Another difference that can be seen when comparing the easy scenarios to the medium and difficult scenarios is the spread in chosen heading solutions. This spread is larger for the easy scenarios, whereas the medium and difficult scenarios show a strong preference for one particular heading.

What furthermore stands out in Figure 37a is that during the training, participants seem to make use of more different solutions when compared to the test phase, as a result of the SSD being removed after the transfer. Participants are not able to see the solution space anymore and will thus be more conservative when choosing heading commands. For the medium scenarios, participants from the SSD group seem to have a large preference for heading 270 during the training. During the test, this preference seems to have shifted away toward one of the other main compass directions (90) and to giving DCT commands as a first command. This shows a change in strategy as well as a more conservative approach,

TABLE X
DIVISION OF ANSWERS FOR THE FINAL STRATEGY QUESTION AND RULE OF THUMB QUESTION

	Strategy			Rule of Thumb		
Answer Category	SSD	No SSD	Sum	SSD	No SSD	Sum
1: corect: slow aircraft behind fast aircraft	7	5	12	8	2	10
2: ambiguous: 1 preferred aircraft, no preferred direction	3	2	5	0	3	3
3: incorrect: slow aircraft in front of fast aircraft	4	7	11	6	9	15
Sum	14	14	28	14	14	28

as participants chose a relatively extreme heading compared to the minimum heading that was required to resolve the conflict. Participants thus prefer to use larger or more extreme heading commands to ensure safety rather than taking efficiency into account, as was also assigned to them. Finally, for the difficult scenarios, Figure 37c shows that most participants show a strong preference for headings 90 - 100. Again, a larger variation in chosen solutions is seen before the transfer when comparing this to the variation after the transfer, as well as an even more conservative approach.

D. *Results of Questionnaire: Strategy*

In analyzing the final strategies and rules of thumb that participants stated in the questionnaire, three categories were defined: (1) correct strategy / rule of thumb (i.e., vectoring the slower aircraft behind the faster aircraft), (2) ambiguous or slightly (in)correct strategy and (3) incorrect strategy (i.e., vectoring the faster aircraft behind the slower aircraft). The second category contains for example strategies where participants stated that they would divert the slow aircraft or pick the slowest aircraft to change the heading. However, they did not state anything about the direction of their solution which means that it is unsure whether they fully understood the rule of thumb but that they were at least starting to think in the correct direction.

As explained before, participants were explicitly asked for their strategy for detecting and resolving conflicts. The results for the answers that were given to the question on what strategy they used to resolve conflicts are shown in Table X on the left. It can be seen that half of the participants of the SSD group wrote down a correct strategy while, 5 participants of the control group answered correctly. When being asked about the rule of thumb (or implicitly asked about their strategy) for resolving conflicts that participants thought they had learned during the experiment, 8 participants of the SSD group wrote down the correct rule of thumb, while only 2 participants of the control group wrote down a correct answer. 3 participants from the control group were, however, thinking in the right direction, but did not answer completely correct.

It can be noticed that the largest values for the correct answer to both questions can be found in the column corresponding to the SSD group. The largest value for incorrect answers is found in the column for the control group for both questions. It was found that there was not a significant association between the type of display and whether or not participants answered the strategy question correct ($\chi^2(2)$ = 1.399, p = 0.605). However, a significant association was found between the display type and the answers to the rule of thumb question ($\chi^2(2)$ = 6.685, p = 0.034), supporting the third hypothesis. It should be noted that due to the small sample size, Fisher's Exact Test was applied, as there were two cells (for both questions) that contained an expected count less than five, indicating that the results of the Chi-Square test could be inaccurate.

VI. DISCUSSION

The goal of this paper was to investigate how EID could possibly promote or support discovery learning processes in the ATC domain through the application of an instructional design method. A custom-made CD&R training was developed to investigate whether an ecological interface could contribute to novices picking up an industry rule of thumb by means of discovery learning and to further explore the effects of this interface on novice knowledge, performance and strategy development. A between-participants experiment was performed with a group of 28 novice participants that were divided over two groups: the SSD group, that performed part of the experiment with the availability of the SSD, and a control group, that performed the same part without the SSD. The knowledge and performance of the two groups of participants was tested and compared before and after a transfer manipulation where the SSD support was unexpectedly removed.

The aim of this study was to further evaluate the short-term impact of ecological interfaces on performance, knowledge and strategy development of novices when being engaged in discovery learning and to overcome several limitations from previous research [10]: (1) intertwined learning curves or processes due to the fact that no instructional design method was applied, (2) participants' learning progress was not tracked or monitored, and (3) participants were told the learning objective or 'best practice' to solve the scenarios in advance due to which no strong effects of the display type were found.

In overcoming these limitations, several adaptations were made: (1) participants were guided through the experiment by means of a step-by-step script, designed according to the Four-Component Instructional Design model, in which several learning objectives have been defined and separated in order to separate the learning curves, (2) participants' learning progress could be monitored as a result of the setup of the script, and (3) participants were not told the 'best practice' in advance but had to obtain this knowledge by means of discovery learning, as it was implicitly integrated in the experiment and script design.

A. *Training Script Effectiveness*

First, by structuring the experiment according to the Four-Component Instructional Design model and guiding all participants through the experiment by means of the step-by-step script, it was expected that sequential occurrence of the learning curves would take place rather than a mixed occurrence. When looking at the training script effectiveness, it is seen that participants, especially from the SSD group, might still have experienced mixed learning curves.

1) Simulation Environment: Both groups had several participants that experienced a number of input errors during the training and test phase, while this was actually expected to be zero or close to zero when participants have become fully familiarized with the simulation environment. Additionally, it was found that the SSD group experienced more input errors after the transfer manipulation, indicating that participants did not obtain the required level of expertise and were possibly not able to process the simulation environment information at an automated level yet, resulting in difficulties during the interaction with aircraft. Additional part-task practice in the first step of the practice phase might thus be required for the interaction with aircraft.

It should also be noted that for several participants, the simulation did not seem to accept all commands. In some cases, this led to conflict situations, more stress and (many) extra (incomplete) corrections to resolve the situation (e.g., participant 10 from the control group as discussed in Section V). Participants from the SSD group furthermore indicated that after the SSD was no longer available, it was difficult to interpret whether a command was correctly sent to an aircraft. Sometimes this led to extra (incomplete) corrections to check whether a command was successful and whether the aircraft was heading to the desired direction. Both effects might thus also have contributed to the larger number of errors experienced by the SSD group after the transfer manipulation.

2) Think-aloud Protocol: A verbal protocol was defined to structure the think aloud data and be able to gain an insight into participants' thinking patterns. As explained in Section V, not all participants adhered to the protocol in the same way and not all participants seemed to perform the thinking-out-loud on a relatively automated level. It was found that a number of participants, from both groups, did not state their reasons for making some decisions out loud. In case they did, it was often mentioned very quickly after performing a command: "HDG X.. *..for efficiency*". It should therefore be questioned whether participants really considered the higher-order goals or whether they mentioned this reason because it was the first thing that came to mind. It is thus recommended to investigate whether a relatively 'strict' think-aloud protocol should be applied in this setting, as potentially important information is left out by some participants, compared to when they are not asked to adhere to a protocol, but to simply mention all their thoughts out loud.

Additionally, a difference in response times, of first defining an approach and actually executing it, was found. Although participants were able to practice the thinking out loud during 7 scenarios before the data would be measured for analysis, additional practice might thus be needed. Next to this, it is recommended to include the timing of executing a command into the order of the thinking steps in the protocol, as this might further reduce the differences between the two response times.

The part-task practice was kept relatively short for this experiment due to a limited time available per participant for the experiment. As it is argued that additional part-task practice is required for the simulation environment as well as the think-aloud protocol, it is recommended to increase the length of the experiment and possibly spread it over multiple days or longer periods of time. In doing so, the long-term impact and potential benefits of the SSD as a cognitive tool can be determined as well as what the retention effects are on the individual learning curves.

3) Rule of Thumb in Still Scenarios: Looking at the still scenario questions that had to be answered by participants, it was found that the control group performed better during the training and the test phase in that more participants from this group answered the still scenario questions correctly. As stated before, the questions had to be answered without the SSD by both participant groups, such that the questions could serve as baseline measurements throughout the experiment. It was expected that having to perform the still scenarios without the SSD already, could possibly already prepare the participants from the SSD group on the transfer manipulation, thereby reducing the impact that the (dynamic) transfer would have on their overall performance. However, it seemed that participants from this group mostly found this difference confusing, resulting in less participants answering the questions correctly. Some participants were also found to draw the SSD on the still scenarios. It is, however, unsure whether these drawings added to the results and the participants' overall understanding of the rule of thumb in a positive or negative way. In addition, all participants indicated that they found the still scenarios harder than the dynamic scenarios as the lack of motion made it more difficult to assess the situation.

Structuring the learning process according to an instructional design model was expected to structure the internal mental model of participants [14]. It was therefore hypothesized that performing the experiment while being guided by the script would be effective and would result in participants successfully obtaining the knowledge about the rule of thumb (**H1**). Even though mixed or simultaneous learning curves were still observed to some degree, it was found that participants from both groups indeed made significantly more use of optimal strategies compared to sub-optimal strategies. This does, however, not say anything about whether the participants were actually aware of the rule of thumb or simply thought the chosen solutions looked more favorable at that exact moment.

The question during the debrief about what participants thought the rule of thumb was they had learned during the experiment, showed that participants from the SSD group were in fact more often aware of this rule of thumb and making use of it (**H1**, **H3**). Additionally, after telling the participants (from both groups) the correct answer to the rule of thumb question after the experiment, in case they answered it incorrectly, some participants indicated that they were not considering the different speeds at all or had a very different strategy in mind. This could indicate that their initial discovery learning skills were possibly not as strong or that these participants would benefit more from different learning styles as will be elaborated further on in this section.

Second, applying the Four-Component Instructional Design model to the experiment and script design was expected to result in being able to better track and monitor participants' learning curves, as they were subsequently expected to become more prominent (**H2**). Translating the model to the experiment design, resulted in scaffolds in the step-by-step script (i.e., task consisting of four items that each had a different learning objective and were added one-by-one until they made up the complete task) as well as in the scenarios (i.e., three different conflict geometries/categories ordered in a sawtooth pattern). The scenario scaffolds can also be observed in the data that were logged by the simulation environment. Following from this, the learning curves that were observed could be categorized in two ways. First, maintaining the original scenario order shows that the three scenario categories result in a sawtooth pattern in the performance data. Second, grouping the scenarios based on scenario category allows for the analysis of the conflict geometry on the performance and indeed shows different levels of performance that each contain learning curves as well. The setup of the experiment thus allowed for the analysis of the learning progress of participants in multiple ways.

B. Solution Space Diagram as Cognitive Tool

The SSD was embedded as a cognitive tool in the experiment in an effort to investigate whether it could support discovery learning and could contribute to novices picking up an industry rule of thumb. It was hypothesized that training with the SSD would be beneficial and would result in more participants picking up the rule of thumb compared to training without the availability of this tool (**H3**). It was furthermore hypothesized that participants from this group would continue to use the same strategies after the transfer manipulation due to their structured knowledge-base and increased understanding (**H7**). As stated before, participants from both groups in general showed to use optimal strategies (i.e., correct rule of thumb) significantly more often than the sub-optimal strategies. The SSD group chose the optimal solutions slightly more often in two out of the three scenario categories, but a significant difference was, however, not present between the groups. In addition, when comparing the number of correctly chosen strategies before and after the transfer, both groups of participants surprisingly showed to make less use of the optimal strategies. However, no significant difference was found between the two experiment phases.

Looking at the question at the end of the experiment that probed the participants' understanding of the rule of thumb showed, however, that there was indeed a significant (positive) association between the use of the SSD during the training phase and the answer to the questions: 8 out of 14 participants of the SSD group indicated to make use of the correct rule of thumb compared to only 2 out of 14 participants for the control group (**H3**).

It was furthermore expected that the three scenario categories or conflict geometries would have an effect on the correct execution of the rule of thumb (**H4**). The results of the strategy analysis showed that an opposite trend was present when comparing the groups. The SSD group chose optimal solutions more often during both experiment phases for the easy scenarios, whereas this was not the case in the medium and difficult scenarios. When comparing, for example, the number of optimal strategies per scenario category in the training phase, it was found that for the SSD group this number decreased as the level of difficulty increased, whereas this number increased for the control group as the level of difficulty increased. No significant effects were found of the scenario category on the type of strategy that was chosen. It was, however, found that the scenario category had a significant effect on part of the performance data, namely the track deviation. As explained in Section V, the conflict geometry directly influenced the operating space of the participant. The easy scenarios contained larger conflict angles and larger CPAs, meaning that participants could make smaller deviations to operate both safely and efficiently, compared to the small conflict angles and smaller CPAs in difficult scenarios, resulting in larger heading deviations and thus larger track deviations.

Although the strategy analysis was mostly based on logged simulation data, these data were sometimes not conclusive, meaning that the video recordings were used to complete the analysis. The strategy analysis was thus (partly) subjective to the interpretation of the researcher. In order to increase the objectivity of this analysis, the analysis should preferably be performed (blindly) by one or more evaluators, similar to how ATC students are also assessed by multiple coaches during the ATC training. Analyzing the logged simulation data in a blind fashion is possible as these data do not hint at the availability of the SSD. However, in case of inconclusive data, the video recordings do hint at the group of a participant as the SSD is visible on the screen in case it was available. A possibility is to first use the transcript from the video before looking at the video as a last resort.

As a risk existed that the SSD would be used as a rule-based tool and would induce participant dependency on the interface, it was expected that participants from the SSD group would experience a 'setback' or delta in performance after removal of the SSD (**H5**). The data showed that participants from this group did indeed experience a significantly larger delta in performance when looking at the minimum separation distance. However, for the track deviation the difference in delta between the groups was not significant.

In addition, it was expected that designing the experiment and script according to an instructional design model would help to structure the mental model and reduce surface learning in combination with the use of the SSD as a cognitive tool. It was therefore hypothesized that although participants from the SSD group would experience a setback in performance after the transfer manipulation, they would also quickly recover and (**H6**) experience steeper learning curves and (**H7**) would continue to use the correct strategies. Although a slight trend in the minimum separation distance data could be seen toward steeper learning curves after the transfer manipulation, the difference was not significant. For the track deviation, most learning curves were found to be less steep after the transfer, but again, no significant difference was present (**H6**).

While the SSD might indeed have contributed to an increased understanding of the rule of thumb, it also showed a negative effect on performance when comparing the data from before and after the transfer manipulation. While all participants indicated to have been surprised by the removal of the SSD in the test phase, this could also be traced back in the data. Participants were found to perform worse in the test phase compared to the training phase, especially when looking at control task efficiency. After the transfer, participants showed to make more input errors, as stated before, and to use more extreme and a smaller variety in heading changes, resulting in larger separation distances and track deviations, suggesting that the SSD might indeed have been used as a rule-based tool by most participants of this group.

Preliminary research showed that when participants were told in advance that the SSD would not be available during a test or measurement phase, they adapted their strategies and approach to the scenarios in such a way that they would make very little to no use of the SSD. Although better performance would then be seen due to no or a small setback being present, the effect that the SSD would have on knowledge development or obtaining the correct rule of thumb could not be determined. A balance should thus be found between the two approaches and additional investigation is required to determine how the dependency on the SSD could be decreased.

This preliminary research furthermore investigated whether phased visual elimination of the elements of the decision-support tool could aid in decreasing the dependency on the SSD. Because participants showed to make little to no use of the SSD during this preliminary research and to also experience mixed learning curves due to a too complex task, the effect of the visual elimination on their performance could not be determined. Combining the structured approach of the step-by-step script, that separates the learning curves and provides students with the required background knowledge in a structured manner, with the phased visual elimination of the SSD, might, however, decrease the dependency on the SSD. It is thus recommended to further investigate this combination.

C. Human Factors

The questionnaire that was filled in by each participant before the start of the experiment was used to identify differences in background, knowledge and skills between participants by assigning scores to the answers to each question, such that the participants could be evenly distributed over the two study groups in an attempt to balance variation. Although some participants indicated to have some knowledge of the subject and thus got a higher score in the questionnaire compared to participants that indicated to have no knowledge of the subject at all, they were found to have less insight into the subject than expected. The effects of less balanced groups due to wrong self-evaluation of participants could be reduced by integrating a more elaborate selection procedure into the experiment design. However, a new question then arises whether this selection procedure would introduce a bias and already influence the initial knowledge or performance level of the novice participants.

As mentioned earlier, human creativity in approaching and solving problems is an important factor in the ATC domain, but also in general. Especially for experiments in which humans are involved, this is an important factor to keep in mind. Although the range of available solutions and strategies was tried to be kept as small as possible in the experiment, personal differences in approaching problems or motivation, could still lead to the development of very different strategies and approaches to the scenarios and thus to differences in the obtained data.

Participants could additionally have different preferred learning styles. Some participants might be more visually-oriented, whereas other participants prefer to write information down. Another difference in learning styles could be related to discovery learning and the initial presence of discovery learning skills [34]. While the discovery learning method works very well for some participants, who are successful in obtaining the required knowledge in a short amount of time, other participants might not be as successful, and would require more time or an explicit instruction before being able to obtain the required amount of knowledge. Although the function of the SSD as a cognitive tool in the experiment was to implicitly serve as a guide towards the right strategy, it is possible that the SSD was not explicit enough for some participants. Only after they are explicitly told the rule of thumb, they are able to connect the dots and successfully perform the task. As stated before, telling participants the correct rule of thumb after the experiment, after they incorrectly answered the rule of thumb question, indeed resulted in participants connecting the dots. Although the SSD did show to be beneficial for part of the participants, it might be more beneficial to tailor the training of participants to their individually-preferred learning styles.

For all scenarios, an orange STCA was scheduled to occur after 30 seconds into the simulation if the participant would not intervene. While this was enough time to detect conflicts and define a resolution plan for many participants, for some participants this was not enough time. The stress resulting from the STCA after waiting too long to give an initial command could have incited different types of behavior or strategies than when participants would have had enough time to define a plan to resolve the situation. A solution could be to increase the length of a scenario. However, a balance has to be found in how much extra time would be required, as longer scenarios could also lead to boredom, which in turn could lead to performance decrements (e.g., due to loss of concentration). In fact, after resolving the conflict in an optimal way, there was often about 30 seconds (or longer) left, which already made some participants feel bored according to the questionnaire. Other participants, however, used this time to actively reflect on the situation, which may have helped them in acquiring the right rule of thumb.

Whereas this research focused on only a small part of instructional design, it is worth looking into how additional elements could be translated to and integrated into the experiment. Reflecting on one's own actions, for example, could help to increase deep knowledge and structure the mental model. While it was decided to not incorporate this into the research for now, it is recommended to investigate how this could be incorporated into future research. Next to active reflection, feedback from assessors is also considered a factor in the ATC training that positively influences the learning process. This does, however, also depend on the individual preferred learning styles. For some participants, the effect of feedback might be different on the acquired knowledge and skills than for others. Both examples are found in the current ATC training and can in the future thus be used to further extend the research to potentially increase the positive influence that the SSD has on strategy development.

VII. Conclusions

This research investigated how EID could possibly promote or support discovery learning processes in the ATC domain through the application of an instructional design method. A human-in-the-loop experiment was performed through which 28 novice participants, divided over two groups, were guided by means of a step-by-step script, developed according to the Four-Component Instructional Design model. The custom-made CD&R training was developed to investigate whether an ecological interface, the SSD, could contribute to novices picking up an implicit industry rule of thumb by means of discovery learning. One group was trained with the SSD serving as a cognitive tool and one group was trained without this tool, after which both groups had to perform the test without the SSD.

Results showed that applying an instructional design method to the experiment design leads to a more structured learning process and participants making significantly more use of optimal strategies, corresponding to the execution of the rule of thumb, compared to sub-optimal strategies. While no significant difference was found between the groups in how often an optimal strategy was used, it was found that the group that was trained with the SSD showed increased awareness of the rule of thumb. Although the SSD is thus believed to have a positive influence on obtaining knowledge about the rule of thumb, it was also found that participants that were trained with this tool showed to have become dependent on it. In order to reduce this dependency, it is recommended to, for example, extend the design of the experiment by including phased visual elimination of the elements of the SSD. Additionally, the experiment only investigated short-term development during a very limited number of scenarios. Further research should point out what the long-term impact and benefits of the SSD are when it is used as a cognitive tool and whether learning tasks or processes tailored to the individual might be more beneficial.

References

[1] FAA, "Aeronautical Information Manual - Official Guide to Basic Flight Information and ATC Procedures," 2017.

[2] ICAO, "Long-Term Traffic Forecasts - Passenger and Cargo," 2018.

[3] SESAR Joint Undertaking / Federal Aviation Administration, "NextGen - SESAR - State of Harmonisation," 2018.

[4] SESAR Joint Undertaking, "European ATM Master Plan: The Roadmap for Sustainable Air Traffic Management," 2012.

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

[6] T. B. Sheridan and R. Parasuraman, "Human-automation interaction," *Reviews of Human Factors and Ergonomics*, vol. 1, no. 1, pp. 89–129, 2005.

[7] G. K. van de Merwe, E. A. P. B. Oprins, D. J. F. Eriksson, and A. H. van der Plaat, "The Influence of automation support on performance, workload, and situation awareness of Air Traffic Controllers," 2010.

[8] 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.

[9] J. M. Flach, F. Tanabe, K. Monta, K. J. Vicente, and J. Rasmussen, "An Ecological Approach to Interface Design," *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, pp. 295–299, 1998.

[10] C. Borst, R. M. Visser, M. M. van Paassen, and M. Mulder, "Exploring Short-Term Training Effects of Ecological Interfaces: A Case Study in Air Traffic Control," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 6, pp. 623–632, 2019.

[11] K. Christoffersen, C. Hunter, and K. J. Vicente, "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.

[12] N. Naikar, "Cognitive work analysis: An influential legacy extending beyond human factors and engineering," *Applied Ergonomics*, vol. 59, pp. 528–540, 2017.

[13] 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: Systems and Humans*, vol. 38, no. 6, pp. 1221–1233, 2008.

[14] F. G. W. C. Paas and J. J. G. van Merriënboer, "Instructional Control of Cognitive Load in the Training of Complex Cognitive Tasks," *Educational Psychology Review*, vol. 6, no. 4, pp. 351–371, 1994.

[15] J. J. G. van Merriënboer, *Training Complex Cognitive Skills - A Four-Component Instructional Design Model for Technical Training*. Englewood Cliffs, NJ: Educational Technology Publications, 1997.

[16] M. D. Roblyer, J. Edwards, and M. A. Havriluk, *Integrating educational technology into teaching*. Upper Saddle River, N.J.: Prentice Hall, 1997.

[17] W. Schneider, "Training high-performance skills: Fallacies and guidelines," *Human Factors*, vol. 27, no. 3, pp. 285–300, 1985.

[18] E. A. P. B. Oprins, E. Burggraaff, and H. van Weerdenburg, "Competence-Based Assessment Design for Air Traffic Control Training," in *Proceedings of the International Symposium of Aviation Psychology*, pp. 546–552, 2005.

[19] FAA, "Plan for the Future - 10-Year Strategy Plan for the Air Traffic Control Workforce 2012-2021," 2012.

[20] LVNL, "Jaarverslag 2018," 2018.

[21] M. Schuver-van Blanken and J. J. G. van Merriënboer, "Air traffic controller strategies in operational disturbances: An exploratory study in radar control," in *Proceedings of the 30th EAAP Conference*, 2012.

[22] ICAO, "Doc 10056, Manual on Air Traffic Controller Competency-based Training and Assessment," 2017.

[23] ICAO, "Annex 2 to the convention on international civil aviation: Rules of the air," 2005.

[24] ICAO, "Doc 4444, Procedures for Air Navigation Services: Air Traffic Management," 2016.

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

[26] M. R. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors Journal*, vol. 37, no. 1, pp. 32–64, 1995.

[27] J. J. G. van Merriënboer, P. A. Kirschner, and L. Kester, "Taking the Load Off a Learner's Mind: Instructional Design for Complex Learning," *Educational Psychologist*, vol. 38, no. 1, pp. 5–13, 2003.

[28] P. A. Kirschner and J. J. G. van Merriënboer, "Ten steps to complex learning: A new approach to instruction and instructional design," in *21st Century Education: A Reference Handbook* (T. L. Good, ed.), ch. 26, pp. 224–253, Thousand Oaks, CA: SAGE Publishing, 2004.

[29] K. Shabani, M. Khatib, and S. Ebadi, "Vygotsky's Zone of Proximal Development: Instructional Implications and Teachers' Professional Development," *English Language Teaching*, vol. 3, no. 4, 2010.

[30] S. P. Lajoie, "Extending the scaffolding metaphor," *Instructional Science*, vol. 33, no. 5-6, pp. 541–557, 2005.

[31] R. Azevedo and M. J. Jacobson, "Advances in scaffolding learning with hypertext and hypermedia: a summary and critical analysis," *Educational Technology Research and Development*, vol. 56, no. 1, pp. 93–100, 2008.

[32] J. van De Pol, M. Volman, and J. Beishuizen, "Scaffolding in Teacher–Student Interaction: A Decade of Research," *Educational Psychology Review*, vol. 22, no. 3, pp. 271–296, 2010.

[33] B. J. Reiser, "Scaffolding Complex Learning: The Mechanisms of Structuring and Problematizing Student Work," *Journal of the Learning Sciences*, vol. 13, no. 3, pp. 273–304, 2004.

[34] W. van Joolingen, "Cognitive tools for discovery learning," *International Journal of Artificial Intelligence in Education*, vol. 10, no. 3, pp. 385–397, 1999.

[35] R. D. Pea, "Beyond Amplification: Using the Computer to Reorganize Mental Functioning," *Educational Psychologist*, vol. 20, no. 4, pp. 167–182, 1985.

[36] C. M. Burns, G. A. Jamieson, G. Skraaning, N. K. Lau, and J. Kwok, "Supporting Situation Awareness Through Ecological Interface Design," in *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, pp. 205–209, 2007.

[37] C. Borst, C. Westin, and B. Hilburn, "An Investigation into the Use of Novel Conflict Detection and Resolution Automation in Air Traffic Management," in *Proceedings of the Second SESAR Innovation Days*, 2012.

[38] P. Hermes, M. Mulder, M. M. van Paassen, J. H. L. Boering, and H. Huisman, "Solution-Space-Based Analysis of the Difficulty of Aircraft Merging Tasks," *Journal of Aircraft*, vol. 46, no. 6, pp. 1995–2015, 2009.

Part II

Preliminary Thesis Report

NOTE:
This part has already been graded under AE4020

Chapter 1

Introduction

1-1 Background

Many, sometimes drastic, changes have occurred in the international airspace system since it first became necessary to inform pilots about the runway conditions, wind directions and other nearby aircraft to avoid potential collisions. The mission statement, to ensure a safe, orderly and expeditious flow of air traffic, and the means with which Air Traffic Control Officers (ATCOs) ensure the safe separation of aircraft, have remained largely unchanged over the past decades [1]. However, with today's rapidly advancing technologies as well as the annually increasing number of flights and air traffic densities, a modernisation of the systems used by ATCOs is inevitable to cope with the increased workload for ATCOs [2]. The current air traffic system is already reaching its limits of capacity while further growth of the air traffic is still expected. The International Civil Aviation Organization (ICAO) has predicted a worldwide growth in air traffic of 4.1% per year between 2015 and 2045 [2]. In order to facilitate this increase, the capacity of the current system will have to be increased, meaning that changes in the tools and procedures nowadays used in Air Traffic Control (ATC) are unavoidable.

Single European Sky ATM Research (SESAR) and Next Generation Air Transportation System (NextGen), two programs currently being developed and implemented by Europe and the United States of America, respectively, focus on a redesign of the tools, systems, methods and procedures used in ATC to improve overall aviation and Air Traffic Management (ATM) system performance, especially in the areas of climate impact and flight efficiency [3, 4]. Within this redesign, both programs aim at better data quality, global digital data links and the introduction of advanced automation systems in order to better support controllers in accomplishing their mission of ensuring a safe and expeditious flow of air traffic.

Several challenges arise with the introduction of increasing levels of automation in a system. The human operator should be kept actively aware of system performance. By introducing systems with increased automation, it becomes inevitable that the level of autonomy and authority of the systems will increase with it and thus the controller is likely to take on a supervising role within the ATC system. Over time, controllers will find they are no longer necessarily involved in controlling every single flight: they will be managing air traffic by exception and intervening in case of unexpected situations only and will learn to trust the automation. In most cases, the automated system will operate perfectly well without the intervention or involvement of humans. However, there might be circumstances in which automation does not work as intended or in which a situation suddenly requires manual operation. Although automation generally improves safety during normal operations, over-reliance on automation as well as deteriorated knowledge and skills of controllers might actually result in an unsafe situation during abnormal operating conditions.

The operator should in that case thus be provided with "right-time" access to the appropriate information in a format that supports the operator in accurately assessing the situation and effectively stepping in when necessary. In a highly complex system such as the ATC system, there is the probability that certain scenarios or problems cannot be anticipated in the design of systems with increased automation. When it comes to this uncertainty, human operators will remain an important factor in the system as humans are known to be very creative and flexible in unanticipated situations. To make optimal use of this creativity and flexibility, the design of the human-machine interface and the presentation of information is of critical importance.

With the aforementioned changes being introduced within ATC and ATCOs taking on a more supervisory role, some of the ATCO's skills required for controlling the air traffic will be down-graded or might even disappear, while other skills, such as conflict detection and resolution, are likely to increase in importance. It thus becomes necessary to introduce a corresponding change in attitude towards the role, function and skills of an ATCO in this changing environment. The selection and training of ATCOs should subsequently be critically evaluated in order to align ATC training with the changing role of ATCOs in future ATC systems. Current learning methodologies used within the ATC training are considered to be subjective and no longer meet the demands of modern technology [5]. The learning methodologies should either be adapted or new techniques should be devised in which students will learn to work with and on the cognitive aspects of future systems.

1-2 Problem Statement

This section elaborates on the motivation for this research as well as the scope and objective of this research. Additionally, the research questions are stated that are to be answered in this preliminary thesis.

Motivation

With the expected increase in air traffic, discussed in the previous section, a shift in ATC and thus in the attitude towards learning methodologies used in ATC is required that allows for human operators to work in a more complex environment. During current-day training, high dropout rates occur, even during later stages of the training as the required level of expertise is not reached [6, 7]. A steep learning curve is required and assessment of the ATCO expertise level is still mostly done on a subjective basis.

In adapting to the increasing levels of automation and in increasing the number of graduating ATCOs, it is thus of importance that the training and selection procedures and tools for ATCOs are critically evaluated to align the training with the changing role of ATCOs [7]. One direction of research is aimed at the tools used by ATCOs to interpret the complex ATC environment, specifically the display interface and decision-support tools [8]. Ecological Interface Design (EID) aims at supporting controllers in gaining a deeper understanding of a complex work domain such as the ATC work domain [9]. However, little research has yet been conducted on training with EID tools in the ATC work domain [10].

The changing role of ATCOs, high dropout rates during the different stages of the ATC training as well as the subjective assessment are reasons for aiming this research at improving current teaching methods and specifically the knowledge development of novices when being trained with an EID interface. The knowledge gained during this research can be used in future development of both training and assessment methods used in ATC.

Research Scope

This research will be an effort to further evaluate the short-term impact of ecological interfaces on performance and knowledge development of novices. As stated before, little research has been done on training with EID tools in the ATC environment [10]. It has previously been shown that EID can indeed lead to better knowledge development, a functionally organized knowledge base and increased performance after an exposure of six months [11]. Research on short-term effects of EID on knowledge development also suggests that EID encourages Knowledge-Based Behavior (KBB) and generates goal-oriented thoughts and can therefore play an important role in the early stages of knowledge development. However, an instructional design model was not yet applied [12]. The Solution Space Diagram (SSD), an example of a decision-support tool designed according to EID principles, could be helpful in developing expert-like behavior [13]. The tool is argued to be useful for shaping the internal mental model and hence gaining a deeper understanding of the system [12].

For this research, novice participants will be trained in an ATC task with gradually fading support of an ecologically-designed decision-support tool, more specifically the SSD. This gradually fading support is also known as scaffolding and is part of the Four-Component Instructional Design model on which the approach to training novices for the preliminary experiment, described in this report, is based [14]. Using the scaffolding approach might help to structure the mental model during the training and improve the knowledge development. A key takeaway of this instructional design model is that learning tasks should be based on concrete real-life experiences and situations to increase the chances of a successful transfer of learning [15]. From the various domains and tasks that exist within the ATC work domain, it was decided to focus on the Area Control (ACC) domain, as the tasks related to this domain are considered to be very cognitively demanding. More specifically, the merging task has been chosen for this experiment. From literature as well as from interviews with Luchtverkeersleiding Nederland (LVNL) employees and an ATC student, it was found that the merging task is often considered to be one of the most difficult and challenging tasks during the training and thus, a decision-support tool might be beneficial in such scenarios [16]. Furthermore, it might be beneficial to include a selection procedure in the experiment to minimize the differences among participants that could influence the experiment results. However, due to the large number of candidates required, the limited time available for this research and the fact that the candidates' full learning curves are desired for analysis, a selection procedure will not be part of the scope of this research.

Finally, the scope of this preliminary thesis is to provide an overview of the relevant literature and current knowledge about training novices with ecologically-designed interfaces, specifically for the ATC work domain. A preliminary experiment, of which the design and findings are also discussed in detail in this preliminary thesis, is conducted in order to select the right approach to scaffolding and refine the design of the final experiment.

Research Objective

The objective of this research is to determine the short-term impact of scaffolding as a support method during ATC training with an ecological interface on the performance and knowledge development of novices by conducting a human-in-the-loop experiment in which novice participants will be trained in an ATC task by means of a scaffolded ecologically-designed decision-support tool. This research contributes to improving learning methods and tools within the ATC domain and can provide new empirical insights in the benefits of EID on knowledge development. The objective of the preliminary thesis is to provide a context for this research objective by conducting a review of relevant literature and performing a preliminary experiment to refine the final experiment.

Research Questions

The main research question to be answered is:

> *To what extent does visual scaffolding of an ecologically-designed decision-support tool (SSD in particular) improve novice learning in the ATC domain, when compared to learning without the application of scaffolded support?*

To be able to successfully answer this question and to better structure the literature review, the research question has been divided into several sub-questions. The sub-questions are:

1. How is knowledge development measured in an objective manner and how can it be visualized?
2. How is the ATC training and assessment currently structured and which training elements are most suitable and/or required to be included in the experiment in order to increase the chances of a successful transfer of learning?
3. Which learning task would be most suitable for analyzing the performance and knowledge development of novices in the ATC work domain?
4. Which elements are important in designing a learning task in general to increase the chances of a successful transfer of learning?
5. Which approaches to scaffolding exist and how can such an approach be combined with the SSD and be integrated in the experiment?

To answer these questions as well as achieve the research objective, several sub-goals have been set:

- Perform a more in-depth literature review on ATC training, EID, Complex Learning and investigate their interrelations;
- Explore the different methods to objectively measure and map novice knowledge development;
- Design an ATC learning task for the human-in-the-loop experiment;
- Determine the right approach for implementing scaffolding as a support method during a preliminary experiment;

- Train and test novice participants on a specific ATC task;
- Analyze the data to determine whether a learning curve is present and whether knowledge development has improved for novices that were trained with scaffolded support.

1-3 Research Approach

To investigate the effects of visual scaffolding of the SSD on novice learning during ATC training, an experiment will be conducted in the ATM Laboratory at the Faculty of Aerospace Engineering at the Technical University in Delft. The experiment data will be analyzed using Python and Matlab, while the statistical part of the analysis will be performed using SPSS. The preliminary research phase consists of performing a literature review and conducting a preliminary experiment. Using the results of the preliminary experiment, the final experiment design can be finalized.

Literature is gathered using online databases such as ResearchGate, Scopus, IEEE Xplore, AIAA, APA, Delft University of Technology Control & Simulation Reference Database and the Library of Delft University of Technology. Key words used during the collection of relevant literature are: Solution Space Diagram, Air Traffic Control, (ATC) training, (ATC) assessment, Conflict Detection & Resolution, competences, Cognitive Work Analysis (CWA), Ecological Interface Design, learning tasks, instructional design, complexity, scaffolding.

1-4 Report Outline

This preliminary research report is outlined as follows: the report starts with a literature review on ATC training, EID and complex learning. Chapter 2 elaborates on ATC background information, the structure of the training and relevant cognitive models used in the ATC domain. Next, Chapter 3 discusses the EID principle. First, the origin and theoretical foundations are discussed, after which the CWA for the ATC domain and specifically the Conflict Detection & Resolution (CD&R) and rerouting task is discussed. The last chapter of the literature review, Chapter 4, discusses the theory behind complex learning and cognitive load management. Subsequently, Chapter 5 presents the design, results and conclusions of the preliminary experiment. This chapter also elaborates on recommendations for the final experiment. Finally, Chapter 6 shows the proposal for the final experiment, based on the findings and recommendations from the preliminary experiment. The last part of the report consists of a set of appendices, used to provide additional information concerning the experiments.

Chapter 2

Air Traffic Control Training

This chapter provides the necessary background information about ATC and the training of an ATCO. First, Section 2-1 and Section 2-2 provide background information about ATC and CD&R in general. Section 2-3 provides an overview of the ATC selection, training and assessment procedures as well as a discussion on the competences important for ATCOs. Furthermore, several cognitive models are presented that are used to assess ATCO competences and identify strategies. Finally, Section 2-4 discusses several methods to define ATC sector complexity.

2-1 Air Traffic Control

The mission statement of Air Traffic Control is to safely and efficiently organize and expedite the flow of air traffic from origin to destination and to provide information and other support to pilots [1]. In order to accomplish this mission, the airspace is divided into several regions and aircraft fly predetermined routes. Such a region is defined as a Flight Information Region (FIR). FIRs are in turn subdivided into sectors. Each FIR sector contains controlled and uncontrolled airspace as well as Special Use Airspace (SUA). Examples of SUA are military training zones, airspace above royal or government buildings and war zones. An overview of the different zones that exist within a sector can be found in Figure 2-1.

Figure 2-1 shows that the controlled airspace is further divided into the Control Zone (CTR), the Terminal Control Area (TMA), the Control Area (CTA) and the Upper Control Area (UTA). The UTA starts at Flight Level (FL) 195 and is controlled by Upper Airspace Control (UAC). Control Areas are generally controlled by ACC. The TMA is controlled by Approach Departure Control (APP) and APP thus controls approaching as well as departing traffic. Finally, the CTR is controlled by the Tower (TWR). The TMA is a very complex area, usually with a high density of in- and outbound traffic. To reduce the controller workload, ensure safety and to also comply with noise, environmental and other constraints set by the airport or authorities, fixed approach and departure routes were defined. A Standard Instrument Departure (SID) connects a departing aircraft to the en-route phase after take-off, while a Standard Terminal Arrival Route (STAR) does the same for arriving aircraft: it connects an arriving aircraft from the en-route phase to the final landing phase.

In controlling the airspace, two main control strategies can be distinguished: strategic control and tactical control. Strategic control can often be referred to as long-term control and comes down to ATC managing traffic in such a way that air routes do not become overloaded and

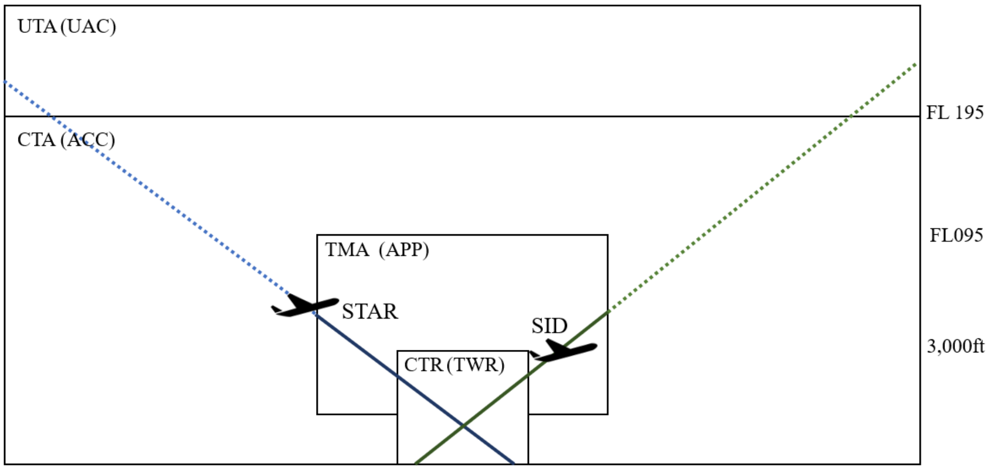


Figure 2-1: Airspace Divided into Different Zones

hence safety will not be at stake. Tactical control is usually called short-term control. This strategy is used to ensure that the aircraft remain at a safe distance from each other at all times. In order to keep the aircraft at a safe distance from each other, it is important to know the locations of the aircraft at all times as well as their speeds and accelerations. This is done by using Communication, Navigaton and Surveillance systems and sometimes happens at the cost of flight efficiency, as aircraft are usually not able to fly the most optimal routes.

To improve the flight efficiency within the European airspace and to manage the expected increase in air traffic, the SESAR program was founded by the European Union [3, 4]. As stated in the introduction, this program focuses on a complete redesign of the tools, systems, methods and procedures that are used in ATC. The aim is to reduce the cost of ATM services and increase the airspace capacity and efficiency. Furthermore, the current airspace consists of many FIRs with corresponding control centers. The borders of the FIRs are generally based on the borders of countries or regions within a country. To reduce the fragmentation of the airspace and thereby increase airspace capacity, SESAR proposes Functional Airspace Blocks (FABs) [4]. One such a block might contain several countries which share the airspace and are controlled by a single control center. Besides introducing more direct routes and thereby increasing air traffic efficiency, the tools, systems and procedures can be further standardized, which in turn allows for ATCOs to easily switch between control centers and for training procedures to be standardized among FABs.

2-2 Conflict Detection & Resolution

One of the main responsibilities of ATC is to ensure safe separation of air traffic [1]. By introducing the radar, it became considerably easier for ATCOs to determine the location of aircraft within the airspace from ground-based stations and thus to successfully separate air traffic. With the current developments that are leaning towards more automation within ATC, it becomes evident that the workload and role of the ATCOs will change again and skills related to CD&R will become more and more important.

To have aircraft fly at a safe distance from each other, separation minima have been defined. Aircraft should be separated by at least 5 NM horizontally and 1,000 ft vertically [17]. Each aircraft thus has a so-called (cylindrical) Protected Zone (PZ) around it, which is visualized in Figure 2-2.

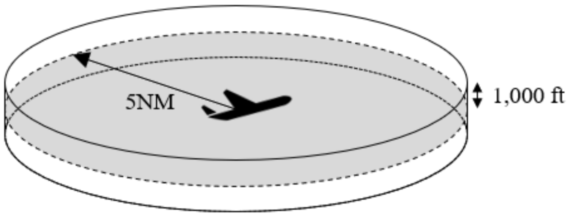


Figure 2-2: Graphic of the Protected Zone of an Aircraft

When an aircraft enters another aircraft's PZ, this is called a Loss of Separation (LoS) [13, 18]. A conflict arises when an actual LoS occurs but also when a potential LoS within a predetermined time window presents itself, thus when two or more aircraft follow trajectories that are set to cross in the near future [13]. This predetermined time window is often referred to as the look-ahead time.

There are several parameters that define a conflict. Important parameters are the time to LoS, t_{LoS}, and the Closest Point of Approach (CPA), which is defined as the moment in a conflict at which the involved aircraft will not approach any closer [13]. The minimum distance between the two aircraft has thus been reached at this particular moment in time. This distance is defined as the distance at CPA, d_{CPA}, which is depicted in Figure 2-3. The figure shows a simple graphic of a conflict that involves two aircraft. Other important parameters that are shown in the figure are the velocity vectors of the involved aircraft, V_A and V_B, the relative velocity vector V_R and the conflict angle σ.

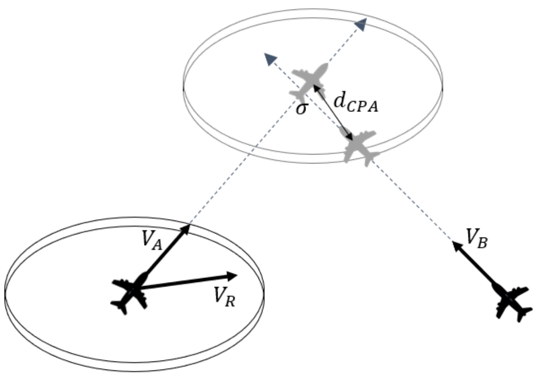


Figure 2-3: Conflict Geometry Along with Important Parameters

In order to assist ATCOs in detecting conflicts, several tools are available. Two examples are the Short-Term Conflict Alert (STCA) and Medium-Term Conflict Detection (MTCD). They are both ground-based safety nets that help ATCOs detect potential or actual violations of the minimum separation distances in a timely manner [19, 20]. As the names suggest, the tools provide ATCOs with short-term and medium-term warnings, meaning they provide warnings up to two and twenty minutes before a conflict arises, respectively.

2-3 Selection, Training and Assessment of Air Traffic Controllers

The two central concepts in ATC, safety and efficiency, are leading throughout the selection, training and assessment of air traffic control trainees. This section elaborates on the competencies important for ATCOs as well as the selection, training and assessment procedures for ATCOs.

2-3-1 Competencies

Since ATCOs are subject to strict safety regulations, there is little to no room for (human) errors or incompetence and thus performance standards for the ATC task are high. The task is considered to be both dynamic and highly complex as it requires processing of large amounts of constantly changing information [21, 22]. The current trends of growing traffic intensity, stricter environmental regulations and increasing safety standards, increase the workload for ATCOs even more. Several competencies have been defined to determine whether someone is capable of conducting the ATC task [22]. According to ICAO Doc 10056 [23], a competency in the context of ATC training can be described as:

> **Competency:** "A combination of skills, knowledge and attitudes required to perform a task to the prescribed standard." [23, p. 6]

Only when the standards for a specific competency are met, that person is said to have successfully obtained this competency. When a competency is successfully obtained, it allows people to come up with solutions for new and unexpected events or situations. As these unexpected events or situations are not uncommon within ATC, it is of great importance that ATCOs learn to deal with these situations in an effective manner while ensuring safety. Because of the complex nature of the task, there is only a minority of people that is able to acquire the competencies within a predetermined training period. Although skills and knowledge, and thus competencies, are the result of a learning process, they differ in trainability. Some components, which are referred to as consistent components, improve after practicing, while non-consistent components do not necessarily improve, similar to recurrent and non-recurrent skills [15]. Knowledge, one of the outcomes of the learning process, can be described as:

> **Knowledge:** "Specific information required to enable a learner to develop and apply the skills and attitudes to recall facts, identify concepts, apply rules or principles, solve problems, and think creatively in the context of work." [23, p. 13]

Another outcome of the learning process is a skill. A skill is developed over time and with practice. It is defined as:

> **Skill:** "Ability to perform an activity or action." [23, p. 13]

Skills are often divided into three types: motor skills, cognitive skills such as reasoning and perception, and meta-cognitive skills, such as planning, which are skills that relate to the ability of the learner to monitor and direct his/her own learning process [23]. Complex and new tasks are often seen as cognitively demanding. But, as the learner practices such a task more often, the cognitive process becomes less demanding and might even be performed without conscious control. When no conscious control is needed for a certain task of skill, it becomes easier for the controller to find solutions in new and complex situations.

Finally, attitude plays an important role in achieving a competence and ensuring safety [23]. It has affective components, cognitive aspects and behavioral consequences and can be described as:

> **Attitude:** "A persisting internal mental state or disposition that influences an individual's choice of personal action toward some object, person or event and that can be learned." [23, p. 14]

During both the selection and training of ATCOs, a reliable assessment system is very important. The time-critical, complex and dynamic nature of the ATC task however, makes it difficult to develop a reliable assessment system [22]. Skills that are ranked as cognitive skills are invisible to observers, which make it skills that are very difficult to assess. Such skills are usually assessed based on over-the-shoulder observations made during simulations or On-the-Job Training (OJT), which makes the assessment rather subjective [22]. Rather than assessing only the skills of the controllers, a solution is to base the assessment on competencies, which are, as stated before, a combination of skills, knowledge and attitude [23]. LVNL, the agency in charge of air traffic control of the Dutch airspace and responsible for ATCO training, developed a model of the competencies that are of importance to ATCOs, based on the ATC Performance Model [22]. This model is found in Figure 2-4 and serves as a framework for the identification and design of performance criteria, which are used in the competence-based assessment of ATCOs that is discussed in Section 2-3-4.

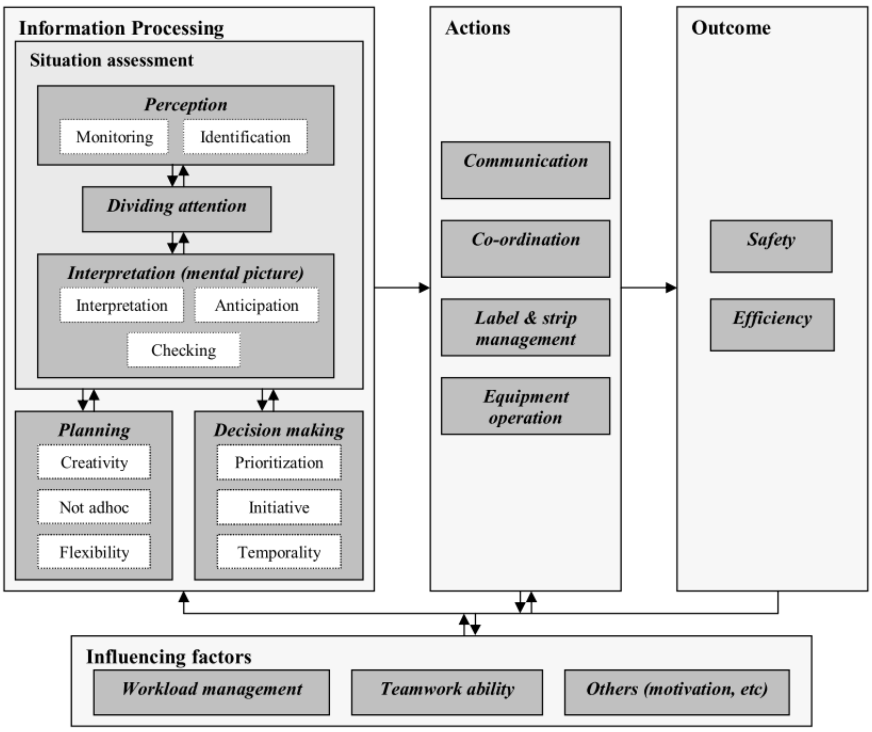


Figure 2-4: ATC Performance Model [22]

The figure shows that the ATC Performance Model roughly consists of four parts: information processing, actions, influencing factors and outcome. The competencies related to information processing are split into different categories or skills and hence the dominant role of the information-processing component within the ATC task is shown [22]. Important for the cognitive process, or the processing of information, are competencies such as situation assessment, planning and decision making. The situation assessment component can be divided further into the elements of perception, dividing attention and interpretation, which are elements that are also found in the Situation Awareness (SA) theory [24]. Information processing subsequently forms the basis for the actions. The action component consists of communicating and coordinating with pilots and other parties to handle air traffic. Furthermore, flight information should be kept up to date and hence the competency label and strip management is important. Finally, the equipment is used to execute the previously mentioned tasks. Both information processing and executing tasks can be influenced by external (personal) factors such as dealing with the workload, teamwork and attitude, whereas actions can in turn also influence the influencing factors, for example, when label and strip management is not performed well, it could influence the experienced workload. The three components information processing, taking action and influencing factors, lead to the outcome. The outcome corresponds to the mission statement and highest goal of ATC: safely and efficiently organizing and expediting the flow of air traffic.

Next to providing information about the competencies that are most important to ATCOs, the model provides information on how the competencies can be assessed [22]. It forms the basis for the performance criteria in the competence-based assessment, which is elaborated on in Section 2-3-4. It can be seen that the model separates criteria that can be measured subjectively and objectively. The objective criteria are the competencies and criteria part of the outcome and action blocks. The competencies and criteria belonging to information processing are cognitive processes and thus invisible. These are criteria that can only be measured subjectively. Furthermore, the model provides information about the trainability of competencies. In contrast to most of the competencies that are related to information processing, the competencies related to actions are trainable as they improve by practicing.

While the competencies presented in the ATC Performance Model might seem straightforward, they will only really become clear or visible in abnormal operating conditions (i.e., when the environment is disturbed due to an unexpected event). Examples of disturbing factors to the normal operational proceedings are a variability in aircraft performance or technical difficulties, weather, traffic numbers or (un)expected changes in operating modes [25]. A disturbed situation can be described as a situation in which the standard traffic handling needs to be adjusted to a different traffic handling method [25]. These unexpected events or situations are exactly what make the ATC task (cognitively) complex and why high standards for ATCO expertise are required [22]. To ensure safety and efficiency while managing this cognitive complexity, a large range of strategies is used by the ATCOs. It was found that based on the characteristics of a situation, controllers adapt their strategies to still comply with the operational constraints and to ensure safety while not sacrificing efficiency [25]. Section 2-3-3 further elaborates on the different strategies used by ATCOs in normal and disturbed operating conditions.

2-3-2 Air Traffic Controller Selection

The ATCO selection procedure of the LVNL is based on the ATC Performance Model, described in the previous section. The competencies that have been defined in this model are used as criteria during the selection procedure. They optimally reflect what is required of ATCOs and are therefore a good predictor of the future performance of candidate ATCOs. With the most important competencies for ATCOs identified, the selection for ATCOs can be set up in such a way that only the people that can acquire these competences with a sufficient likelihood within the time frame set for the ATC training, are selected.

The selection procedure consists of several rounds. During a phone interview conducted on May 23rd, 2019, an LVNL selection expert stated that for each of these rounds, the competencies are considered and assessed. As can be expected, the competencies related to safety and efficiency are most important and thus weigh heavily in this assessment. The selection rounds consist of standardized tests, individual and group assignments and interviews. Important during each of the rounds is how candidates react to stress, behave in teams, make decisions and communicate. The final selection round is a three-day course in which candidates have to complete exercises that are similar to exercises from the actual ATC training. Next to the learning curve that can be identified during this course, candidates are assessed based on how many mistakes they make as well as how quickly these errors are recognized and solved, similar to the assessment of the actual ATC training.

2-3-3 Air Traffic Controller Training

As described in the introduction, ATC has changed significantly over time and with that, the corresponding training programs have changed. However, ATC training programs around the world are often still based on 'older' systems such as old simulators that are being used in the simulator training. Furthermore, since ATC is considered to be a complex task in which unexpected events are not uncommon, the adaptivity and creativity of the human controllers remain important factors in both the training of ATCOs as well as the interface design. It is exactly this creativity and adaptivity that makes it difficult to redesign the ATC interface and the training program.

ATCO Training Structure

Guidelines for the ATC training are provided by ICAO Annex 2, Doc 4444 and Doc 10056 but neither of them provides a predetermined lesson plan [17, 23, 26]. The SESAR program however, proposes the introduction of FABs among which procedures, methods, tools and even training programs can be standardized [4]. This proposal might evolve into a set lesson plan for ATCOs in the future.

The workload during the ATC training is considered to be high and the learning curve that is expected of trainees is steep. This leads to trainees ending their training prematurely (often already quite far in the program) as they are not able to meet the high standards set for the training [6, 7]. During a phone interview on May 15th, 2019, an LVNL training expert explained that the LVNL has reviewed its training program and has restructured it in order to increase the number of successful admissions and increase the number of students that successfully finish the training. Nowadays, the training program at LVNL consists of several phases.

All students start by learning basic practical skills, such as how to use the ATC equipment and how to handle flight strips in the right way. Furthermore, students take several theoretical courses about subjects such as meteorology, aircraft science and performance, radio communication, aircraft recognition, air traffic law, equipment, human factors and navigation. After the basic training, students are admitted to a unit, based on their own preferences and how well their skills/competencies match the competencies that are required for a specific unit. The LVNL developed so-called blueprints for the competencies of an ATCO for the different units and airports in the Netherlands. These blueprints can be found in Appendix A.

After being admitted to a unit, to start the training for Approach or Area Control, respectively, students start training at a basic simulator and follow unit-specific exercises during the so-called Initial Training. The Initial Training consists of sets of exercises related to inbound flights, outbound flights, neighboring fields and a consolidation phase. For each of these sets, the complexity is increased throughout the exercises. The training usually consists of either two exercises a day plus a demo that is performed by a licensed ATCO or three exercises a day. Each exercise takes about 25-45 minutes and complexity and difficulty are increased every exercise by increasing the traffic density, the types of aircraft involved in the scenario or by creating unknown situations. An exercise can be focused on a specific task of air traffic control, for example, only merging traffic. During the adjacent fields phase, extra focus is put on protocols and agreements with the adjacent fields for a specific sector, such that students learn to perform a correct transfer of control. However, the goal of all exercises is to expedite and maintain a safe and orderly flow of air traffic. It is furthermore expected that the trainee indicates the competencies or skills he/she likes to improve or work on during an exercise, as discovery learning and self-reflection are important aspects during the training.

Once the student successfully finishes the Initial Training, the student obtains his/her Student Controller License and is allowed to move on to the next phase of the training: Unit Training. During this phase of the training, students are taught subjects such as air structure, classification, aircraft recognition and routes specific for that unit and are further familiarized with all rules, regulations, procedures and protocols specific to the sectors controlled by the unit. The final stage of the ATC training is the On-the-Job Training. During the OJT, the student will follow and watch professional ATCOs perform their job but will also get the chance to work as an ATCO in that specific sector while still being guided by professional ATCOs. More detailed information on the structure of the training can be found in Appendix B.

Control Strategies

ATCO control strategies are crucial in ATC performance and appear to be one of the key elements to success during the ATC training [25]. Having a range of strategies available generally reduces the risk of performance being compromised during disturbed operational situations. A strategy can be described as a method or class of ATC activities that achieves one or more objectives such as safety and efficiency within a certain time [25].

Which strategy is used by a controller usually depends on the characteristics of a situation as well as operational constraints. Although strategies seem to be a key element to success during the ATC training, they are not specifically taught to trainees during the training, which was confirmed by a training coach during a visit to the LVNL on April 24th, 2019. There is no predetermined lesson plan regarding control strategies. Students however, obtain the strategies by means of discovery learning. Coaches might steer the trainees ad-hoc towards the use of certain strategies depending on the scenarios and the complexity of these scenarios. This does, however, cause the development of these strategies and the associated competencies to be heavily influenced by the personal preferences of instructors as also discussed in Section 2-3-4. By creating solutions ad-hoc or by directing the trainees ad-hoc towards certain strategies, robust knowledge is developed and a single solve-all strategy is discouraged as such a strategy is usually not sufficient to resolve an unexpected situation.

The main control strategies of ATCOs that have been identified are discussed below and can be categorized according to the cognitive processes as described by the ATC Cognitive Process & Operational Situation (ACoPOS) model which is elaborated upon in Section 2-3-5 [25].

- **Perception Strategies:** These strategies mainly consist of extracting flight information such as flight level, heading and speed, especially of inbound aircraft, to be able to detect any present conflicts.
- **Interpretation Strategies:** Strategies used to identify patterns in the operational situation. Aircraft are grouped based on destination or potential conflict and standard and non-standard traffic flows are identified. Finally, critical hot-spots are identified in an early stage.
- **Anticipation Strategies:** Possible threats that may lead to future disturbances are continuously being identified and proactively mitigated. Furthermore, situations are mentally played out to identify the progression of events.
- **Attention and Workload Management Strategies:** These strategies are used to manage monitoring and to save attentional resources. This is done by solving situations rather than monitoring them; monitoring one situation instead of simultaneously monitoring multiple situations and taking action if monitoring takes up too much time and attention. Monitoring is done to verify whether a solution takes effect or different actions need to be taken. In case of a non-standard situation, traffic is kept on standard routing as much as possible in order to focus attention on the disturbance. This is also called 'always keeping spare time' in order to be able to focus on unexpected disturbances or complex situations. Finally, another attention and workload management strategy is to consider the workload of the next ATCO. An ATCO might decide to increase his/her own workload to maintain safety in that specific sector and the adjacent sector and thereby decrease the workload of another ATCO.
- **Solving Strategies:** Any existing conflicts are usually solved by level separation first: a flight level is searched for that is safe and available and has the smallest impact on efficiency. After level separation comes vectoring or further climbing/descending to keep efficiency. By doing this, time is created to gain a better understanding of the evolving situation and hence being able to fine-tune for a higher efficiency later. Other identified solving strategies can be characterized as 'creating space' for solutions and maneuvers, where space is created by applying non-standard routing or by moving crossing points. This strategy is considered crucial to maintain efficiency, create time, avoid domino-effects and to prevent the increase of attentional resources.
- **Planning Strategies:** These strategies include the identification of escape possibilities, which are used in case the initial plan does not work or the situation does not develop as expected. ATCOs try to stick to standard routings or routinized working methods as much as possible to ensure safety and to manage attention and workload. Another planning strategy relates to creating temporary standard patterns: non-standard routings

are temporarily made standard to manage attention and workload. Finally, a strategy that is emphasized during the entire training is 'avoiding becoming reactive' but rather stay proactive.

- **Decision-Making Strategies:** ATCOs usually immediately react to disturbed situations. The situation is either solved partially or completely to prevent any additional problems and hence maintain safety and efficiency as much as possible. In some occasions, when there is enough time and space, ATCOs indicate to wait and see how the situation develops before they initiate any actions. Another decision-making strategy is characterized as 'reflection on action'; a decision is adapted or withdrawn when the action or solution takes too long to take effect or the situation still develops differently than initially expected.

Conflict Detection & Resolution

Little is known about the ways ATCOs are taught to detect and resolve conflicts. As stated before, guidelines are provided by ICAO and the Federal Aviation Administration (FAA) but no set lesson plan exists on how to solve conflicts in an effective and expeditious manner. ATCOs usually obtain the required knowledge by discovery learning and by input and feedback from the coaches during the ATC training, as was confirmed by a training coach during a visit to the LVNL on April 24th, 2019. Research has shown however, that there are three factors that are of influence to the decision-making strategies of ATCOs [27]. These three factors are discussed below.

- **Expediency:** Maneuvers that resolve conflicts more rapidly are preferred over maneuvers that are more time-consuming. Expediency is influenced by control order, gravity and lateral geometry. Especially lateral maneuvers such as turns are executed less rapidly than vertical maneuvers that exploit gravity, such as descending or leveling off.
- **Preservation of airspace structure:** Maneuvers that are least disruptive to the overall traffic flow and environment are preferred. Lateral maneuvers tend to be more disruptive to the overall traffic flow than vertical maneuvers. Turns, while resolving one conflict, are more likely to put aircraft in potential conflict with other aircraft at the same altitude. Besides, vertical maneuvers are easier to implement and require the least amount of attention and caution in the usually busy airspace. Vertical movements can however also be difficult and require caution when they cross levels of other aircraft.
- **Visualization:** Maneuvers that, after initiation, can be perceived more rapidly on the radar screen are preferred. Visualization of conflicts and their resolutions on the Plan View Display (PVD) of a controller is influenced by two factors. First, lateral maneuvers are easier to visualize as their effects are immediately visible compared to vertical maneuvers which are visualized by numerical data that first have to be interpreted by the controller. Second, level-offs are easier to visualize as the altitude displayed remains constant, while for climbs or descends, the controller has to direct his/her attention to constantly changing altitude data.

It can be noticed that for expediency as well as for preservation of airspace structure vertical movements are preferred as they require the least amount of attention. For visualizations and conflict resolutions however, lateral movements are preferred as the effect is immediately visible to the controller.

2-3-4 Air Traffic Controller Assessment

As described in Section 2-3-1, cognitive skills are difficult to measure and assess as they cannot be observed directly (e.g., situation awareness). Furthermore, competencies can generally be difficult to log compared to performance data. Besides, the actual performance of a controller can be influenced by external factors which can be environmental or personal, such as stress or fatigue. Such external influencing factors make it even more difficult to measure or observe a specific competence.

Competence-based Assessment

Assessment of ATCOs during training is currently done by performing so-called over-the-shoulder observations in either simulations or during the OJT, as was confirmed by an LVNL training expert during a phone interview on May 15th, 2019. At LVNL, the assessors, usually licensed ATCOs or coaches depending on the training phase, rate the controllers' performance on a 6-point rating scale:

1. The student has demonstrated that this competency has not (yet) been developed: the coach needs to help the student a lot while the student hardly shows the required behavior. The student makes many mistakes and does not notice them or notices them too late.
2. The student has demonstrated at varying moments to have developed the competency, but performs predominantly inadequate: the coach still needs to help the student a lot. The student does show the required behavior in varying ways but still makes many mistakes.
3. The student has repeatedly demonstrated to have developed the competency, but is sometimes unable to recognize and/or correct mistakes in time: the coach only needs to help in some occasions (especially during complex exercises). The student repeatedly shows the required behavior but this behavior is still not stable. The student still makes mistakes and does not always correct them in time.
4. The student repeatedly shows the required behavior and is able to recognize and correct mistakes in time: only little help is needed from the coach. The student often shows the required behavior and sometimes makes mistakes but also corrects these.
5. The student has shown to possess the competency and works almost flawlessly. If (minor) errors are made, they are always recognized and corrected in time: practically, no help is needed from the coach and the student always shows the required type of behavior during exercises. The student only occasionally makes a mistake, which the student also corrects for.
6. The student has shown to possess the competency and works almost flawlessly. If (minor) errors are made, they are always recognized and corrected in time. The student has overcapacity and shows to have reserves: no help from the coach is required, the student performs better than required for the specific step. The student furthermore shows exemplary behavior, works very smoothly and shows a performance that is more extensive/better than what is expected according to the requirements for that specific step.

In general, it holds that the better the performance of a student for a specific competency is, the more often the student shows the required type of behavior and the more often the student recognizes and corrects any present errors. Because the competences can for a large part be observed through behavior, the assessment or performance criteria are usually formulated as behavior descriptions or behavioral markers. An example of the behavior descriptions for the competence *strip management* that is used by the LVNL is shown in Table 2-1.

Strip management	Assessment Indicators
Correctly manages the strips	The student notes the correct and required information on the strips in the correct manner, without falling behind: • The student is ready for the next action.
Places the strips at the right time at the right place on the flight progress board	The student places the strips correctly and at the right time according to the operations manual and work agreements: • The strips correspond to the current situation. • The placement of the strips cannot lead to safety issues.

Table 2-1: Example of LVNL Assessment Indicators for the Competence Strip Management

Combining performance data with the types of behavior recognized by the assessors, leads to a more or less complete picture of the progress of the controller during the execution of a certain task. As competencies are a combination of several factors, it is difficult to analytically split them up in skill sets and knowledge markers that are easier to assess. Assessment should take place at a higher level, as skills and certain pieces of knowledge might be mastered in different orders or in different amounts of time and social, emotional and environmental factors should be taken into account during the assessment [22].

Assessment at higher levels allows for distinguishing learning curves of these skills and parts of knowledge. Learning curves are important in monitoring the students' progress: they can serve as an indicator for whether students are still in the process of learning or whether they have already reached a learning plateau [22]. This information is useful in deciding whether a student can continue his/her training and whether the training should be adapted to the student's needs. On top of this, it is important to keep the trainability of competencies in mind, as consistent components can be improved more easily than non-consistent components that tend to be innate (e.g., planning or teamwork) [22].

Non-consistent components are usually emotional and social components or skills. They are also known as non-technical skills and include skills such as teamwork, management capabilities, leadership, planning and decision making [22]. The assessment of these skills, which are usually cognitive processes, is very important in ATC and is also part of the reason why the assessment of ATCOs has only been partially automated [5]. Observations and performance data alone are usually not enough as the processes are not visible to the assessors and hence should be complemented with additional information about the thinking patterns and strategies of the students. Methods often applied to obtain this information are think-aloud protocols which are applied during the test, critical incident analysis, interviews and re-runs of training scenarios [22]. This information was also confirmed by a training coach during a visit to the LVNL on April 24th, 2019. Not only can the assessments be performed better by applying these methods, it also becomes easier for assessors to give feedback when they have insights into the students' thinking patterns. As can be read in Chapter 4, feedback is a very important part of the learning process. At LVNL, three types of feedback or evaluation forms are used: exercise reports, coach reports and trainee log files, which can be found in Appendix C. These forms are usually filled out by the coach and trainee together, which leads to a better insight in the student's points for improvement and learning progress. A trainee has multiple coaches during a training phase. By having multiple coaches assess a student, the assessment becomes less subjective and hence more reliable.

During the different training phases, the required competencies remain the same. However, the complexity of the scenarios that students are dealing with is increased over time. Complexity is usually characterized by the amount of traffic, the number of safety violations or the degree of efficiency that is required, which is further elaborated on in Section 2-4. For simulator training scenarios it is easy to increase the complexity over time. During the OJT on the other hand, this is not possible as scenarios cannot be planned in advance but simply happen. Therefore, the assessment during OJT is slightly different. Students are assessed based on the degree of safety and efficiency they maintain, the traffic complexity they have to deal with and their independence of the coach. The increase in complexity during ATC training is what allows coaches to monitor the students' progression and learning curves.

Continuous Assessment

As stated before, when insight is gained in the student's learning curves of different skills/competencies, appropriate measures can be taken if it turns out a student lacks certain skills or parts of knowledge. During a visit to the LVNL on April 24th, 2019, a training coach explained that continuous assessment is therefore applied during the training phase of ATCOs. The assessors continuously interact with the trainees during and after the training exercises. During the training, the students are often asked questions about their decisions or to probe their situation awareness. This is what allows the assessors to gain insight in the cognitive processes and strategies applied by the students and hence give appropriate feedback on strategies that were well- or ill-chosen. During the debriefing, the student and coaches discuss the exercise and sometimes do a re-run of the scenario to analyze what could be improved. Assessments usually take place every two weeks. However, after each exercise an evaluation report is filled in by the coach and trainee, which is used to discover learning curves in the student's progression.

As stated before, multiple coaches are involved in assessing the students which increases the reliability of the assessment. Furthermore, the coaches receive dedicated training beforehand about the assessment system, behavior descriptions and the interpretation of these descriptions or criteria and avoidance of errors during the rating of students. Although this training is meant to reduce the subjectivity of the assessment, objective assessment is impossible due to the cognitive processes that need to be assessed and the personal influences of the coaches. The continuous interaction between coaches and students during the training phases influences the students preferred work style, behavior and thus also performance.

To counterbalance the effect of the coaches' influences on the students' performance, the continuous assessment reports are combined with performance data from, for example, simulator training scenarios. Opinions from both coaches as well as students about this combined assessment system have been positive [22], as was also confirmed by an LVNL training expert during a phone interview on May 15th, 2019. Coaches have indicated to better understand the requirements, recognize the required types of behavior and be able to validate their 'gut feeling' while students indicate to better understand the complete picture and know what is expected from them because of the well-defined competencies and repeated reviews of their learning progress [22].

2-3-5 Cognitive Models for the ATC Task

As the ATC task is considered to be a complex and dynamic task, cognitive processes play an important role. This section discusses two models for the cognitive processes taking place related to the ATC task, in addition to the ATC Performance Model that was presented in Section 2-3-1.

The ACoPOS Model

As stated in Section 2-3-3, the strategies often used by ATCOs are based on the cognitive processes that relate to the ATC task. To identify these cognitive processes and to reduce cognitive complexity, the LVNL developed the ATC Cognitive Process & Operational Situation model. This model is a combination of the competencies identified by the ATC Performance Model, that was described in Section 2-3-1, and elements from the operational situation. The model can be seen in Figure 2-5 and is recommended to be used as a basis for setting requirements for training design to keep the cognitive complexity acceptable [28].

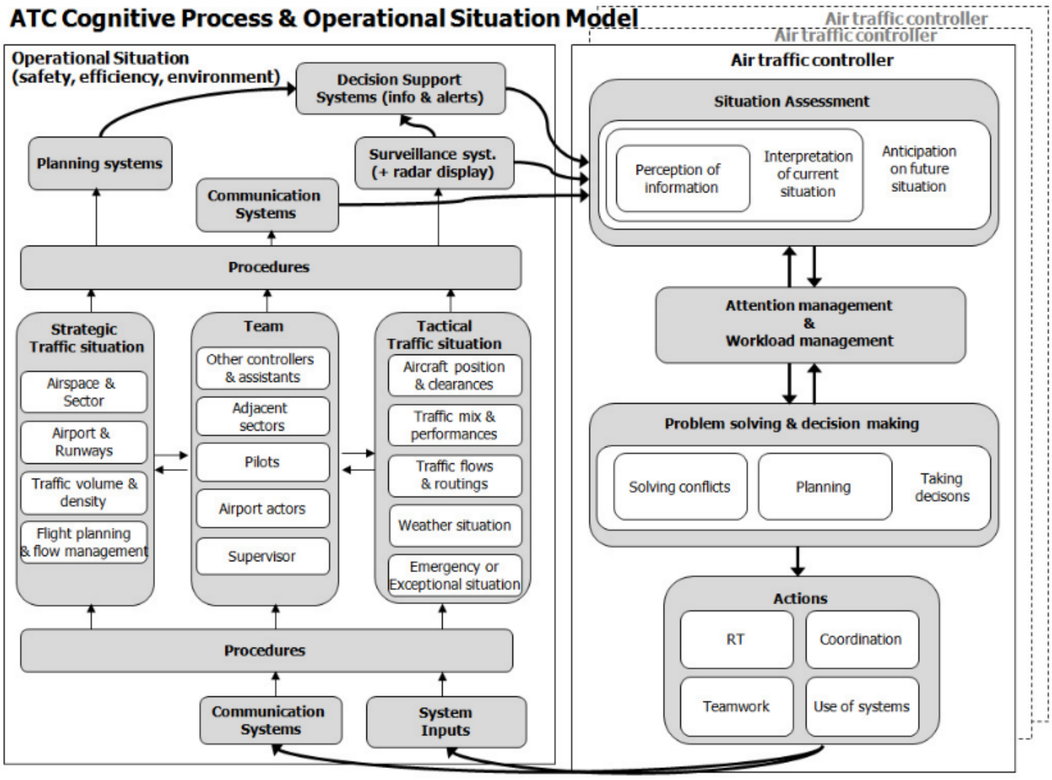


Figure 2-5: The ATC Cognitive Process & Operational Situation Model [28]

In identifying the cognitive processes and complexity during ATC tasks, the model makes the (operational) factors that influence and cause the cognitive complexity, visible at a single glance [28]. The model depicts the ATCO on the right and the operational situation on the left. It can be seen that the blocks inside the Air Traffic Control Officer (ATCO) part are similar to the ATC Performance Model. Three cognitive processes can be distinguished, namely Situation Assessment, Attention and Workload management, and Problem solving and decision making. Each of the cognitive processes is further divided into competences, again similar to the ATC Performance Model.

Situation assessment involves the three levels of SA: perception of information, interpretation of the current situation and anticipation on the future situation [24]. Acquiring and maintaining SA can be cognitively demanding due to the dynamic and interactive elements in the operational situation and the interaction with other cognitive processes [28].

Problem Solving and Decision Making is a cognitive process used by humans to stay in control. ATCOs search for patterns and relevant cues of problems or possible disturbances that are often encountered and from there, determine a particular course of action. In case a solution to a new problem does not exist, ATCOs use problem-solving techniques to develop a new solution. Besides, ATCOs continuously develop and adjust (back-up) plans in case solutions or plans do not succeed and the situation develops differently than expected.

Attention and Workload Management involves continuously setting priorities and directing attention towards different sources of information to monitor situations and thus maintain SA. To avoid losing SA, managing workload, mainly by setting priorities, is crucial.

The last item that can be distinguished within the ATCO part of the model is *actions*. These are the result of the cognitive processes and are performed by the ATCO to interact with the operational environment and implement solutions according to predetermined plans. Actions include radio-telephony, coordination, teamwork and use of operational systems.

The model furthermore displays the information flow between the ATCO and these cognitive processes on one side and the operational situation on the other side and hence provides insight in the impact operational factors have on cognitive processes. The operational situation is also divided into several different parts that all contribute to the requirements of the ATC task that have to be met and balanced: safety, efficiency and environment.

Strategic traffic situation can be described as the long-term traffic situation. It can vary in airspace structure and sectors, layout and runway, traffic volume, etc. The configuration of a strategic situation influences the complexity.

Tactical traffic situation can be described as the actual or short-term traffic situation. It includes aircraft positions, clearances given, traffic mix, aircraft performances, etc. and is characterized by the dynamic nature of the situation.

Teamwork means working in different team situations. Teams usually consist of ATCOs, supervisors, assistants and colleagues from adjacent sectors and centers. Changes in traffic situations means positions can be combined or split resulting in different team configurations and operation procedures. At the same time, interaction with pilots must constantly be maintained.

Procedures describe the formal operating procedures and rules for performing the ATC task. In case more (complex) procedures are used, complexity of the situation and traffic handling is increased.

Technical systems are used to perform the tasks related to ATC. This includes systems for communication, planning, navigation and surveillance and decision support. The design of these systems can influence complexity in case a mismatch arises between system characteristics and human information processing. A transparent system that is adjusted to the requirements of the controller is therefore of high importance.

The ECOM model

Rather than displaying the competences related to certain cognitive processes and determining assessment criteria from there, the Extended Control Model (ECOM) structures the strategies used by ATCOs (discussed in Section 2-3-3) in cognitive processes and translates them to assessment criteria. The model, found in Figure 2-6, identifies patterns of control in terms of feedback loops underlying the interaction between perceiving and acting [29].

Figure 2-6: Extended Control Model [29]

Figure 2-6 shows that the control patterns take the form of four control loops. The four loops that can be distinguished are:

- Targeting: controllers choose goals and set priorities depending on how they expect the situation to develop;
- Monitoring: controllers monitor the situation while keeping the goals in mind from higher layers and adjusting to feedback from lower layers;
- Regulating: actions and resources are managed in correspondence to the objectives that were set in guiding plans;
- Tracking: within this loop controllers keep key parameters within specific limits.

Each of the four control loops is affected by the operational constraints and interactions between the loops can lead to reactive or anticipatory problem solving [29]. How the loops work and interact can vary per situation and per controller. The loops may succeed each other in a small time window which means the model acknowledges that human behavior can take place at several loops simultaneously. Additionally, controllers might focus on, for example, the lower loops only when responding to threats or disturbances or on the higher loops when assessing the situation and prioritizing goals.

The four loops ensure, among others, that key parameters are kept within limits, work progress is tracked, mental models are updated and reframed, desired goals are updated and prioritized and plans are updated accordingly [29]. Transfer of control can take a top-down or bottom-up approach or even a mixed approach. In case of a bottom-up approach, as in the original version of ECOM, feedback from the tracking loop is used as input to the situation assessment in the regulation loop, similar to how information from the monitoring loop is used as input to situation assessment in the target layer [29].

The tracking control loop simply describes strategies required to keep traffic within the bounds and operational constraints of a specific sector. The controller has direct influence on the ATC environment. Tasks or strategies can include scanning of the radar screen, managing flight strips, issuing instructions to pilots and communicating with adjacent sectors. The goals and criteria for the tracking loop are derived from the regulating loop. Within ATC, regulating concerns the position changes of aircraft with respect to other aircraft and weather phenomena. It involves recognizing patterns, maintaining safe and orderly flows and structuring traffic information, which corresponds to the perception and interpretation strategies discussed in Section 2-3-3. Strategies taking place at the regulating loop refer to specific plans and targets from the monitoring loop. The monitoring loop and its strategies are concerned with setting targets or objectives, activating plans and monitoring whether situations are developing as expected. Instead of position changes of aircraft with respect to other aircraft as for the regulating loop, this loop is concerned with the positioning of aircraft relative to the traffic environment. Finally, the targeting loop focuses on the overall goal of the task and the system. This loop thus takes the criteria for expediting a safe and orderly traffic flow into consideration. Trade-offs have to be made in order to meet the overall goal and to comply with operational constraints and personal factors such as workload. For example, when the workload is high, safety criteria get the highest priority and consequently the efficiency criteria are sacrificed.

2-4 ATC Sector Complexity

During the phone interview with an LVNL employee, conducted on May 15th, 2019, and during a visit to the LVNL on April 24th, 2019, complexity was often mentioned as a factor that increases during training or even during scenarios. Complexity can be very subjective and in order to make it more objective, it is worth looking into methods how to objectify sector complexity. Complexity is often directly related to the mental workload controllers are experiencing [30, 31]. Moreover, the term workload often refers to the individual load that is experienced, due to which it is difficult to assess in an objective manner. To circumnavigate this problem, several studies have tried to objectify workload and hence ATC (sector) complexity [16, 18, 32]. The methods that were looked into to define workload and hence complexity in an objective manner are Aircraft Count, Dynamic Density, Traffic Load Index and the Solution Space method and are discussed below.

The Aircraft Count Method uses, as the name suggests, the number of aircraft to be controlled as a metric for complexity [16, 18, 32]. The method is very simple and easy to implement: when more aircraft are active in a sector, the workload experienced by the controller will be higher. A major disadvantage of this method however, is that it does not take the aircraft interaction or flight characteristics into account. Different situations can lead to different amount of workload, e.g., aircraft all flying in the same direction or many routes at the same flight level that are crossing [18].

The second method, that was looked at is Dynamic Density [18, 33]. Where the Aircraft Count method lacks the contribution of aircraft interaction and flight characteristics, the Dynamic Density method was designed to take these (dynamic) characteristics into account. It is defined as "the collective effect of all factors, or variables, that contribute to sector-level air traffic control complexity or difficulty at any given time" [33]. Each variable contributing to the complexity, is given a weight based on subjective workload ratings and is calibrated for a specific sector and scenario. Since the weights for the Dynamic Density are based on subjective data, it makes the method very case specific and not extrapolatable to different sectors or situations [16, 18]. It is therefore hard to predict the complexity level and workload for different situations and sectors.

To improve the Dynamic Density method, the Traffic Load Index was developed, which also assigns weights to aircraft in order to estimate the workload experienced by the controller [16]. By default, every aircraft is assigned a weight or load of 1. Based on whether an aircraft is involved in a possible conflict or part of a high workload scenario, the weight can be increased up to 3.5. The summed total of weights assigned to aircraft in the sector results in a Traffic Load Index that corresponds to a certain workload in the sector at a specific time. Although the method takes dynamic behavior and uncertainties into account, it still relies on subjective ratings for the load factor for specific sectors [16, 30]. It is therefore also very case specific and hard to extrapolate.

Finally, a recently new method for assessing the complexity and workload is the Solution Space [16, 18]. According to this method, ATC workload and thus complexity, is inversely correlated to the space consisting of all available solutions for the controlled aircraft [16, 18]. This is referred to as the solution space. When the solution space is decreased, the controller has less maneuvering space to resolve a conflict, hence the complexity and workload for such a situation is higher. The method is based on research involving the SSD, which is discussed in Section 3-3 [13]. The Solution Space method is a promising (objective) method to assess complexity as it takes aircraft count into account as well as the dynamics, flight characteristics and interactions between different aircraft. The main disadvantage of this method is that it does not take 3D complexity into account as it was developed in 2D.

Chapter 3

Ecological Interface Design

A good way to analyze and understand the requirements for an ATCO training system, is to perform a CWA. CWA is a framework for analyzing, designing and evaluating work in complex sociotechnical systems [10, 34]. As stated in Section 2-3-1, ATC is considered to be a complex cognitive or high performance task [22]. Only when one really understands the nature of such a complex task, it is possible to design interfaces that are required for the operators to successfully perform the required tasks. An example of an approach to a CWA is the ecological approach [35]. This chapter elaborates on EID, a framework, used for designing interfaces for complex human-machine systems, that focuses on visualizing and supporting the complete range of possibilities and activities that an operator might be faced with [9, 35].

Next to the theoretical foundations and developments of EID, important tools such as the Abstraction Hierarchy (AH), the Skills, Rules and Knowledge (SRK) taxonomy and the Decision Ladder (DL) are discussed in Section 3-1. Section 3-2 describes the CWA that was performed for the ACC ATC task. Furthermore, the SSD, an ecologically-designed decision-support tool, is presented in Section 3-3. Finally, as little research has currently been conducted on the effect of training with ecologically-designed tools in the ATC environment, the application of EID in the ATC training is discussed in Section 3-4, as well as the identification of cognitive processes in order to assess novice knowledge development.

3-1 Origin and Theoretical Foundations

EID was first introduced by Vicente and Rasmussen and is a design framework that focuses on the specific problem of designing human-machine interfaces for complex sociotechnical systems [9]. It was first introduced in process control in order to increase safety but can be applied in a wide range of domains, including ATC. The term ‘Ecological’ was derived from the ecological approach to psychology which states that psychology should not be concerned with only the organism but also with its environment and the relationships between these two [36]. Translating this to the interface design problem leads to a design philosophy that focuses on making constraints and relationships in the work domain visible to the end-user. By supporting the end-user in adapting to change and novelty, it allows them to limit their core activities to higher-order problem solving and decision making [9].

Within the design framework, events in complex human-machine systems can be classified based on their degree of novelty from the perspective of the operator: familiar (routine) events, unfamiliar but anticipated events and finally, unfamiliar and unanticipated events

[9]. As traditional interface design practices, however, do not provide the right support to operators during familiar, unfamiliar and unanticipated events, a different approach was required to be able to design an interface that supports the operator in all three types of events [9]. This is where the second goal of EID comes in: the interface should support the entire range of activities an operator might be faced with [9].

Two theoretical concepts that originate from research done by Rasmussen in the process control domain are used to structure the approach to interface design. These are the AH, a framework useful for representing a work domain in a relevant way for the interface design problem, and the SRK-taxonomy, a framework that can be used for describing human behavior and processing of complex information [37, 38]. The SRK-taxonomy distinguishes three different categories of human behavior (Skill-Based Behavior (SBB), Rule-Based Behavior (RBB) and Knowledge-Based Behavior (KBB)) based on whether information is processed as a signal, sign or symbol and thus on the corresponding level of cognitive effort required to perform a task [38]. The last type of behavior, KBB, requires the largest amount of cognitive effort and usually takes place during unfamiliar situations for which no specific set of rules is available from previous experience. It is based on a symbolic representation and requires a good understanding of the underlying principles to be able to perform the correct action. This type of behavior is especially important for this research since ATC is considered to be a complex work domain in which unfamiliar and unanticipated events are not uncommon. It is therefore of importance that all required information is readily available to the ATCOs in order to construct a correct mental model, which is a representation of the operational situation and contains information such as the locations of neighboring aircraft in the sector, their directions and exit waypoints. With a correct and complete mental model ATCOs can perform the right actions during these unfamiliar events.

As the operator is considered to be a vital part of a control system, it is important to analyze the decision-making behavior of the operator. The DL, developed by Rasmussen, is one of the most common tools for describing decision-making activities and is often used as part of the CWA to analyze the control task [39]. When developing the DL, Rasmussen observed different behavior for experts compared to novices executing the same task. Where novices usually follow a linear sequence, experts are often observed to take shortcuts within the linear sequence [39]. The number of steps an operator thus strongly depends on the skills of the operator and his/her familiarity with the task. As the operator becomes more skilled, he/she is more likely to act according to RBB and follow the shortcuts within the ladder. Only occasionally will the skilled operator move through the entire sequence. Next to taking these shortcuts, a skilled operator is often inclined to start the sequence at a different point of entry instead of starting at the activation block, depending on the operator's previous experience and 'feel' for the system. More general information about EID, the AH, the SRK-taxonomy and theDL can be found in Appendix D.

3-2 Cognitive Work Analysis for Air Traffic Control

This section elaborates on the CWA for the ATC task. It consists of the analysis of the work domain, control tasks, strategies, social organization and worker competencies.

3-2-1 Work Domain Analysis

A work domain analysis is performed to identify the functional structure of the system and is independent of activities and who performs the activities [34]. Chapter 2 already discussed the main purpose of ATC, the airspace structure and the structure of the training. Taking the mission statement of safely and efficiently organizing and expediting air traffic into account, the primary tasks and job characteristics of the ACC ATC task are:

- Adhering to the airspace structure. The airspace and hence the work is organized in sectors. Within each sector, airways, routes, waypoints and stacks can be found.
- Respecting separation minima and maintaining vertical as well as horizontal separation while considering aircraft characteristics (e.g., flight envelope and performance). At all times the movement of traffic is monitored.
- Providing instructions to aircraft such as heading, speed or altitude clearances and transfers to the next sector. These instructions should also take into account pilot and company preferences.

To obtain a structured functional map of the ATC system and the activities described above, the AH, discussed in Section 3-1, can be used. The result can be seen in Figure 3-1.

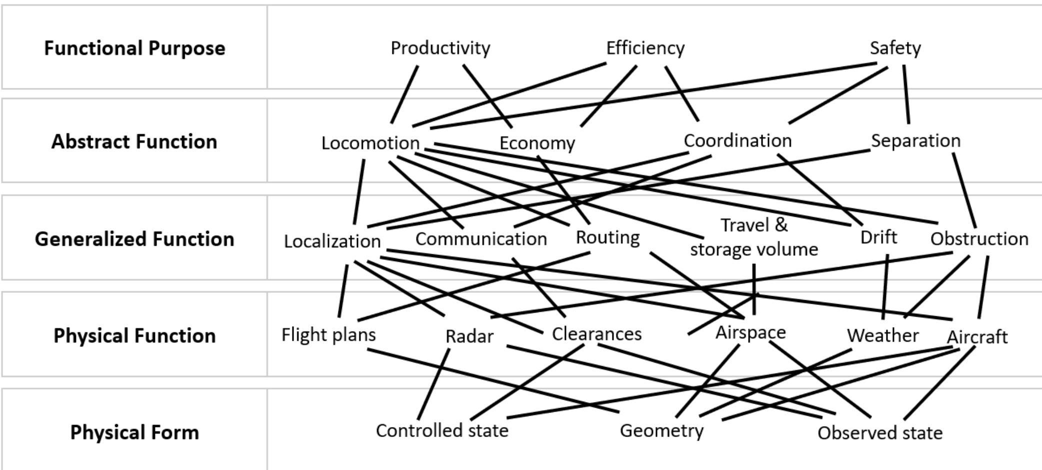


Figure 3-1: Abstraction Hierarchy for Area Control

3-2-2 Control Task Analysis

Once the functional map is constructed, the goals and states that need to be controlled or achieved can be identified by means of a Control Task Analysis [34]. This analysis only identifies and describes which tasks should be performed. It does not describe how these tasks should be performed nor who should perform them. Looking at the activities described in Section 3-2-1, the following tasks are identified for the ATC work domain:

- Welcoming aircraft when they enter the sector;
- Monitoring air traffic;

- Routing aircraft from entry to exit waypoint and reordering traffic when necessary;
- Detecting possible conflicts;
- Resolving conflicts, by means of giving clearances and rerouting aircraft;
- Temporarily holding aircraft in stacks;
- Handing over aircraft to adjacent sectors when they leave the sector.

A tool that can be used to structure and visualize these tasks is the DL (Section 3-1), as it describes which tasks should be performed in order to achieve the functional purposes identified in the Work Domain Analysis. Figure 3-2 shows the DL for the CD&R task: detecting that two aircraft are on converging flight paths and resolving this conflict by rerouting (one of) them.

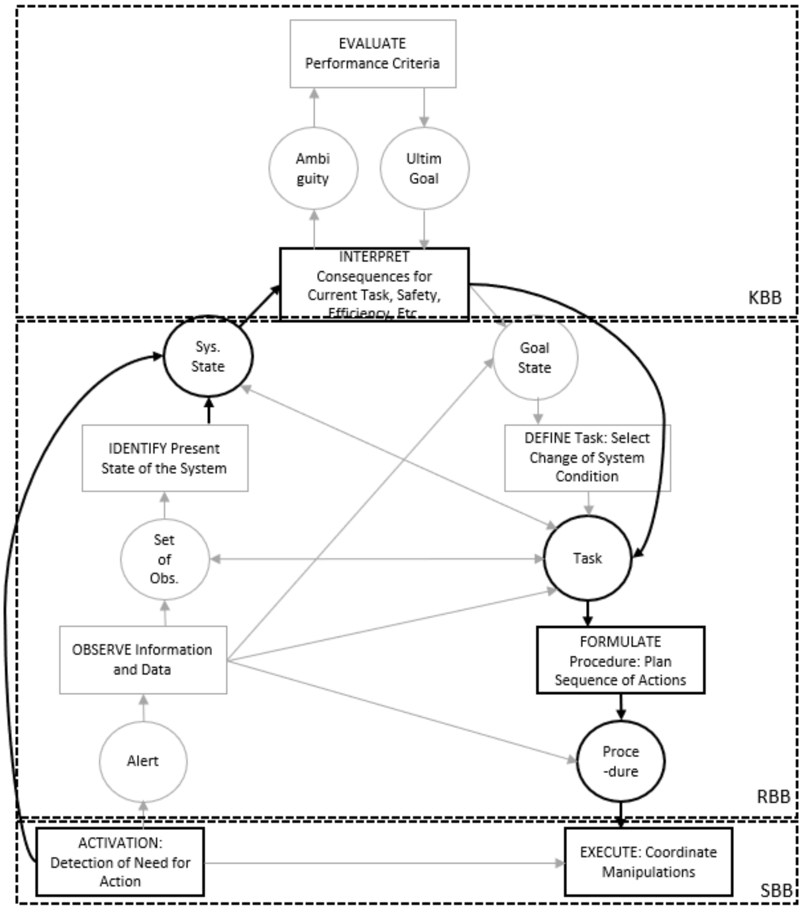


Figure 3-2: The Decision Ladder with States and Activities Highlighted that are Associated with Conflict Detection & Resolution and Rerouting an Aircraft (adapted from [39])

It can be seen in Figure 3-2, that once the knowledge exists that multiple aircraft within a sector have converging flight paths, the controller moves from the *Activation* activity to the knowledge state *System State*. From there, the controller needs to process the information to determine the criticality of the situation. This happens in the *Interpret* activity. Once the information has been interpreted, the controller leaps to the next knowledge state: *Task*. This state corresponds to the knowledge of what the task entails and hence which flight paths must be adapted to correct the situation and avoid a LoS. The controller is able to make a leap to this knowledge state based on previous experience that flight paths should be modified in case of (potential) conflicts. Next, the controller must construct a plan of action to achieve the goal, namely a prevention of a LoS. This step involves selecting a strategy and constructing a sequence of actions to accomplish the rerouting, also *Formulating a procedure*. It naturally follows that the next knowledge state is *Procedure* and the controller possesses the knowledge of the right strategy to prevent a LoS. Finally, in the *Execute* activity, the controller executes the modifications to the flight paths to correct for the converging paths and completes the task.

Although the DL is depicted here, the controller also moves down through the different abstraction levels of the AH. The ATCO starts at the Functional Purpose level with the obtained knowledge about the system state. By interpreting this information and determining the consequences for the current situation, the ATCO moves to the Abstract Function level. Answering the questions of how critical the situation is and how best to solve this situation leads to a new knowledge state and a new abstraction level, namely the Generalized Function. From there the Physical Function abstraction level is reached when the controller obtains the knowledge about the desired strategy to solve the potential conflict. Moving to the last level of abstraction, Physical Form, corresponds to the activity of executing the planned modifications.

3-2-3 Strategies Analysis

When the tasks have been identified that need to be performed within the work domain, the next question that arises is the question of how to perform these tasks. Again, the strategies are irrespective of any controller(s) performing the work. A strategy is in this case defined as a category of cognitive task procedures that transform an initial knowledge state into a final knowledge state [34].

Taking the tasks that correspond to CD&R that have been identified in the previous section, these can be described in further detail:

- Monitoring air traffic and detecting possible conflicts:
  - Inspect aircraft altitudes, speeds and headings;
  - Inspect crossing flight paths and merging points;
  - Focus on aircraft pairs that are on converging paths.
- Routing aircraft from entry to exit waypoint and reordering traffic to resolve (potential) conflicts:
  - Modify aircraft altitude, speed and/or heading;
  - Re-organize traffic patterns and routing structure.

Information flow maps are often used to conduct strategies analyses. They are graphical representations of information-processing activities and knowledge states corresponding to certain strategies [34]. Taking the rerouting task as a result of a possible conflict into consideration again, three strategies can be identified for performing this task, namely holding, rerouting or tweaking an aircraft that is on a converging path.

- Holding an aircraft: Out of two aircraft that are flying on converging paths, one aircraft is chosen that is ordered to hold. The other aircraft continues on its original flight path. Once the aircraft are no longer on converging paths, the held aircraft is released. The information flow map for this strategy is found in Figure 3-3 [34]. This strategy will likely result in the least amount of additional cognitive load as one of the two aircraft does not need to be monitored during the holding time window. It is, however, also inefficient as the aircraft that is being held, is not making any progress on its flight path.

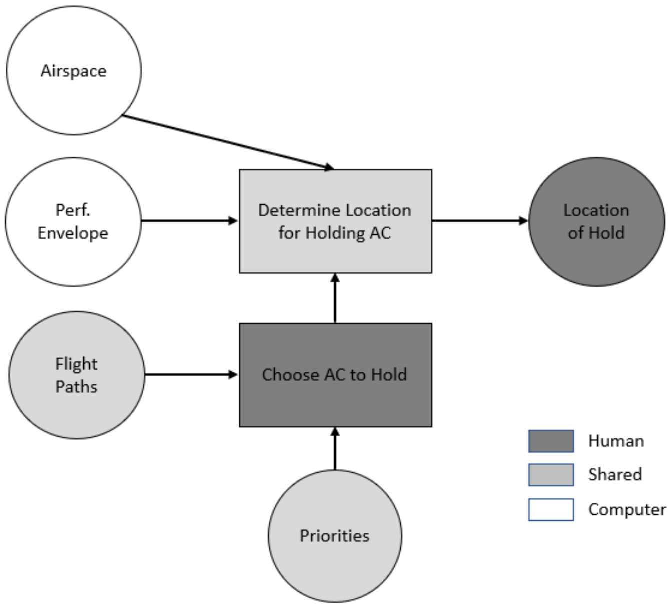


Figure 3-3: Information Flow Map for the Holding Strategy

- Rerouting an aircraft: Out of two aircraft that are flying on converging paths, one aircraft is redirected to a different route while the other aircraft continues on its original flight path. The aircraft that is being redirected is redirected to a path such that the two aircraft will no longer be flying on converging paths. The information flow map for this strategy is found in Figure 3-4 [34]. This strategy is considered to be more efficient than the holding strategy, as both aircraft make progress on their flight paths. It does, however, pose a higher cognitive load on the controller as a new route needs to be selected and both aircraft need to be monitored.

- Tweaking an aircraft: Out of two aircraft that are flying on converging paths, one aircraft is given a series of clearances to slightly alter its flight path in such a way that the potential conflict is eliminated. The information flow map for this strategy is found in Figure 3-5 [34]. This strategy is, similar to the rerouting strategy, considered to be more efficient than the holding strategy, as both aircraft remain on their flight paths and make progress toward their destination. Similar to the rerouting strategy, the tweaking strategy does, however, pose a higher cognitive load on the controller as both aircraft constantly need to be monitored.

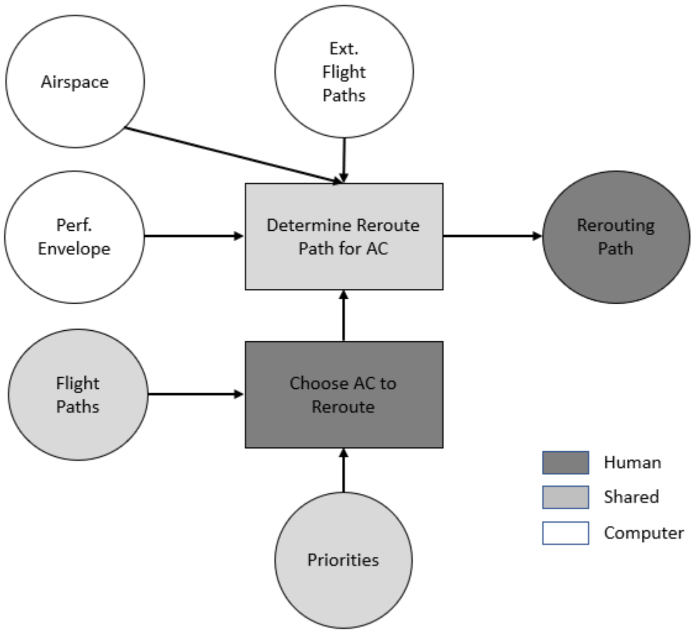


Figure 3-4: Information Flow Map for the Rerouting Strategy

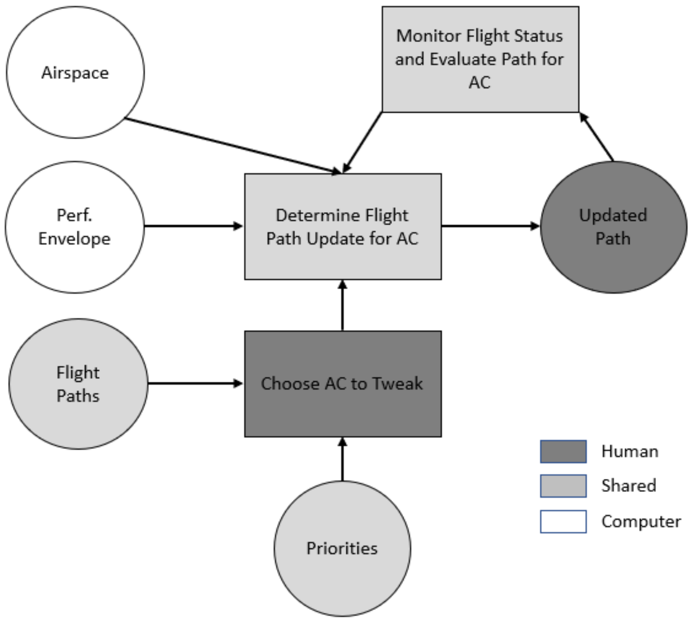


Figure 3-5: Information Flow Map for the Tweaking Strategy

3-2-4 Social Organization

Up until this part of the CWA, only the tasks related to ATC, CD&R and rerouting have been described but not which actors are involved in performing these tasks. The Social Organization describes how responsibility is allocated for the work domain, control tasks and strategies across different actors [34]. These responsibilities can be allocated to both humans and automation. The results of the social organization analysis can also be found in Figures 3-3, 3-4 and 3-5. It can be seen that, similar to the AH and DL, the process moves through different levels. The rerouting task, in Figure 3-4, starts by observing and interpreting information that is collected and presented by computers to the controller. Next, when deciding on the target state and while formulating a procedure, both the human as well as the computer are involved. Finally, executing the task and performing the modifications is done by the human operator.

3-2-5 Worker Competencies

The last part of the CWA is the Worker Competencies Analysis. In this analysis, the constraints and capabilities, associated with the different actors involved in the task, that are required to successfully perform the task and function effectively, are identified [34]. The analysis identifies what SRK behavior is required per step to successfully perform the work. Figure 3-2 shows the division of SRK behavior on the DL.

The first step in the CD&R and rerouting task is scanning for aircraft and determining whether they are flying on converging flight paths. These steps are associated with both SBB and RBB as controllers are constantly monitoring the movement of aircraft and scanning for signs that indicate two aircraft might be in conflict in the near future. Moving from the system state to the activity block of interpreting the consequences, the controller moves from RBB to KBB as he/she is constantly calculating the aircraft states to predict their future location and determining what the criticality of the situation is. Determining the strategy and a set of actions is done while moving down the ladder and therefore corresponds again to RBB. Finally, executing the task is a type of behavior that belongs to the skill-based domain.

3-3 Solution Space Diagram

The SSD is a tool that was designed according to the EID principles [13]. While the interface was originally designed for airborne self-separation by pilots, it was later adapted with a focus on air traffic management capabilities [13, 40]. The EID principles work especially well with the open and dynamic nature of the ATC environment. With regular unanticipated events, the SSD provides an instant overview of all solution possibilities in the 2D plane to the controller. Providing the solutions in this manner allows for shortcuts in Rasmussen's DL [12]. More difficult tasks which require more cognitive effort, considered to be part of KBB, can then be moved towards the quicker Rule- (or Skill-)Based behavior. In short, the constraint-based user interface shows the controller the available control area for the controlled aircraft with respect to other observed aircraft in terms of heading and speed. As stated in Section 2-2, a conflict is defined as the event where an aircraft (potentially) enters another aircraft's PZ, also called a LoS. Figure 3-6a shows an imminent LoS and the construction of the corresponding solution space [12]. The translation from relative to absolute space as well as the construction of the SSD for this scenario, is described below.

The scenario depicted in Figure 3-6 considers two aircraft, a controlled aircraft A with velocity V_1 and an observed aircraft B with velocity V_2. The PZ of aircraft B is depicted by the circle around it and thus represents the minimum separation distance that should be maintained at all times. Drawing tangent lines from aircraft A to both sides of the minimum separation circle of the observed aircraft results in a so-called Forbidden Beam Zone (FBZ) or conflict zone [13]. The gray area between these lines indicates that the aircraft will experience a LoS

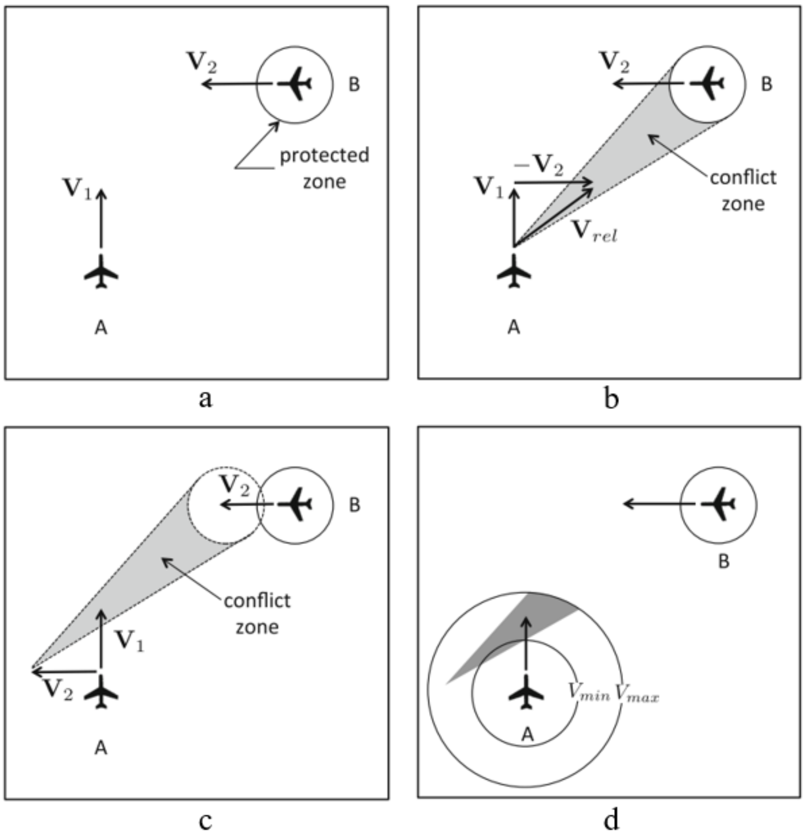


Figure 3-6: Construction of the Solution Space Diagram [12]

in the near future if the relative velocity of aircraft A with respect to aircraft B lies inside this area. Thus, any combination of V_1 and V_2 where the resulting relative velocity vector V_{rel} lies inside this triangle, will lead to a LoS. The gray area is thus a collection of relative velocities that will result in an aircraft entering another aircraft's PZ in the near future. This is shown in Figure 3-6b. Next, the conflict zone can be translated to the SSD of the controlled aircraft by transposing the origin of the conflict zone by V_2, the velocity vector of the observed aircraft. This is shown in Figure 3-6c. Finally, mapping the minimum and maximum velocity, the speed envelope, of the controlled aircraft on the transposed conflict zone results in the complete SSD, as is shown in Figure 3-6d. It can be seen that the velocity vector V_1, showing the direction and the magnitude of the velocity of the controlled aircraft, is within the velocity limits of the aircraft and is directed into the gray conflict area. If aircraft A thus continues flying with its current heading and speed and no action is taken, a LoS will occur.

The SSD thus shows the locomotion constraints imposed by the presence of observed aircraft on heading and speed commands for a controlled aircraft. This visualization allows controllers to detect conflicts and avoid losses of separation by moving the velocity vector of the controlled aircraft outside of the conflict zones by giving heading and speed clearances. Any clearance that directs the speed vector outside the conflict zones results in a safe separation. This may, however, not always be the optimal solution. A clearance that is both safe and efficient would direct the controlled aircraft into a safe area closest to the conflict zone and closest to the destination waypoint. Such a clearance would result in the smallest state change with the least additional miles relative to the original flight path. It is up to the controller to decide on the best strategy to balance safety, efficiency and productivity. When doing this, the controller does not have to determine or consider the relative velocities but sees the complete solution space at a single glance. The SSD furthermore allows the controller to link conflict zones to observed aircraft by looking at the shape and orientation of the conflict zones. By doing this, the controller is able to roughly determine the location and flight direction of the neighboring aircraft. The base of the conflict zone triangle usually points to the involved aircraft at a slight offset and the width of the triangle indicates the proximity of the involved neighboring aircraft. A small width indicates a far-away aircraft, while a triangle with a larger width indicates a nearby aircraft. Finally, by drawing a line from the controlled aircraft toward the tip of the conflict zone triangle, the absolute speed vector of the observed aircraft can be determined. In the example shown in Figure 3-6 a LoS will occur if aircraft A continues flying with its current heading and speed and no action is taken. Aircraft A must modify its speed vector in such a way that it will be placed outside the FBZ of aircraft B. Figure 3-6d shows that a change in heading will in this case resolve the conflict.

By integrating constraints found on the lower levels of the AH and showing how these constraints affect the solution space of the controlled aircraft in terms of heading and speed, the SSD can be referred to as a visualization of the Abstract Function level of the AH. By using the SSD to determine the course of neighboring aircraft, the controller is able to move from higher-level functional information of the AH down toward lower-level objects. Although the SSD might furthermore be a good tool to develop expert-like behavior and to reach higher regions in Rasmussen's DL, a risk of this interface is that surface learning could occur which would lead to the development of shallow knowledge in case the interface would be used as a rule-based tool only [11]. A larger dependency on the interface might be developed when the interface is used in such a way, which is one of the ironies of automation [41].

Although research has been conducted into including altitude in the SSD [42, 43], the simpler and more basic 2D version of the diagram will be used in this research. Including altitude in the diagram results in an increased visual complexity and will thus increase the difficulty of understanding the visualizations. Using the basic diagram means a greater resemblance to the actual display as used by ATCOs and limits the complexity of traffic scenarios. Next to using the SSD as a tool in CD&R, it can also be used to define a metric for complexity, as was discussed in Section 2-4 [18]. The SSD contains only visual information and cannot be applied as a metric directly. However, since flight path, longitudinal and lateral separation, direction of flight and relative velocity can all be seen in or derived from the SSD, the conflict area depicted in the SSD can be used as a metric to define complexity. A larger conflict area indicates a smaller solution space and thus a more complex conflict or scenario.

3-4 EID and ATC Training

Central in this research is determining the effect of EID on performance, skills and knowledge development in an ATC work domain. How EID and the ATC training relate to each other and how the effect of EID on this development can be predicted is discussed in this section.

3-4-1 EID Applied to ATC Training

As the SSD could be helpful in developing expert-like behavior and thus a deeper insight and understanding of the system, it is argued that the tool can be used to shape the internal mental model [12]. As already mentioned in Chapter 1, it has previously been shown that EID can indeed lead to better knowledge development, a functionally-organized knowledge base and increased performance after an exposure of six months [11]. Research on short-term effects of EID on knowledge development also suggests that EID encourages KBB and generates goal-oriented thoughts and can therefore play an important role in the early stages of knowledge development [12]. The SSD, designed according to EID principles, is expected to increase the performance during training without negatively influencing the deep knowledge of the controller [12].

With the mission statement of ATC that is focused on safety and efficiency, the main focus during ATC training is put on expediting traffic in the safest and most efficient way. As discussed in Section 2-3-3, trainees are taught several strategies and procedures to do this, but a large segment of ATC training still consists of discovery learning in which trainees find out what works best by means of trial and error. To develop more robust knowledge of the system, trainees are discouraged from trying to develop a solve-all strategy. By introducing unexpected and unfamiliar events in the training scenarios and assignments, the trainees' versatility and resourcefulness are trained and knowledge-based problem-solving is encouraged. As EID accounts for unanticipated and unfamiliar events or situations by providing the controller with the complete range of solutions, an overlap is found between EID and ATC training.

Next to this, in order to successfully accomplish the mission statement of ATC, it is of importance that an ATCOs develops a correct mental model. As EID contributes to a larger SA [44], another overlap is found between EID and ATC training. Elements from EID such as Rasmussen's AH can be used to contribute to an increased SA by grouping different domain-relevant constraints at different levels of abstraction [37]. This information can then be used to structure the information on a display which can in turn be used as an external mental model to structure the internal mental model. The actions taken by controllers after consulting the SSD can give insight into their thinking pattern and the order of their cognitive steps. The cognitive steps taken during the training scenarios can be visualized on the theoretical ECOM and DL and possible shortcuts between different behavioral domains taken by the controller can be identified at different stages during the training. These shortcuts are also discussed in Section 3-4-2.

As stated before, the means with which ATCOs fulfill their tasks, have remained largely unchanged over the past decades, just as the training of ATCOs has seen little change over the same time span. The ATC training today is still largely based on subjective expert opinions, although significant effort has already been put into objectifying the training and especially the assessment of the trainees by developing cognitive models and visualizing thinking patterns. To objectify the training even more, an ecologically-designed decision-support tool such as the SSD could be a promising tool as coaches are then able to use the tool to base their feedback and instructions on.

3-4-2 Visualizing ATCO Cognitive Processes

As described in the previous section, the actions taken by controllers after consulting the SSD can give insight into their (invisible) thinking patterns and the order of their cognitive steps. The effect of the SSD on knowledge development can be analyzed by making use of Rasmussen's DL. The DL can be used as a template to structure the cognitive processes and thus to determine which information-processing activities are taking place.

While the actions after consulting the SSD can give insight into the thinking patterns of controllers, the action preceding this consultation is just as important in determining which type of behavior is displayed by the controller. An ATCO might have very different reasons for consulting the SSD (e.g., a warning, a predicted possible conflict, monitoring the situation or for no reason at all), which could all result in different types of behavior.

To be able to analyze what effect the EID interface, or the SSD, has on the performance and knowledge development of the controllers, the SSD is split into different elements. Each of these elements corresponds to a certain cognitive shortcut that can be visualized in the DL. The SSD elements are:

- The velocity vector indicating the heading of the aircraft and the magnitude of the aircraft's velocity;
- The inner circle depicting the aircraft's minimum speed;
- The outer circle depicting the aircraft's maximum speed;
- The area between the two circles, also known as the solution space;
- The conflict triangles or gray areas within the diagram.

The first three elements are all related to the aircraft's velocity. They present the controller with information about the aircraft's current velocity as well as a full overview of the aircraft's velocity range. They thus provides the controller with instant information on the possibilities for speed vectoring without the need to first identify the aircraft type and altitude to make estimated guesses about these possibilities. In terms of the DL this means that the information of the system state is visually presented by the SSD and the controller does not need to process the information first to come to the same conclusion. A leap can thus be taken in the DL from the knowledge state *Set of Observations* to the knowledge state *System State*. This is depicted by the blue arrow in Figure 3-7.

The last two elements also provide the controller with different types of information. The shape, size and number of the conflict triangles provide the controller with information about the number of neighboring aircraft and their locations. The conflict triangles furthermore represent areas that will result in a conflict when the velocity vector of the controlled aircraft is directed into this area. The triangles thus provide information about the future state of the aircraft in terms of conflict detection, which the controller normally has to estimate by him/herself without the use of the SSD. Additionally, more conflict triangles mean a busier airspace and a smaller solution space. The solution space between the two circles does however not only show how busy the airspace is but also what the complexity of the solution that is being sought is. The SSD furthermore assists in conflict resolution as it shows the solution possibilities close to the speed vector of the controlled aircraft and thus hints on the amount of deviation from the original route that is needed to resolve a conflict. The conflict areas show exactly where the boundaries lie between a conflict and no conflict and thus commands can be given with great accuracy. Without the SSD, the distinction between solution space and conflict area is very hard to estimate and thus less precise commands are given that result in larger deviations from the original path. This is another example of information that becomes directly available to the controller by using the SSD, rather than first scanning the airspace and the aircraft states of neighboring aircraft to reach the same conclusion. In terms of the DL, the solution space and conflict triangles can be visualized by means of two shortcuts. The SSD provides the controller in a single glance with all required information

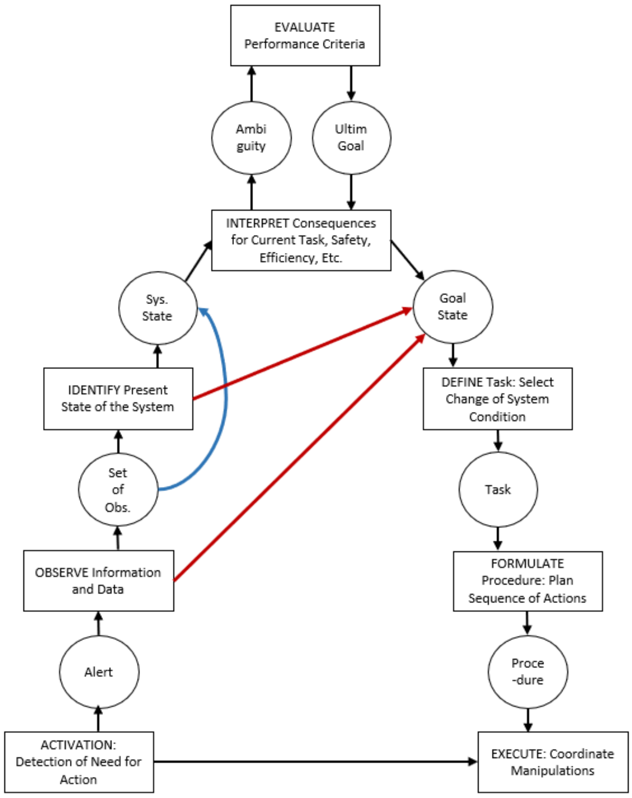


Figure 3-7: Examples of Leaps and Shunts Visualized on the Decision Ladder as a Result of Using the Solution Space Diagram

on the current system state but also shows the possibilities for the future or goal state. Two shunts can thus occur from the actions *Observe* and *Identify* to the knowledge state *Goal State*, as is depicted by the red arrows in Figure 3-7.

The SSD elements thus present visual solutions that correspond to shortcuts in the DL. Because of these shortcuts, the controller does not have to mentally create these solutions by him/herself and thus the cognitive load is reduced. Visualizing the mental shortcuts, induced by the different SSD elements, on the DL allows for the analysis of behavior types demonstrated by the novice controllers, such as more goal- or task-oriented behavior. The most important behavior types that are focused on in this research are:

- System-Sate- and Goal-State-Oriented behavior;
- Task-oriented behavior;
- Procedure-Oriented behavior;
- Action-Oriented behavior.

System-State- and Goal-State-Oriented Behavior

In between problem analysis and the implementation of a solution and without making use of the SSD, ATCOs usually compare the system state to the goal state by interpreting the situation in the top loop of the DL. An experienced ATCO will thus first consider several goal-oriented solutions in the upper iterative (KBB) loop of the DL while keeping the current system state in mind. When making use of the SSD, a situation assessment can be conducted by one or multiple consultations of the SSD. A shortcut is then created from either the *Observe* or *Identify* action to the knowledge state *Goal State*, as is depicted by the red arrows in Figure 3-7. This type of behavior corresponds to more abstract and higher-level thinking and usually involves forming plans or monitoring how certain situations develop. ATCOs can, for example, analyze the current situation and look for any potential conflicts, prepare for a peak in the workload or plan ahead to increase efficiency. System State and Goal State oriented behavior corresponds with proactive behavior that is usually seen from expert controllers.

Task-Oriented Behavior

A type of behavior that is less abstract than system-state- and goal-state-oriented behavior, occurs when a task is evident from a certain situation and can be selected without further evaluation. An example is when an aircraft enters the sector and after consulting the SSD, it becomes evident that the aircraft needs to be vectored to avoid a conflict in the future. The ATCO selects the task of vectoring the aircraft and can proceed to the action of finding the right procedure to safely execute this task. When a controller is able to select a task without further delay, this corresponds to RBB. However, in case a new procedure is required, the task calls upon KBB to evaluate different procedures and their effect on the final system state. In terms of the DL, Task-oriented behavior can be visualized by shortcuts from the actions *Observe* and *Identify* to the knowledge state *Task*, as can be seen in Figure 3-8.

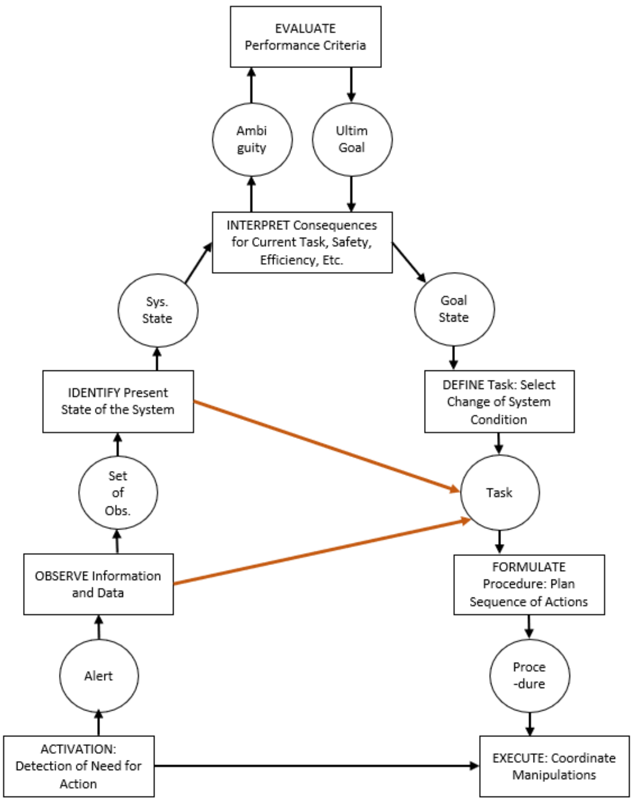


Figure 3-8: Task-Oriented Behavior Visualized on the Decision Ladder

Procedure-Oriented Behavior

The type of behavior that is least abstract and leads to a quick selection and execution of procedures is called Procedure-oriented behavior. For this type of behavior, consulting the SSD can lead to the quick selection and execution of a vectoring procedure in case of an imminent conflict, without further consulting or evaluating the surrounding environment. In the DL this is represented by two shortcuts from the *Observe* and *Identify* actions to the knowledge state *Procedure*, as is visualized in Figure 3-9.

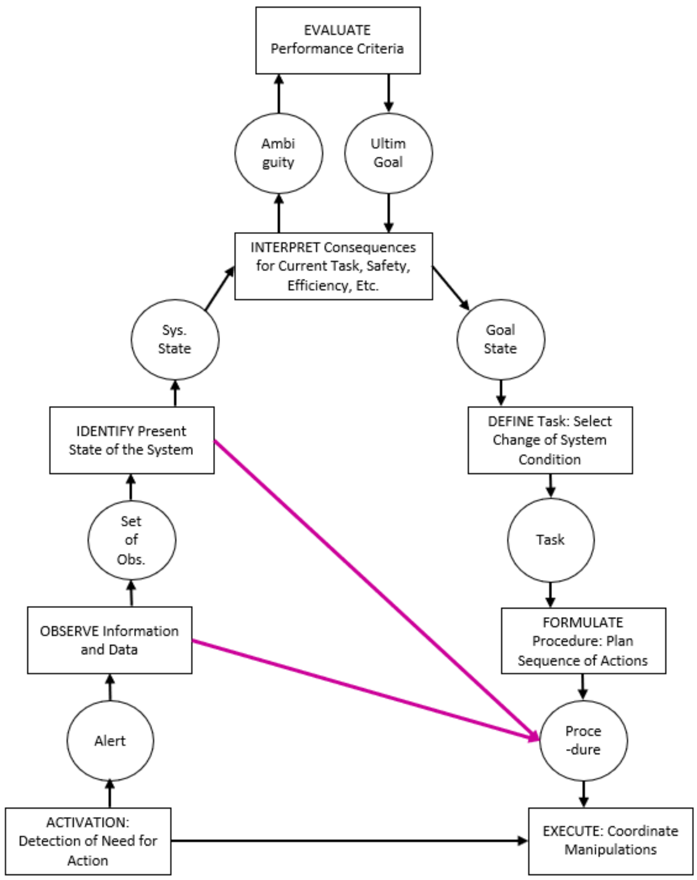


Figure 3-9: Procedure-Oriented Behavior Visualized on the Decision Ladder

Action-Oriented Behavior

It should be noted that a fourth type of behavior can also be distinguished. This type of behavior can be visualized on the DL by a shortcut between the actions *Activation* and *Execute* and corresponds to SBB. In case of a LoS, the controller is activated by an alert either via sound or visuals and immediately vectors the aircraft to a safe state without consulting the environment. This type of behavior is therefore disregarded for this research This shortcut in the DL is visualized in Figure 3-10.

Section 5-2-8 further elaborates on how the theory about the different types of behavior is linked to reality in the experiment. It is expected that by using the SSD in ATC training,

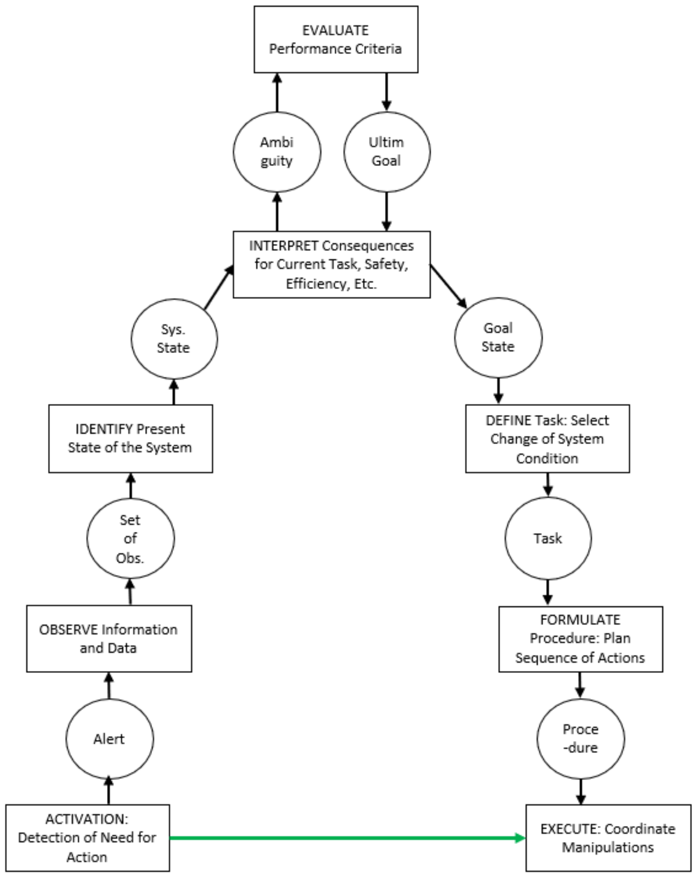


Figure 3-10: Action-Oriented Behavior Visualized on the Decision Ladder

more proactive and expert-like behavior is seen from novices. As stated before, a concern does, however, still exist that the experiment will see 'lazy' use of the SSD, characterized by performing rule-based shortcuts only. From the experiment, certain strategies and behavioral markers need to be identified in order to match the behavior of the participants to the DL. This is done using a think-aloud protocol to give more insight into the thinking patterns of the controllers. Next to this subjective data, objective performance data from the logging of experiment scenarios will further help to identify the different types of behavior found among the novice participants.

Chapter 4

Complex Learning

This chapter provides the necessary background information concerning complex learning. First, an approach to designing a complex learning task, the Four-Component Instructional Design model, is elaborated on in Section 4-1. Subsequently, Section 4-2 discusses cognitive load reduction by means of scaffolding. Finally, the application of scaffolding in EID and ATC training is discussed in Section 4-3.

4-1 The Four-Component Instructional Design Model

There are many approaches when it comes to learning complex and dynamic tasks such as the ATC task. Complex learning can be defined as the integration of knowledge, skills and attitudes, also referred to as competencies [15], as can be read in Section 2-3-1. Some examples of models developed to promote complex learning are cognitive apprenticeship, collaborative problem solving, learning by doing and the Four-Component Instructional Design model [15]. When the standards or performance criteria for a certain competence have been met, the controller is said to have successfully obtained said competence and should thus be able to combine skills and transfer what was learned during training to new and sometimes unexpected situations. This last learning goal, transfer of learning, is often difficult to achieve.

According to instructional design, which can be referred to as "the systematic and professional planning and implementation of education or training" [45, p. 196], a holistic design approach is necessary to increase the chances of a successful transfer of learning [15]. Holistic design approaches deal with the learning domain or system as a whole and thus, do not lose sight of separate elements and the interactions between them [15]. It is the opposite of atomistic design, where complex tasks are generally reduced to their simplest elements, which are easily transferable to learners. Atomistic design approaches work very well, given that there are few interactions between the different elements of a task. A holistic design, however, solves the problems of compartmentalization, fragmentation and the transfer paradox [15]:

- Compartmentalization: separating a task into different parts or categories, such as the cognitive, affective or psychomotor learning domains. Holistic design integrates these learning domains to facilitate an integrated knowledge base that increases the chances of a transfer of learning.
- Fragmentation: breaking tasks of a learning domain further down into small, incomplete and isolated parts (e.g., recalling facts, applying a certain procedure). For each of these

parts, a different instructional model is applied with a different objective. Students learn one part or sub-skill at a time, until finally, the student practices the complete and complex skill. However, this fragmentation fails to take the interactions between the elements into account and therefore, students are not able to integrate the different elements in a transfer situation.

- The Transfer Paradox: in instructional design, efficiency is often one of the main goals. To this end, instructional models are often selected that minimize the number of practice items, time spent on a task and students' investment or effort to achieve the set of objectives. Although some methods might be more efficient than others, they often result in a low transfer of learning, as the models encourage students to construct very specific knowledge. In the long run, models are preferred that are sometimes less efficient but encourage students to construct general and abstract knowledge that can be used in unfamiliar situations and thereby increase the chances of a successful transfer of learning. This is also known as the Transfer Paradox.

As stated before, one of the models that follows the holistic design principles and can be applied for a complex learning task is the Four-Component Instructional Design model. This model assumes that blueprints for complex learning can always be described by four components: learning tasks, supportive information, procedural information and part-task practice [14]. A schematic of the model is shown in Figure 4-1 [15]. Each of the components is discussed below.

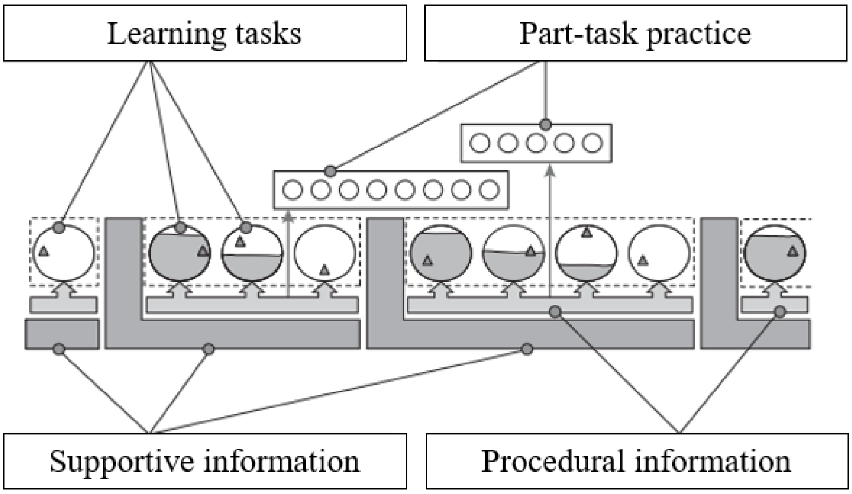


Figure 4-1: Schematic of the Four-Component Instructional Design Model [15]

Learning Tasks

Students should preferably work on tasks that aid in developing an integrated knowledge base by inductive learning [15]. This means that tasks should be based on concrete and real-life experiences. Furthermore, each learning task should be a so-called whole task, meaning that all or nearly all constituent skills, including the corresponding knowledge and attitude, that are important to be able to successfully complete a task, should be introduced to the student. This way, students develop a holistic view and approach early on in the learning process. Learning tasks should furthermore be part of easy-to-difficult task classes, have a high variability and have a decreasing level of support when moving through the task sequence. These requirements are elaborated on below.

- *Task classes:* a task class is defined as a category of learning tasks that represents a version of the task with the same level of difficulty. This is indicated by the circles (learning tasks) that are grouped together in Figure 4-2 and means that all tasks within a certain class can be performed with the same amount and type of knowledge. Students should start to work on relatively easy whole tasks and from there continue with learning tasks with increasing difficulty and complexity. Determining the complexity of a task can, however, be very difficult and often requires the input from domain experts to discover which factors influence complexity.

Figure 4-2: Schematic of a Learning Task, Including Task Classes

- *Variability:* to construct a more general and abstract knowledge base, it is important that the chosen learning tasks contain a high variability on all dimensions that also vary in real-life situations. Variability of practice has repeatedly been shown to be the most important factor in predicting a transfer of learning [14]. Variability of practice is indicated in Figure 4-3 by the triangles that are changing location within the learning task and within the task classes.

Figure 4-3: Schematic of a Learning Task, Including Task Classes and Variability

- *Support and Guidance:* it is of great importance that students receive support and guidance when starting on a new (more difficult) task class. Task support is aimed at providing students assistance with the products involved in the training (e.g., given information, goal and information that gets the student from the given information to the goal). Guidance, however, focuses on providing the students with assistance on the processes taking place to successfully complete the learning task. The level of support and guidance generally follows a sawtooth pattern according to the method of scaffolding: at the start of a new task class, the level of support is high and decreases as the students' expertise increases and a conventional task can eventually be carried out without any support. The decreasing level of support is indicated by the decreasing gray area within the circles, depicted in Figure 4-4.

Figure 4-4: Schematic of a Learning Task, Including Task Classes, Variability and Support

Supportive Information

Supportive information is information that allows students to perform non-routine aspects of a task, explains how a learning domain is organized and how domain-specific problems should be approached [14, 15]. The purpose of this type of information is to assist in the process of schema construction such that students are able to deeply process new information and to connect this to already existing information and schemes in their memory. The information is usually presented when students start working on a new task class and should always be available during that task class in order for students to go back and forth between the task and the information. This is usually most effective for novices [14]. The more complex a task class, the more knowledge and information of a domain is usually required.

Procedural Information

Procedural information is information that allows students to perform routine aspects of a task [14, 15]. It specifies how to perform these routine aspects during a task and is preferably presented just in time as step-by-step information (e.g., presented by a supervisor, posters on a wall, a pop-up on a screen). By having information presented on how to perform the task while performing it, cognitive rules are constructed on how to perform the task and rule automation is encouraged.

Part-Task Practice

Although, as stated earlier, learning tasks should preferably be whole tasks based on real-life situations and experiences in order to prevent compartmentalization and fragmentation, situations may arise where part-task practice is necessary. This is the case when a certain aspect is required to be performed on a highly-automated cognitive level. For these aspects, additional part-task practice should be provided (e.g., practicing musical scales). Part-task practice involves significant amounts of repetition (strengthening) and should, however, only be provided after the aspect has been introduced in the context of the whole task [14, 15].

4-2 Cognitive Load Management by Means of Scaffolding

When students are learning complex tasks, such as the ATC task, they can sometimes be overwhelmed by the amount of information and the complexity of the task. For students to successfully obtain a skill or competence, it is therefore of importance to manage the cognitive load during the learning process [46]. One method for managing cognitive load, that has also briefly been mentioned in the previous section, is scaffolding [47].

Depending on the context, scaffolding can have different definitions. A scaffold can be referred to as a temporary framework that is used to support workers during the construction or modification of a building. Once the construction has been completed, the scaffold is removed. Translating this definition to instructional design, a scaffold can be defined as a temporary framework to support learners in accomplishing a task and removing this framework when the support is no longer required [48].

As mentioned in the previous section, support can be varied throughout a task class. Initially, the support allows a student to perform a task or achieve a goal that would not have been achievable without the support. As the student's expertise increases, support gradually decreases until the student no longer needs the support and is able to perform the task independently [48, 49]. This is also referred to as scaffolding and is based on Vygotsky's concept of the Zone of Proximal Development (ZPD) which is defined as "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with a more capable peer" [49, p. 238]. The ZPD thus refers to the actual level of development of the student and the next level attainable through the use of support via tools or human guidance [49].

It is important to balance the amount of support as well as to determine the right type of support as excessive or insufficient support can obstruct the learning process [47, 50]. The support generally guides the student during the execution of a task and should not direct the student. Another important factor to take into consideration when designing the support for a task, is the timing and rate at which the level of support is decreased. Next to these considerations, it is critical that the support, especially for novice learners, is fully integrated or embedded in the task environment. When this is not the case, split-attention effects could occur and the cognitive load could be increased, while support is meant to decrease this load [47]. A perfect example of embedded support are training wheels on children's bikes, which are considered to be more effective than non-integrated support, such as the parent running next to the child while directing the child to keep the handlebar straight. In scaffolding whole-task practices, two complementary approaches can be distinguished according to the Four-Component Instructional Design model: simple-to-complex sequencing, that aims to decrease the intrinsic cognitive load for novices, and varying types of learning tasks, that focuses on decreasing extraneous cognitive load [47].

Simple-to-Complex Sequencing

The fact that students can sometimes feel overwhelmed when learning a new and complex task, already indicates that starting a course with highly-complex learning tasks can result in negative learning effects. A solution that is commonly implemented is to let students start with relatively simple tasks and increase the complexity as the course progresses [47].

For the part-task approach to a task with components that are hardly interacting, this works especially well. Part-task learning approaches to sequencing work very well when trying to prevent an increase in intrinsic cognitive load for novices as the load of a distinct part is considered significantly lower than the cognitive load associated with the whole task [47]. However, when learning complex tasks that require much coordination between separate highly-interacting parts, this approach does not work as well. By reducing the cognitive load according to this approach, the chances of a successful transfer of learning are decreased as students are not able to integrate the separate parts of knowledge and apply it in new and unfamiliar events, as was also discussed in Section 4-1.

It is therefore often stressed that students should obtain a holistic vision of the whole task from the start of the learning process, which corresponds to the 'global before local skills' principle [47]. Ideally, the learning process starts with the most simple and authentic case that a professional could encounter in a real-life situation and from there, the complexity can be increased by decreasing the number of simplified factors. Task classes with a lower complexity are associated with a lower cognitive load because the cognitive schemes do not yet contain all elements required to solve a task when simplifications are removed or relaxed. For more complex task classes, the number of elements and especially interactions between them, increases, requiring a larger working memory and hence a higher cognitive load is experienced. In the whole-task approach, each learning task class represents a different level of complexity and thus contains a set of learning tasks that can be performed with the same amount and type of knowledge (e.g., mental models, cognitive strategies or other cognitive schemes). A task class that is more complex thus requires more knowledge or more elaborate knowledge for effective performance than a preceding, simpler class.

Types of Learning Tasks

Learning tasks are generally associated with conventional problems that require the student to apply (weak) problem-solving methods, leading to a high cognitive load while hardly encouraging schema-construction processes [15, 46].

An alternative learning task format is, for example, studying worked-out examples [47]. This format provides students with the given and goal state and the example solution. Studying these examples helps students to build up generalized knowledge and construct cognitive schemes and has furthermore proven to be more effective for a successful transfer of learning than actually solving the problems [47]. A disadvantage however, is that students tend to not study the examples carefully and only briefly look at them or only consult them when necessary. Studying the example while carrying out a task, however, actually leads to an increased cognitive load. To increase the level of involvement of students, a different effective method is the use of completion tasks. These tasks contain only a partial solution and encourage students to be active learners, as students are required to finish the rest of the problem themselves [47]. Finally, another type of learning task is a reverse task [47]. Reverse tasks ask the student for which situations a particular solution may be helpful in reaching the given goal state.

Decreasing the level of support for learning tasks according to the various types of learning tasks can easily be realized and has been proven to show positive learning effects as it decreases the extraneous cognitive load [47]. One can, for example, start out with worked-out examples, move to completion assignments for which larger and larger parts should eventually be completed and finally move to conventional problems. Within one task class, the cognitive load is thus controlled by starting out with learning tasks with a high level of built-in support, progressing with learning tasks with intermediate levels of support and ending with conventional tasks without any support.

4-3 Scaffolding Applied to Cognitive Tools

Scaffolding is traditionally used to refer to the process in which a teacher or more knowledgeable person assists the learner in accomplishing a task that would otherwise be out of reach. Central to this definition is that a second person intervenes the learner at appropriate times and what the learner can actually accomplish by means of these interventions [47, 51, 52]. Although scaffolding is traditionally applied in classroom settings in which a teacher will support the students less and less throughout a learning task until the students are able to perform the task on their own, the use of scaffolding is increasing in instructional design [52]. It is important to stress the two most important goals of scaffolding in instructional design: accomplishing a task that would otherwise be out of reach and learning from this experience

[47, 52]. If students are able to perform the task but are not able to understand how they accomplished this, the scaffolding will not have provided the right support for learning.

Recent instructional design research, focused on current-day learning environments, has been aimed at applying scaffolding to software [48, 52]. Rather than teachers or peers supporting a learner, a software tool can support learners by changing the task in such a way that learners can accomplish tasks that would otherwise be out of reach [52]. Examples of software being used to support learners in accomplishing a task are providing prompts or reminders to learners to perform a certain step, or graphical organizers that help learners structure and organize their problem solving [52]. Scaffolding approaches to cognitive tools include scaffolding by computer and a human tutor combined or by means of a fully-embedded cognitive tool in a computer-based learning environment [48]. Cognitive tools that are used in this way can lead to a deeper understanding of the system by actively helping to organize a controller's knowledge and by helping learners to reflect on their own problem-solving processes and skills [53].

The SSD can be considered such a cognitive tool in the context of this research. As already mentioned in Section 3-4, the SSD could aid in shaping the controller's internal mental model and hence, contribute to a deeper understanding of the system [12]. It helps organizing the controller's knowledge by showing the different constraints imposed on the system and the corresponding solutions. An overlap is thus found between scaffolding and EID as both concepts help to organize the controller's knowledge and thereby decrease the controller's cognitive load. Applying the scaffolding approach to the SSD in an ATC learning task could help to further organize and structure the controller's knowledge and increase his/her performance during training without negatively influencing the development deep knowledge.

Chapter 5

Preliminary Experiment

As stated in Chapter 1, this research focuses on the effect of scaffolding of an ecologically-designed decision-support tool, the SSD, on novice learning. As multiple approaches to scaffolding exist, two methods have been chosen and are tested during a preliminary experiment in order to refine the final experiment.

This chapter starts with the objective of the preliminary experiment, described in Section 5-1. A detailed description of the experiment design is given in Section 5-2. The results of the experiment are presented and discussed in Section 5-3. The chapter ends with the preliminary experiment conclusions and recommendations for the final experiment in Section 5-4.

5-1 Preliminary Experiment Objective

As this research focuses on learning to perform a merging task, as is explained in Section 5-2-2, several merging scenarios have been constructed according to different levels of complexity. The preliminary experiment is first of all used to refine the scenarios and possibly modify the levels of complexity in case of unsuccessful identification of learning curves and a transfer of learning. As it is uncertain how many experiment runs are required for a learning curve to be identified and for a transfer of learning to take place, the preliminary experiment is additionally used to determine the length of the experiment and the number of experiment runs.

From the literature review it was found that support during a learning task can consist of various components and that several methods exist to integrate this support into a task. The Four-Component Instructional Design model describes how scaffolding elements of the learning task can have a positive effect on the overall learning experience and results [47]. Scaffolding, however, can be applied in different, often complementary ways, as was described in Section 4-2. The preliminary experiment is thus also used to test two scaffolding methods in order to select one method for the final experiment.

Summarizing the above, the objectives for the preliminary experiment are:

Preliminary Experiment Objectives

Objective 1: Determine whether the constructed air traffic scenarios in combination with the SSD provide data suitable for identifying learning curves and whether they result in a successful transfer of learning.

Objective 2: Determine the number of experiment runs that should be performed in order to be able to identify learning curves and to increase the chances of a successful transfer of learning.

Objective 3: Determine which SSD scaffolding sequence results in better novice performance and behavior when learning a merging task and which sequence results in a more evident learning curve.

5-2 Preliminary Experiment Design

The design of the preliminary experiment consists of the following elements: scaffolding sequence, ATC merging task, independent variables, control variables, apparatus, procedure, dependent measures and hypotheses. Each of these elements is discussed below.

5-2-1 Scaffolding Sequence

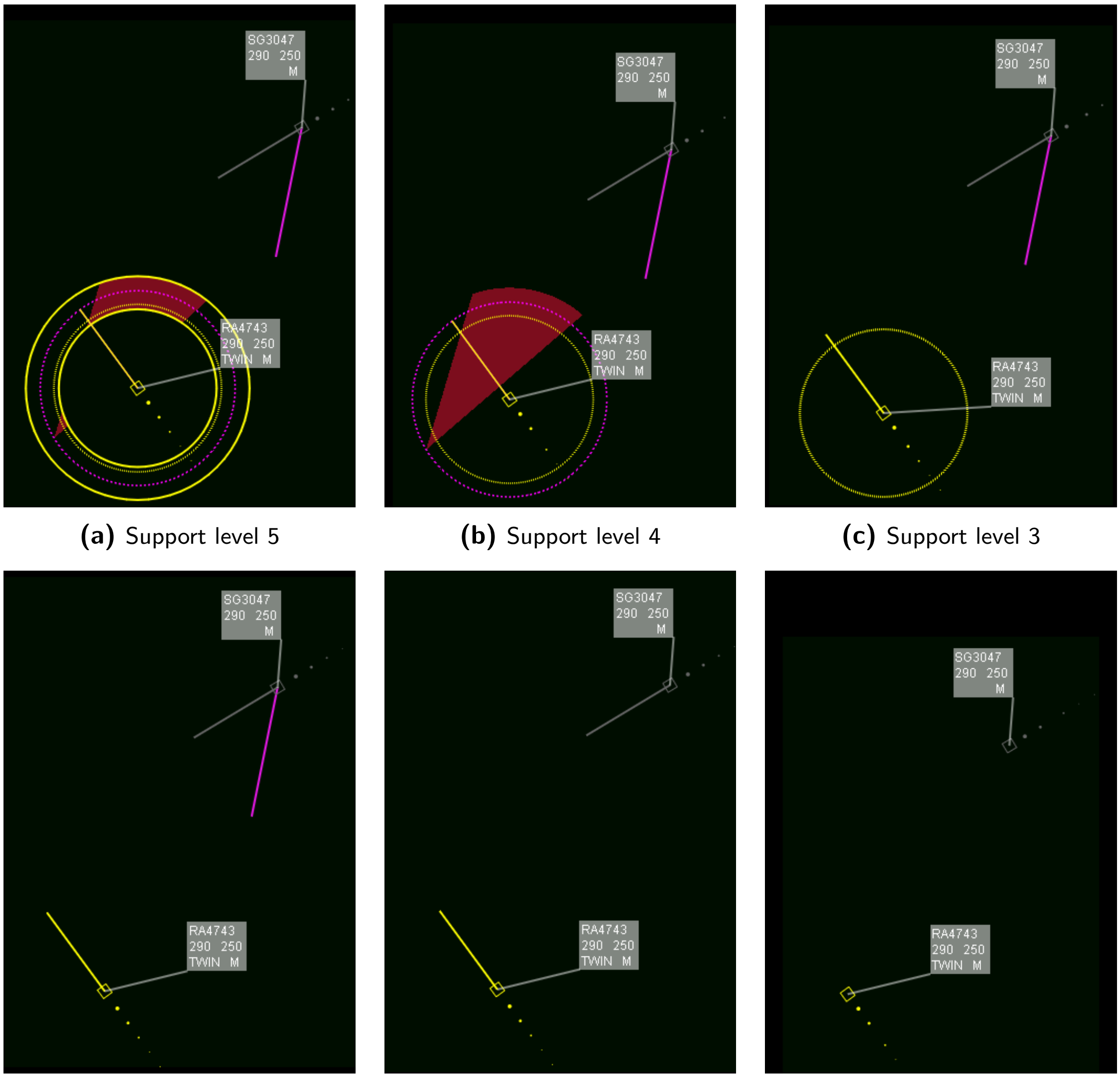
As stated before, this research is focused on scaffolding of the SSD. Section 3-3 discussed that the SSD is constructed by means of several elements. These elements can also be resembled in the simulation environment and are depicted in Figure 5-1. By using the different elements as a basis for the scaffolding or support levels, the participant is shown how the SSD is constructed and hence it is expected that the participant moves up through the different levels of the AH while developing system knowledge.

As stated in Section 4-2, scaffolding is a form of support and guidance that diminishes as the learner's expertise increases and as he/she progresses through a certain training task class. Two approaches to scaffolding have been discussed in the literature review: simple-to-complex sequencing and different types of learning tasks. As the simple-to-complex sequencing method is better applicable to the SSD, it was decided to only focus on this method during the preliminary experiment. It can be seen in Figure 5-1 that by reducing the number of elements of the SSD, the situation evolves from simple to complex as the student has less and less support until he/she has to perform the task without any support.

However, within the simple-to-complex sequencing approach and by taking the Four-Component Instructional Design model, that was discussed in Section 4-1, into consideration, again multiple options arise for the scaffolding sequence that can be applied to the SSD. The two options that are tested during the preliminary experiment are discussed below.

- **Method 1:** Strictly adhering to the Four-Component Instructional Design model (Figure 4-1) leads to the first option for a scaffolding sequence. This first method is visualized in Figure 5-2.

  The schematic in Figure 5-2 shows that for this method several levels of complexity are tested, indicated by the dotted-line boxes, that each correspond to a different task class.



(a) Support level 5 **(b)** Support level 4 **(c)** Support level 3

(d) Support level 2 **(e)** Support level 1 **(f)** Support level 0 (baseline)

Figure 5-1: Scaffolding Steps of the Solution Space Diagram

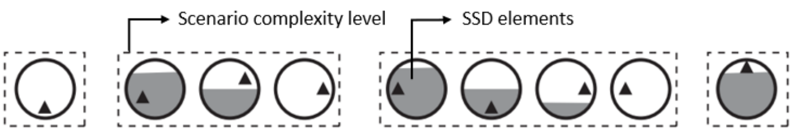


Figure 5-2: Schematic of the First Approach to Scaffolding

The level of complexity thus increases per task class, as is described in Section 5-2-2. During one task class or one level of complexity, the number of elements of the SSD decreases until the participant has to perform the task without the help of any of the SSD elements. This corresponds to the level of support that is decreasing until a conventional task has to be performed without any support, as was described in Section 4-2. After completing a task class, the next task class with a higher complexity is started, in which the number of SSD elements again gradually decreases. The scaffolding pattern of this method can therefore also be described as a sawtooth pattern.

- **Method 2:** The second approach to scaffolding strictly adheres to the definition of scaffolding: support that gradually decreases until the learner can independently perform a conventional task. Each SSD element now corresponds to a single task class. No sawtooth pattern is present but the number of SSD elements gradually decreases by moving through the task classes. Within a single task class, thus per support level as defined in Figure 5-1, the complexity level increases.

Each of these methods will affect the preliminary experiment procedure in a different way. How they affect the procedure is further explained in Section 5-2-7.

5-2-2 Control Task

Merging Task

A key takeaway of the Four-Component Instructional Design model is that learning tasks should be based on concrete real-life experiences and situations to increase the chances of a successful transfer of learning [15]. From literature as well as from interviews with LVNL employees and an ATC student, it was found that the merging task is often considered to be one of the most difficult and challenging tasks in the ATC training and thus, a decision-support tool might be beneficial when learning this task [16]. This task is therefore chosen as the main task for the experiment.

Complexity in Traffic Scenarios

The level of complexity was often mentioned during the interviews mentioned above as a factor that increases throughout the ATC training and sometimes even within training scenarios. To replicate this increasing complexity, different levels of complexity are defined by varying the following factors that, according to literature, have a large impact on the complexity level of a scenario or traffic situation [16, 30, 31, 33, 54]:

- Number of merging streams;
- Aircraft horizontal proximity;
- Aircraft count;
- Conflict angle;
- Number of conflicts.

Three levels of complexity are defined, numbered 1 to 3, that are stated in Table 5-1. To avoid recognition of the scenarios by the participants, two different scenarios have been defined per complexity level, both for training and for measurement purposes. These scenario groups have been set up in such a way that the scenarios show similarities within a group but each still has a distinct character to increase the variability of practice.

	1	2	3
Number of Streams	3	4	4
Horizontal Proximity	low	low	high
Stream Angles [deg]	(-)30, (-)40	30, -40, (-)60	30, -40, (-)60
Number of Conflicts	1	2	3

Table 5-1: Characteristics of the Complexity Levels of the Six Training Scenarios

Variability of the learning tasks is found to be one of the most important indicators for a successful transfer of learning and hence, unexpected events are regularly introduced during ATC training in order for ATCOs to develop more robust knowledge and to increase their versatility and resourcefulness [15, 22]. Next to creating different scenarios within a complexity group by varying traffic stream angles, the scenarios are rotated during the experiment to resemble this variability of practice and thereby encourage the novices' resourcefulness.

Control Sub-Tasks

During the experiment, the participants are given several tasks:

- *Separation assurance:* participants are instructed to ensure the minimum horizontal distance of 5 NM between neighbouring aircraft at all times to prevent losses of separation.
- *Clearing aircraft to their Cleared Sector Exit Point (COPX) after merging them into an existing traffic stream or airway:* participants are instructed to merge aircraft arriving in the sector into an existing traffic stream without introducing conflicts; to clear aircraft to their respective COPX with an exit speed of 240 kts and finally perform a Transfer of Control (TOC) to hand over aircraft to the next sector.
- *Communicating thoughts and strategies according to a "think-aloud" protocol:* participants are asked to think aloud while performing the different scenarios in order to gain an insight into their strategies and thinking patterns. More specifically, in case of a conflict, they are asked to mention the two aircraft involved in the conflict, the aircraft used to resolve the conflict and the chosen solution.

Next to these tasks, the participants are given several rules that apply to the sector:

- Only traffic between FL 280 and 300 needs to be controlled;
- Inbound traffic needs to be merged into the nearest existing traffic stream indicated by an airway, while maintaining the minimum separation distance (5 NM), and should subsequently be guided to its assigned COPX;
- Outbound traffic has to leave the sector at its respective COPX with an exit speed of 240 kts;
- Aircraft have to be given a TOC before they leave the sector;
- When aircraft are given a TOC, they have to be separated (at least 5 NM) from any other aircraft and should not be involved in any conflicts (i.e., they should not be involved in an STCA.

To obtain more information about the dynamics of the scenarios and events and to avoid underachievement from the controllers, the simulations are run at 2 times real-time speed to still stay as close as possible to the actual task and prevent negative side effects that influence the experiment results. With this simulation speed, every scenario is run for 5 minutes real-time.

5-2-3 Independent Variables

As stated before, scaffolding can be applied in various, often complementary ways. For the final experiment, a single scaffolding method will be chosen. The preliminary experiment is used to determine which method will be used in the final experiment.

The independent variable for this preliminary experiment is therefore the scaffolding method and with that, the experiment procedure or sequence. Two methods are chosen and tested, as discussed in Section 5-2-1. Differences in participant performance and behavior are identified and used to determine which method will be used in the final experiment.

5-2-4 Control Variables

In order to prevent the introduction of confounds, several measures are kept constant throughout this preliminary experiment. The control variables for this preliminary experiment are:

- Decision-Support Tool: Solution Space Diagram;
- Control Task and Conflict Type: Merging traffic;
- Flight Level: All traffic is limited to the horizontal 2D plane at FL 290;
- Aircraft Type: Three aircraft types are used in the experiment: light, medium and heavy aircraft;
- Overflying Traffic: All overflying traffic in the scenarios consists of the same number of aircraft and is limited to the same FLs outside FL 280 to 300;
- Experiment Briefing: The students are given a set of instructions about apparatus, the SSD and its visualization and the control task, which can be found in Appendix E (it should be noted that the procedure section is different for both groups);
- Display Layout: Conform industry standards. Sector size and shape of 60 NM by 60 NM, display colors and waypoint locations are kept constant throughout the experiment;
- General Procedure: The experiment is spread over two days with one day in between and contains several predefined scenarios that are presented to all groups. Note that the specific procedure per group, related to the order of the scenarios and the scaffolding sequence, is different per group;
- Feedback: No feedback is given as discovery learning is encouraged as much as possible.

5-2-5 Apparatus

The preliminary and final experiment are performed on a desktop computer in the ATM Lab of the Faculty of Aerospace Engineering at Delft University of Technology. The software that is used to construct the training scenarios and to conduct the measurements is the Java application MUFASA [55]. The traffic motion in the simulator has been simulated by simple, linear kinematic equations, no wind conditions are taken into account and all aircraft velocities are given in knots Indicated Airspeed (IAS) [55].

The scenarios contain a 60 NM by 60 NM square airspace for both training and measurement runs in order to be able to rotate the scenarios and thereby increase variability of practice. The square airspace contains eight waypoints and one main traffic stream in the middle of the sector. Figure 5-3 shows an impression of the simulation environment.

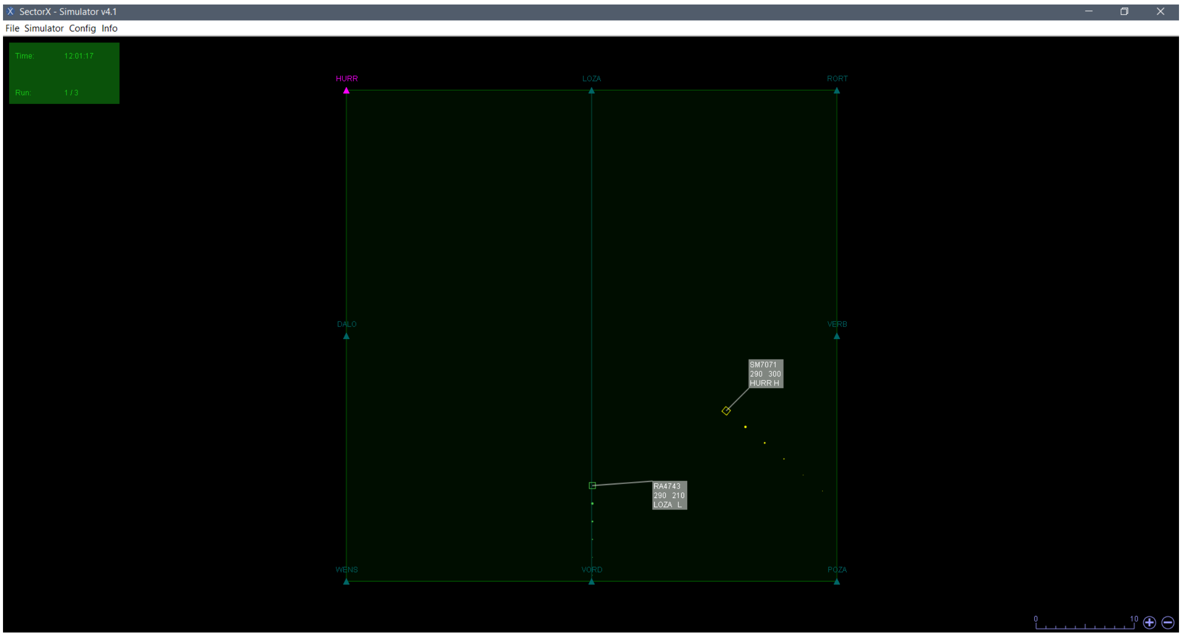


Figure 5-3: Simulation Environment

On the top left of the simulation environment the time that the current experiment run is running and the number of the current experiment run are displayed. On the bottom right a scale of 10 NM and buttons for zooming in and out are displayed. Aircraft in the simulation environment can have different colors:

- Gray: default color;
- Green: the aircraft is on course to its assigned COPX, such as aircraft RA4743 in Figure 5-3;
- Yellow: the aircraft turns yellow after it has been selected by the user to inspect its SSD or to give a command. Furthermore, its corresponding COPX in the sector furthermore turns magenta, as can be seen in Figure 5-3;
- Orange: STCA indicating a time to LoS of 120 seconds real-time for all involved aircraft;
- Red: STCA indicating a time to LoS of 60 seconds real-time for all involved aircraft.

In its most basic form, each aircraft on the radar screen is accompanied by a flight label and history dots. The history dots indicate the track that was followed by the aircraft as well as its current velocity. A larger distance between the dots indicates a higher velocity.

Figure 5-4 shows an example of an aircraft accompanied by its flight label with the following fields: callsign (SM7071), FL (290), IAS in kts (250), assigned COPX (COZA) and the type of aircraft, in this case medium (M).

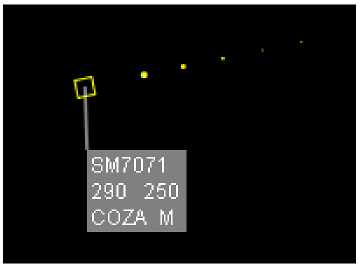


Figure 5-4: Impression of an Aircraft Accompanied by its Flight Label

Interaction with aircraft (i.e., giving clearances) takes place via a command window, shown in Figure 5-5, by using a computer mouse. No voice communication is required to command heading or speed changes to the aircraft, as this would interfere with the "think-aloud" task of the participants. A speed change can be checked by looking at the aircraft's flight label as the speed will temporarily change color, the aircraft's history dots or the aircraft's speed vector in case this vector is present, whereas a heading clearance can only be checked by the aircraft's history dots or speed vector that changes direction. A TOC command can be check by the aircraft symbol that changes from a square to an asterisk/star. The controller will no longer be able to give the aircraft commands after a TOC has been given to that aircraft.

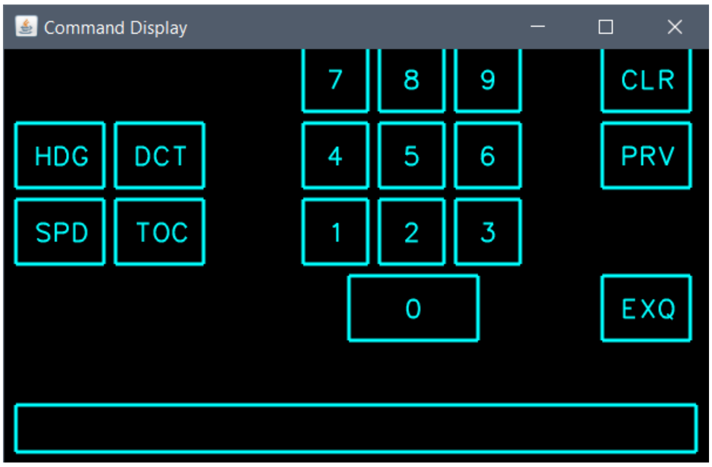


Figure 5-5: Command Display

When an aircraft is selected, its SSD appears, given that the SSD is available during that scenario. Figure 5-6 gives an impression of the SSD in the simulation environment. The four circles depict, ranging from largest to smallest, the aircraft's maximum velocity in yellow, the aircraft's current velocity in magenta, the aircraft's minimum velocity in yellow and the aircraft's protected zone as a yellow dotted line. The conflict zone is furthermore depicted in red.

In this example, the aircraft's speed vector is directed just outside the red conflict zone and hence the aircraft can continue flying at its current speed and heading without being in conflict with other aircraft. It should be noted that when an aircraft is selected to inspect its SSD,

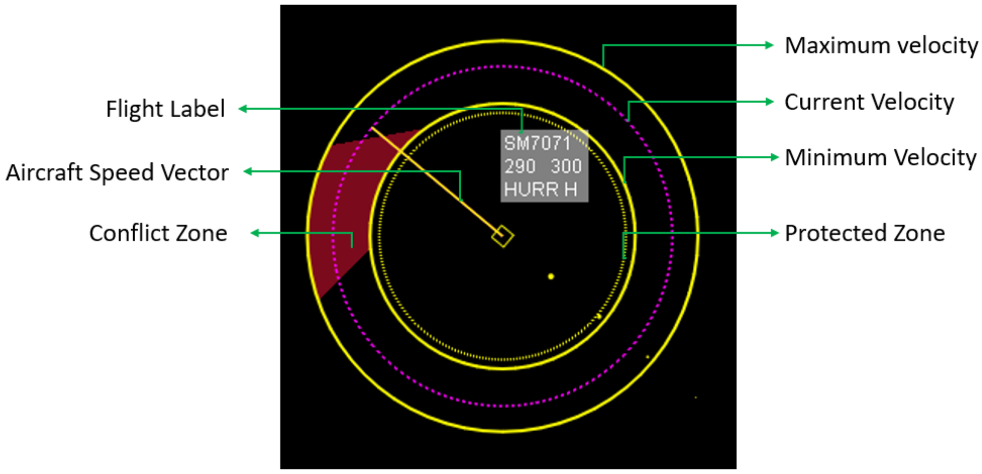


Figure 5-6: Impression of the Solution Space Diagram in the Simulation Environment

the relative speed vectors of other aircraft in the sector with respect to the selected aircraft appear in magenta, as can be seen in Figure 5-1a.

During the experiment, three types of aircraft occur in the simulations. The aircraft types and their speed envelopes are stated in Table 5-2 and are also mentioned on the Experiment Info Sheet, which can be found in Appendix F.

Aircraft Type	Label	Minimum IAS [kts]	Maximum IAS [kts]
Light	L	180	250
Medium	M	200	290
Heavy	H	230	350

Table 5-2: Aircraft Types used in the Simulation Environment

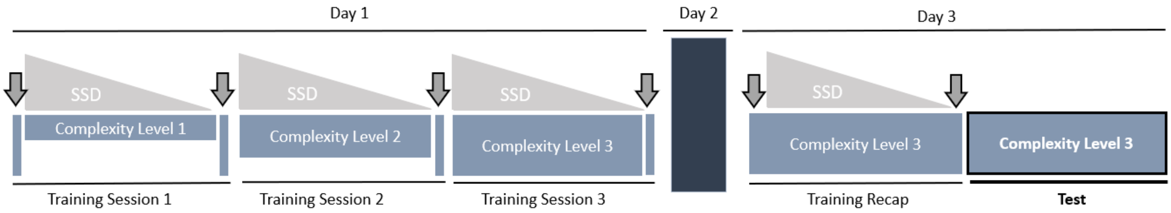
5-2-6 Participants

The preliminary experiment was performed by four students of the Delft University of Technology, specifically MSc students from the MSc profile Control and Simulation of the Faculty of Aerospace Engineering. As the research is aimed at the effect of EID on novice learning in ATC training, mostly task-naive participants were asked to voluntarily participate in the experiment. The two scaffolding methods, proposed in Section 5-2-1, were tested by two students each.

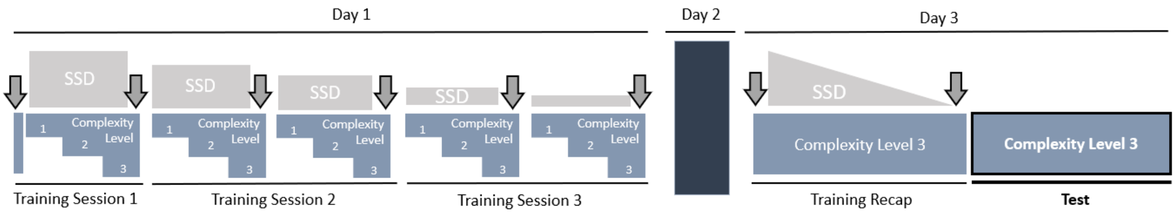
5-2-7 Experiment Procedure

As stated before, the independent variable for this preliminary experiment is the scaffolding sequence. As the experiment procedure is largely dependent on the sequence of the scenarios and the scaffolding that is applied to these scenarios, it means that each scaffolding method results in a different experiment procedure. Two experiment procedures are thus tested. Both procedures are described below.

As the research focuses on short-term knowledge development, the experiment consists of two separate sessions, with one day in between, for both procedures. The two sessions are conducted at similar times during the day to prevent the introduction of a confound. Figure 5-7 shows the schematics of both scaffolding methods. It can be seen that both procedures consist of three training sessions on day 1 and a training recap and test on day 3.



(a) Preliminary Experiment Procedure 1



(b) Preliminary Experiment Procedure 2

Figure 5-7: Preliminary Experiment Procedures

For the first method, depicted in Figure 5-7a, each training session corresponds to a single level of complexity in which two five-minute scenarios are alternated and rotated multiple times to increase the variability of the task. Each training block contains five training scenarios and one measurement scenario, which is indicated by the arrows in the schematic. At the start of a training session, the SSD appears in its complete form and subsequently decreases in number of elements until the final scenario of the training session has to be performed without any support of the SSD. This is also characterized as a sawtooth pattern, as was also described in Section 5-2-1. The second part of the experiment, performed on day 3, consists of two sessions: a training recap and a test. The training recap is similar to the last training session on day 1 and the test consists of four eight-minute measurement scenarios in which the participants have to perform the merging task without any SSD support.

For the second method, depicted in Figure 5-7b, the SSD support gradually decreases over the course of all three training sessions. Each complexity block or stairway corresponds to a different number of SSD elements and thus to a different support level. As the participant moves through the training sessions and thus through the support levels, the SSD will for each block decrease by one element. For each support level, three five-minute scenarios have to be performed with increasing complexity from level 1 to 3, after which a five-minute measurement scenario takes place at the highest complexity level. Similar to the first method, no SSD support is available during measurement scenarios, which is again indicated by arrows, as can be seen in Figure 5-7b. The second part of the experiment, performed on day 3, is similar to the second part of the first procedure. A training recap takes place during which the complexity level is equal to the highest level of complexity that was experienced on day 1 and during which the SSD will decrease from the full SSD to no SSD at the end of the training recap. The last block consists of the test in which four eight-minute measurement scenarios are performed without any SSD support.

The time schedule for the two preliminary experiment procedures can be found in Figure 5-8. Both training and measurement scenarios are ended with a question on how the participant would rate the workload experienced during that scenario. Rating the workload is done by indicating the experienced workload on a sliding bar, where sliding it to the left indicates a lower workload and sliding it to the right indicates a higher workload. After each training set there is the possibility for a short break. During all scenarios, thus both training and test scenarios, the control task remains the same: safely and efficiently merging incoming aircraft into an existing traffic stream without introducing conflicts and while adhering to the think-aloud protocol.

	Day 1 – Session 1				**Day 2 – No experiment**	**Day 3 – Session 2**		
Experiment Part	Briefing / practice	Training Session 1	Training Session 2	Training Session 3		Training Recap	Test	Debrief
Time	~15 min	35 min	30 min	30 min		35 min	40 min	10 min

(a) Time Schedule for Group 1

	Day 1 – Session 1				**Day 2 – No experiment**	**Day 3 – Session 2**		
Experiment Part	Briefing / practice	Training Session 1	Training Session 2	Training Session 3		Training Recap	Test	Debrief
Time	~15 min	25 min	40 min	40 min		35 min	40 min	10 min

(b) Time Schedule for Group 2

Figure 5-8: Time Schedule for both Preliminary Experiment Procedures

Prior to the experiment, all participants received a set of instructions consisting of an explanation of the procedure, the control task, the equipment, the SSD and the scaffolding. Any additional questions that participants had about the experiment and the procedure, that were not believed to be able to influence the results, were answered during the briefing at the start of the experiment. After discussing the instructions and making sure the participant knew what was expected of him/her, the participant started the experiment with three simple practice scenarios to become familiar with the display, the equipment, the dynamics and vectoring the aircraft. During the practice runs as well as during the training and measurement scenarios of the experiment, participants could always make use of the Experiment Info Sheet, presented in Appendix F.

The last part of the experiment, which is also indicated in Figure 5-8, is the debrief. When the test was finished and all data had been collected, a final debrief took place during which the participant were asked a number of questions, depending on the behavior and remarks of the participants made during the experiment, to gain insights in their thinking patterns, strategies and experiences.

5-2-8 Dependent Measures

To quantify the performance development, aircraft positions and states as well as the given commands were logged every 5 seconds. Furthermore, all measurement scenarios were recorded with a camera to obtain the think-aloud data that is used to identify the knowledge or behavior types, discussed in Section 3-4-2.

To be able to meet the preliminary experiment objectives stated in Section 5-1, several dependent measures are defined. These can be related to performance or to behavior or to both. An overview of the most important dependent measures is provided below, after which these dependent measures are discussed as well.

Dependent Measures

Performance-Related

- STCA count
- STCA duration
- Minimum separation distance
- Workload rating

Behavior-Related

- Behavior type (described in Section 3-4-2)
- Decisiveness and confidence
  - Number of commands
  - Type of commands
  - Timing of commands
  - Number of clicks
  - Type of clicks

Measuring and Assessing Performance

Performance is foremost measured in terms of safety and can additionally be measured in terms of how well the merging task is performed. In terms of safety, performance is characterized by the number and duration of (potential) PZ violations as well as the minimum separation distance between aircraft.

Other measures that are related to performance in terms of how well the merging task is performed are:

- Number of successful TOCs;
- Root Mean Square (RMS) of the exit speed of all aircraft;
- RMS of the exit heading of all aircraft;
- RMS of the exit separation distance of all aircraft;

Measuring and Assessing Knowledge Development

As described in Section 3-4, cognitive processes are difficult to assess as they are invisible processes. The actions taken by controllers after consulting the SSD can however give insight into their thinking patterns and the order of their cognitive steps. These insights can be analyzed by making use of the DL.

The different behavior types that were discussed in Section 3-4 have to be identified during or after the experiment by means of the recordings of the measurement scenarios during which participants are instructed to think aloud while performing the experiment. The thoughts help in identifying the decision-making strategies of the participants and mapping them onto Rasmussen's DL. Mapping the participant's strategy onto the DL allows for classification of the decision-making behavior in terms of the SRK-taxonomy.

Identifying the behavior types and strategies based on the experiment recordings is, however, a subjective process. Performance measures, such as described above, that can be logged during the experiment, can, however add to the identification of the behavior types in an objective manner. Examples of such performance measures are:

- Number, types and timing of commands;
- Number of SSD inspections or types of clicks;

These performance measures can be used to determine the decisiveness or confidence of the participant in a certain situation as they can hint on the number of corrections that is performed before the desired state is reached.

Learning Curves

As measurement runs are conducted throughout the entire experiment, the dependent measures above can be used to identify learning curves as performance or specific cognitive processes might change and/or improve during the training. These performance data can also be used to determine the decisiveness and confidence with which the participants are performing the task throughout the experiment, an important competence for ATCOs.

5-2-9 Hypotheses

Keeping the literature in mind, the following hypotheses are constructed for the preliminary experiment:

- H1: Scaffolding the SSD according to method 1 (see Section 5-2-1) will lead to a better transfer of learning, meaning better performance results for novices such as fewer conflicts, fewer violations of the PZ and traffic streams being merged successfully.
- H2: Scaffolding the SSD according to method 1 (see Section 5-2-1) will lead to better decision-making behavior and a higher SA, meaning more goal-oriented than task-oriented behavior will be demonstrated and the novice participants will show a higher decisiveness.
- H3: Scaffolding the SSD according to method 1 (see Section 5-2-1) will improve the learning process as learning plateaus can be distinguished earlier in the process.

5-3 Preliminary Experiment Results and Discussion

This section discusses several results and findings of the preliminary experiment. The findings presented in this section are discussed below based on the preliminary experiment objectives stated in Section 5-1. Additional results can be found in Appendix G.

Objective 1

The first objective to be completed was to determine whether the constructed air traffic scenarios in combination with the SSD provide suitable data to identify learning curves and whether they result in a successful transfer of learning. During and after conducting the experiment, several observations were made concerning this first objective.

- A strategy-related learning process was noticeable throughout the experiment for all participants. This learning process is, however, related to the participants becoming more familiar with the task and developing a strategy on how to best handle traffic scenarios like the ones that were presented to them. The participants eventually developed a strategy that contained, among others, the steps described below. It should be noted that not all participants performed the steps in the same order.

  - Determine which aircraft fly at FL 290 and should thus be controlled. Aircraft flying at other flight levels were given a TOC or the flight labels of these aircraft were dragged outside the sector to indicate that these aircraft were not of importance.
  - Give the first aircraft that is already on track the right exit speed, a Direct To (DCT) command and a TOC command.
  - Give the aircraft first entering the sector a larger speed to create space in the back of the sector for aircraft entering the sector at a later point in time.
  - Determine the order in which aircraft will be flying along the airway to their concurrent waypoint. Keeping this order in mind, a so-called 'train' was built and space was created within this train to merge the aircraft in the traffic stream at the spot the participants first had in mind.

  At the start of the experiment, participants often forgot to give aircraft a TOC command before they exited the sector and forgot to merge several aircraft into the traffic stream which led to aircraft exiting the sector at the wrong side. However, as the participants progressed through the experiment, more sub-tasks of the control task were completed successfully. It should be noted that the SSD did not influence the development of the strategies described above.

- Different strategies were developed among participants to handle the traffic in the scenarios. Some participants, for example, allowed the aircraft to get very close to the airway before giving a DCT command to merge the aircraft into the traffic stream, while others gave DCT commands soon after aircraft entered the sector or let aircraft fly parallel to the main airway for a while, even after being told to let aircraft fly on their original tracks for as long as possible. Because of these different strategies, the data, gathered from the experiment, from the different participants cannot be compared or used to determine any differences in performance and the development of this performance.

- Participants did not make use of the SSD or made only very little use of the SSD. During the experiment, participants indicated that the SSD simply contained too much information. Because of this and because they knew the SSD would eventually disappear, they chose to rather not make use of it or not spend too much time on understanding what kind of information the tool provided, in order to still be able to successfully perform the different sub-tasks of the control task. One participant in particular did, however, highlight that he/she sometimes used the SSD to be able to determine at which point in time aircraft flying parallel to the main traffic stream, could best be merged into this main traffic stream. Since participants did not or hardly make use of the SSD, the data gathered from the experiment cannot be used to evaluate the effect of scaffolding of the SSD on the performance and knowledge development of novices.

Objective 2

The second objective was to determine the number of experiment runs required to be able to identify learning curves and to increase the chances of a successful transfer of learning. As stated before, learning curves could be identified but these could not be attributed to the influence of the SSD.

Including the briefing and questions prior to the experiment and the breaks throughout the experiment, all participants took nearly three hours to complete the scenarios on day 1. Although the identified learning process was not related to the SSD, a learning process was still present and noticeable. It is therefore recommended to not increase the length of the first part of the experiment any further as to avoid negative influences on performance and behavior due to a reduced attention span or fatigue.

Objective 3

The third objective for the preliminary experiment was to determine which SSD scaffolding sequence and hence which experiment procedure results in better performance and behavior when learning a merging task and to determine if one of the two sequences results in a more evident learning curve. As stated above, the participants did not make use of the SSD or made only very little use of it. Several observations could be made related to this fact and to the scaffolding sequences of the SSD. These are stated below.

- All participants indicated that the SSD contained too much information (i.e., "too much red" and too many overlapping conflict triangles). They furthermore indicated that they were too busy completing the control task and could therefore not spend enough time on fully understanding the SSD and its capabilities.
- The participants of the first group (sawtooth method) furthermore indicated that the elements of the SSD started disappearing too soon in their opinion, due to which they were not able to fully understand the SSD's capabilities and learn from it. This, in combination with the fact that participants indicated that the SSD contained too much information and the fact that the participants knew the SSD would eventually completely disappear, contributed to the fact that the participants would rather not make use of the SSD, or the elements that remained, as it took too long to process the information of the SSD during a task which already required a significant amount of cognitive

effort. Instead, they preferred to put more focus on successfully completing the control task.

- All participants indicated that especially the second scaffolding step, depicted in Figure 5-1b, contained too much information due to which they would rather not make use of the SSD at this time as it took too long to process the SSD information in an already very busy scenario.

- Three out of the four participants indicated that the scaffolding step depicted in Figure 5-1c was the most useful and that out of all the scaffolding steps, this was the step where they did make use of the available support. It should however be noted that these participants only started using the PZ and the relative speed vector to determine whether two aircraft would be in conflict after they were asked if they were making use of these elements.

Other Findings

Next to findings related specifically to one of the objectives, some more observations were made that are discussed below.

- During the experiment, the participants were asked to rate the workload they experienced during a scenario. The participants, however, each had a different understanding of workload. Some participants defined workload as how much physical work they had to perform (e.g., the number of clicks or how busy they were most of the time), while other participants defined workload as how hard or easy the scenario was to perform or solve. The data collected for experienced workload from both groups can therefore not be compared.

- As described in Section 5-2-8, safety is measured in terms of number and duration of STCA alerts. Participants were specifically instructed to prevent these alerts at all times and if they did occur, to solve them as quickly as possible. During the experiment it was found that participants did not always solve these alerts as soon as possible as they sometimes gave priority to the merging task. In several instances, the participant gave a small correction after an STCA alert knowing it would take longer to resolve the issue but also knowing it would be a more efficient solution when keeping the merging task in mind. This "wrong" prioritization should thus be kept in mind when looking at the STCA alert data.

- Participants were instructed to think aloud during the experiment and more specifically to mention the aircraft involved in a conflict; the aircraft they chose to give a command to in order to resolve the conflict; the commands with which they resolved the conflict and the reasoning behind this. Aircraft call signs had the form XX-YYYY, where X represents a letter and Y a number. Participants usually referred to aircraft as 'this one' or only mentioned the first two letters of the call sign instead of naming their actual full call sign, as this required less cognitive effort. It should be kept in mind that because participants had to spent cognitive capacity on naming aircraft and on thinking out loud in general, less capacity was available for performing the control task and hence, some participants indicated they would normally have reacted sooner in certain situations.

5-4 Conclusions and Recommendations

Based on the findings stated in Section 5-3, several conclusions can be drawn from the preliminary experiment. These conclusions, as well as the recommendations for the final experiment following these conclusions, are stated in this section.

5-4-1 Conclusions

The conclusions drawn from the preliminary experiment are:

- The control task, presented to the participants, is too complex and does not provide suitable data to identify learning curves or differences in performance or behavior related to the use of the SSD.
- Participants have too much freedom in how they approach a scenario which leads to the development of different strategies.
- The strategy-related learning process that was observed throughout the preliminary experiment can be attributed to the scenario complexity levels and not to the use of the SSD.
- Participants had too little knowledge of the SSD and its capabilities or added value which caused them to make little to no use of the tool during the experiment.
- Scaffolding according to method 1 (sawtooth method) contributed to the fact that participants are less likely to make use of the SSD or its available elements as the elements start disappearing before the participant has a good understanding of them.
- The second scaffolding step, depicted in Figure 5-1b, was experienced as overwhelming and least useful.
- The third scaffolding step, depicted in Figure 5-1c, was experienced as most useful and was indicated as step where most use was made of the available support.

5-4-2 Recommendations

Following the conclusions stated above, several recommendations are made for the final experiment. Each of the recommendations is stated and elaborated on below.

Control Task and SSD Use

First and foremost, it is recommended that the control task be simplified. Because the control task consists of several sub-tasks, there are simply too many factors to take into account while also trying to understand the SSD. Hence, in the preliminary experiment participants did not discover the full potential of the SSD due to the overload of information. To make the use of the SSD more attractive, the SSD should first of all contain less information and more specifically a smaller conflict area and no or little overlapping conflict triangles for the participants to be able to trace the present conflict triangles back to the aircraft that is causing the conflict. This can be accomplished by reducing the number of aircraft present in a scenario. Secondly, the control task should consist of a task for which the added value of the SSD is very clear. An example was already given by one of the participants, namely, using the SSD to determine the point in time at which a parallel-flying aircraft can successfully be merged into the main traffic stream without introducing conflicts. During the interview with an ATC student on March 19th, 2019, about the ATC training, it was also discussed that timing this specific merging action is one of the more challenging tasks to become proficient in during the ATC training.

The control task should furthermore be adapted in such a way that participants have less degrees of freedom in completing a scenario. This can, for example, be done by focusing on a specific use case of the SSD. For the preliminary experiment, participants were able to give different types of commands to aircraft. Limiting the solution space by, for example, only focusing on heading, limits the number of strategies that can be developed and carried out by participants and also lowers the threshold to make use of the SSD as less information needs to be processed. The ideal situation would be that the control task in combination with the SSD is used to teach the participant a specific strategy. In this case it is important and therefore recommended that the higher learning goals or implicit learning objective of the experiment be more clearly defined, for example, what the strategy is participants are taught by means of the SSD and what strategy they should be able to apply in new or unexpected situations. Once this implicit learning objective has been defined, all constructed scenarios should be tested against this implicit learning objective to verify that the constructed scenarios will indeed contribute to this objective.

Scaffolding Sequence

It is recommended that the scaffolding steps be simplified, which can be accomplished by using different scaffolding elements. Simplifying the steps is expected to result in less cognitive effort required to understand the applicability of the SSD.

It is furthermore recommended that more time be allocated to each scaffolding step and hence, scaffolding method 2 (gradually decreasing support) be used. This method allows for participants to become familiar with the different scaffolding steps and thus SSD elements over a longer period of time. By giving the participants more time to understand the different elements of the SSD, it is expected that they are more likely to understand the benefits of the tool and are able to derive the correct strategy from it.

In Chapter 4, several scaffolding examples were given with respect to the Four-Component Instructional Design model and recent instructional design research. The scaffolding examples that were focused on current-day learning environments were, however, all aimed at directing participants' attention at a certain point on the screen by, for example, prompting questions or highlighting certain elements on the screen to aid the learners in developing the right strategy or knowledge of a certain procedure. Scaffolding of a decision-support tool has not earlier been applied in the way it has in this preliminary experiment. It is thought that this type of scaffolding, in combination with a complex ATC merging task, does not match with the theory of the Four-Component Instructional Design model. Where this theory stresses that learning tasks should be as realistic as possible, it was found that a realistic merging task is too complex for the SSD to have a positive influence or any influence at all on the development of participants' strategies. It is therefore recommended to look into different instructional design models as well as a different control task that better fit the Four Component Instructional Design model and better match the scaffolding approach applied to the SSD.

Other

Finally, related to the other findings discussed in Section 5-3, several additional recommendations are made. It is recommended that:

- the definition of workload be better defined so that participants all rate the situation based on the same definition;
- the prioritization of the experiment rules be stressed more. By stressing the order these rules more, participants are directed more towards a single strategy and the development of different strategies by different participants is discouraged;
- a score be added to the experiment. When participants know that their score is affected by certain actions, they are more likely to behave in a certain way or develop a certain strategy. Defining the score should, however, be done very carefully as it could also induce unwanted behavior when participants start focusing more on the score than on the actual task;
- the aircraft be given different and especially easier call signs to reduce the mental effort for the think-aloud protocol. Aircraft can, for example, be given a name with a single letter so it becomes easier to name the aircraft while thinking out loud.

Chapter 6

Summary and Future Steps

The objective of the preliminary thesis was to provide a context for the research objective stated in Chapter 1 by conducting a review of relevant literature and performing a preliminary experiment to refine the final experiment. This research will be an effort to further evaluate the impact of ecological interfaces on short-term knowledge and performance development of novices. The knowledge gained during the experiments and during the entire research can be used to contribute to improving learning methods and cognitive tools within the ATC domain and can provide new empirical insights in the benefits of EID on knowledge development.

This report includes a literature review as well as the design and evaluation of a preliminary experiment. By combining the findings from the literature review and the recommendations and lessons learned from the preliminary experiment, several changes are proposed for a final experiment.

After stating the final research objective and question in Section 6-1, Section 6-2 gives an overview of the key takeaways from the literature review regarding the ATC training, EID and complex learning. After this, a summary of the preliminary experiment is given in Section 6-3. Section 6-4 describes the proposed changes for the final experiment, based on the findings from the preliminary experiment. Finally, the future steps for this research are stated in Section 6-5.

6-1 Research Question

Using the findings from the literature review and the preliminary experiment, the final research objective and question(s) can be defined by using the objective and question(s) that were defined for the preliminary experiment. The final research objective is:

Final Research Objective

The objective of this research is to determine the impact of scaffolding as a support method during ATC training with an ecological interface on the short-term knowledge and performance development of novices by conducting a human-in-the-loop experiment in which novice participants will be trained on an ATC task using a scaffolded, ecologically-designed ATC decision-support tool.

The final research question is stated below.

Final Research Question

To what extent does using visual scaffolding of an ecologically-designed decision-support tool (SSD in particular) improve novice learning in the ATC domain by producing better performance results and higher-order decision-making behavior, when compared to learning without scaffolding?

6-2 Literature Review

An extensive literature study was performed in order to gain insight into the research field. The literature review can be divided into three parts. First of all, the ATC training was looked into. Secondly, a closer look was taken at the ecological approach for interface design and a CWA was performed for the CD&R task within ACC. Finally, complex learning, and specifically the scaffolding of support, was investigated. The results for each of these subjects are summarized below.

Air Traffic Control Training

Since ATCOs are subject to strict safety regulations, there is little to no room for (human) errors or incompetence and thus performance standards for the task are high. The ATC task is considered to be both dynamic and highly complex as it requires processing of large amounts of constantly changing information [21, 22]. Several competencies have been defined to determine whether someone is capable of conducting the ATC task [22]. A competency in the context of ATC training can be described as "a combination of skills, knowledge and attitudes required to perform a task to the prescribed standard"[23, p. 6] and is embodied in and observed through behavior. Guidelines for the ATC training are provided by ICAO (Doc 10056, Annex 2, Doc 4444 [23, 26, 17]) but a predetermined lesson plan is not provided. During the different training phases, the required competencies remain the same. However, the complexity of the scenarios that students are dealing with is increased over time. Complexity is usually characterized by the amount of traffic, the number of safety violations or the degree of efficiency that is required [18, 30].

Because no predetermined lesson plan exists, there is not much known about the ways ATCOs are taught to detect and resolve conflicts or how they are taught other strategies such as perception, interpretation, anticipation, workload management and planning strategies. [25]. Strategies are crucial in ATC performance and appear to be one of the key elements to success during the training [25]. From interviews with LVNL employees and an ATC student it was found that ATCOs usually obtain the required knowledge by self discovery of what works and what does not and by ad-hoc input and feedback from the coaches during the training depending on the scenarios and their complexity. By creating solutions ad-hoc or by directing the trainees ad-hoc towards certain strategies, a single solve-all strategy is discouraged as such a strategy is usually not sufficient to solve unexpected situations. Research has furthermore shown, that there are three factors that are of influence to the decision-making strategies of ATCOs: expediency, preservation of airspace structure and visualization [27].

The time-critical, complex and dynamic nature of the ATC task as well as the fact that cognitive skills (e.g., information processing) cannot be measured or observed directly, makes it difficult to develop a reliable assessment system and is also part of the reason why the assessment of ATCOs has only been partially automated [5]. Combining performance data with information about the thinking patterns of students, obtained from thinking-out-loud protocols and so-called over-the-shoulder observations during simulations or during OJT, leads to a more or less complete picture of the skills and competencies of the student during the execution of a certain task. As competencies are a combination of several factors, the assessment should take place at a higher level and social, emotional and environmental factors should be taken into account during this assessment [22]. Continuously assessing at higher levels allows for distinguishing learning curves of these skills and competencies, which are important in monitoring the students' progress and reducing the subjectivity of the training and assessment [22].

Several models have been developed to map the cognitive processes taking place inside an ATCO's head. While the ATC Performance Model [22] and the ACoPOS model [28] can serve as a framework for the identification and design of objective performance criteria used in the competence-based assessment of ATCOs, the ECOM [29] can be used to structure strategies and subsequently translate these to assessment criteria.

Ecological Interface Design

A good way to analyze and understand the requirements for an ATCO training system, is to perform a CWA [10, 34]. An example of an approach to CWA and interface design is the ecological approach [35]. EID is a design framework that focuses on the specific problem of designing human-machine interfaces for complex socio-technical systems by visualizing and supporting the complete range of possibilities and activities that an operator might be faced with [9]. By supporting the end-user in adapting to change and novelty, it allows them to limit their core activities to higher-order problem solving and decision making [9]. Part of the CWA are two theoretical concepts that originate from research done by Rasmussen in the process control domain: the AH [37] and the SRK [38]. Rasmussen's DL, one of the most common tools for describing decision-making activities and also part of the CWA, can be used as a template to structure the cognitive processes and thus to determine which information-processing activities are taking place [39].

The SSD is a tool that was designed according to the EID principles [13]. The EID principles work especially well with the open and dynamic nature of the ATCs environment. By integrating constraints found on the lower levels of the AH and showing how these constraints affect the solution space of the controlled aircraft in terms of heading and speed, the SSD provides an instant overview of the solution possibilities in the 2D plane to the controller and thus allows the controller to move between higher-level functional information of the AH and lower-level objects.

During ATC training, unexpected events are regularly introduced for ATCOs to develop more robust knowledge and to increase their versatility and resourcefulness. As the SSD accounts for unanticipated and unfamiliar events or situations by providing the controller the complete range of solutions, an overlap is found between EID and ATC training. Next to this, in order to successfully accomplish the mission statement of ATC, it is of importance that the ATCOs learn to develop a correct mental model. As EID contributes to SA, another overlap is found [24, 44]. Tools from the EID principle such as Rasmussen's AH can be used to contribute to this increased SA by grouping different domain-relevant constraints at different levels of abstraction [37]. This information can then be used to structure the information on a display. The actions taken by controllers after consulting the SSD can give insight into their thinking patterns and the order of their cognitive steps, which can be visualized on the DL, and possible shortcuts between different behavioral domains taken by the controller can be identified at different stages during the training. Finally, as the assessment during ATC training is still largely based on subjective expert opinions, an ecologically-designed decision-support tool such as the SSD could be a promising tool in an effort to objectify the training and assessment.

Although the SSD might be a good tool to aid in the development of expert-like behavior and to reach higher regions in Rasmussen's DL, a risk of this interface is that surface learning could occur which would lead to the development of shallow knowledge in case the interface would be used as a rule-based tool only [11]. Furthermore, a larger dependency on the interface might be developed when the interface is used as a rule-based tool only.

Complex Learning

To increase the success of a transfer of learning of a complex and dynamic task such as the ATC task, a holistic design approach is necessary according to instructional design, that does not lose sight of the separate elements and their interrelations but deals with the system or learning domain as a whole [15]. One such an approach is the Four-Component Instructional Design model which assumes that blueprints for complex learning can always be described by four components: learning tasks, supportive information, procedural information and part-task practice [14, 15]. Learning tasks subsequently consist of easy-to-difficult task classes, a high variability and a decreasing level of support and guidance when moving through the task sequence.

When students are learning complex tasks, such as the ATC task, they can sometimes be overwhelmed by the amount of information and the complexity of the task. Hence, it is of importance to manage the cognitive load during the learning process by providing the right amount and type of support and guidance that is fully integrated in the task [47]. Initially, the support allows a student to perform a task or achieve a goal that would not be achievable without the support. As the student's expertise increases, support is gradually decreased until the student no longer needs the support and is able to perform the task independently [49, 48]. This is also referred to as scaffolding and is based on Vygotsky's concept of the ZPD [47, 49]. Scaffolding support can be provided in many ways [50, 51]. In scaffolding whole-task practices according to the Four-Component Instructional Design model, two complementary approaches can be distinguished: simple-to-complex sequencing, that aims to decrease the intrinsic cognitive load for novices, and varying types of learning tasks, that focuses on decreasing the extraneous cognitive load [47]. Scaffolding approaches to cognitive tools include scaffolding by computer and a human tutor combined or by means of a fully-embedded cognitive tool in a computer-based learning environment [48]. Cognitive tools that are used in this way can lead to a deeper understanding of the system by, for example, actively helping to organize a controller's knowledge, or helping to reflect on their own problem-solving processes [53].

As stated before, the SSD could be helpful in shaping a controller's internal mental model and hence, in contributing to a deeper understanding of the system [12]. It helps organizing the controller's knowledge by showing the different constraints imposed on the system and the corresponding solutions. An overlap is thus found between scaffolding and EID as both concepts help to organize the controller's knowledge and thereby decrease his/her cognitive load. Applying the scaffolding approach to the ATC learning task and combining it with the various elements of the SSD could help to further organize and structure the controller's knowledge and increase his/her performance during training without negatively influencing the development of deep knowledge.

6-3 Preliminary Experiment

As stated in Chapter 1, this research focuses on the effect of scaffolding of an ecologically-designed decision-support tool, the SSD, on novice learning. Based on interviews and literature, the merging task was found to be considered one of the most difficult and challenging tasks during the ATC training and thus, a decision-support tool might be beneficial when learning this task [16]. Furthermore, the level of complexity was often mentioned as a factor that increases throughout the training and sometimes even within training scenarios. To replicate the increasing complexity, different levels of complexity have been defined by varying the number of merging streams, the aircraft horizontal proximity, the aircraft count, the conflict angle and the number of conflicts, that, according to literature, have a large impact on the complexity level of a scenario or traffic situation [16, 30, 31, 33, 54]. For the preliminary experiment, several merging scenarios have been constructed with different levels of complexity to increase the reality of the task, which is, according to the Four-Component Instructional Design model, important for a successful transfer of learning [15].

During the preliminary experiment, the participants were instructed to merge aircraft into an existing traffic stream while ensuring a minimum horizontal separation distance of 5 NM, to subsequently clear aircraft to their respective COPX with an exit speed of 240 kts and perform a TOC and finally, to think aloud during the entire experiment. The preliminary experiment was performed on a desktop computer in the ATM Lab of the Faculty of Aerospace Engineering at Delft University of Technology. The software that is used to construct the training scenarios and to conduct the measurements is the Java application MUFASA [55].

As stated before, one of the main focuses of this research is the application of the scaffolding to the SSD. Two complementary approaches to scaffolding have been discussed in the literature review: simple-to-complex sequencing and different types of learning tasks. As the simple-to-complex sequencing method is better applicable to the SSD, it was decided to only focus on this method during the preliminary experiment. The SSD is made up of different elements, as discussed in Section 3-3. By using these different elements as a basis for the scaffolding or support levels, the student is shown how the SSD is constructed and hence, it is expected that the student moves up through the different levels of the AH while developing system knowledge. Two options following from the simple-to-complex sequencing method are tested during the preliminary experiment: a saw-tooth method and a gradually-decreasing support method. The two methods, and hence two experiment procedures, were each tested by two, mostly novice, students from the Delft University of Technology.

The preliminary experiment was used to determine (1) whether the constructed scenarios in combination with the SSD provide data suitable for identifying learning curves, (2) whether the proposed number of experiment runs was sufficient to identify a learning curve, and (3) which of the SSD scaffolding sequences resulted in a more evident learning curve when learning a merging task.

Several conclusions could be drawn from the preliminary experiment of which the most important conclusions were that the control task was too complex and that participants had too much freedom in how they approached and solved a scenario. This all led to the fact that the participants made very little to no use of the SSD, due to which the effect of the scaffolded SSD on novice learning during an ATC merging task could not be determined.

Following the conclusions, several recommendations were made for the final experiment. First of all, the control task should be simplified in order to lower the threshold for participants to make use of the SSD and thereby discover its full potential. Furthermore, participants should have less degrees of freedom in approaching and solving the scenarios to reduce the number of strategies that can be developed by participants throughout the experiment and thereby be able to compare the data of the different groups. It is therefore recommended to better define the implicit learning objective for the participants of the final experiment and test all constructed scenarios against this implicit learning objective. Additionally, regarding scaffolding, it is recommended that the scaffolding steps be simplified and the information overflow per step is reduced in order to further lower the threshold for participants to make use of the (available) SSD elements. Finally, more time should be allocated to each scaffolding step to give participants more time to understand the available SSD elements. It should be noted that although it is recommended to allocate more time to each step, it is also recommended to not increase the length of the experiment any further to prevent negative influences on the results due to a reduced attention span or fatigue.

6-4 Proposed Changes for the Final Experiment

Following the findings, conclusions and recommendations from the preliminary experiment, several changes are proposed for the final experiment. These changes, related to the implicit learning objective, the control task and the scaffolding steps and sequence, are discussed below.

6-4-1 Control Task

The following subsections elaborate on the implicit learning objective as well as on a proposal for a redefined control task.

Implicit Learning Objective

As described above, for the use of the SSD to become apparent to the user, it is important to define what the function of the SSD is within the control task and to which higher learning goal it is contributing.

It was decided to focus on the following rule of thumb: *In case two aircraft are in conflict, the slower aircraft should be steered behind the faster aircraft. To make sure this action resolves the conflict, the controller should give the slower aircraft a heading command in the direction of the faster aircraft.*

This rule of thumb is illustrated in Figure 6-1. This figure shows that the slower aircraft can indeed best be steered behind the faster aircraft as there is more room on this side of the solution space. It also shows that steering the slower aircraft in the direction of the faster aircraft does result in a safe separation with a large separation margin.

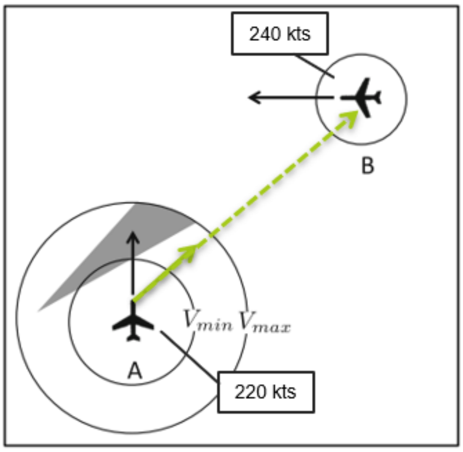


Figure 6-1: Schematic of the Rule of Thumb

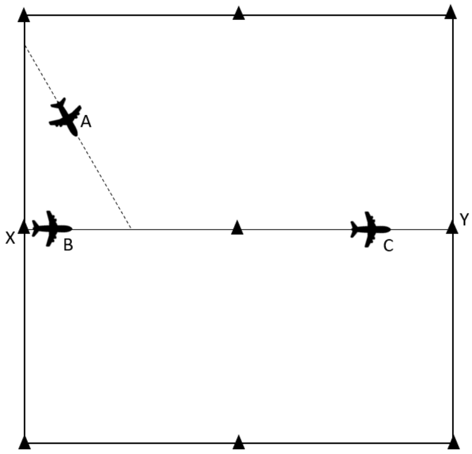
During the training phase of the final experiment, participants should thus implicitly be taught the rule of thumb described above in, for example, a merging situation. During the test, the participants will be expected to be able to apply this obtained knowledge in new situations. It will thus be tested to what extent participants are able to apply this newly learned rule of thumb in new situations, for example, in conflict types other than a merging conflict.

Merging Task

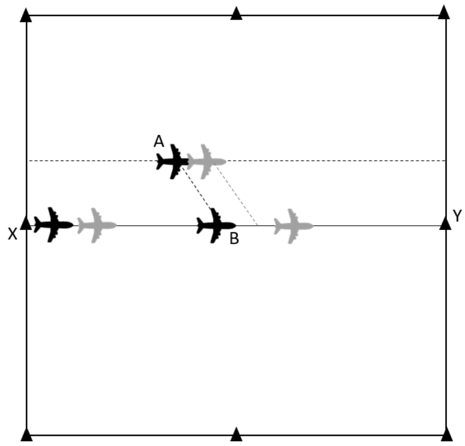
Now that the implicit learning objective for the final experiment has been defined, the control task can be redefined to match this objective and to meet the recommendations that were made based on the results of the preliminary experiment.

For the final experiment, it is proposed to still focus on the merging task, but to zoom in further on one specific sub-task of this merging task, and more specifically, the timing of the merging command. As stated earlier, the ATC student from the interview on March 19th, 2019, also indicated that timing this specific merging action is one of the more challenging tasks to become proficient in during ATC training. During the preliminary experiment, one participant already focused on this specific part and it was found that in this situation, the SSD could aid in developing a 'feeling' for when to best give this command. Figure 6-2 gives an impression of the new merging task compared to the previous merging task, where a black aircraft indicates the aircraft's current position while a gray aircraft indicates the aircraft's predicted future position.

For the proposed merging task, aircraft will be flying on parallel tracks. The sector will contain a single main traffic stream into which the aircraft flying on parallel routes, have to be merged. Figure 6-2b shows that, to merge slower aircraft A into stream X-Y, the controller has to wait before aircraft A can be given a heading command to safely be merged into stream X-Y, behind faster aircraft B. The timing of this command is especially tricky as giving the command too early will result in a conflict between aircraft A and B and hence, trigger an STCA, whereas giving the command too late might result in a conflict with other aircraft in the sector. The SSD is beneficial in this situation as it will show the controller the moment a certain heading becomes available, which corresponds to the moment that that aircraft can be merged into the main stream. This is indicated in Figure 6-3.

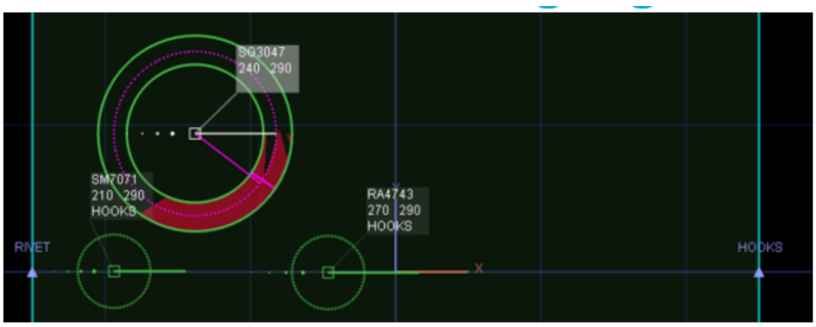


(a) Schematic of the Merging Task of the Preliminary Experiment

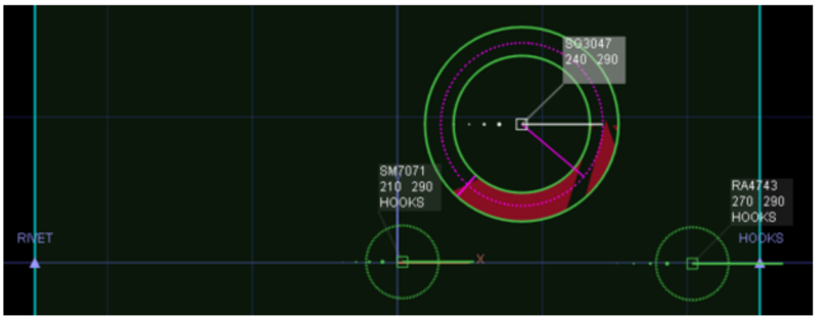


(b) Schematic of the Merging Task of the Final Experiment

Figure 6-2: Schematics of the Merging Tasks of the Preliminary and Final Experiment



(a) The Desired Heading to Safely Merge the Aircraft into the Main Stream is not Available



(b) The Desired Heading to Safely Merge the Aircraft into the Main Stream is Available

Figure 6-3: Example of the Added Value of the Solution Space Diagram in the New Control Task

Finally, to make the SSD easier to use and understand, it was recommended to reduce the number of aircraft in a scenario. It was therefore decided to construct simple scenarios with either three or four aircraft in total. The 'red area' in the SSD is then decreased which makes it easier to process and use the information provided by the SSD.

Control Sub-Tasks

In order to meet the recommendations made regarding the number of sub-tasks of the control task, it is proposed that the participants are given the following sub-tasks:

- *Separation assurance:* participants are instructed to ensure the minimum horizontal separation distance of 5 NM between neighboring aircraft at all times to prevent losses of separation.
- *Clearing aircraft to their COPX after merging them into the main traffic stream or airway:* participants are instructed to merge aircraft arriving in the sector into a main traffic stream without introducing conflicts and clear aircraft to their respective COPX.
- *Communicating thoughts and strategies according to a "think-aloud" protocol:* participants are asked to think aloud while performing the different scenarios in order to gain an insight into their thinking patterns.

Participants thus no longer need to give all aircraft leaving the sector a certain exit speed and a TOC. The scenarios will furthermore contain less aircraft and no overflying traffic.

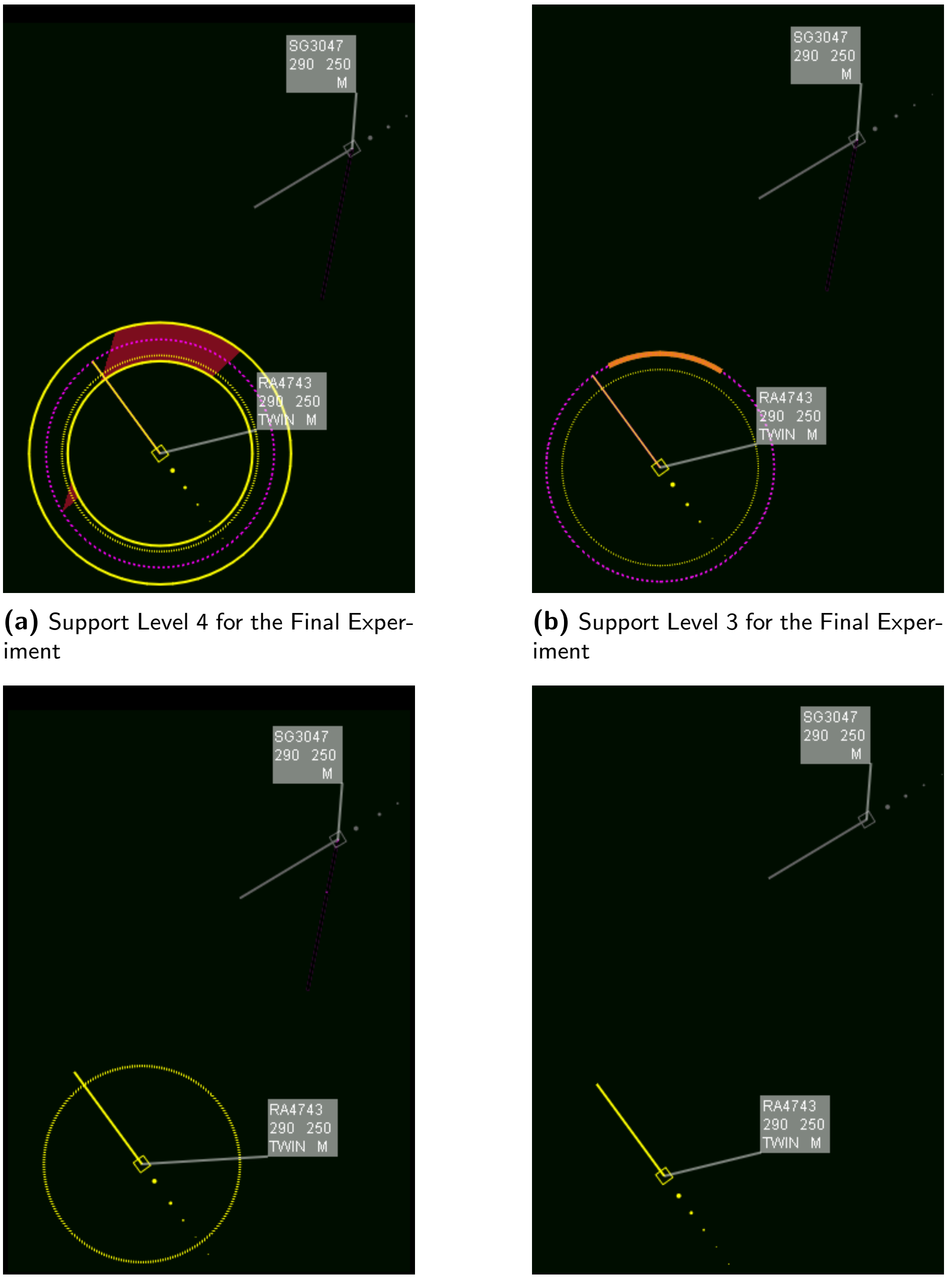
Altogether, the control task and the scenarios will largely be simplified by applying the proposed changes and hence, the information displayed to the participants by means of the SSD will also be reduced. Less sub-tasks will result in more cognitive capacity available to understand the SSD that already contains less information. It is thus expected that the threshold for making use of the SSD is significantly lowered.

6-4-2 Scaffolding Sequence

Based on the findings related to scaffolding, stated in Section 5-3, several recommendations were made. Most important is that the information overflow per step should be reduced and more time should be allocated to each of the steps.

Starting with the first recommendation, the scaffolding steps of the SSD should be changed for the final experiment. The proposed scaffolding steps can be found in Figure 6-4. The figure shows several differences compared to the scaffolding sequence of the preliminary experiment, presented in Figure 5-1.

First of all, a new scaffolding step is introduced, which is depicted in Figure 6-4b. This step is also known as the Heading Band (HB), which is a simplification of the SSD [56]. While the SSD shows the solution space for the complete speed envelope of the controlled aircraft, the HB shows the solution space for the current speed only. The magenta circle in Figure 6-4b reflects the current speed the aircraft is flying at and the part that is highlighted in orange represents a set of headings that will result in an orange STCA with a time to LoS between 60 and 120 seconds real-time. While the controller is thus provided less information in compared to the full SSD, the most important information for successfully completing the control task, is still provided (i.e., the moment the desired heading becomes available to be able to merge the controlled aircraft safely into the main traffic stream). By looking at the proposed merging task in the previous subsection, it becomes apparent that especially the heading-functionality of the SSD contributes to the implicit learning objective. By introducing the HB step into the sequence, the participants are thus exposed to the heading functionality for a longer period of



(a) Support Level 4 for the Final Experiment

(b) Support Level 3 for the Final Experiment

(c) Support Level 2 for the Final Experiment

(d) Support Level 1 for the Final Experiment (Baseline)

Figure 6-4: Scaffolding Sequence of the Solution Space Diagram for the Final Experiment

time and thus implicitly encouraged to focus on heading to resolve a conflict, as was defined in the learning objective.

Secondly, the relative speed vector is removed from all steps. It was decided to remove this element as the relative speed vector does not provide the controller with useful information when aircraft are flying parallel. In the situation where aircraft are flying parallel, all speed vectors, including the relative speed vector, will be parallel and hence, no useful information can be derived from this relative speed vector in the case of the proposed control task.

Finally, the last step depicted in Figure 5-1f, has been removed. This step had been added to increase the reality of the control task of the preliminary experiment to match the key takeaways of the Four-Component Instructional Design model. As no conclusion could be drawn regarding the combination of this model with the scaffolded SSD, it was decided to no longer apply this model to the experiment and hence, the extent to which scenarios have to be realistic, could be reduced. This last step was therefore deemed unnecessary.

The second recommendation was related to the (amount of) time allocated to each scaffolding step. By reducing the number of steps, as was explained above, more time can be allocated to each scaffolding step. Furthermore, by applying the second scaffolding method (gradually-decreasing support), described in Section 5-2-1, more consecutive time can be allocated to each scaffolding step. This method still complies with the basic definition of scaffolding: a form of support that gradually diminishes as the student's expertise increases, until the student has to perform the task on his/her own.

6-5 Future Steps

This preliminary thesis presented the results of an extensive literature review concerning ATC training, EID and complex learning, as well as the findings, conclusions and recommendations from a preliminary experiment. The future research steps needed to complete this research consist of the application of the proposed changes to the final experiment. New scenarios need to be constructed that all need to be tested against the implicit learning objective that was defined in Section 6-4. Furthermore, as the proposed scaffolding steps have changed, compared to the steps of the preliminary experiment, a new experiment procedure needs to be defined. Once the new procedure and scenarios have been defined, novice participants need to be recruited to partake in the experiment. From the data obtained from the experiment, both simulation data as well as video and audio recordings, the development of behavior can be analyzed. It needs to be determined whether learning curves can be identified in the shown behavior as well as in the performance data.

The next steps for this research, depending on the outcome of the data analysis of the final experiment, will be to look into further extending the complexity of the task and hence increasing the reality. By increasing the reality, the research is one step closer to implementing the SSD as a cognitive tool during the actual ATC training at LVNL.

Bibliography

[1] FAA, “Aeronautical Information Manual - Official Guide to Basic Flight Information and ATC Procedures,” 2017.

[2] ICAO, “Long-Term Traffic Forecasts - Passenger and Cargo,” 2018.

[3] SESAR Joint Undertaking / FAA, “NextGen - SESAR: State of Harmonisation,” 2018.

[4] SESAR Joint Undertaking, “European ATM Master Plan: The Roadmap for Sustainable Air Traffic Management,” 2012.

[5] C. D. Wickens, A. S. Mavor, and J. P. McGee, *Flight to the Future: Human Factors in Air Traffic Control*. Washington DC: National Academy Press, 1997.

[6] FAA, “Plan for the Future - 10-Year Strategy Plan for the Air Traffic Control Workforce 2012-2021,” 2012.

[7] LVNL, “Jaarverslag 2018,” 2018.

[8] G. K. van de Merwe, E. Oprins, D. J. F. Eriksson, and A. H. van der Plaat, “The Influence of automation support on performance, workload, and situation awareness of Air Traffic Controllers,” 2010.

[9] 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.

[10] N. Naikar, “Cognitive work analysis: An influential legacy extending beyond human factors and engineering,” *Applied Ergonomics*, vol. 59, pp. 528–540, 2017.

[11] K. Christoffersen, C. Hunter, and K. J. Vicente, “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.

[12] C. Borst, R. M. Visser, M. M. v. Paassen, and M. Mulder, “Exploring Short-Term Training Effects of Ecological Interfaces: A Case Study in Air Traffic Control,” *IEEE Transactions on Human-Machine Systems*, 2019. Unpublished.

[13] 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: Systems and Humans*, vol. 38, no. 6, pp. 1221–1233, 2008.

[14] J. J. G. van Merrienboer, *Training Complex Cognitive Skills - A Four-Component Instructional Design Model for Technical Training.* New Jersey: Educational Technology Publications, 1997.

[15] P. Kirschner and J. J. G. van Merrienboer, "Ten steps to complex learning: A new approach to instruction and instructional design," in *21st Century Education: A Reference Handbook* (T. L. Good, ed.), ch. 26, pp. 224–253, Thousand Oaks, California: SAGE Publishing, 2004.

[16] P. Hermes, M. Mulder, M. M. Van Paassen, J. H. L. Boering, and H. Huisman, "Solution-Space-Based Analysis of the Difficulty of Aircraft Merging Tasks," *Journal of Aircraft*, vol. 46, no. 6, pp. 1995–2015, 2009.

[17] ICAO, "Doc 4444, Procedures for Air Navigation Services: Air Traffic Management," 2016.

[18] S. M. B. Abdul Rahman, *Solution Space-based Approach to Assess Sector Complexity in Air Traffic Control.* PhD thesis, Delft University of Technology, 2014.

[19] EUROCONTROL, "Guidelines for Short Term Conflict Alert - Part I," 2017.

[20] EUROCONTROL, "Specification for Medium-Term Conflict Detection," 2017.

[21] W. Schneider, "Training high-performance skills: Fallacies and guidelines," *Human Factors*, vol. 27, no. 3, pp. 285–300, 1985.

[22] E. Oprins, E. Burggraaff, and H. van Weerdenburg, "Competence-Based Assessment Design for Air Traffic Control Training," in *Proceedings of the International Symposium of Aviation Psychology*, pp. 546–552, 2005.

[23] ICAO, "Doc 10056, Manual on Air Traffic Controller Competency-based Training and Assessment," 2017.

[24] M. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors Journal*, vol. 37, no. 1, pp. 32–64, 1995.

[25] M. Schuver-van Blanken and J. J. G. Van Merrienboer, "Air traffic controller strategies in operational disturbances: An exploratory study in radar control," in *Proceedings of the 30th EAAP Conference*, (Villasimius, Sardinia, Italy), 2012.

[26] ICAO, "Annex 2 to the convention on international civil aviation: Rules of the air," 2005.

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

[28] M. Schuver-van Blanken, "The ATC Cognitive Process and Operational Situation Model - A model for analysing cognitive complexity in ATC," in *Proceedings of the 29th EAAP Conference*, (Budapest, Hungary), 2010.

[29] E. Hollnagel and D. D. Woods, *Joint Cognitive Systems: Foundations of Cognitive Systems Engineering.* Boca Raton, FL: CRC Press, 2005.

[30] S. Athenes, P. Averty, S. Puechmorel, D. Delahaye, and C. Collet, "ATC complexity and controller workload: Trying to bridge the gap," in *Proceedings of the International Conference on Human-Computer Interaction in Aeronautics*, (Cambridge, MA), pp. 56–61, AAAI Press, 2002.

[31] S. M. B. Abdul Rahman, M. Mulder, and M. M. Van Paassen, "Using the Solution Space Diagram in Measuring the Effect of Sector Complexity During Merging Scenarios," in *AIAA Guidance, Navigation, and Control Conference*, (Portland, OR), AIAA, 2011.

[32] B. Hilburn, "Cognitive complexity in air traffic control: a literature review," report, EUROCONTROL, 2004.

[33] P. Kopardekar and S. Magyarits, "Dynamic density: measuring and predicting sector complexity," in *Proceedings of the 21st Digital Avionics Systems Conference*, (Irvine, CA), pp. 2C4–1 – 2C4–9, IEEE, 2002.

[34] A. M. Bisantz and C. M. Burns, *Applications of Cognitive Work Analysis*. Boca Raton, FL: CRC Press, 2008.

[35] J. M. Flach, F. Tanabe, K. Monta, K. J. Vicente, and J. Rasmussen, "An Ecological Approach to Interface Design," in *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, (Chicago, IL), pp. 295–299, 1998.

[36] K. J. Vicente and J. Rasmussen, *A Theoretical Framework for Ecological Interface Design*. 1988.

[37] J. Rasmussen, "The Role of Hierarchical Knowledge Representation in Decisionmaking and System Management," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 15, no. 2, pp. 234–243, 1985.

[38] J. Rasmussen, "Skills, Rules, and Knowledge; Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 13, no. 3, pp. 257–266, 1983.

[39] J. Rasmussen, "Outlines of a Hybrid Model of the Process Plant Operator," in *Monitoring Behavior and Supervisory Control* (T. B. Sheridan and G. Johannsen, eds.), (New York, NY), pp. 371–383, Plenum Press., 1976.

[40] G. A. Mercado Velasco, M. Mulder, and M. M. Van Paassen, "Air Traffic Controller Decision-Making Support using the Solution Space Diagram," in *IFAC Proceedings Volumes*, vol. 43, (Valenciennes, France), pp. 227–232, 2010.

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

[42] B. Beernink, B. C., J. Ellerbroek, M. M. van Paassen, and M. Mulder, "Toward an Integrated Ecological Plan View Display for Air Traffic Controllers," in *International Symposium on Aviation Psychology*, (Dayton, OH), pp. 55–60, 2015.

[43] J. Lodder, J. Comans, M. M. van Paassen, and M. Mulder, "Altitude-Extended Solution Space Diagram for Air Traffic Controllers," in *International Symposium on Aviation Psychology*, (Dayton, OH), pp. 345 – 350, 2011.

[44] C. Burns, G. Jamieson, G. Skraaning, N. Lau, and J. Kwok, “Supporting Situation Awareness Through Ecological Interface Design,” in *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, (Baltimore, MD), pp. 205–209, 2007.

[45] F. Schott and N. M. Seel, “Instructional Design,” in *International Encyclopedia of the Social & Behavioral Sciences* (J. D. Wright, ed.), pp. 196–200, Oxford, UK: Elsevier, second ed., 2015.

[46] F. G. W. C. Paas and J. J. G. Van Merriënboer, “Instructional Control of Cognitive Load in the Training of Complex Cognitive Tasks,” *Educational Psychology Review*, vol. 6, no. 4, pp. 351–371, 1994.

[47] J. J. G. van Merrienboer, P. Kirschner, and P. Kester, “Taking the Load Off a Learner's Mind : Instructional Design for Complex Learning,” *Educational Psychologist*, vol. 38, no. 1, pp. 5–13, 2003.

[48] S. P. Lajoie, “Extending the scaffolding metaphor,” *Instructional Science*, vol. 33, no. 5-6, pp. 541–557, 2005.

[49] K. Shabani, M. Khatib, and S. Ebadi, “Vygotsky's Zone of Proximal Development: Instructional Implications and Teachers' Professional Development,” *English Language Teaching*, vol. 3, no. 4, 2010.

[50] R. Azevedo and M. J. Jacobson, “Advances in scaffolding learning with hypertext and hypermedia: a summary and critical analysis,” *Educational Technology Research and Development*, vol. 56, no. 1, pp. 93–100, 2008.

[51] J. Van De Pol, M. Volman, and J. Beishuizen, “Scaffolding in Teacher–Student Interaction: A Decade of Research,” *Educational Psychology Review*, vol. 22, no. 3, pp. 271–296, 2010.

[52] B. J. Reiser, “Scaffolding Complex Learning: The Mechanisms of Structuring and Problematizing Student Work,” *Journal of the Learning Sciences*, vol. 13, no. 3, pp. 273–304, 2004.

[53] R. D. Pea, “Beyond Amplification: Using the computer to Reorganize Mental Functioning,” *Educational Psychologist*, vol. 20, no. 4, pp. 167–182, 1985.

[54] S. M. B. Abdul Rahman, C. Borst, M. Mulder, and M. M. Van Paassen, “Sector Complexity Measures: A Comparison,” *Jurnal Teknologi*, vol. 76, no. 11, pp. 131–139, 2015.

[55] C. Borst, C. Westin, and B. Hilburn, “An Investigation into the Use of Novel Conflict Detection and Resolution Automation in Air Traffic Management,” in *Proceedings of the Second SESAR Innovation Days*, (Braunschweig, Germany), 2012.

[56] 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,” *Air Traffic Control Quarterly*, vol. 22, pp. 115–136, 2014.

Part III

Preliminary Thesis Report Appendices

NOTE:
This part has already been graded under AE4020

Appendix A

ATCO Competences - Blueprints

As mentioned in 2-3-3, students are admitted to a unit, based on their own preferences and how well their skills/competencies match the competencies that are required for a specific unit. The LVNL developed so-called blueprints which illustrate to what extent an ATCO of a specific unit should possess certain competencies.

Figure A-1 shows the competencies that are important to ATCOs within ACC. The figure shows that cognitive capacity, flexibility and being able to anticipate are especially important. As unexpected events in area control are not uncommon, controllers need to be able to anticipate these events, come up with plans on how to solve these situations, and be flexible in which solution they choose and how these solutions affect their other plans. The SSD matches especially well with supporting these competencies, as presenting the full range of solutions reduces cognitive load and helps controllers anticipate and be more flexible.

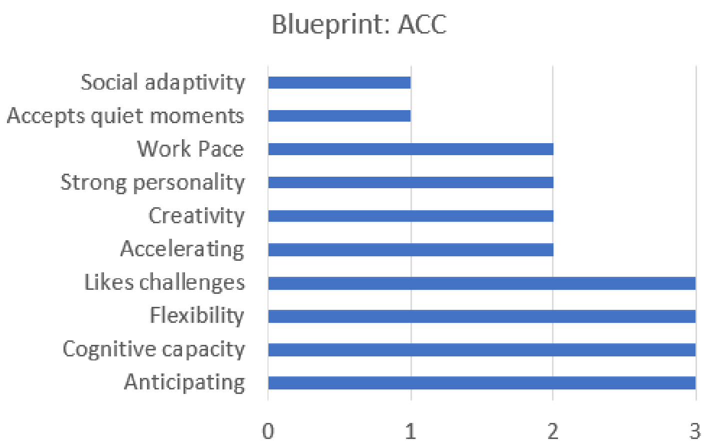
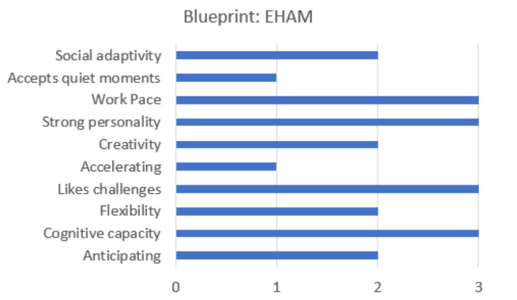


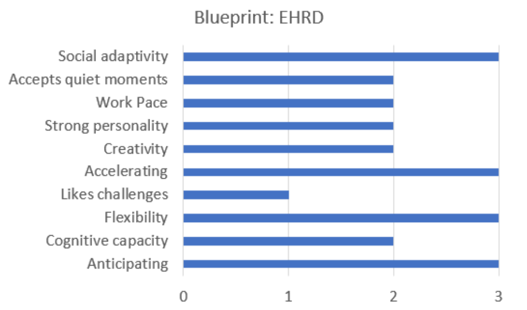
Figure A-1: Blueprint for Air Traffic Control Officers at Area Control [1]

Within APP, a distinction is made for the different airports as ATCOs need to possess very different competencies and skills per airport. The resulting blueprints for Amsterdam Schiphol

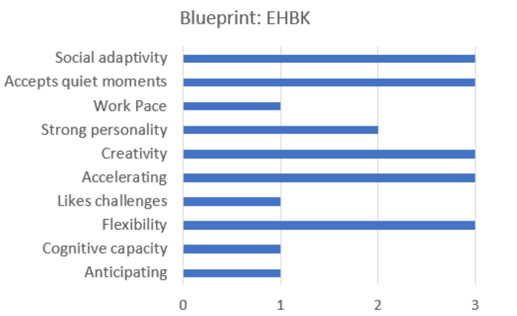
Airport, Rotterdam The Hague Airport, Maastricht Aachen Airport and Groningen Aiport Eelde are shown in Figure A-2.



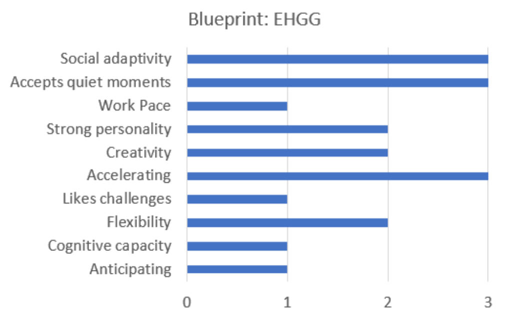
(a) Blueprint for ATCOs at APP/TWR of Amsterdam Airport Schiphol



(b) Blueprint for ATCOs at Approach Departure Control of Rotterdam The Hague Airport



(c) Blueprint for ATCOs at Approach Departure Control of Maastricht Aachen Airport



(d) Blueprint for ATCOs at Approach Departure Control of Groningen Airport Eelde

Figure A-2: Blueprint for Air Traffic Control Officers at Approach Departure Control [1]

Appendix B

ATC Training Structure

The structure of the ATC training was shortly discussed in Section 2-3-3 and is further elaborated on in this appendix. The structure of the ATC training, as it is nowadays, is shown in Figure B-1 and discussed below.

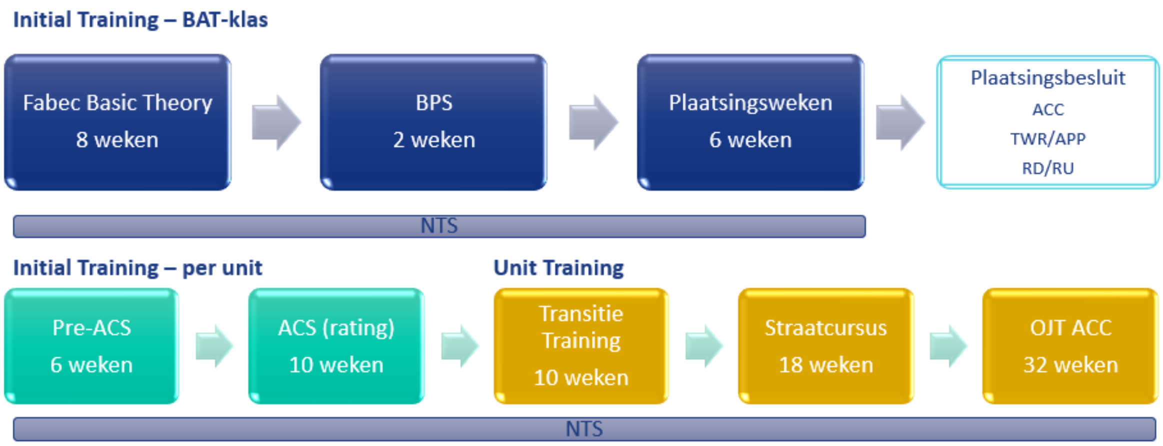


Figure B-1: Initial Training vs Unit Training [1]

- **Basic Training:** The basic training takes eleven weeks in total. The goals of the basic training are for the trainee to obtain the background knowledge required for ATCOs and the so-called Basic Practical Skills (BPSs). During the BPS phase, students are trained in practical skills such as how to use the equipment in the right way and how to handle flight strips. Furthermore, the students take 8 weeks of theoretical courses about subjects such as meteorology, aircraft science and performance, radio communication, aircraft recognition, air traffic law, equipment, human factors and navigation.
- **Placement Weeks:** During the placement weeks, which consist of six weeks, students visit the three different ATC units for two weeks each: tower, approach and area control. The modules consist of visits to the work floor during which students get the opportunity to talk to professionals and get a feel of the work atmosphere and simulator training

for that specific unit. Once a student expresses his/her preference for a certain unit, in this case Area Control Surveillance (ACS), and is admitted to this unit, the student continues with Pre-ACS training.

- **Pre-ACS training:** The Pre-ACS training is conducted at a basic simulator in preparation of the ACS rating training. The instructors in this phase can be licensed ATCOs as well as training specialists. The six weeks of basic simulator training consist of two weeks of inbound flights, two weeks of outbound flights, a week of neighboring fields and finally a week in which everything is combined. For each of these themes, the complexity is increased throughout the training.
- **ACS (Rating Training):** The ACS rating training is structured in a similar way as the pre-ACS training but takes 10 weeks. The structure for both phases can be found in Figure B-2. First, inbound flights are practiced after which outbound flights are practiced. After this, the special agreements with, for example, neighboring fields are practiced as well as special procedures and teamwork. The training ends with a block in which all the previous is combined.

Plaatsing

Module	Pre-ACS – 6 weken				Rating Training ACS – 10 weken					
Thema taakklasse	Inbounds	Outbounds	Neven-velden	Consoli-deren	Inbounds	Outbounds	Neven-velden	Teamwork	Bijzondere procedures	Consoli-deren
Stap	1	2	3	4	5	6	7	8	9	10
Duur stap	2 weken	2 weken	1 week	1 week	2 weken	2 weken	1 week	2 weken	1 week	2 weken
Coaching	IT + 1 OPS boven de sterkte				OPS					
Aantal oefeningen per dag	2-3	2-3	2-3	2-3	2-3	2-3	2-3	2-3	3	3
Duur oefening	35 min.	35 min.	35 min.	35 min.	35-45 min.	35-45 min.	35-45 min.	45 min.	45 min.	60 min.
Trainee als pilot	Deels								Nee	
Oefeningsverslag	X	X	X	X	X	X	X	X	X	X
Coachverslag	X	X	X		X	X	X	X	X	
Progressierapport				X	X	X		X		X

Student Controller License

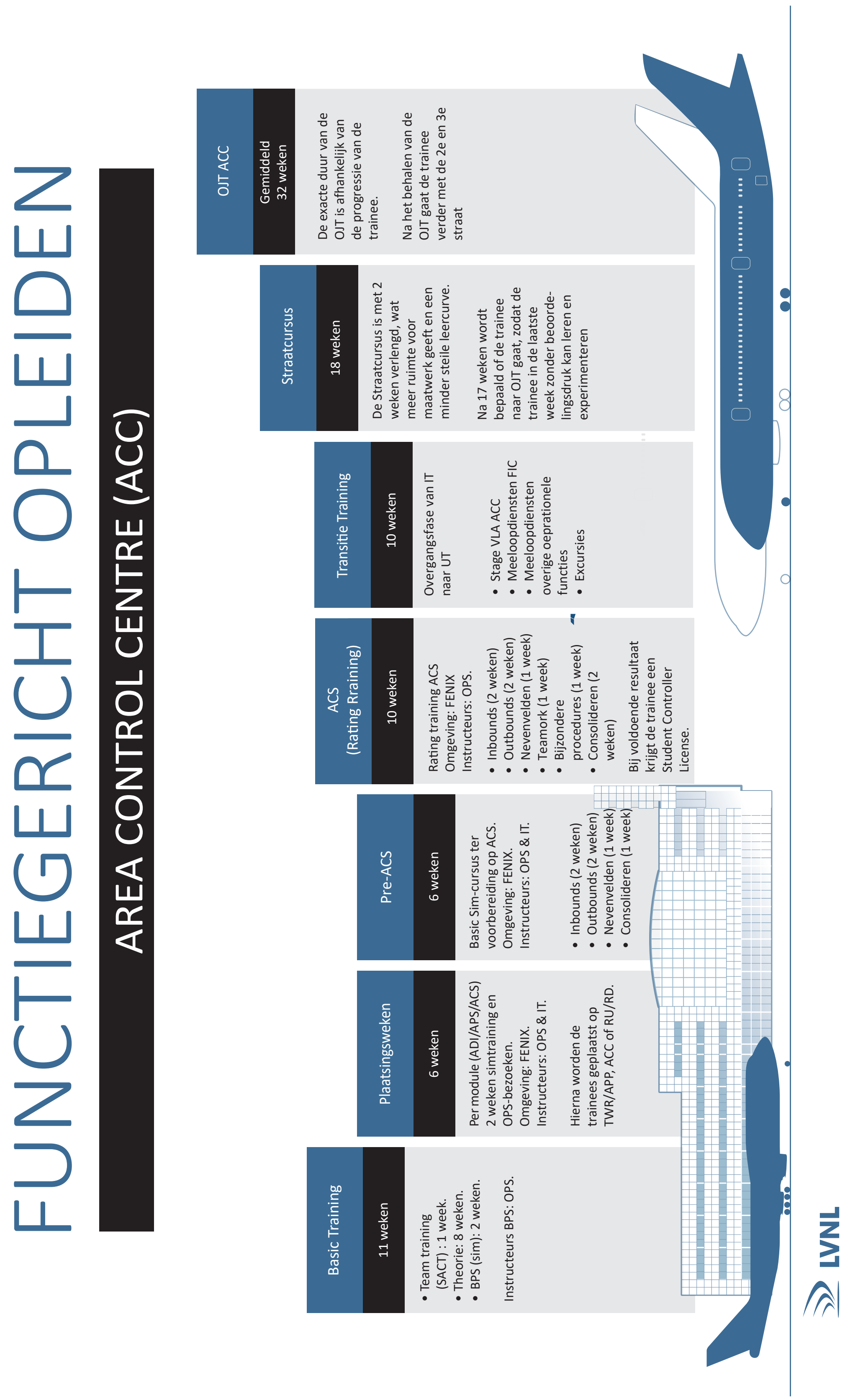
Figure B-2: The Structure of the Pre-ACS and ACS Training [1]

During the pre-ACS training as well as during the ACS rating training, the training consists of either two exercises a day plus a demo that is performed by a licensed ATCO or three exercises a day. Each exercise takes about 35-45 minutes and complexity and difficulty are increased every exercise by increasing the traffic density, the types of aircraft involved in the scenario or by creating unknown situations. An exercise can be focused on a specific task of air traffic control. During the adjacent fields phase, extra focus is put on protocols and agreements with the adjacent fields for a specific sector. However, the goal of all exercises is to expedite and maintain a safe and orderly flow of air traffic. It is furthermore expected that the trainee indicates the competencies or skills he/she likes to improve or work on during an exercise. The instructors for the rating training are all licensed ATCOs. Once the student successfully finishes the ACS rating training and hence the Initial Training, the student obtains his/her Student Controller License and is allowed to move on to the next phase of the training: Unit Training.

- **Transition Training:** The transition training is where the students continue from Initial Training to Unit Training. During this part of the training, the students perform a so-called internship at ACC and shadow several people with different functions at this unit. The transition training is where the students are taught subjects such as air structure, classification, aircraft recognition and routes specific for that unit.

- **Sector Training:** During the sector training the students get further familiarized with all rules, regulations, procedures and protocols specific to a single sector controlled by the unit. An example of this is the familiarization with the agreements with adjacent sectors. The sector training is guided and assessed by ATCOs. After 17 out of 18 weeks total, it is determined whether the students are allowed to continue to the OJT. The last week of the sector training is used for learning and experimenting without the pressure of being assessed.
- **On-The-Job Training:** The final stage of the ATC training takes about 32 weeks. The exact duration is specific to the progression of the student. During the OJT, the student will follow and watch professional ATCOs perform their job but will also get the chance to work as an ATCO in that specific sector while still being guided by licensed ATCOs. After successfully completing the training for one sector, the student continues with the sector training and OJT for the other sectors.

Another overview of the ACC training structure is given on the next page.[1]


FUNCTIEGERICHT OPLEIDEN
AREA CONTROL CENTRE (ACC)
Basic Training
11 weken
• Team training (SACT) : 1 week.
• Theorie: 8 weken.
• BPS (sim): 2 weken.
Instructeurs BPS: OPS.
Plaatsingsweken
6 weken
Per module (ADI/APS/ACS) 2 weken simtraining en OPS-bezoeken.
Omgeving: FENIX.
Instructeurs: OPS & IT.
Hierna worden de trainees geplaatst op TWR/APP, ACC of RU/RD.
Pre-ACS
6 weken
Basic Sim-cursus ter voorbereiding op ACS.
Omgeving: FENIX.
Instructeurs: OPS & IT.
• Inbounds (2 weken)
• Outbounds (2 weken)
• Nevenvelden (1 week)
• Consolideren (1 week)
ACS (Rating Rraining)
10 weken
Rating training ACS
Omgeving: FENIX
Instructeurs: OPS.
• Inbounds (2 weken)
• Outbounds (2 weken)
• Nevenvelden (1 week)
• Teamork (1 week)
• Bijzondere procedures (1 week)
• Consolideren (2 weken)
Bij voldoende resultaat krijgt de trainee een Student Controller License.
Transitie Training
10 weken
Overgangsfase van IT naar UT
• Stage VLA ACC
• Meeloopdiensten FIC
• Meeloopdiensten overige oeprationele functies
• Excursies
Straatcursus
18 weken
De Straatcursus is met 2 weken verlengd, wat meer ruimte voor maatwerk geeft en een minder steile leercurve.
Na 17 weken wordt bepaald of de trainee naar OJT gaat, zodat de trainee in de laatste week zonder beoordelingsdruk kan leren en experimenteren
OJT ACC
Gemiddeld 32 weken
De exacte duur van de OJT is afhankelijk van de progressie van de trainee.
Na het behalen van de OJT gaat de trainee verder met de 2e en 3e straat
LVNL

Appendix C

ATC Assessment - Reporting Forms

As can be read in Section 2-3, feedback is a very important part of the learning process. Students are encouraged to indicate on what skills and competencies they want to work during an exercise, keep track of their own progress, and actively reflect on these skills and competencies during and after exercises. This is also illustrated in Figure C-1.

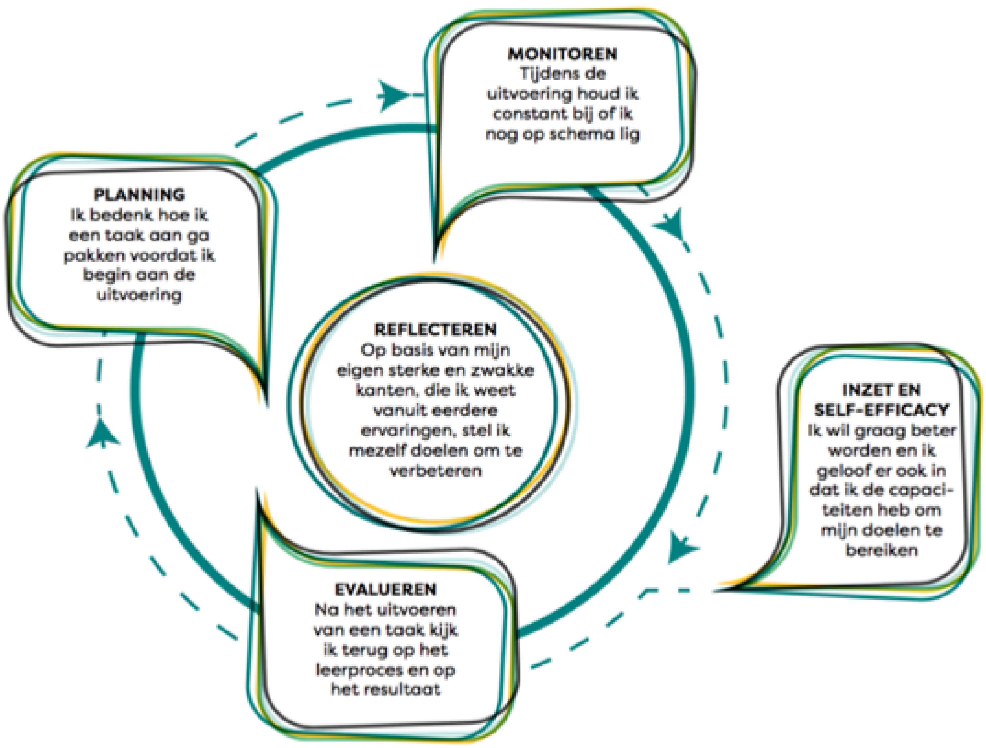
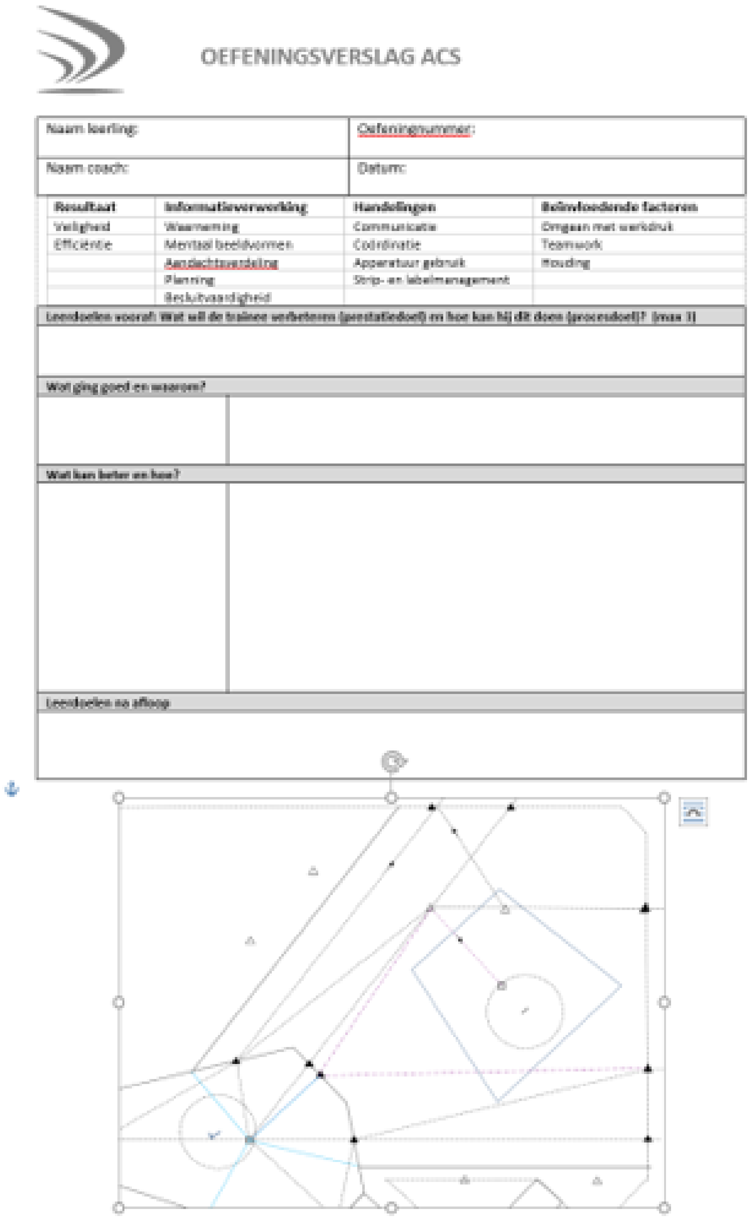


Figure C-1: Elements of Feedback Important During the Air Traffic Control Training [1]

At LVNL, three types of feedback or evaluation forms are used: exercise reports, coach reports and trainee log files, which can be found in Appendix C. These forms are usually filled out by the coach and trainee together, which leads to a better insight in the student's points for improvement and learning progress.



OEFENINGSVERSLAG ACS

Naam leerling:	Oefeningnummer:
Naam coach:	Datum:

Resultaat	Informatieverwerking	Handelingen	Beïnvloedende factoren
Veiligheid	Waarneming	Communicatie	Omgaan met werkdruk
Efficiëntie	Mentaal beeldvormen	Coördinatie	Teamwork
	Aandachtsverdeling	Apparatuur gebruik	Houding
	Planning	Strip- en labelmanagement	
	Besluitvaardigheid		

Leerdoelen vooraf: Wat wil de trainee verbeteren (prestatiedoel) en hoe kan hij dit doen (procesdoel)? (max 3)

Wat ging goed en waarom?

Wat kan beter en hoe?

Leerdoelen na afloop

Figure C-2: Exercise Report Form [1]

Coachverslag

Periode:

Coach:

Trainee:

Waar heb je aan gewerkt (**prestatiedoel**)?

Hoe is aan de prestatiedoelen gewerkt: welke **procesdoelen** zijn geformuleerd?

Wat heeft de trainee goed gedaan? Waarin heeft de trainee **progressie** laten zien? (minimaal 3)

Waaraan moet nog worden gewerkt? Wat zijn de doelen voor de komende periode?

Welke coaching werkt goed bij de trainee?

Welke tips heb je voor de volgende coach (algemeen en specifiek voor coaching)?

Figure C-3: Coach Report Form [1]

Trainee logboek

Datum:

Coach:

Beschrijf je prestatiedoel (wat?) en je procesdoel (hoe?). Beperk je tot maximaal 3 prestatie en 3 procesdoelen.

Doel	**Wat** is mijn **prestatiedoel**? **Wat** wil ik doen? **Wat** wil ik verbeteren?	Wat is mijn **procesdoel**? **Hoe** ga ik dit doen? Waar ga ik op letten? Welke aanwijzingen of triggervragen stel ik mijzelf?
1		
2		
3		

Geef cijfers van 1 – 10 op de volgende vragen (1 = slecht, 10 = uitstekend):

Ging het goed met mijn doel vandaag?	Doel 1: Doel 2: Doel 3:
Had ik controle over hoe ik me voelde (emoties)?	
Hoe positief was ik vandaag over mijzelf?	
Hoe was mijn concentratie?	
Hoe goed accepteerde ik mijn fouten?	

Evaluatie

Hoe is het gegaan? Op welke manier heb ik aan mijn doel gewerkt?

Doel	Hoe heb ik aan mijn doel gewerkt?	Hoe is het gegaan?
1		
2		
3		

Wat ging goed tijdens de dienst? Noem 3 dingen!

1	
2	
3	

Wat wordt mijn volgende doel? (indien van toepassing)

Figure C-4: Trainee Log Form [1]

Appendix D

Ecological Interface Design - Origin and Theoretical Foundations

As described in Section 3-1, EID is a design framework that focuses on the specific problem of designing human-machine interfaces for complex sociotechnical systems and thereby on making constraints and relationships in the work domain visible to the end-user [9].

During the development of the EID framework, the first step was to identify important challenges when designing complex human-machine systems. First of all, the classification of events in complex human-machine systems can be based on the degree of novelty of these events from the perspective of the operator: familiar (routine) events, unfamiliar but anticipated events and finally, unfamiliar and unanticipated events [9]. When taking a closer look at unfamiliar and in particular unanticipated events, human factor problems that usually arise are characterized as mistakes (errors of intention), while human factor problems in routine events are usually characterized as slips (errors of execution) [9]. Traditional interface design practices however, do not provide the right support to operators during familiar, unfamiliar and unanticipated events and hence a different approach was required to be able to design an interface that supports the operator in all three types of events [9]. This is where the second goal of EID comes in: the interface should support the entire range of activities an operator might be faced with [9]. Finally, when designing an interface, it is important to keep in mind that an interface is part of a control system which is inherently bound by the laws of control theory [9]:

- Law of Requisite Variety: a complex system requires complex controllers;
- Physical systems can be described by a set of constraints;
- Every controller must possess a model of the system being controlled.

Following these observations and keeping these fundamental laws in mind, a generic structure for the interface design problem was established by Vicente and Rasmussen [9]. This structure can be found in Figure D-1.

The generic structure, shown in Figure D-1, consists of a minimal set of questions to which any approach to interface design should provide answers [9]. First of all, how can the work domain complexity be described in a psychologically relevant way and second, how can this information be communicated to the operator in the most effective way? The first question

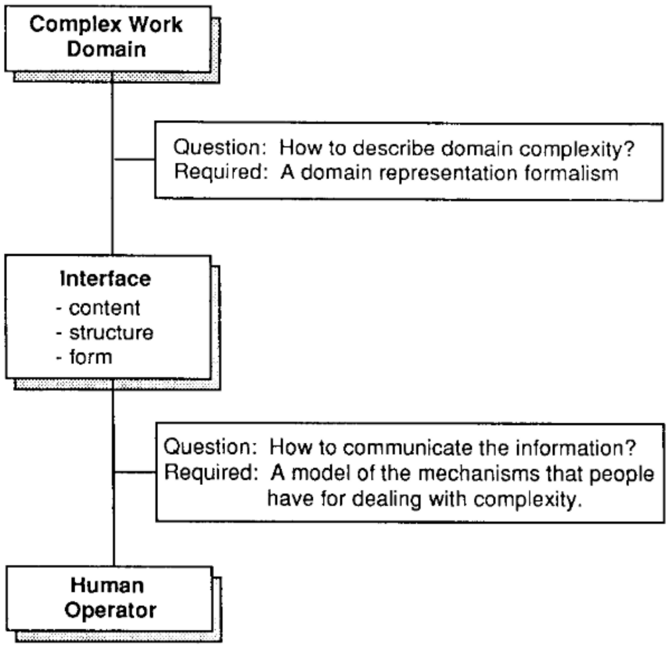


Figure D-1: Structure of the Interface Design Problem [9]

is related to the domain's characteristics, while the second question is related to those of the operator. Two theoretical concepts that originate from research done by Rasmussen in the process control domain are used to answer these questions. These are the AH, a framework useful for representing a work domain in a relevant way for the interface design problem, and the SRK-taxonomy, a framework that can be used for describing human behavior and processing of complex information [37, 38].

The Abstraction Hierarchy

The AH is a tool that was developed to model the work domain and analyze complex socio-technical systems [37]. Within interface design, the AH is used to identify the different bits of information that are required to be displayed by the interface, and how this information can best be arranged. It represents the functional properties of a technical system in several levels of functional abstraction using means-end relationships between objects, functions and system purposes. In total, five different levels of abstraction can be identified [37]:

1. Functional purpose: general goal and purpose of the complete system;
2. Abstract function: laws and principles governing the goal of the system;
3. Generalized function: processes involved in these laws and principles;
4. Physical function: physical components associated with the processes identified at the generalized function level;
5. Physical form: appearance, condition and location of these physical components.

As stated before, the different levels of the AH are connected by means-end relationships and each level is a complete but unique description of the system. The top level specifies the higher-level purpose and goal of the system or work domain. Moving down the levels answers *how* certain states or elements are achieved while moving up answers the question *why* these elements exist. The bottom level represents the physical form of the system. The AH can be a useful framework to describe the different control tasks that are required to maintain satisfactory system operation [37].

The SRK-Taxonomy

The second theoretical concept used to answer the second question of the generic interface design structure of how the information can best be presented to the operator, is the SRK-taxonomy. It was first developed by Rasmussen and can be described as a framework for describing the various mechanisms that people have for processing information [9]. The taxonomy distinguishes three different categories of human behavior based on whether information is processed as a signal, sign or symbol and thus on the corresponding level of cognitive effort required to perform a task [38].

The first or lowest level of cognitive control is the basic motor performance of the human body. This level corresponds to SBB. Rasmussen states that SBB represents sensory-motor performance during acts or activities that take place without conscious control [38]. The behavior is thus characterized as smooth, automated and highly integrated, e.g., riding a bike or walking. RBB covers interaction or activities that make use of stored rules or procedures from experience. The operator recognizes and follows a set of rules or cues in a certain sequence that does not require a significant amount of cognitive effort nor requires the operator to understand the underlying principles of the system. The last type of behavior, KBB, requires the largest amount of cognitive effort and usually takes place during unfamiliar situations for which no specific set of rules is available from previous experience. It is based on a symbolic representation and requires a good understanding of the underlying principles to be able to perform the correct action.

The Decision Ladder

As described in Section 3-1, the DL, developed by Rasmussen, is one of the most common tools for describing decision-making activities and is often used as part of the CWA to analyze the control task [39]. The processes taking place during the development of a decision can be described as "a sequence of standard subroutines linked by heuristic rules" '[39, p. 373]. When developing the DL, Rasmussen observed different behavior for experts compared to novices executing the same task. Where novices usually follow a linear sequence, experts are often observed to take shortcuts within the linear sequence [39]. The resulting DL can be found in Figure D-2.

It can be seen that the ladder consists of several elements, namely states of knowledge and information-processing activities. Moving up the ladder corresponds to the observation and analysis of a system or a situation, for example, moving up from the information-processing activity *observe information and data*, results in the knowledge state *set of observations*. Moving down the ladder on the right side, corresponds to the steps required to plan and execute a task to achieve the target system state. The number of steps an operator has to follow in this sequence strongly depends on the skills of the operator and his/her familiarity

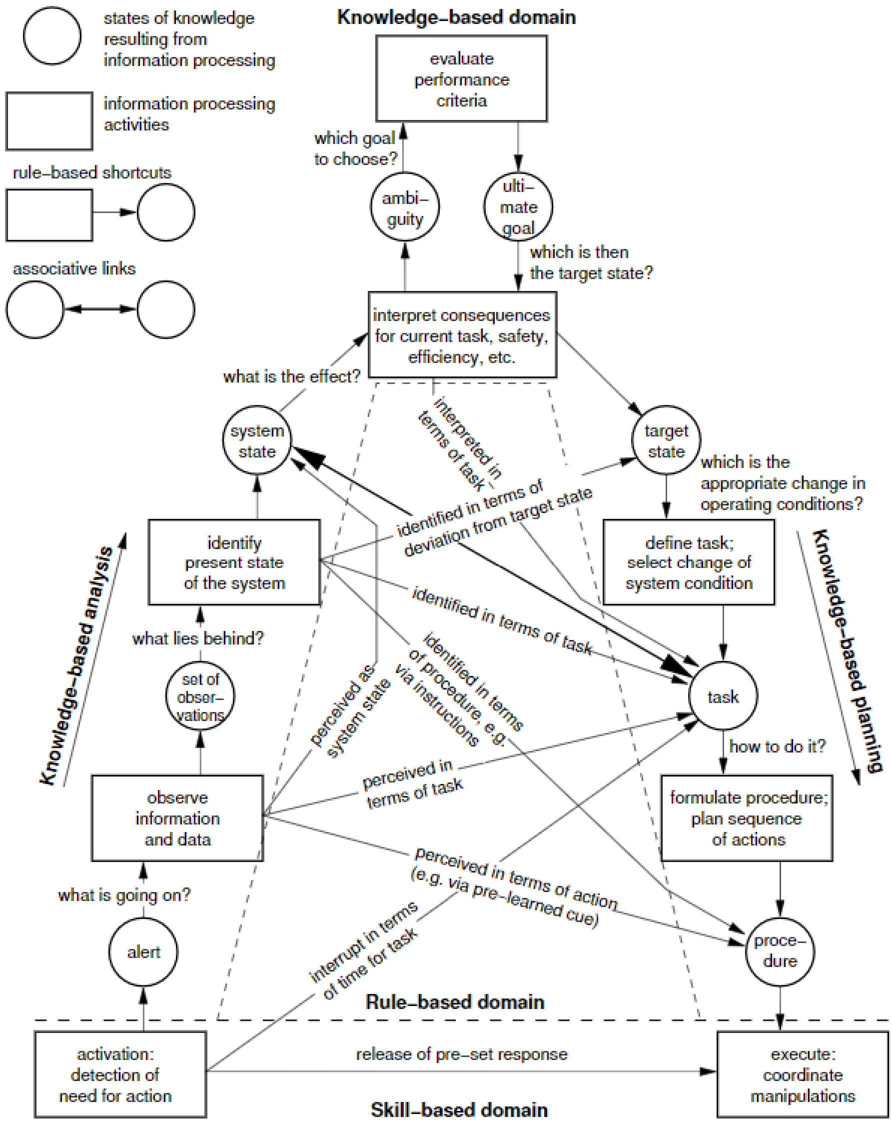


Figure D-2: The Decision Ladder, as Developed by Rasmussen [39]

with the task. As the operator becomes more skilled, he/she is more likely to act according to RBB and follow the shortcuts within the ladder. Only occasionally will the skilled operator move through the entire sequence. The lines in the diagram connecting the different elements on the left- and right-hand side, represent shortcuts and thus sequences followed by experts during control tasks [39]. Next to taking these shortcuts, a skilled operator is often inclined to start the sequence at a different point of entry instead of starting at the activation block, depending on the operator's previous experience and 'feel' for the system.

Two types of shortcuts can be identified. The first type, called shunting, connects information-processing activities (boxes) to knowledge states (circles). Shunting occurs when mental activity takes place at a symbolic level and directly leads to a certain state of knowledge [39]. An example is when the operator moves from the activity *observe information and data* directly to the knowledge state *system state*. In this case, processing information of a complete system takes place rather than observation of separate items that need to be connected first, hence the shunt to a higher knowledge state. The second type of shortcut is called an associative leap and happens when association of a knowledge state directly leads to a leap to a different state based on previous experience [39]. It thus connects two knowledge states to each other (circles). When an operator recognizes a certain situation, often the planning of an action is not necessary as the operator can still recall the corresponding action that has been correctly executed in the past and hence leaps forward in the DL sequence. This leap is also shown in Figure D-2: moving from the knowledge state *system state* to *task*.

Appendix E

Preliminary Experiment Briefing

On the following pages, the preliminary experiment briefing is presented. Participants were given this set of instructions prior to the preliminary experiment and should have a reasonable understanding of the preliminary experiment after reading it. It should be noted that the briefing presented here, is the briefing that was provided to participants of group 1 (sawtooth method). The section Experiment Procedure is slightly different for group 2.

Experiment Briefing

The influence of EID on short-term knowledge development

First of all, thank you for taking part in this experiment! You will be participating in an experiment that is taking place in the Air Traffic Management (ATM) Lab at TU Delft, in which short-term knowledge development is investigated using an ecologically-designed decision-support tool to perform an Air Traffic Control (ATC) task. This document will provide you with a short introduction to the experiment and states what is expected of you as a participant.

Goal of the Experiment

The mission statement of ATC is "to safely and efficiently organize and expedite the flow of air traffic from origin to destination and to provide information and other support to pilots". With the current developments that are leaning towards more automation within ATC, it becomes evident that a redesign of the airspace and the methods and tools used by Air Traffic Controllers (ATCos) and hence their training, is necessary.

Ecological Interface Design (EID) is a design framework that focuses on the complete range of solutions of a specific task and was used to develop the Solution Space Diagram (SSD). The SSD, developed at the TU Delft, is a conflict detection and resolution decision-support tool for ATC, that shows the complete solution space to the controller in a single glance. The goal of this experiment is to investigate short-term knowledge development and performance when being trained in an ATC task with this ecologically-designed decision-support tool as a method of support.

Your participation in this experiment is completely voluntary and you have the right to withdraw from the experiment at any moment without giving an explanation. The data from the experiment is made anonymous and is used for project-related purposes only.

Background Information

Terminology

Due to several uncertainties in, for example, weather, turbulence effects and determining the location of an aircraft with radar systems, aircraft cannot fly in too close proximity of each other. To have aircraft fly at a safe distance from each other, separation minima have been defined. Aircraft should be separated by at least 5 NM horizontally and 1,000 ft vertically. Each aircraft thus has a so-called (cylindrical) Protected Zone (PZ) around it with a radius of 5 NM and a height of 2,000 ft.

When an aircraft enters another aircraft's PZ, this is called a Loss of Separation (LoS). A conflict arises when an actual LoS occurs but also when a potential LoS within a predetermined time window presents itself, thus when two or more aircraft follow trajectories that are set to cross in the near future. This predetermined time window is often referred to as the look-ahead time. A conflict is thus not defined as a collision between two aircraft but rather as a loss of minimum separation or a violation of the PZ.

Aircraft and their Flight Labels

Each aircraft on the radar screen is accompanied by a flight label and history dots. The history dots indicate the track that was followed by the aircraft as well as its current velocity. A larger distance between the dots indicates a larger velocity. Figure 1 shows an example of an aircraft accompanied by its flight label with the following fields: callsign (SM7071), flight level (290), Indicated Airspeed (IAS) in kts (250), assigned Cleared Sector Exit Point (COPX) (COZA) and the type of aircraft, in this case medium (M). Note that in the experiment, you will be able to drag the aircraft labels.

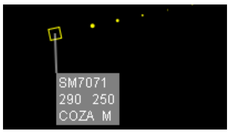


Figure 1. Impression of an Aircraft Accompanied by its Flight Label

After issuing a speed clearance (command) to an aircraft, the new speed appears in magenta in the flight label. After issuing a heading clearance, this becomes clear by the changing track of the history dots.

Solution Space Diagram

As stated before, the SSD is a decision-support tool that was developed for ATC. With unanticipated events that are not uncommon in the ATC work domain, the SSD provides an instant overview of all solution possibilities in the 2D plane to the controller. Figure 2 shows an imminent LoS and the construction of the corresponding solution space.

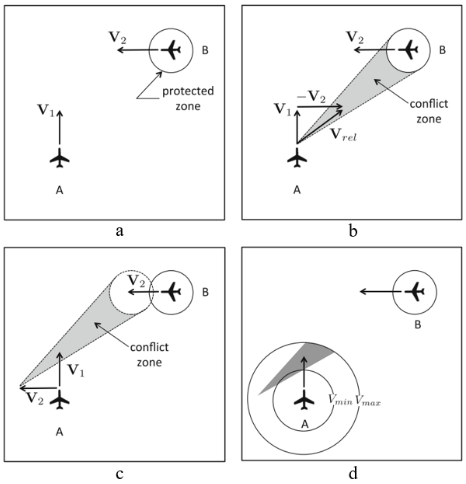


Figure 2. Construction of the Solution Space Diagram

The scenario depicted in Figure 2 considers two aircraft, a controlled aircraft A with velocity V_1 and an observed aircraft B with velocity V_2. The PZ of aircraft B is depicted by the circle around it and thus represents the minimum separation distance that should be maintained at all times. Drawing tangent lines from aircraft A to either sides of the minimum separation circle of the observed aircraft, as is depicted in Figure 2b, results in a so-called Forbidden Beam Zone (FBZ) or conflict zone. The gray area between these lines indicates that the aircraft will experience a LoS in the near future if the relative velocity of aircraft A with respect to aircraft B lies inside this area. Thus, any combination of V_1 and V_2 where the resulting relative velocity vector V_{rel} lies inside this triangle, will lead to a LoS. The gray area is thus a collection of relative velocities that will result in an aircraft entering another aircraft's PZ in the near future. Next, the conflict zone can be translated to the SSD by transposing the origin of the conflict zone by V_2. This is shown in

2/8

Figure 2c. Finally, mapping the minimum and maximum velocity or the speed envelope of the controlled aircraft on the transposed conflict zone results in the complete SSD, as shown in Figure 2d. It can be seen that the velocity vector V_1, showing the direction and the magnitude of the velocity of the aircraft, is within the velocity limits of the aircraft. It can furthermore be seen that a LoS will occur if aircraft A continues flying with its current heading and speed and no action is taken. Figure 2d shows that a change in heading will in this case resolve the conflict.

The SSD furthermore allows the controller to link conflict zones to observed aircraft by looking at the shape and orientation of the conflict zones. By doing this, the controller is able to roughly determine the location, flight direction and proximity of the neighboring aircraft. The end of the conflict zone triangle, the side opposite of the origin of the triangle, usually points to the involved aircraft at a slight offset. The width of the triangle indicates the proximity of any neighboring aircraft: A small width indicates a far-away aircraft while a triangle with a larger width indicates a nearby aircraft. Finally, by drawing a line from the controlled aircraft toward the tip of the conflict zone triangle, the absolute speed vector of the observed aircraft can be determined.

Apparatus

The experiment will be conducted on a desktop computer in the ATM Lab, which is located on the second floor of the SIMONA building, room SIM 2.03. The software used for the experiment is the Java application MUFASA. Figure 3 shows an impression of the simulation environment. In the figure you can see a square-shaped sector of 50 NM by 50 NM with one airway in the middle of the sector. When an aircraft is on course to the correct COPX, it turns from gray to green. Selecting an aircraft to inspect its SSD or to give a command, turns it yellow. When selecting an aircraft, its corresponding COPX in the sector turns magenta.

On the top left of the simulation environment, you can furthermore see two elements: the time that a particular experiment run is running and the number of that specific run.

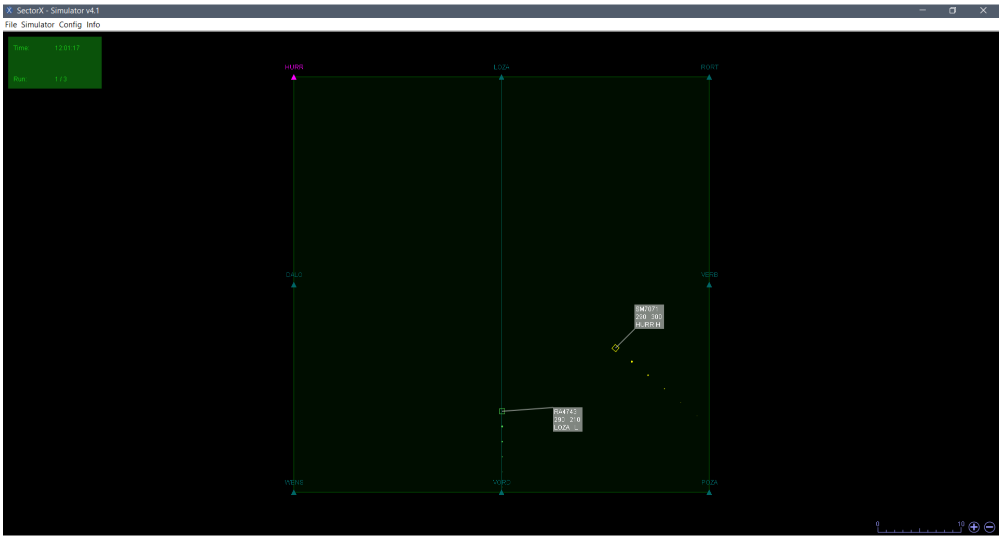


Figure 3. Simulation Environment

3/8

Figure 4 shows an impression of the SSD in the simulation environment. The SSD can be inspected by simply clicking on an aircraft in the sector. The four circles depict, ranging from largest to smallest, the maximum velocity, the current velocity, the minimum velocity and the protected zone. The conflict zone is furthermore depicted in red.

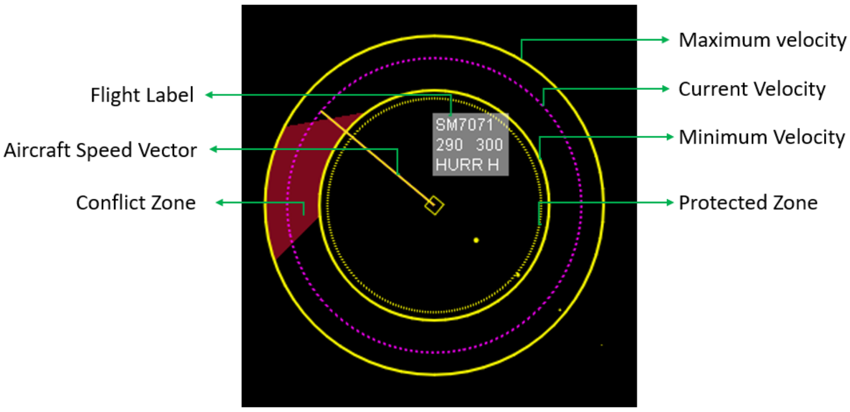


Figure 4. Impression of the Solution Space Diagram in the Simulation Environment

During the experiment, three types of aircraft occur in the simulations. The aircraft types and their speed envelopes are stated in Table 1 and are mentioned on the Experiment Info Sheet as well.

Table 1. Aircraft Types used in the Simulation Environment

Aircraft Type	Label	Minimum IAS (kts)	Maximum IAS (kts)
Light	L	180	250
Medium	M	200	290
Heavy	H	230	350

Controller Task

For this experiment you will perform several traffic scenarios. The SSD will be available as a support tool throughout the experiment but will, in certain scenarios in the experiment, gradually decrease in number of elements until you have to perform the tasks described below without the SSD and thus without any form of support.

For each scenario you are given the following tasks:

- *Separation assurance*: ensuring the minimum horizontal distance of 5NM between neighboring aircraft at all times to prevent losses of separation.
- *Clearing aircraft to their Cleared Sector Exit Point (COPX) after merging them into an existing traffic stream:*
    - Merge aircraft into an existing traffic stream without introducing conflicts;
    - Clear aircraft to their respective COPX with an exit speed of 240 kts;
    - Perform a Transfer of Control to hand aircraft over to the next sector.

4/8

- *Communicating thoughts and strategies according to a think-aloud protocol:* During the experiment you are asked to think aloud while performing the different scenarios. The introduction video gives an impression of how to do this.

Separation Assurance

The experiment is a 2D experiment, which means that altitude clearances are not part of this experiment and only speed and heading clearances can be given. The minimum vertical separation distance does therefore not have to be taken into account. However, keep in mind that traffic can occur at different flight levels during the experiment.

During the experiment you are tasked with ensuring the minimum horizontal distance of 5NM between neighboring aircraft at all times to prevent losses of separation. The horizontal separation distance of 5 NM can be determined by means of the scale in the bottom right corner of the simulation environment, which shows the scale of 10 NM.

In case two or more aircraft follow trajectories that are set to cross in the near future and will thus result in a LoS, a Short-Term Collision Avoidance (STCA) alert will be given by the system. The involved aircraft will first turn orange and then red. Both colors indicate a different time to LoS, namely 120 and 60 seconds, respectively. The 60-second red STCA alert is furthermore accompanied by an alarm sound.

As stated before, any clearance that directs the speed vector outside the conflict zone in the SSD, thus results in a safe separation. This may however not always be the optimal solution. A clearance that is both safe and efficient would direct an aircraft into a safe area closest to the conflict zone and closest to its original path or its destination waypoint. Such a clearance would result in the smallest state change with the least additional miles relative to the original flight path. It is up to you to decide on the best strategy to balance safety and efficiency. It should be noted that above all, safety is most important and PZ violations and STCA alerts are to be avoided at all times.

Merging Air Traffic

The rules that apply to the sector are:

- Only traffic between Flight Level (FL) 280 and 300 needs to be controlled;
- Inbound traffic needs to be merged into the nearest existing traffic stream, indicated by an airway, while maintaining the minimum separation distance (5 NM);
- Outbound traffic has to leave the sector at its respective COPX with an exit velocity of 240 kts;
- Aircraft have to be given a Transfer of Control (TOC) before they leave the sector;
- When aircraft are given a TOC, they have to separated (at least 5 NM) from each other and should not be involved in any conflicts.

During the experiment you will execute different traffic scenarios where the sector rules, as described above, apply. Figure 5 shows a schematic of the merging task. In this example, the existing traffic stream follows airway A-B. The aircraft entering the sector from the top left should be merged safely into this traffic stream while continuing its route (the dotted line – this line is not visible in the simulation) for as long as possible and should subsequently be guided to their target waypoint, in this case Waypoint B. Once the aircraft is directed to the correct waypoint and has the correct heading, the aircraft will turn green. The final step of the task is to perform a TOC in which control of the aircraft is handed over to the next sector.

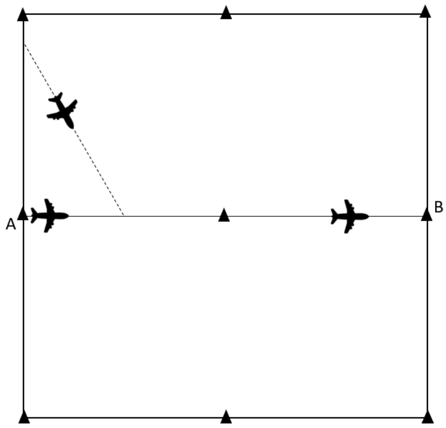


Figure 5. Schematic of the Merging Task

You control each aircraft and can thus give commands from the moment the aircraft enters the sector until it leaves the sector. Interaction with the aircraft takes place via the command window shown in Figure 6. The following commands can be distinguished:

- Execute (EXQ): clicking on this button results in executing a command.
- Speed (SPD): select an aircraft by clicking on it or clicking on its label, click *SPD*, click on a numerical value for the new speed, click on *EXQ*.
- Heading (HDG): select an aircraft by clicking on it or clicking on its label, click *HDG*, click on a numerical value for the new heading, click on *EXQ*.
- Direct To (DCT): select an aircraft by clicking on it or clicking on its label, click *DCT,* click on *EXQ*. This command can be given to send an aircraft immediately in the direction of its assigned COPX.
- Transfer of Control (TOC): select an aircraft by clicking on it or clicking on its label, click *TOC,* click on *EXQ*.
- Clear (CLR): CLR clears the command in case of a wrongly entered number or clearance. The button
- Preview (PRV): this button can be ignored during this experiment.

Note that multiple commands can be given at once, for example: "SPD240, TOC, EXQ" changes the aircraft's speed to 240 and transfer the control of the aircraft to the next sector in a single command.

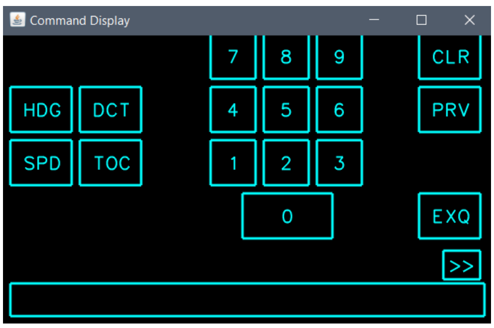


Figure 6: Command Window

6/8

Heading (HDG) clearances can vary from 000 to 360 degrees. It is however custom to give heading clearances rounded to 5 degrees. A compass card is given on the Experiment Information Sheet that can be used as a reference during the experiment. The aircraft speed envelope ranges are also stated on this information sheet.

Experiment Procedure

The experiment consists of different parts: an initial training phase on Day 1, a rest day on Day 2 and a recap and test on Day 3. The experiment will start with a briefing after which three short practice runs take place to get familiar with the interface and with giving commands to aircraft. When you indicate to have a good understanding of both the interface and what is expected of you during the experiment, the experiment will start.

Besides the briefing, the first day will consist of three training sets of approximately 30 minutes each during which the SSD support will gradually decrease. Each training set consists of six scenarios of five minutes that are all ended with a question on how you would rate the workload you experienced during that scenario. Rating the workload is done by indicating the experienced workload on a sliding bar. After each training set there is the possibility for a short break. The first day will be ended with a debrief during which you are asked to fill in a short questionnaire and get the opportunity to give comments on the scenarios or your performance.

Day 3 will start with a recap of the training sessions. This training recap is similar in length to a single training session that you will have performed on the first day. After the recap training set during which the SSD support will again gradually decrease, there's an opportunity for a short break. Next, the test, which takes approximately 40 minutes, will take place. You will perform five scenarios without any support. During all scenarios, thus both training and test scenarios, your task will remain the same: safely and efficiently merging aircraft into an existing traffic stream without introducing conflicts and while adhering to the think-aloud protocol. Similar to the training scenarios, each scenario is ended with a question on how you rate your workload during that scenario. When the test is finished and all data has been collected, a final debrief will take place. At this point you will fill out another short questionnaire and are furthermore encouraged to give comments, feedback or simply share your experiences.

The complete time schedule for the experiment is found in Table 2.

Table 2. Time Schedule of the Experiment

	Day 1 – Session 1						Day 3 – Session 2		
Experiment Part	Briefing / practice	Training Session 1	Training Session 2	Training Session 3	Debrief	*Day 2 – No experiment*	Training Recap	Test	Debrief
Time	~15 min	30 min	30 min	30 min	10 min		35 min	40 min	10 min

Experiment Execution

For each training session and for the test, the procedure described below will be followed:

1. The researcher applies / checks the settings for the next set of scenarios.
2. The researcher checks whether the participant is ready to start on the next set of scenarios.

7/8

3. The participant starts the experiment set when ready by clicking on *start experiment* on the screen.
4. The participant performs the experiment set.

Some other small remarks:

- Make sure to be well-rested before the experiment;
- Please do not discuss any of the scenarios or the procedure in general with other participants.

Thank you for participating and do not hesitate to get in touch in case of questions or remarks.

Contact information researcher:	Contact information research supervisor:
Jennifer Stoof	dr. ir. Clark Borst
j.stoof@student.tudelft.nl	c.borst@tudelft.nl
+31 6 21402028	+31 15 2789099

Appendix F

Preliminary Experiment Information Chart

On the following page, the preliminary experiment information chart is presented that was provided to the participants during the preliminary experiment. Participants were allowed to make use of the information sheet throughout the entire preliminary experiment.

EXPERIMENT INFORMATION CHART

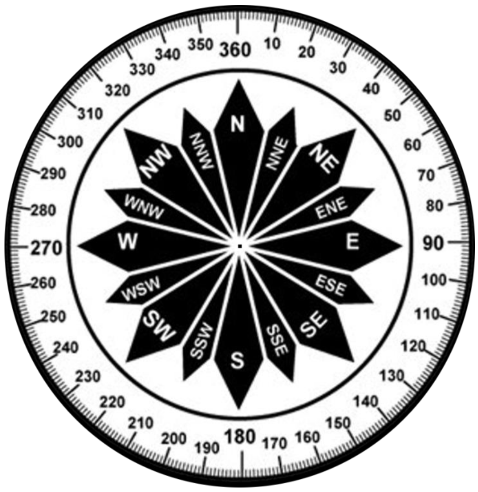
AIRCRAFT LABEL

AIRCRAFT CALL SIGN

FLIGHT LEVEL (FL) CURRENT SPEED (kts)

EXIT WAYPOINT (COPX) AIRCRAFT TYPE

COMPASS CARD



INPUT CONTROLS

LEFT MOUSE CLICK	A/C SSD INSPECTION / COMMAND WINDOW INPUT
CTRL + SCROLL WHEEL	ZOOM IN/OUT

AIRCRAFT TYPES

AIRCRAFT TYPE	LABEL	MINIMUM IAS (KTS)	MAXIMUM IAS (KTS)
LIGHT	L	180	250
MEDIUM	M	200	290
HEAVY	H	230	350

Appendix G

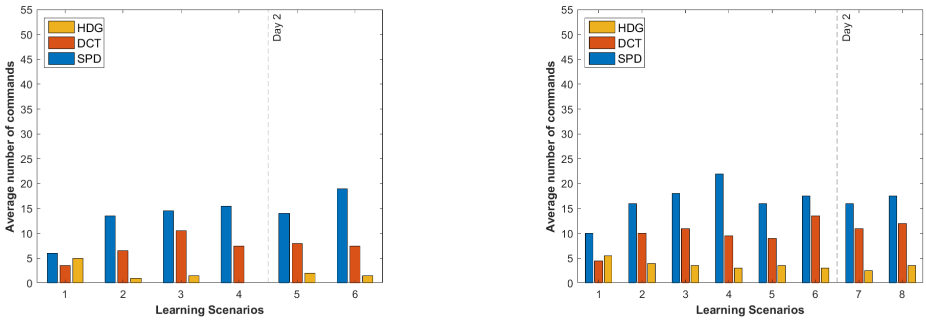
Preliminary Experiment Results

This appendix presents additional results of the preliminary experiment. The results that are shown in this appendix are the results for the type of commands, STCA count, STCA duration and minimum horizontal separation distance.

It should be noted that group 1, corresponding to the first scaffolding method presented in Section 5-2-1, performed less measurement scenarios on Day 1 in the training phase, which can also be seen in Figure 5-7. The dotted vertical line in the training phase figures indicates the separation between the measurement runs that were taken on Day 1 and on Day 2, whereas the dotted line in the test phase figures indicates a difference in complexity level. Scenarios 1 and 2 were of equal complexity as well as scenarios 3 and 4.

Commands

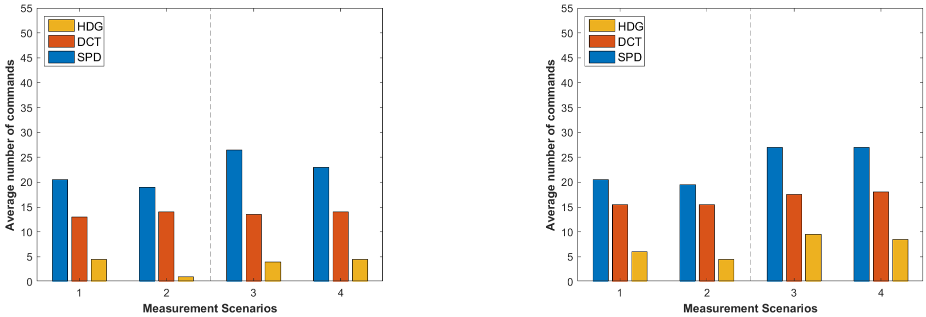
Figure G-1 shows the different type of commands that were used in the measurement scenarios of the training phase, whereas Figure G-2 shows the different type of commands that were used in the scenarios of the test phase. The commands that are shown are the Speed (SPD), Heading (HDG) and DCT commands.



(a) Group 1 **(b)** Group 2

Figure G-1: Average Number of Commands per Scenario during the Training Phase

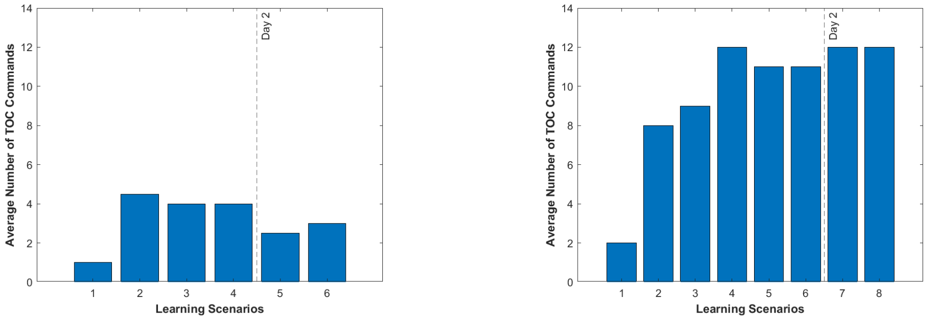
Figure G-3 shows the number of TOC commands that were used in the measurement scenarios during the training phase, whereas Figure G-4 shows the number of TOC commands that were used in the scenarios in the test phase. The number of TOC commands gives an indication of the strategies used by participants. TOC commands could either be given very early in



(a) Group 1 **(b)** Group 2

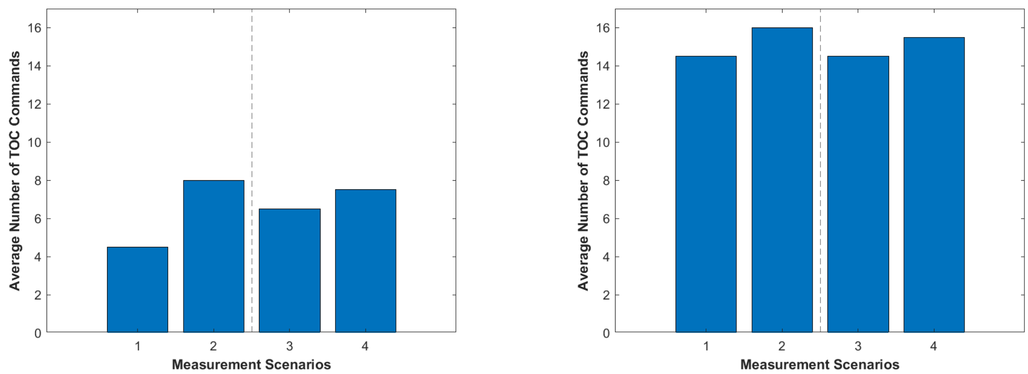
Figure G-2: Average Number of Commands per Scenario during the Test Phase

the sector or just before the aircraft would leave the sector. Hence, it gives an indication of their level of confidence when executing the task and how far ahead the participants were planning. The large difference in number of TOC commands between the two groups can furthermore be explained by participants from group 2 also giving the TOC command to aircraft at different flight levels to make a distinction between aircraft that should and should not be controlled.



(a) Group 1 **(b)** Group 2

Figure G-3: Average Number of TOC Commands per Scenario during the Training Phase



(a) Group 1 **(b)** Group 2

Figure G-4: Average Number of TOC Commands per Scenario during the Test Phase

STCA Count

Figures G-5 and G-6 show the STCA count in the measurement scenarios of the training phase, whereas Figures G-7 and G-8 show the STCA count in the scenarios of the test phase.

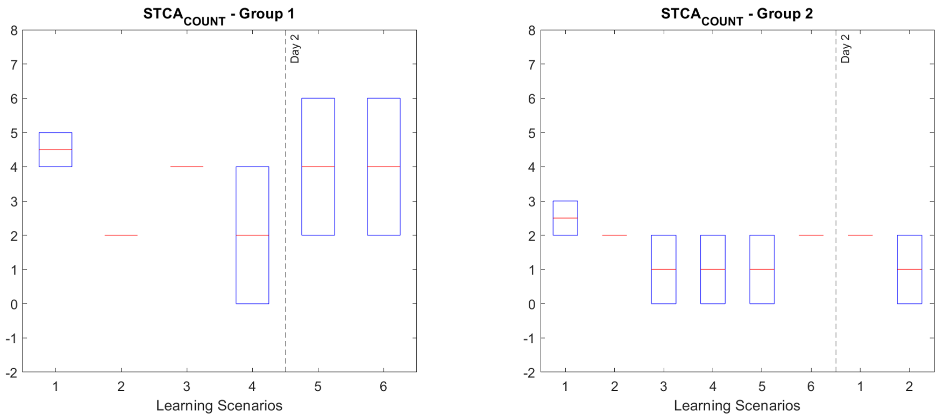


Figure G-5: STCA Count during the Training Phase

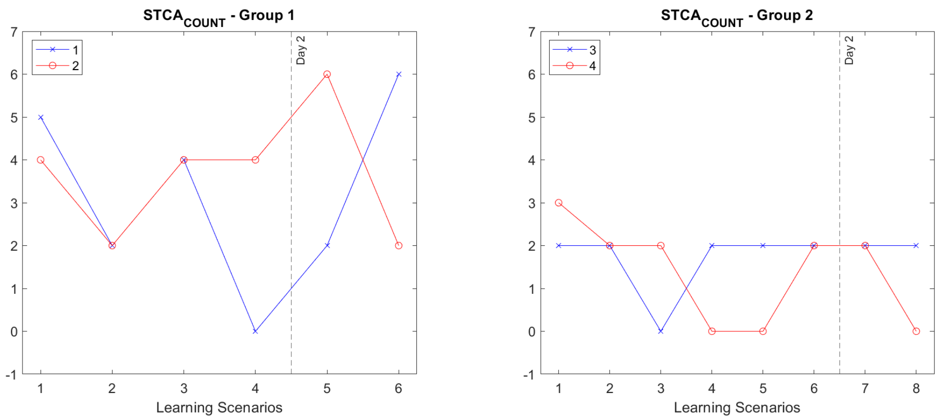


Figure G-6: STCA Count per Participant during the Training Phase

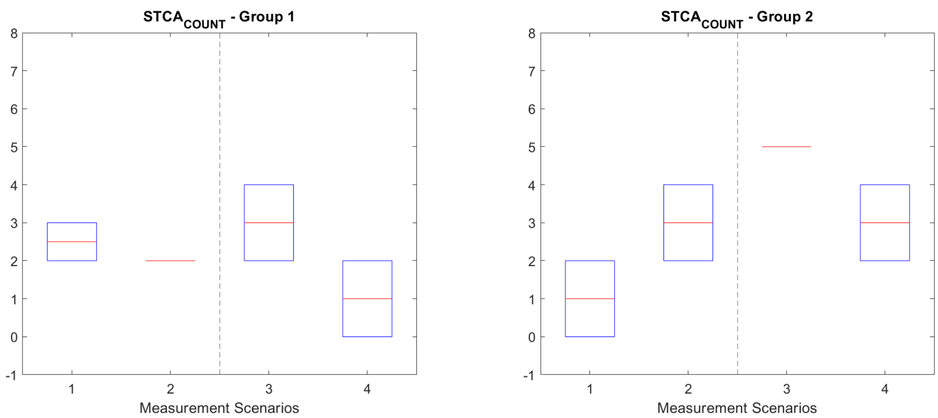


Figure G-7: STCA Count during the Test Phase

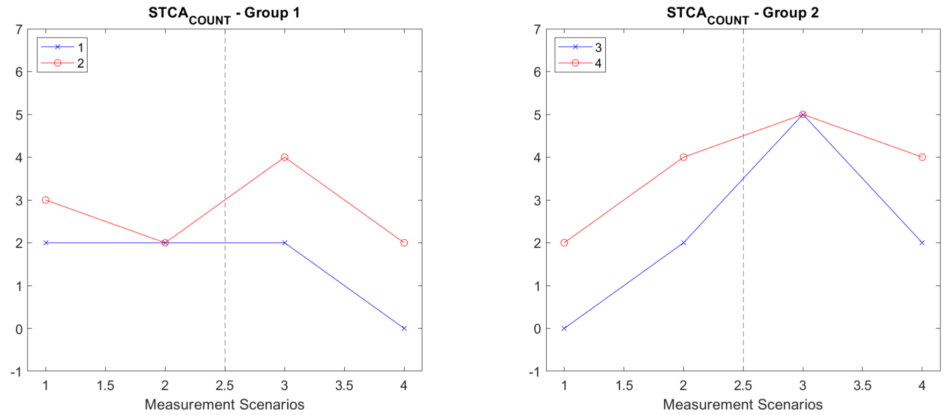


Figure G-8: STCA Count per Participant during the Test Phase

STCA Duration

Figures G-9 and G-10 show the STCA duration in the measurement scenarios of the training phase, whereas Figures G-11 and G-12 show the STCA duration in the scenarios of the test phase.

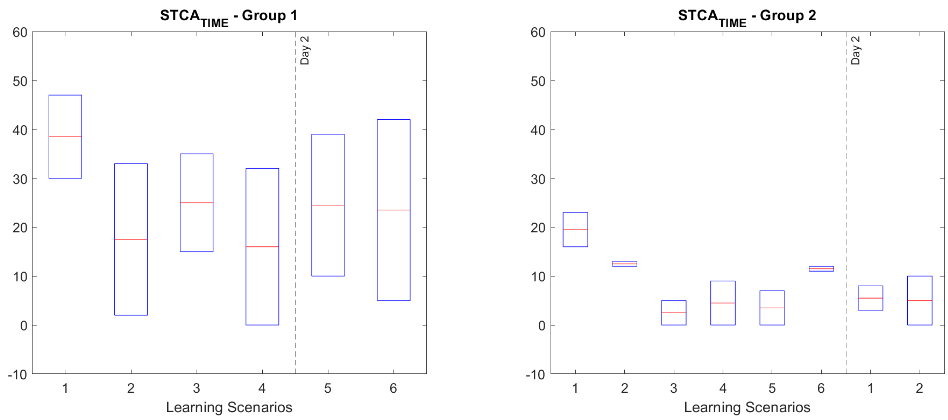


Figure G-9: STCA Duration during the Training Phase

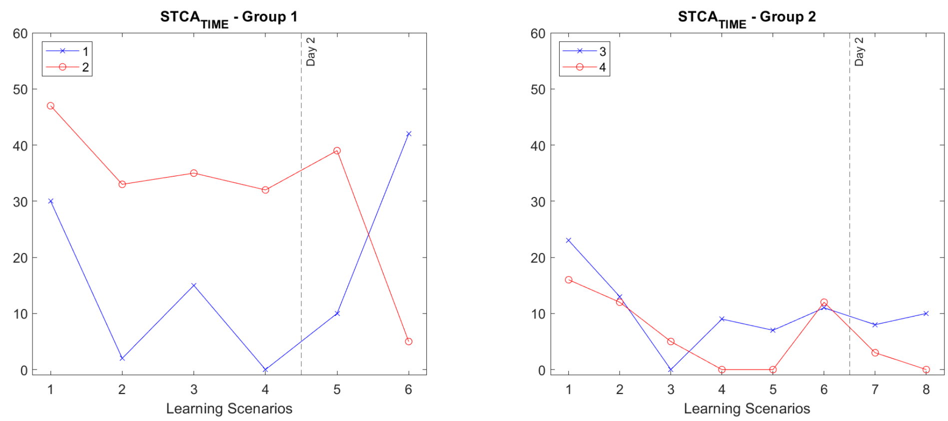


Figure G-10: STCA Duration per Participant during the Training Phase

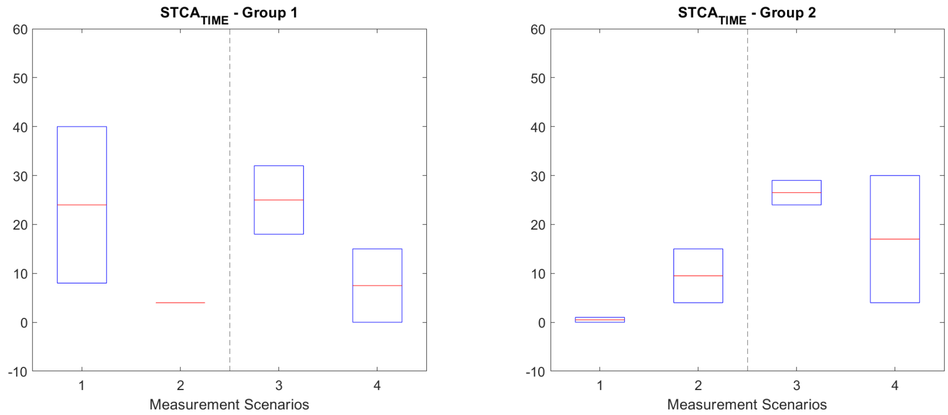


Figure G-11: STCA Duration during the Test Phase

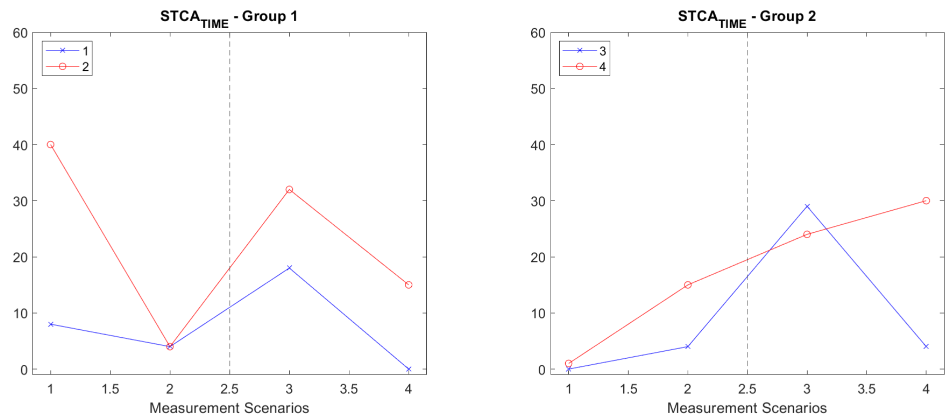


Figure G-12: STCA Duration per Participant during the Test Phase

Minimum Separation Distance

Figures G-13 and G-14 show the minimum horizontal separation distance in the measurement scenarios of the training phase, whereas Figures G-15 and G-16 show the minimum horizontal separation distance in the scenarios of the test phase.

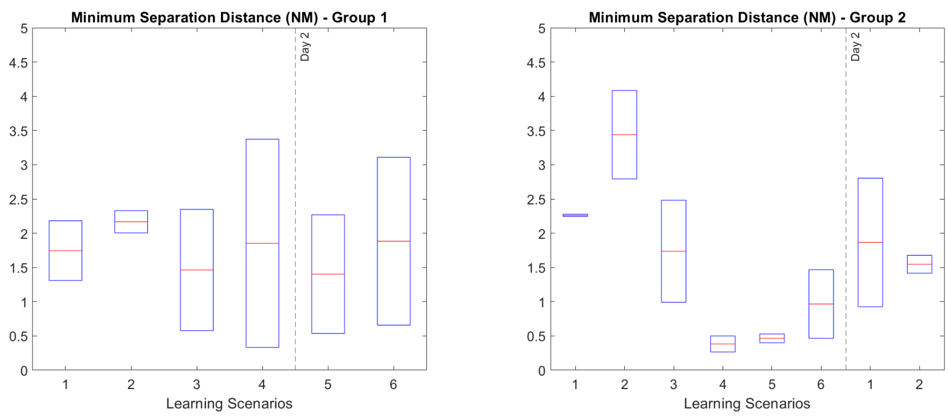


Figure G-13: Minimum Horizontal Separation Distance during the Training Phase

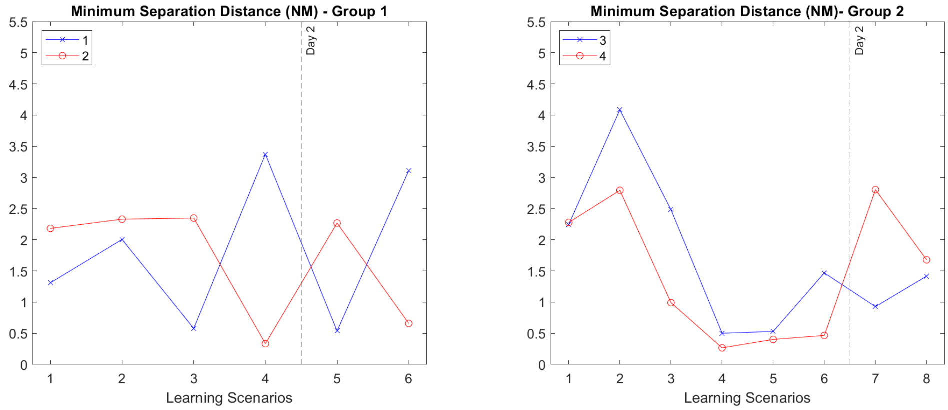


Figure G-14: Minimum Horizontal Separation Distance per Participant during the Training Phase

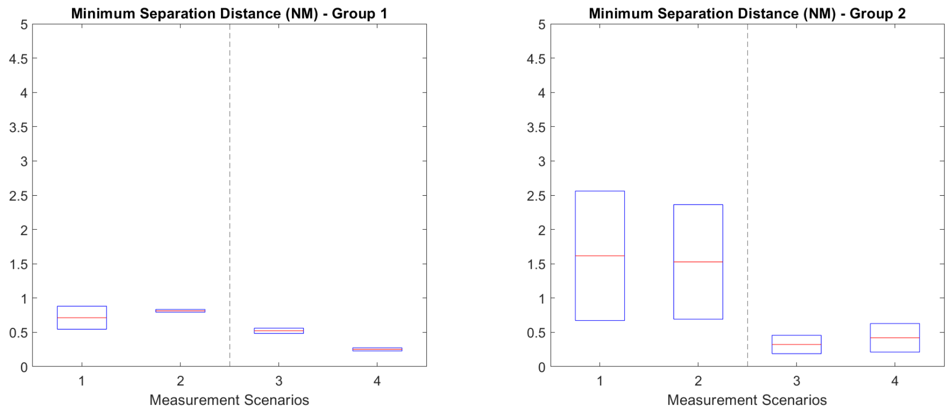


Figure G-15: Minimum Horizontal Separation Distance during the Test Phase

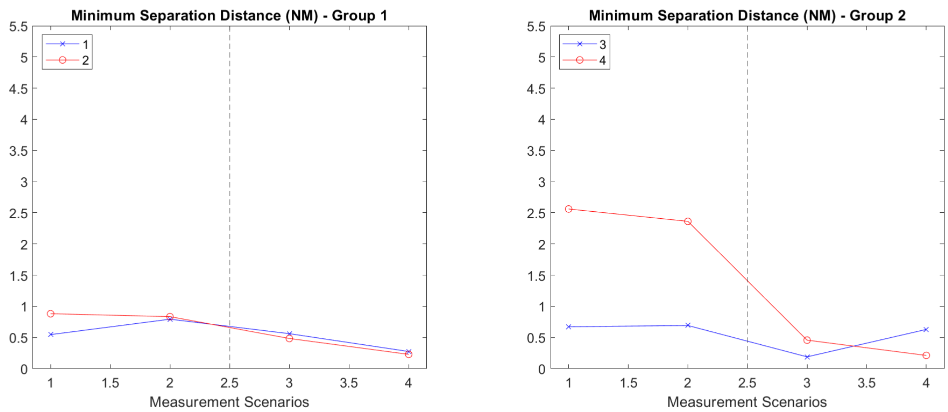


Figure G-16: Minimum Horizontal Separation Distance per Participant during the Test Phase

Part IV

MSc Thesis Report Appendices

Appendix H

Recruitment of Experiment Participants

The poster that is presented here was used to recruit experiment participants. It was shown during a lecture to recruit first-year MSc students of the Control and Simulation profile of the MSc Aerospace Engineering for participation in the experiment and was put up on several places around the department and faculty.

Looking for Experiment Participants!

For my MSc thesis into short-term knowledge development during Air Traffic Control (ATC) training, I am looking for participants to take part in my experiment in the ATM Lab at the Faculty of Aerospace Engineering in the **week of 9 December or 16 December**.

Why you want to participate in this experiment:

- No preparation necessary;
- No previous experience required;
- Experiment takes only 2 – 2,5 hours;
- Great opportunity to experience a human-machine systems experiment!

Participant requirements:

- No experience in ATC / little knowledge about ATC;
- Did not take the elective course Supervisory Control (yet);
- Did not take part in other ATC related experiments.

Interested? Sign up via: https://forms.gle/x4PNKzn7m1zj4rFc7
Questions? Send an email to: **j.stoof@student.tudelft.nl** or find me in SIM0.08

Appendix I

Pre-Experiment Questionnaire and Group Division

The following page presents the pre-experiment questionnaire. This questionnaire was filled out by all participants prior to receiving the experiment briefing. This questionnaire contained questions about the background and education of the participants and was meant to gain an insight into participants' activities and initial knowledge related to ATC. The information gained from this questionnaire was used to make a balanced division of two groups of participants: one group that was trained with the SSD and one group that was trained without.

First, the questionnaire is shown, after which the results are presented that were used for the group division. Finally, the group division with assigned scores per participant is presented.

Pre-Experiment Questionnaire

Thank you for taking part in this experiment! Before you receive more information about the experiment, I kindly ask you to complete the following questionnaire.

In case of questions, do not hesitate to contact me!

j.stoof@student.tudelft.nl

*Required

1. Email address *

   ____

2. Name *

   ____

3. Age *

   ____

4. I am a *

   Mark only one oval.

   - ( ) 1st year MSc student
   - ( ) 2nd year MSc student
   - ( ) 3rd(+) year MSc student
   - ( ) PhD Student
   - ( ) Other: ______________________________

5. MSc Track/Profile *

Mark only one oval.

- C&S
- ATO
- AWEP
- ASM
- Other: ______________________

6. Phone number (in case of delays/cancellations) *

____

Air Traffic Control (ATC) Knowledge

7. Do you have any knowledge on the ATC work domain in general? (e.g., how pilots and Air Traffic Controllers guide aircraft safely during their flight from gate to gate.) *

Mark only one oval.

- Yes, I have extensive knowledge on the ATC work domain.
- Yes, I have followed a course on ATC and have quite some knowledge about the subject.
- Yes, I have some knowledge on the ATC work domain.
- Yes, I have little knowledge on the ATC work domain
- No, I do not have any knowledge on the ATC work domain.

8. Do you have any knowledge on the rules and strategies that Air Traffic Controllers (ATCOs) use to guide aircraft as safely and efficiently as possible? *

Mark only one oval.

- Yes, I know what the ATCOs learn during their training and I am familiar with the rules and strategies.
- Yes, I know a few of these rules but do not exactly know how they are applied.
- Yes, I have a good idea of these rules but do not know any of them by heart.
- No, I do not know any of these rules or strategies.

9. Do you have any (professional) ATC experience? *

Mark only one oval.

- Yes, I have completed an ATC training course.
- Yes, I have completed part of an ATC training course but have not finished it.
- Yes, I have had some instructions / lessons on ATC.
- Yes, I participated in an ATC experiment.
- No, I do not have any (professional) experience.

10. Estimate your expertise in ATC: *

Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
No expertise											Expert Controller

11. Have you worked or do you currently work in the ATC field? (e.g., for study, research or any other parttime of fulltime job related to ATC) *

Mark only one oval.

- Yes, I work(ed) as an Air Traffic Controller or in the support staff.
- Yes, I study / do(/did) research in the field of ATC.
- Yes, I work(ed) with ATC systems.
- No, and I have not worked in the ATC field in the past.

12. Do you (currently) play any video games or apps related to ATC? *

Mark only one oval.

- Yes, on a(n) (almost) daily basis.
- Yes, on a weekly basis.
- Yes, on a monthly basis.
- Only a few times per year.
- Yes, in the past.
- No.

Other	Please fill in your availability in the following sheet by filling in your first and last name in ONE of the time slots: https://docs.google.com/spreadsheets/d/1txmqKvas_xfJosG8LBfCA8pIAYSnPGnENBdkviSo_XA/edit?usp=sharing

13. I have chosen a time slot to take part in the experiment. *

Mark only one oval.

- Yes
- No

14. Have you participated in an ATC-related experiment before? If yes, how long ago and whose experiment was this? *

15. Do you have any other remarks worth mentioning for this research? *

Google

Below, the results are shown per question. For each question, the score that was assigned per answer is also stated.

Figure I-1 shows the results for the question where participants were asked to indicate whether they had any knowledge about the ATC work domain. The answers were scored 0, 1, 1.5, 2 and 2.5, respectively, where the lowest score was assigned to 'No, I do not have any knowledge on the ATC work domain' and the highest score was assigned to 'Yes, I have extensive knowledge on the ATC work domain'.

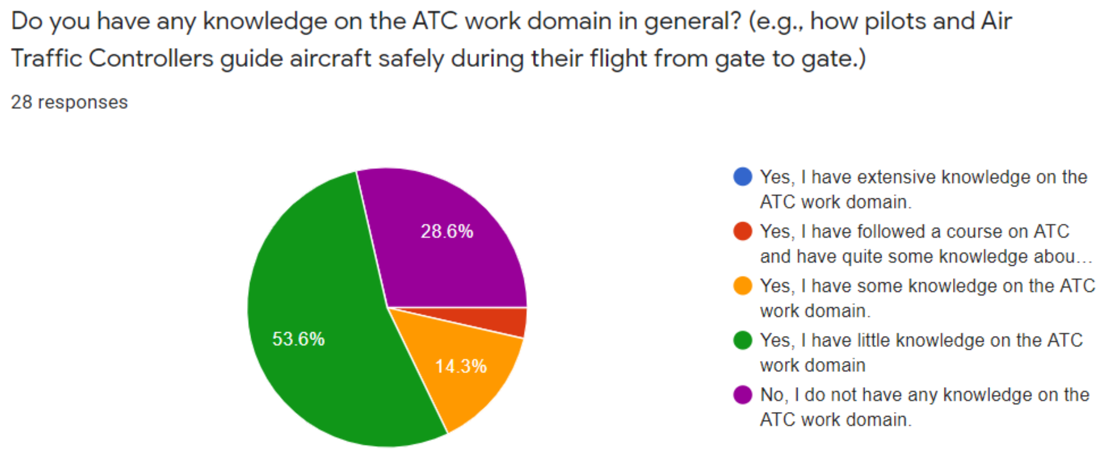


Figure I-1: Results for Question on Air Traffic Control Knowledge

Figure I-2 shows the results for the question where participants were asked to indicate whether they and any knowledge about ATC strategies. As this knowledge was considered to be of great influence to the participants' performance in the experiment, the answers were assigned scores with larger weights. Having no knowledge was assigned a score of zero, the following three answers were assigned scores of 2, 3 and 4, respectively.

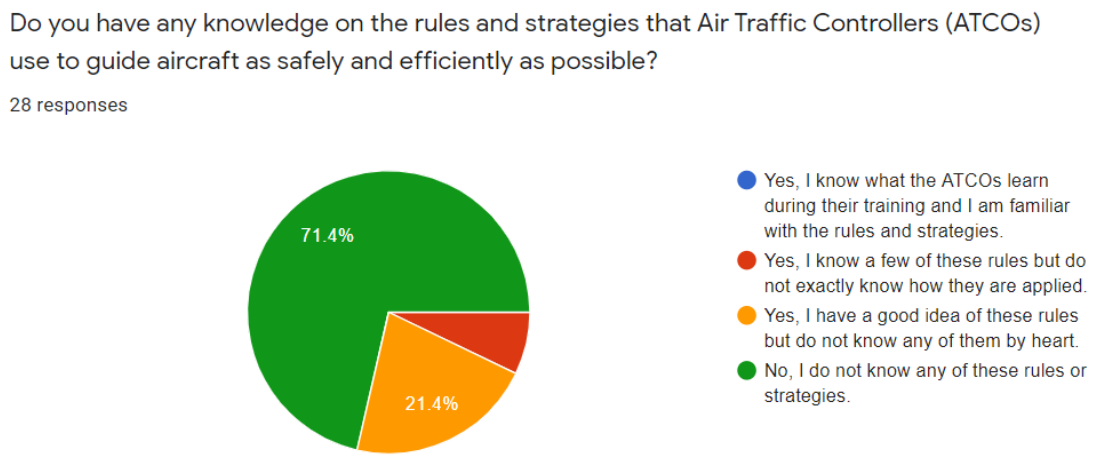


Figure I-2: Results for Question on Air Traffic Control Strategy Knowledge

Figure I-3 shows the results for the question where participants were asked to indicate whether they had any (professional) ATC experience. Having followed a 'professional' course was considered to be of great influence to the participants' performance in the experiment as they would have probably learned about ATC strategies during such a course. These two answers were thus assigned a score of 3 and 5. Having followed some instructions was considered to be of less influence as the university course on ATC generally does not get into controller strategies. This answer was thus assigned a score of 1. In case participants participated in an experiment before, they were first asked what kind of ATC experiment this was. In case this experiment did not cover controller strategies, the participant was allowed to participate in the current experiment and this answer was assigned a score of 0.5. Finally, having no experience was assigned a score of zero.

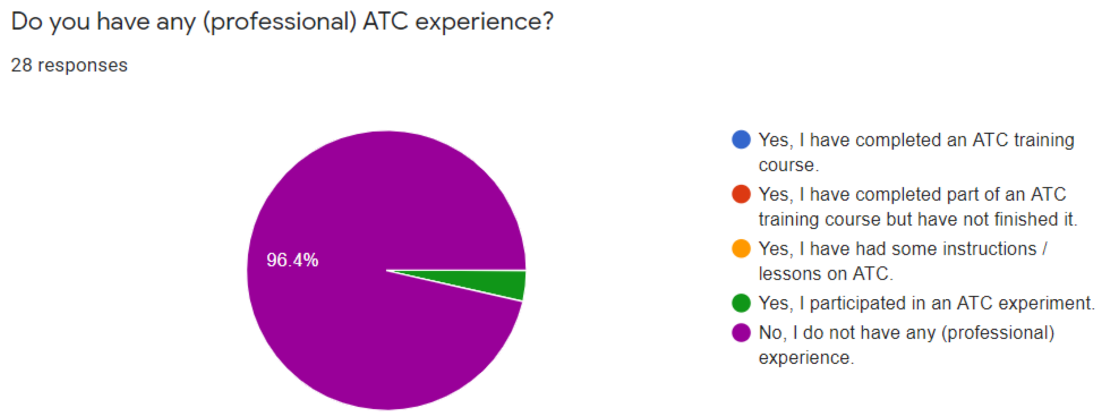


Figure I-3: Results for Question on (professional) Air Traffic Control Experience

Figure I-4 shows the results for the question where participants were asked to indicate whether they had any work experience in the ATC field. The answers to this question were assigned scores of 0, 1, 1.5 and 2, respectively, where a score of 0 was assigned to the answer 'No, and I have not worked in the ATC field in the past' and a score of 2 was assigned to 'Yes, I work(ed) as an Air Traffic Controller or in the support staff'.

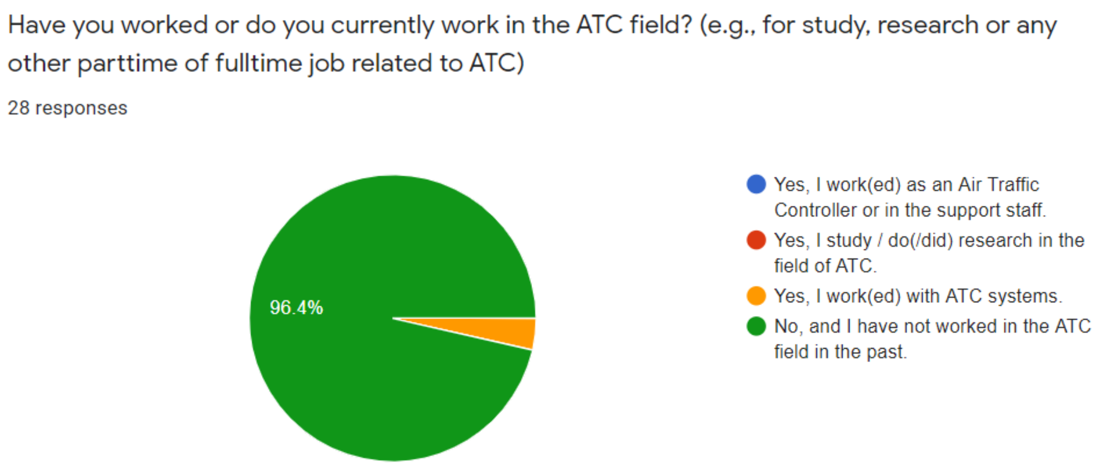


Figure I-4: Results for Question on Work Experience in the Air Traffic Control Domain

Figure I-5 shows the results for the question where participants were asked to indicate whether they previously or currently played any video games related to ATC as this was believed to potentially be of influence to participants strategies. The answers to this question were assigned scores of 0 to 2.5 with a 0.5-point increment, where a score of 0 was assigned to the answer 'No' and a score of 2.5 was assigned to the answer 'Yes, on a(n) (almost) daily basis'.

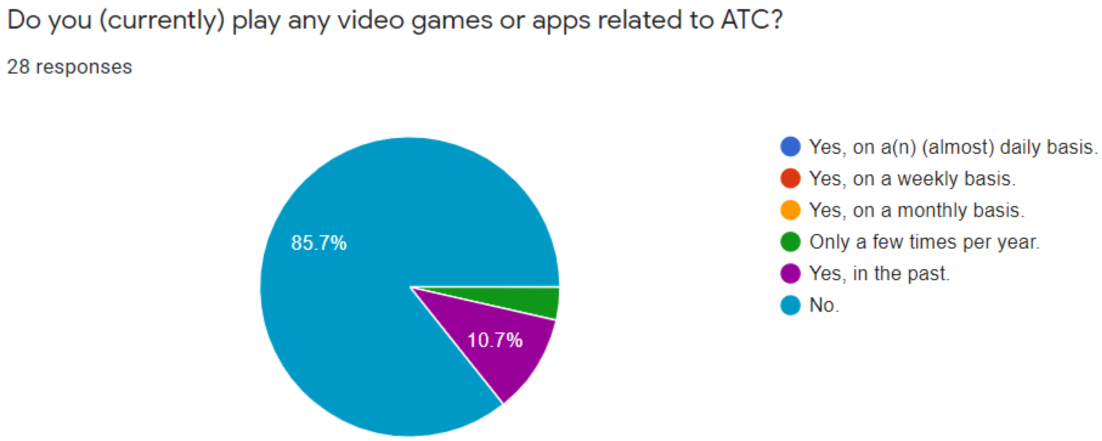


Figure I-5: Results for Question on Air Traffic Control Game Experience

The scores per participant were summed in order to make matching pairs, of which one participant would be assigned to the SSD group and one participant would be assigned to the control group. Next to the score, the pairs were also matched on master track and the rating of their own level of expertise. Figure I-6 shows the results for the question where participants were asked to rate their own level of expertise in ATC. The answers to this question were not assigned scores but were only used for checking (and possibly slightly adjusting) the matched pairs in the final group division.

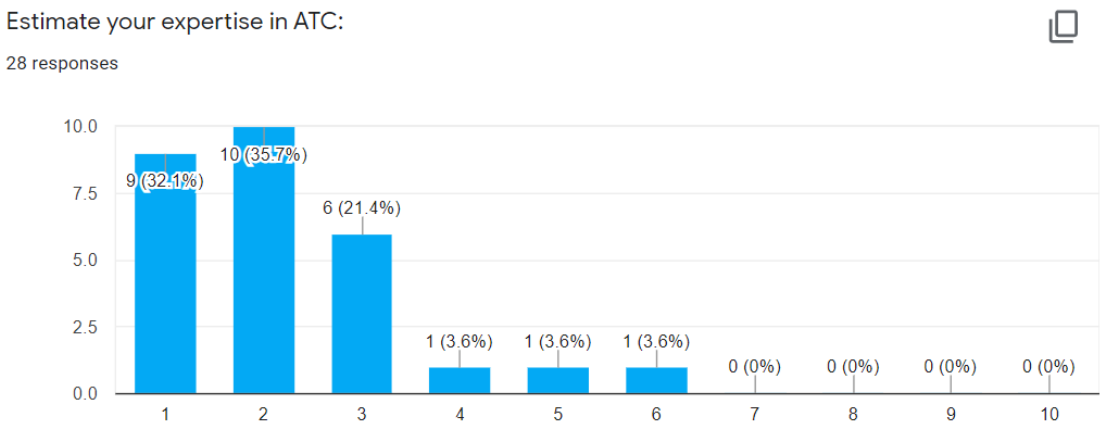


Figure I-6: Results for Question about One's Own Air Traffic Control Level of Expertise

Table I-1: Group Division including Assigned Scores per Participant

SSD Group			NO SSD group		
Participant	MSc Track	Score	Participant	MSc Track	Score
1	ATO	4,5	15	C&S	5
2	C&S	3,5	16	C&S	3,5
3	AWEP	0	17	AWEP	0
4	C&S	1,5	18	C&S	2,5
5	C&S	1,5	19	S&C	1,5
6	C&S	1	20	C&S	1
7	AWEP	1	21	AWEP	1
8	AWEP	1	22	AWEP	1
9	Ind. Ecology	1	23	C&S	1
10	C&S	0	24	C&S	0
11	C&S	1	25	C&S	1
12	C&S	3,5	26	C&S	4
13	C&S	1	27	ATO	1
14	C&S	0	28	Space	0
		20,5			22,5

Appendix J

Experiment Briefing

On the following pages, the experiment briefing is presented. Before the experiment, each participant was sent this two-page document explaining the goal and motivation for the experiment. Additionally, the participants were given information about the procedure of the experiment and the contact details of the researcher. The experiment briefing was the same for both participant groups.

Experiment Briefing

First of all, thank you for taking part in this experiment! You will be participating in an experiment that is taking place in the Air Traffic Management (ATM) Lab at TU Delft, in which short-term knowledge development is investigated when learning to perform an Air Traffic Control (ATC) task. This document will provide you with a short introduction to the experiment and states what is expected of you as a participant.

Goal of the Experiment

Current developments within ATC are leaning towards more automation and hence, Air Traffic Controllers (ATCOs) are taking on a more supervisory role. This means that some of the ATCO's skills required for controlling the air traffic and accomplishing the mission statement - "to safely and efficiently organize and expedite the flow of air traffic from origin to destination and to provide information and other support to pilots"- will be down-graded or might even disappear, while other skills, such as Conflict Detection and Resolution (CD&R), are likely to increase in importance. It thus becomes evident that a redesign of the airspace, the methods and tools used by ATCOs and hence their training, is necessary in this changing environment. In the current ATC training, trainees are expected to learn a large part of the rules of the air, the ATC strategies and rules of thumb to control the air traffic by means of discovery learning. In this experiment, you will learn to perform an ATC task in a similar fashion in an effort to gain new empirical insights in short-term knowledge development in the ATC domain.

Your participation in this experiment is completely voluntary and you have the right to withdraw from the experiment at any moment without giving an explanation. The data from the experiment are made anonymous and are used for project-related purposes only.

Experiment Procedure

The total experiment will take approximately two and a half hours. During the experiment, you will work with an interactive step-by-step script that guides you through the experiment during a number of scenarios, that each have a specific learning objective. First, in order to have a good understanding of how to fulfill your role as ATC student in the experiment, all tools and features available to you in the experiment simulator will be described and shown to you during nine practice scenarios. Subsequently, the control task for the experiment is explained after which the training phase, consisting of eleven scenarios, follows during which you will be able to practice the control task. Finally, a test takes place, also consisting of eleven scenarios during which you have to apply your newly-obtained knowledge and skills. You are furthermore required to answer several questions at certain points in the experiment to test your understanding so far.

The scenarios in the practice phase have a varying length from 2 to 3 minutes, whereas the scenarios in the training and test phase all last 2 minutes. Each scenario ends with a question on how you would rate the difficulty of that scenario. Rating the difficulty is done by indicating the perceived level of difficulty on a sliding bar, where sliding it left indicates an easier scenario and sliding it right indicates a more difficult scenario. During the experiment, there is an opportunity for a break between the sessions. The experiment ends with a short debrief during which you are asked several questions and get the opportunity to give comments and feedback on the scenarios or your performance or to simply share your experiences.

The complete time schedule for the experiment is found in Table 1.

Table 1. Time Schedule of the Experiment

Experiment Part	Briefing	Practice Session	Training Session	Test Session	Debrief
Time	~5 min	~70-90 min	~30 min	~30 min	~5 min

Experiment Execution

For each session, the procedure described below will be followed:

1. The researcher applies / checks the settings for the next session/set of scenarios.
2. The researcher checks whether the participant is ready to start with the session/set of scenarios.
3. The participant starts the session when ready by following the steps described in the training manual.
4. The participant performs the steps described in the manual for that session.

Some other small remarks:

- Make sure to be well-rested before the experiment;
- Please do not discuss any of the scenarios or the procedure/experiment in general with other participants as this could influence the results.

Thank you for participating and do not hesitate to get in touch in case of questions or remarks.

Contact information researcher:
Jennifer Stoof
j.stoof@student.tudelft.nl
+31 6 21402028

Contact information research supervisor:
dr. ir. Clark Borst
c.borst@tudelft.nl
+31 15 2789099

Appendix K

Experiment Consent Form

The next page presents the consent form that was approved by the Human Research Ethics Committee of the Delft University of Technology and signed by all participants prior to the experiment.

Experiment Consent Form
The influence of EID on short-term knowledge development

I hereby confirm that

1. I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time, without having to give a reason.

2. I understand that taking part in the study involves performing different ATC scenarios in a simple desktop simulator environment and that both audio and performance data is recorded during the experiment. The researcher has explained that audio recordings will be transcribed as text and will be destroyed afterwards and that all data is anonymized. I confirm that the researcher has provided me with detailed safety and operational instructions for the experiment.

3. I have read and understood the information from the Experiment Briefing, or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.

4. I understand that information I provide will be used for project-related purposes only. The data will be analysed and results will be published in the form of a scientific article and thesis report.

5. I understand that personal information collected about me that can identify me, such as [e.g. my name or where I live], will not be shared beyond the study team.

6. I give permission for the transcribed audio recordings and performance data that I provide to be archived in [*name of data repository*] so it can be used for future research and learning.

7. I understand that this research study has been reviewed and approved by the TU Delft Human Research Ethics Committee (HREC). To report any problems regarding my participation in the experiment, I know I can contact the researchers using the contact information below.

8. I have been given a copy of this consent form.

____________________________	__________________________	________________
Name of participant [printed]	Signature	Date

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

____________________________	__________________________	________________
Researcher name [printed]	Signature	Date

Contact information researcher:
Jennifer Stoof
j.stoof@student.tudelft.nl
+31 6 21402028

Contact information research supervisor:
dr. ir. Clark Borst
c.borst@tudelft.nl
+31 15 2789099

Appendix L

Verbal Briefing Protocol

At the start of the experiment, every participant received the same instructions during a short briefing. The verbal protocol for this briefing is shown below.

- Thank the participant for taking part in the experiment and ask whether they have read the briefing.
- Explain the procedure of the experiment: The experiment consists of three parts: practice phase, training phase and test phase. After the practice phase or the first nine scenarios, there is an opportunity for a short break.
- Explain the documents that participants will need during the experiment:
  - Training script: The participant will be working with the script. Everything they need to know for the experiment, all the steps they need to perform and all the questions they need to answer are stated in this script. Participants are not allowed to write or draw in the script.
  - Information sheet: The participant is allowed to use the information sheet as a 'cheat sheet' throughout the entire experiment.
  - Answer sheet: Participants are asked to write down the answers to the questions from the script on this form. They are also allowed to write / draw in the figures in this form.
  - Consent form: Participants are asked to sign this form after the briefing and before starting the experiment, if they agree to all terms.
- From the training phase onward, the researcher records the computer screen by means of a video camera to record what participants are doing and saying. Participants will not be on camera. The audio recordings from the video will be transcribed to text, which will subsequently be used for the data analysis.
- In case of emergency, show the emergency exit.
- Participant is allowed to ask any questions regarding the experiment.
- In case of no further questions, the participant is asked to sign the consent form.

Appendix M

Training Script - SSD Group

On the following pages, the training script is presented. This script was developed to guide participants through the experiment. It should be noted that the script presented here is the script that was given to the SSD group. The script for the control group contained a different chapter on CD&R, that is presented in Appendix N. The rest of the script was the same for both groups.

Training Script

"Introduction to Air Traffic Control"

Training Script

TU Delft Delft University of Technology

Aerospace Engineering

Part of MSc Thesis - J. Stoof

February 16, 2020

Acronyms

ATC Air Traffic Control.

ATCO Air Traffic Control Officer.

CD&R Conflict Detection & Resolution.

CLR Clear.

COPX Cleared Sector Exit Point.

DCT Direct To.

EXQ Execute.

FBZ Forbidden Beam Zone.

FL Flight Level.

HDG Heading.

IAS Indicated Airspeed.

LoS Loss of Separation.

PRV Preview.

PVD Plan View Display.

PZ Protected Zone.

RT Radio Telephony.

SSD Solution Space Diagram.

STCA Short-Term Conflict Alert.

TOC Transfer of Control.

Contents

1 Introduction

First of all, thank you again for taking part in this experiment! You are participating in an experiment in which short-term knowledge development is investigated when learning to perform an Air Traffic Control (ATC) task. As was mentioned in the briefing, ATC trainees are expected to learn a large part of the rules of the air, the Air Traffic Control Officer (ATCO) strategies and rules of thumb by means of discovery learning. In this experiment, you will learn an ATC task in a similar fashion. This document will guide you through the experiment.

As was also stated in the consent form, your participation in this experiment is completely voluntary and you have the right to withdraw from the experiment at any moment without giving an explanation. The data from the experiment are made anonymous and are used for project-related purposes only.

Please think out loud during the entire experiment and try to motivate your reasoning for the decisions you make during the scenarios and the questions. Read the instructions carefully and do not hesitate to ask questions if something is unclear.

Experiment Setup

As also described in the briefing, this document is written in the form of an interactive step-by-step script to guide you through the experiment during a number of scenarios, that each have a specific learning objective. First, in order to have a good understanding of how to fulfill your role as a future ATC student in the experiment, all tools and features available to you in the experiment simulator will be described and shown to you during nine practice scenarios. Subsequently, the control task is explained after which the training phase, consisting of eleven scenarios, follows during which you will be able to practice the control task. Finally, a test takes place, also consisting of eleven scenarios during which you have to apply your newly-obtained knowledge and skills. You are furthermore required to answer several questions at certain points in the experiment to test your understanding so far.

Airspace and Traffic

The controlled airspace used in all scenarios is an artificial, en-route upper airspace sector of $50NM$ by $50NM$, designed especially for this experiment and shown to you in the following chapter. As the experiment is a 2D experiment, all aircraft will fly at the same altitude. You will be able to manipulate the aircraft's heading, but speed changes and vertical movements / separation are not supported.

2 Simulation Environment

This chapter is used to familiarize you with the simulation environment and the basic information it presents. You are furthermore taught how to give clearances (commands) to aircraft.

You may now click on *Start experiment* on the screen. The simulation is still paused at this point, so please take the time to carefully read the following information about the setup of the simulation environment.

SCENARIO 1

Simulation Environment

Simulator

As you can see on the screen, the experiment simulator consists of two displays:

- Plan View Display (PVD): The screen on the left-hand side represents the PVD and shows a top-down radar view of the sector, the entry and exit waypoints and all aircraft. You can see a square-shaped sector of 50 NM by 50 NM. On the top left of the simulation environment, you can furthermore see two elements: the time that the current experiment run has been running and the number of that specific run out of the total of that session. The bottom right furthermore contains a scale of 10 NM.
- Command Display: The display on the right-hand side represents the Command Display. The following commands can be distinguished:
  - Execute (EXQ): button used to execute a command;
  - Heading (HDG): button used to select a heading command;
  - Direct To (DCT): button used to send an aircraft immediately in the direction of its assigned exit waypoint (Cleared Sector Exit Point (COPX));
  - Clear (CLR): button used to clear a command in case of a wrongly entered number or clearance before clicking the EXQ button;
  - Transfer of Control (TOC): this button can be ignored during this experiment;
  - Preview (PRV): this button can be ignored during this experiment.

  How to interact with aircraft and issue clearances is explained later in this section.

Basic Information on the PVD

The simulation is still paused at this point. Before continuing to the dynamic part, take the time to carefully read and execute each of the following steps to learn more about the information that is presented to you on the PVD.

1. In this scenario, there is one controlled aircraft (call sign: `ALPHA`). The heading of this aircraft is indicated by a speed vector that is aligned with its current route. The tip of the speed vector

indicates the future position of the aircraft, given that the aircraft continues following the current heading at the current speed for 60 seconds. A longer speed vector thus indicates an aircraft with a higher speed. Note that the simulations will run at 3x speed which means that the speed vector indicates the future position after a time of 20 seconds.

2. Each aircraft on the radar screen is furthermore accompanied by a flight label and history dots. The history dots indicate the track that was followed by the aircraft as well as its current speed. A larger distance between the dots indicates a higher speed. Each flight label has the following fields: call sign (ALPHA), Flight Level (FL) (290), Indicated Airspeed (IAS) in kts (250), assigned COPX (WIPP) and the type of aircraft, in this case medium (M). Another example of a flight label is presented on the Experiment Information Sheet, that can be used as a reference during the experiment. Note that it is possible to drag the aircraft labels as they might sometimes overlap.

3. You can see that aircraft are by default gray-colored. Select aircraft **ALPHA** by clicking on the aircraft or clicking on its flight label. The aircraft turns yellow after it has been selected by the user. Its corresponding COPX (assigned exit waypoint) in the sector furthermore turns magenta, as can be seen on the screen. Try to drag the aircraft's flight label. You can deselect an aircraft by clicking anywhere in the PVD, outside the aircraft.

Dynamic Route Manipulation

Now that you know the basic information that is shown to you on the PVD, it is time to practice giving clearances. Before starting the simulation, carefully read all of the following steps first.

1. The route of an aircraft can be modified by making use of the Command Display. The two types of commands that can be given during this experiment are:

   - HDG: a heading command is given by selecting an aircraft in the PVD, either by clicking on the aircraft or clicking on its label, clicking on the HDG button in the command display, clicking on the individual numeric buttons that together form the new heading value, and finally, clicking on EXQ.

     After issuing a heading clearance, this becomes clear by the changing direction of the speed vector and the changing track of the history dots. Note that, depending on the difference with the current heading, it might take some time for the aircraft to take on its new heading.

   - DCT: a direct-to command is given by selecting an aircraft in the PVD, either by clicking on the aircraft or clicking on its label, clicking on the DCT button in the command display, and finally, clicking on EXQ.

     After issuing a DCT command, this becomes visible by the aircraft turning green once it is on track to its assigned COPX. Again note that, depending on the difference with the current heading, it might take some time for the aircraft to take on its new heading and turn green.

   You will thus only need the mouse to give clearances to aircraft and you will not need to make use of the keyboard.

2. After you have started the simulation, you will be giving the aircraft at least the following commands:

   (a) HDG 80

   (b) HDG 270

   (c) DCT

   When the simulator is running, take your time for each of the commands and observe how the heading commands affect the speed vector and history dots and how long it takes for the aircraft to arrive at its new heading. Also observe that after giving a DCT command, the controlled aircraft turns green *when it is deselected.*

Note that at the start of the simulation, the aircraft has not yet entered the sector. You control each aircraft from the moment the aircraft enters the sector until it leaves the sector and can thus only give commands to aircraft when they are inside the sector.

3. When you're done with step 2, continue to practice giving clearances. A compass card is presented on the Experiment Information Sheet that can be used as a reference during the experiment. HDG clearances can vary from 001 to 360 degrees. It is, however, custom to give heading clearances rounded to 5 degrees.
4. When you feel comfortable with manipulating the route of the aircraft, you can wait until the end of a scenario when the 'perceived difficulty' scale pops up.
5. After each scenario, a 'perceived difficulty' scale appears in the center of the PVD. Please indicate your experienced level of difficulty by sliding the bar to the correct position. Sliding the bar left indicates a scenario that was perceived to be easier, whereas sliding the bar right indicates a more difficult scenario.
6. The simulation will start running at 3x speed for three minutes. The time a scenario has been running is indicated in the top left of the PVD.
7. You may now press the play button in the top left corner of the simulator. Perform step 2.a until 2.c and continue to practice giving clearances if you have time left. Remember to first select an aircraft before you can give the aircraft a command.
8. After rating the difficulty at the end of the scenario, click on *Next scenario* and continue with the steps presented in the next chapter.

3 Conflict Detection & Resolution

As the mission statement of ATC is “to safely and efficiently organize and expedite the flow of air traffic from origin to destination and provide information and other support to pilots”, your main task in the experiment will also be to manage traffic as safely as possible while also keeping efficiency in mind.

Due to several uncertainties in, for example, weather, turbulence effects and determining the location of an aircraft with radar systems, aircraft cannot fly in too close proximity of each other. To have aircraft fly at a safe distance from each other, a minimum separation distance has thus been defined. Aircraft should horizontally be separated by at least 5 NM and vertically by 1,000 ft. Each aircraft thus has a so-called Protected Zone (PZ) around it with a radius of 5 NM and a height of 2,000 ft. As this experiment focuses on horizontal separation only, you can disregard the vertical separation distance of 1,000 ft.

When an aircraft enters another aircraft’s PZ, this is called a Loss of Separation (LoS), which is also depicted in Figure 3.1. A conflict arises when an actual LoS occurs but also when a potential LoS within a predetermined time window presents itself, thus when two or more aircraft follow trajectories that are set to cross in the near future. A conflict is thus not defined as a collision between two aircraft but rather as a (potential) LoS.

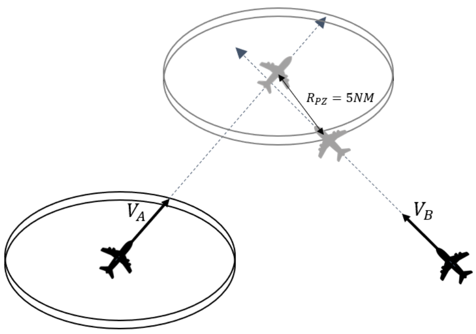


Figure 3.1: Schematic of a Conflict

This chapter is used to familiarize you with conflicts in the simulation environment as well as a decision-support tool that was developed to aid controllers in detecting and resolving conflicts and thus in ensuring a safe separation.

3.1 Solution Space Diagram

The Solution Space Diagram (SSD) is a Conflict Detection & Resolution (CD&R) decision-support tool that was developed for ATC and provides an instant overview of all solution possibilities in the 2D plane to the controller. Figure 3.2 shows an imminent LoS and the construction of the corresponding solution space.

The scenario depicted in Figure 3.2 considers two aircraft, a controlled aircraft A with velocity V_1 and an observed aircraft B with velocity V_2. The PZ of aircraft B is depicted by the circle around it and thus represents the minimum horizontal separation distance that should be maintained at all times.

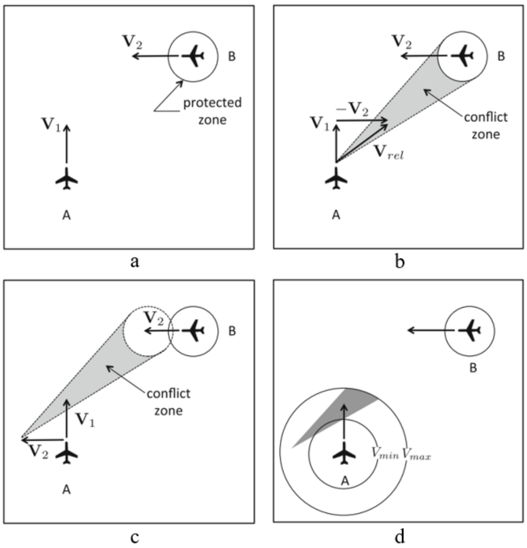


Figure 3.2: Construction of the Solution Space Diagram

Drawing tangent lines from aircraft A to either side of the minimum separation circle of the observed aircraft B, as is depicted in Figure 3.2b, results in a so-called Forbidden Beam Zone (FBZ) or conflict zone. The gray area between these lines indicates that the aircraft will experience a LoS in the near future if the tip of the relative velocity vector of aircraft A with respect to aircraft B lies inside this area. Thus, any combination of V_1 and V_2 where the tip of the resulting relative velocity vector V_{rel} lies inside this triangle, will lead to a LoS.

Next, the conflict zone can be translated from the relative to the absolute (solution) space of the controlled aircraft by transposing the origin of the conflict zone by V_2, the velocity vector of the observed aircraft. This is shown in Figure 3.2c.

Finally, mapping the minimum and maximum velocity or the speed envelope of the controlled aircraft on the translated conflict zone results in the complete SSD, as shown in Figure 3.2d.

It can be seen that the tip of the velocity vector V_1, showing the direction and the magnitude of the velocity of the controlled aircraft, is within the velocity limits of the aircraft and is directed into the gray area. If aircraft A thus continues flying with its current heading and speed and no action is taken, a LoS will occur. Figure 3.2d shows that a change in heading will in this case direct the tip of the velocity vector outside the conflict zone and resolve the conflict.

SCENARIO 2

Simulation Environment

In the following scenario, the SSD as well as a conflict are introduced in the simulation environment. Again, take your time to carefully read and execute the following steps.

1. On the PVD you see two aircraft, both on course to their respective exit waypoints and both surrounded by a circle. The circles around the projected aircraft positions are called the half-PZ circles and have a radius of 2.5 NM, hence, a LoS occurs when these circles overlap.

2. Select aircraft `QUEBEC`. After having selected the aircraft, the SSD for this aircraft appears. An example of the SSD is given in Figure 3.3. The three circles in this figure depict, ranging from largest to smallest, the aircraft's maximum velocity, current velocity and minimum velocity. The conflict zone is furthermore depicted in red, meaning that if the tip of the speed vector of the controlled aircraft is directed inside the conflict zone, a LoS will occur if both aircraft continue flying at their current heading and speed and no action is taken. The aircraft depicted in Figure 3.3 is thus *not* in conflict as the tip of the speed vector is directed *outside* the conflict zone.

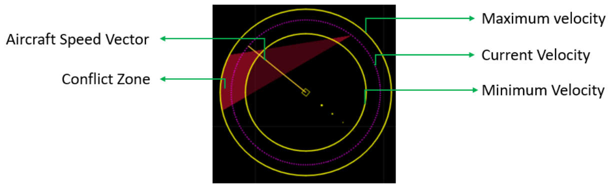


Figure 3.3: Impression of the Solution Space Diagram in the Simulation Environment

3. The SSD allows the controller to link conflict zones to observed aircraft. Hover over the conflict zone. Observe that the conflict zone as well as the (observed) aircraft that is causing this conflict zone, are highlighted in yellow when hovering over the conflict zone or the observed aircraft. Do the same for aircraft `VICTOR`: select aircraft `VICTOR`, hover over the conflict zone and over aircraft `QUEBEC` and see how the conflict zone with its corresponding aircraft are being highlighted. Also observe that the SSDs of both aircraft look different. Each aircraft thus has its own SSD.

Dynamic Conflict Detection and Resolution

Before starting the simulation, carefully read all of the following steps first.

1. Select aircraft `VICTOR`.

2. In case two or more aircraft follow trajectories that are set to cross in the near future and will thus result in a LoS, a Short-Term Conflict Alert (STCA) will be given. When you will start the simulation in step 5, you will notice that the aircraft will soon turn orange. This is an STCA indicating a time to LoS of 120 seconds. If you wait, you will notice that the aircraft turns red. This is an STCA indicating a time to LoS of 60 seconds. The red STCA is also accompanied by an alarm sound. It should be noted that above all, safety is most important and PZ violations and STCA alerts are to be avoided at all times.

   Note that the times mentioned above are different in the simulation. As the simulations run at 3 times the actual speed, the STCAs indicate a time to LoS in the simulation of 40 and 20 seconds, respectively.

3. While waiting for the STCAs, observe how the shape of the SSD conflict zone changes. After you have experienced the orange and red STCA, you can resolve the conflict by giving a heading clearance to one or both aircraft, whichever solution you think is best.

4. The horizontal separation distance of 5 NM can be determined by means of the scale in the bottom right corner of the simulation environment, which shows the scale of 10 NM. It can furthermore be determined by means of the half-PZ circles.

5. You may now put on the headphones and press the start button on the top left corner of the simulator to start the scenario. This scenario lasts 3 minutes. When the scenario is finished, you are again asked to rate the perceived difficulty of the scenario.

6. Click on *Next Scenario* and continue with the next steps.

Questions

To test your knowledge and understanding so far, please indicate for Figure 3.4, 3.5 and 3.6 whether the aircraft depicted in the figure are in conflict or not. For each question, tell your answers to the researcher and explain your reasoning. You will thus not need the answer sheet yet.

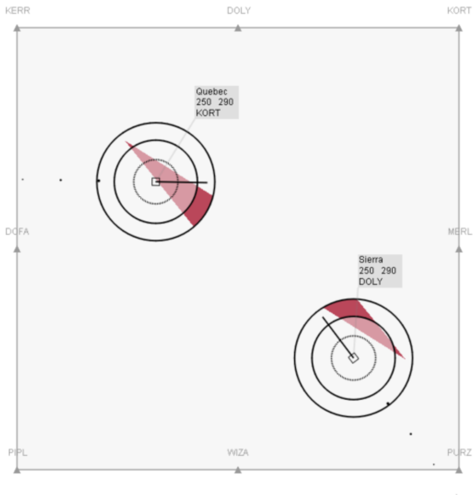


Figure 3.4: Question 3-1

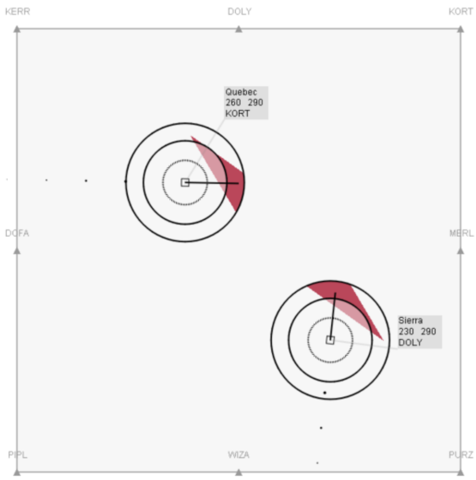


Figure 3.5: Question 3-2

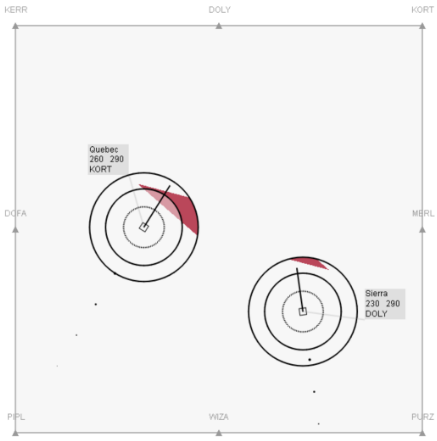


Figure 3.6: Question 3-3

3.2 Other Information Presented by the Solution Space Diagram

Next to simply detecting conflicts and showing the solution space to the controller, the SSD allows the controller to link conflict zones to observed aircraft by looking at the shape and orientation of the conflict zones. By using this information, the controller is able to roughly determine the flight direction and proximity of the neighboring aircraft with respect to the controlled aircraft.

Direction of Conflict Triangle

The base of the triangle, the side opposite of the origin of the triangle, is aimed, at a slight offset, at the aircraft that is causing the conflict. Figure 3.7 shows an SSD with two conflict zones, both aimed at a different aircraft.

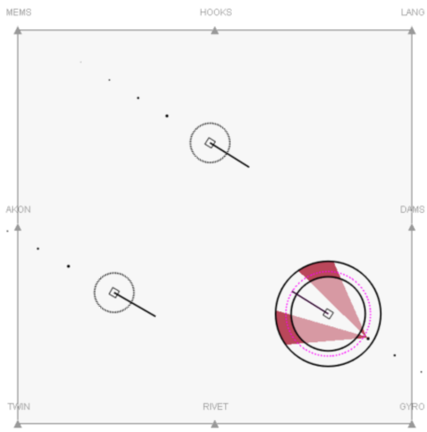


Figure 3.7: Conflict Triangle Direction

Width of Conflict Triangle

The width of the triangle indicates the proximity of the neighboring aircraft: a small width indicates a far-away aircraft whereas a triangle with a larger width indicates a nearby aircraft. Figure 3.8 shows an SSD with two conflict zones, both having a different width. Keeping in mind that the base of the triangle is aimed at the aircraft causing this conflict, it can be seen that the conflict triangle with a small width indicates a far-away aircraft and the conflict triangle with a larger width indicates a nearby aircraft.

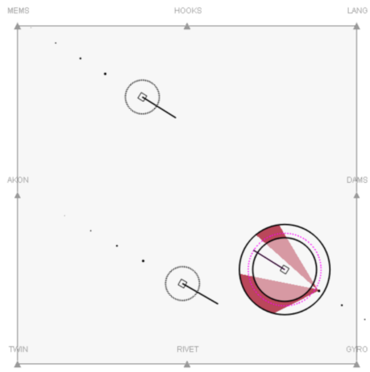
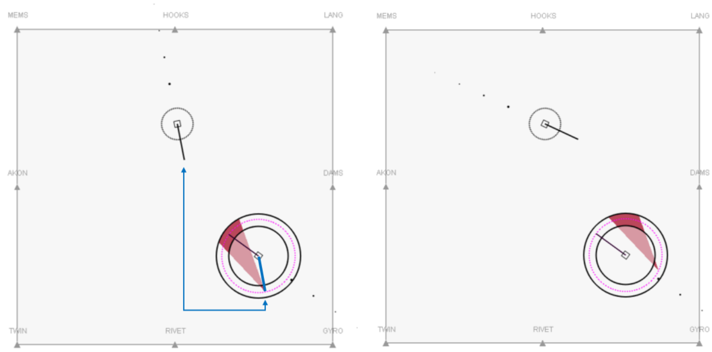


Figure 3.8: Conflict Triangle Width

Location of Conflict Triangle Origin

By drawing a line from the controlled aircraft towards the origin of the conflict zone triangle in the SSD, the absolute speed vector of the observed aircraft can be determined. Note that the origin of the triangle can also be placed outside of the SSD area and may thus sometimes not be visible. Figure 3.9 shows two different SSDs. Observe how the direction of the speed vector of the observed aircraft influences the location of the origin of the conflict zone in the SSD of the controlled aircraft.



(a) Conflict Triangle Origin Location 1 (b) Conflict Triangle Origin Location 2

Figure 3.9: Conflict Triangle Origin Location Based on Direction of Speed Vector

Solution Direction

Next to giving the controller information about the location and proximity of neighboring aircraft, the SSD provides the controller with information such as which solution direction is favorable for solving conflicts. Figure 3.10 shows that aircraft `QUEBEC` can best be steered right and pass behind aircraft `SIERRA` and aircraft `SIERRA` can best be steered right to pass in front of aircraft `QUEBEC`, as the solution space on the right side of both aircraft is larger. Again notice that the aircraft have different SSDs.

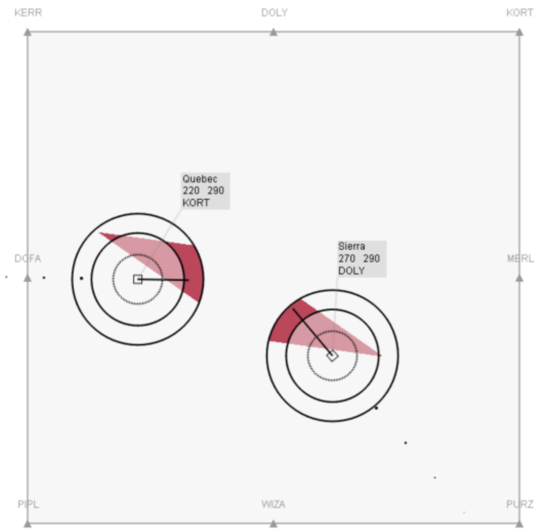


Figure 3.10: Conflict Triangle Indicates Whether to Pass in Front or Behind an Observed Aircraft

Questions

To test your knowledge and understanding of the SSD and the information it provides you with, please answer the following questions. Tell your answer to the researcher and explain your reasoning. You will again not need the answer sheet yet.

1. Which SSD belongs to red aircraft ROMEO in this scenario?

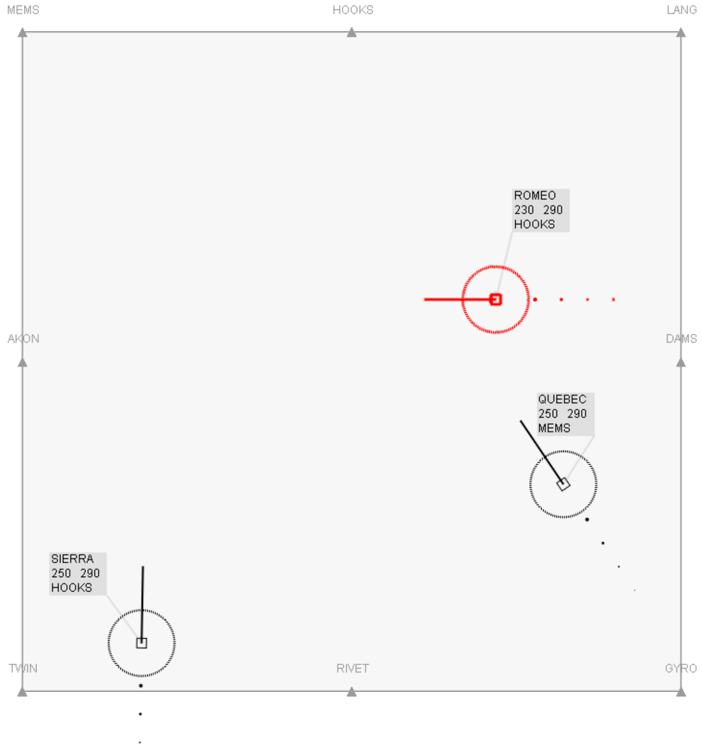
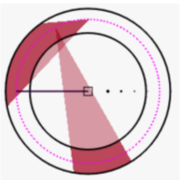
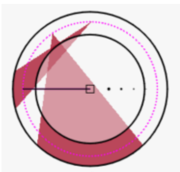


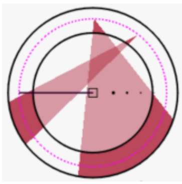
Figure 3.11: Question 3-4



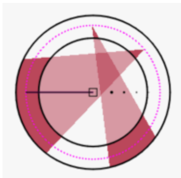
(a) SSD 1



(b) SSD 2



(c) SSD 3



(d) SSD 4

2. Which scenario corresponds to the SSD for aircraft `KILO` (circled in red), shown in Figure 3.13?

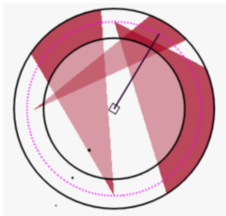
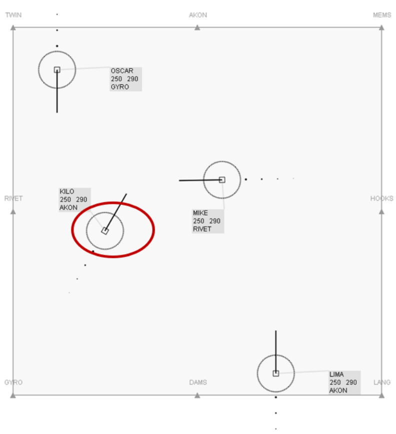
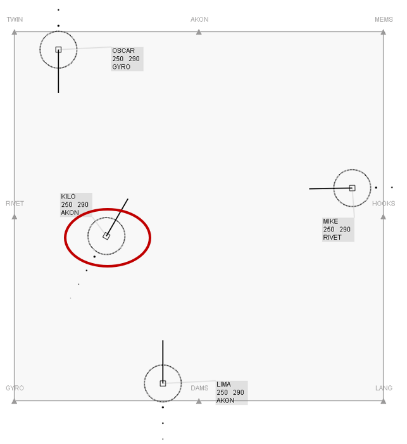


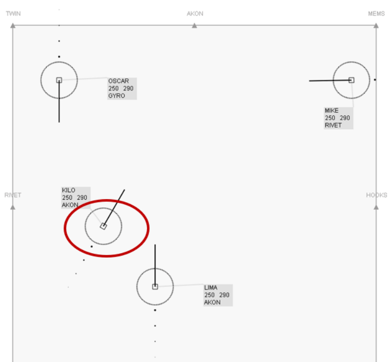
Figure 3.13: Question 3-5



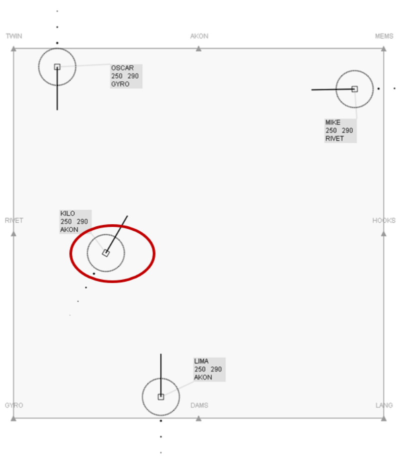
(a) Scenario 1



(b) Scenario 2



(c) Scenario 3



(d) Scenario 4

4 Think-Aloud Protocol

In the previous scenarios, you have been shown the tools and features that are available to you in the experiment simulator during (part of) the experiment. The scenarios that follow are intended as further practice to make you feel comfortable with the simulation environment while performing your task in this experiment.

During the ATC training, trainees are required to follow a course on Radio Telephony (RT). Standard phraseologies are taught to the trainees in order for efficient, clear, concise and unambiguous communications to take place between ATCOs and pilots. To resemble the RT and in order for the researcher to gain an insight in your thoughts and strategies during the experiment, you are asked to follow a think-aloud protocol during the rest of this experiment. This chapter presents the think-aloud protocol and is accompanied by three scenarios to practice the thinking out loud.

For the think-aloud protocol it is important that you only talk about the experiment and what you see, do, think and experience during the experiment runs. More specifically, you are asked to:

- Name the aircraft you want to give a command to by its call sign: e.g., `ALPHA`;
- In case of conflict, name the aircraft involved in the conflict: e.g., `Conflict ALPHA TANGO`;
- Name the chosen solution as well as the aircraft used to resolve the conflict: e.g., `TANGO Heading 240 (two-forty)`;
- Name the reason you want to give a certain command: e.g., `TANGO Heading 240 to resolve conflict; Direct To for sending ALPHA to COZA`;
- Name any other information / observations regarding the chosen solutions and scenarios that come to mind during the scenarios, e.g.,:
  - `More room to send TANGO left/right to resolve conflict, thus Heading 240;`
  - `More room to send TANGO in front/behind to resolve conflict, thus Heading 240;`
  - `Heading 240 to minimize additional track miles / results in less additional track miles;`
  - `Heading 240 to minimize monitoring time / results in less monitoring time;`
  - Etc.

SCENARIOS 3 - 5

1. In the three scenarios that follow, you are free to manipulate the trajectories of the aircraft to practice using the think-aloud protocol and giving HDG and DCT commands.
2. The SSD, that was introduced in the previous chapter, will be available during these and the following scenarios.
3. Similar to the two previous scenarios, you will continue with the next scenario after rating the difficulty of the scenarios and clicking on *Next scenario*. For each scenario, press the play button in the top left corner of the screen to start the scenario.
4. Remember that from now on, it is expected that you adhere to the think-aloud protocol. Mention everything that you think is relevant for the scenarios and the experiment and really try to reason out loud.
5. You may now press the play button in the top left corner of the screen to start the scenario. Each of the three scenarios lasts 2,5 minutes. Once you reach scenario 6, continue to the next chapter.

5 Control Task

Up until now you have been familiarized with the tools and features that are available to you in the experiment simulator during (part of) the experiment and the think aloud protocol. This chapter is used to familiarize you with the control task for the rest of the experiment.

5.1 Introduction to the Control Task

Before continuing to the explanation of the control task, you are asked to answer questions 1 - 3. For each of the scenarios depicted in Figures 5.1 to 5.3, indicate whether a conflict is present, and if so, which aircraft you would give a command to resolve the conflict, and whether you would send this aircraft left or right (from the perspective of the aircraft). You can indicate your answers on the answer sheet. The figures can also be found on the answer sheet.

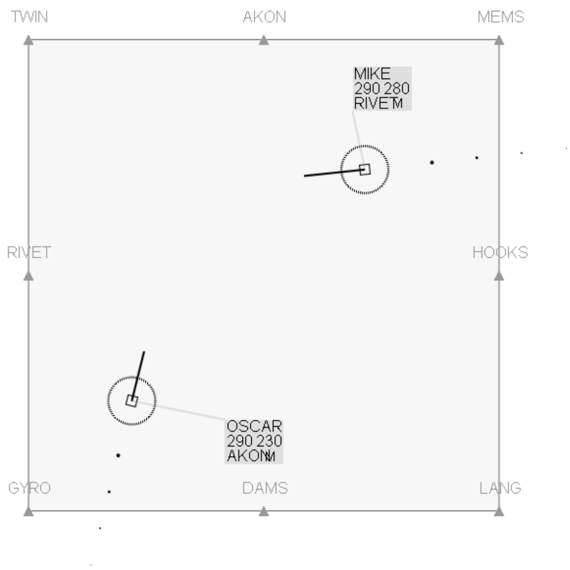


Figure 5.1: Question 1

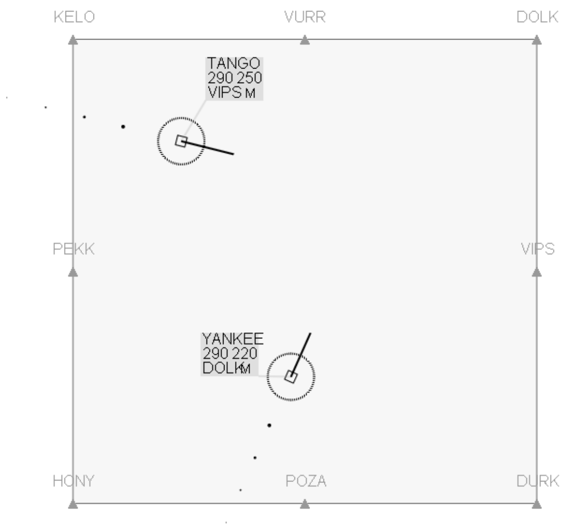


Figure 5.2: Question 2

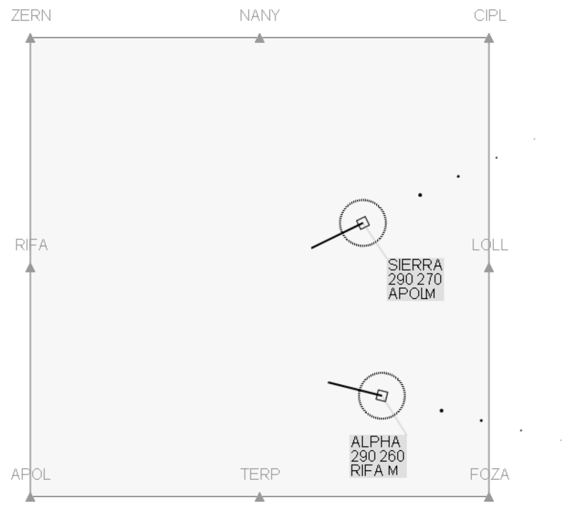


Figure 5.3: Question 3

Once you have answered questions 1 - 3, you will continue with four scenarios in the simulation environment. For each of the scenarios, first read all of the steps below. While performing these steps, observe what happens and remember to think aloud. After these scenarios, the control task is introduced.

SCENARIO 6

1. In this scenario, you see two aircraft with a different speed, that are both on track to their assigned COPX.
2. The aircraft are in conflict. Take a close look at the SSD.
3. After starting the scenario, give aircraft `ROMEO` a HDG command of **115 degrees**.
4. Observe what happens to the track and the SSD.
5. When the conflict is resolved and you think it is safe to clear aircraft `ROMEO` to its COPX, give aircraft `ROMEO` a DCT command.
6. At the end of the scenario, you will again be asked to rate the difficulty of the scenario. After rating this difficulty, click on *Next Scenario*.
7. You may now press the play button in the top left corner. Remember to think aloud according to the protocol as was described in Chapter 4. This scenario lasts 2 minutes.

SCENARIO 7

1. In this scenario, you see two aircraft with a different speed, that are both on track to their assigned COPX.
2. The aircraft are in conflict. Take a close look at the SSD.
3. When aircraft `KILO` enters the sector, give this aircraft a HDG command of **310 degrees**.
4. Observe what happens to the track and to the SSD.
5. When the conflict is resolved and you think it is safe to clear aircraft `KILO` to its COPX, give aircraft `KILO` a DCT command.
6. At the end of the scenario, you will again be asked to rate the difficulty of the scenario. After rating this difficulty, click on *Next Scenario*.
7. You may now press the play button in the top left corner. Remember to think aloud according to the protocol as was described in Chapter 4. This scenario lasts 2.5 minutes.

SCENARIO 8

1. In this scenario, you see two aircraft with a different speed, that are both on track to their assigned COPX.
2. The aircraft are in conflict. Take a close look at the SSD.
3. When aircraft `DELTA` enters the sector, give this aircraft a HDG command of **60 degrees**.
4. Observe what happens to the track and the SSD.
5. When the conflict is resolved and you think it is safe to clear aircraft `DELTA` to its COPX, give aircraft `DELTA` a DCT command.

6. At the end of the scenario, you will again be asked to rate the difficulty of that scenario. After rating the difficulty, click on *Next Scenario*.

7. You may now press the play button in the top left corner. Remember to think aloud according to the protocol as was described in Chapter 4. This scenario lasts 2 minutes.

SCENARIO 9

1. In this scenario, you see three aircraft with different speeds that are on track to their assigned COPX.

2. Two of the aircraft are in conflict. Take a close look at the SSD.

3. Give aircraft `SIERRA` a HDG change to resolve the conflict.

4. Observe what happens to the track and the SSD.

5. When the conflict is resolved and you think it is safe to clear aircraft `SIERRA` to its COPX, give aircraft `SIERRA` a DCT command.

6. At the end of the scenario, you will again be asked to rate the difficulty of the scenario. After rating the difficulty, click on *Next Scenario*. The simulation will be closed after this scenario.

7. You may now press the play button in the top left corner. Remember to think aloud according to the protocol as was described in Chapter 4. This scenario lasts 2 minutes.

5.2 Control Task

As you may have guessed, the control task for this experiment is a CD&R task. Similar to the mission statement of ATC, which is to *safely* and *efficiently* organize and expedite the flow of air traffic from origin to destination, you are given the following two tasks for each scenario:

1. Separation assurance: ensuring the minimum horizontal distance of 5NM between neighboring aircraft at all times to prevent losses of separation. STCAs (orange and red) are thus to be avoided at all times and should be resolved as soon as possible in case they do occur.

2. Clearing aircraft to their COPX in the most efficient way possible, without introducing conflicts or PZ violations.

Furthermore, remember that during each scenario you are asked to communicate your thoughts and strategies according to the think-aloud protocol as was presented in Chapter 4.

Questions

Before continuing to the next section, indicate for the scenarios depicted in Figures 5.4 to 5.6 (questions 4 - 6) whether a conflict is present, and if so, which aircraft you would give a command to resolve the conflict, and whether you would send this aircraft left or right (from the perspective of the aircraft). You can indicate your answers on the answer sheet. The figures can again also be found on the answer sheet.

Finally, answer questions 7 and 8 on the answer sheet about the strategies you used to detect and resolve conflicts in Questions 1 - 6.

When you have answered questions 4 - 8, it is time for a break.

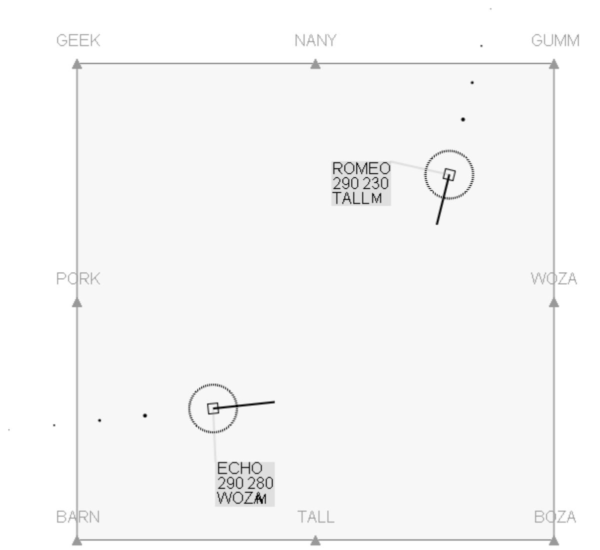


Figure 5.4: Question 4

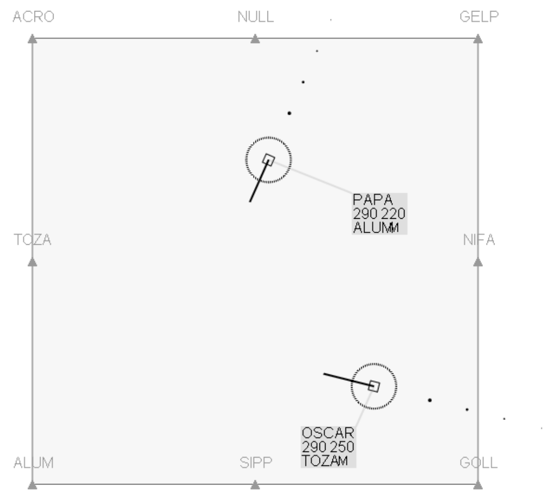


Figure 5.5: Question 5

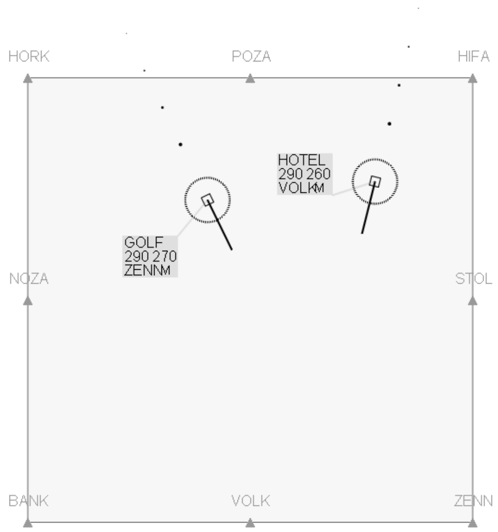


Figure 5.6: Question 6

5.3 Training Phase

In the previous section, you have been introduced to the control task. The eleven scenarios that follow are intended to further increase your experience, and to make you feel comfortable with performing the control task during the rest of the experiment.

SCENARIOS 10-18

1. In the scenarios that follow, you are asked to adhere to the control task as much as possible to increase your experience with the task as much as possible before starting the test in the next chapter.

   Control Task: ensuring the minimum separation distance between neighboring aircraft while efficiently clearing aircraft to their respective COPX and while adhering to the think-aloud protocol.

2. Keep in mind that, above all, safety is most important and PZ violations and STCAs are to be avoided at all times.

3. For this round of scenarios, you will automatically continue with the next scenario after rating the difficulty of a scenario and clicking on *Next scenario*. You will thus not need to press the play button for each of these scenarios. The scenarios last 2 minutes.

4. You may now put on the headphones, press *Start experiment* and perform each of the eleven scenarios. The first scenario will immediately start after pressing this button. When you have performed all scenarios, the simulation will automatically close.

5. When you have performed all scenarios, you are asked to answer questions 9 and 10 on the answer sheet: describe your strategy for detecting and resolving conflicts during the Training Phase scenarios. After describing your strategies, you can continue with the questions in the next section.

Questions

You are asked to answer questions 11 - 13 before continuing to the next chapter. For Figures 5.7 to 5.9, again indicate whether a conflict is present, and if so, which aircraft you would give a command to resolve the conflict, and whether you would send this aircraft left or right (from the perspective of the aircraft). You can indicate your answers on the answer sheet. The figures can again also be found on the answer sheet.

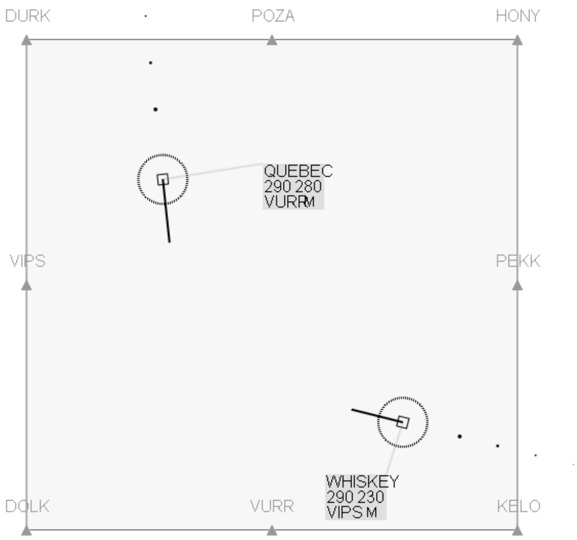


Figure 5.7: Question 11

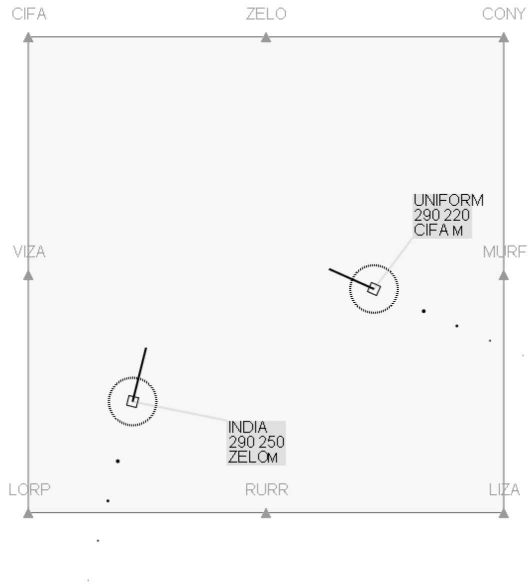


Figure 5.8: Question 12

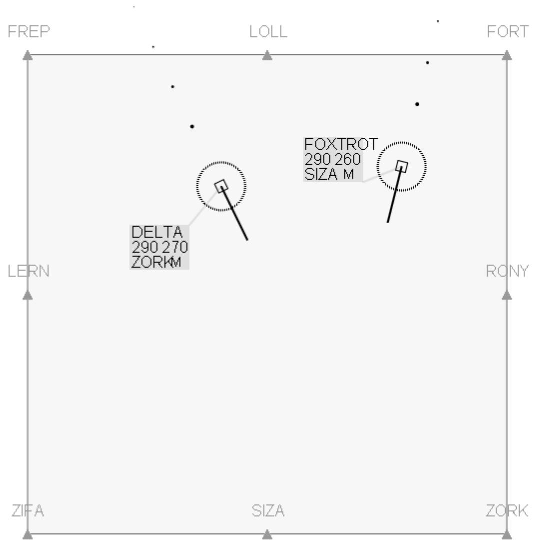


Figure 5.9: Question 13

6 Test Phase

You have now completed the Training Phase of the experiment. As was also explained in the experiment briefing, the rest of this session consists of the Test Phase.

SCENARIOS 19 - 26

1. The eleven scenarios that you will perform during the Test Phase will be similar in length, structure and difficulty to the set of scenarios that you performed in the Training Phase. The control task will be exactly the same.

   Control Task: ensuring the minimum separation distance between neighboring aircraft while efficiently clearing aircraft to their respective COPX and while adhering to the think-aloud protocol.

2. Keep in mind that, above all, safety is most important and PZ violations and STCAs are to be avoided at all times.

3. During the Test Phase, the SSD will *not* be available and you will thus have to perform the control task without any form of support.

4. For this set of scenarios, you are again asked to rate your perceived difficulty after each of the scenarios and you will automatically continue to the next scenario after having indicated this difficulty. You will thus not need to press the play button for each of these scenarios.

5. You may press *Start experiment* when you are ready and perform each of the eleven scenarios. The first scenario will immediately start after pressing this button. Again, the simulation will automatically close after the final scenario. The scenarios last 2 minutes.

6. When you have performed all scenarios, you are asked to answer questions 14 and 15 on the answer sheet: describe your strategy for detecting and resolving conflicts in the Test Phase scenarios. After describing your strategies, you can continue with the questions in the next section.

Questions

You are asked to answer questions 16 - 24. For Figures 6.1 to 6.9, indicate whether a conflict is present, and if so, which aircraft you would give a command to resolve the conflict, and whether you would send this aircraft left or right (from the perspective of the aircraft). You can indicate your answers on the answer sheet. The figures can again also be found on the answer sheet.

After the conflict questions, answer questions 25 and 26: describe the strategies you used to detect and resolve conflicts in questions 16 to 24. Finally, continue with the remaining questions on the answer sheet.

When you have answered all questions, the experiment is finished. You can hand in the answer sheet and the researcher will ask you some final questions.

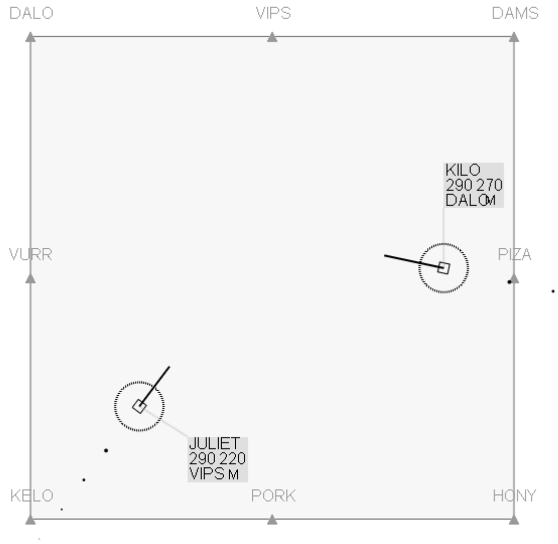


Figure 6.1: Question 16

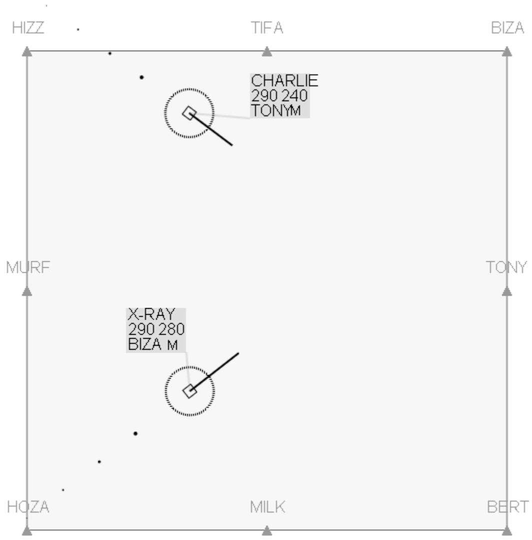


Figure 6.2: Question 17

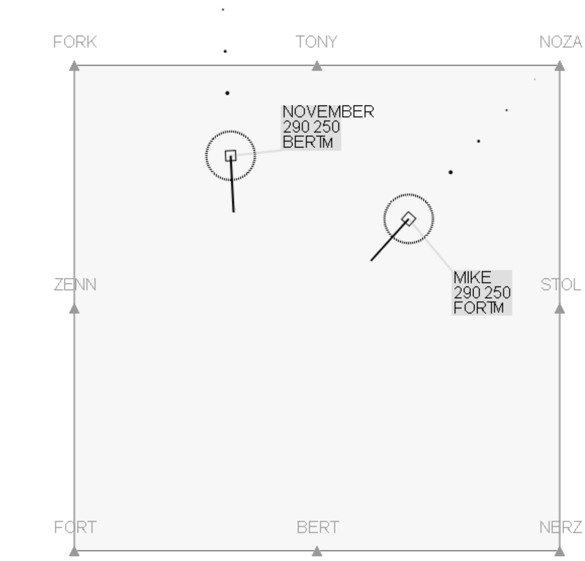


Figure 6.3: Question 18

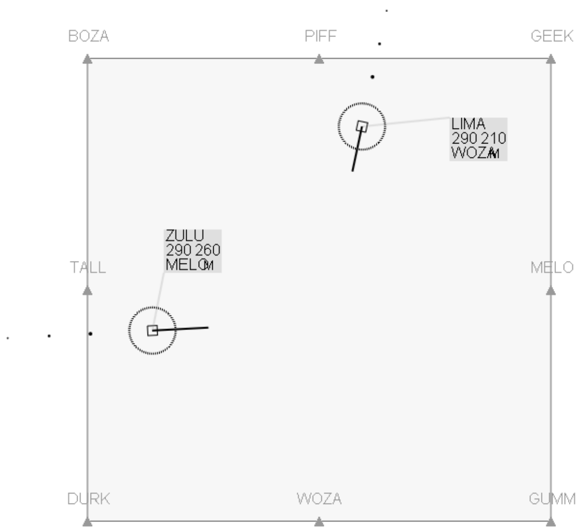


Figure 6.4: Question 19

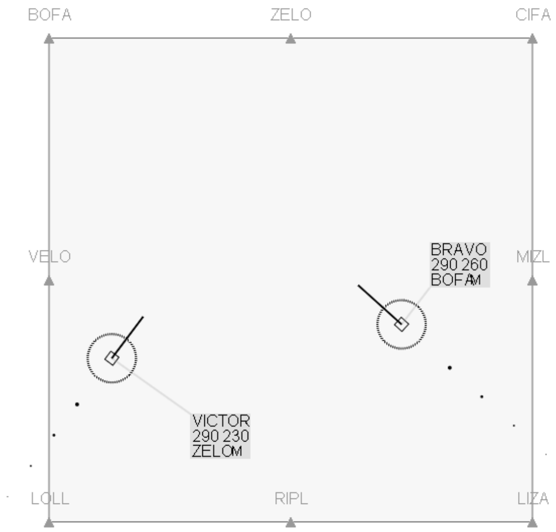


Figure 6.5: Question 20

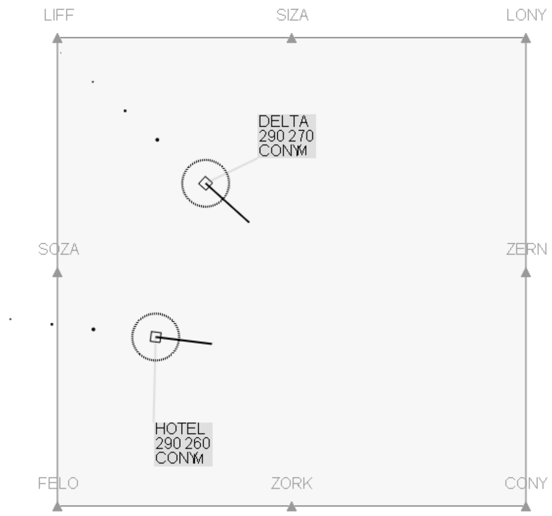


Figure 6.6: Question 21

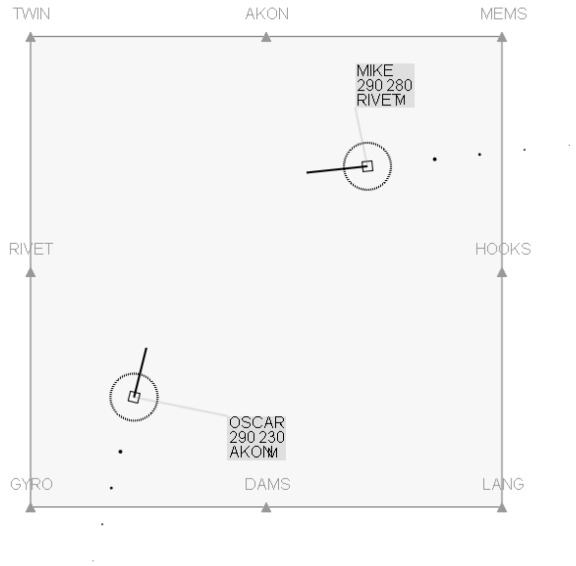


Figure 6.7: Question 22

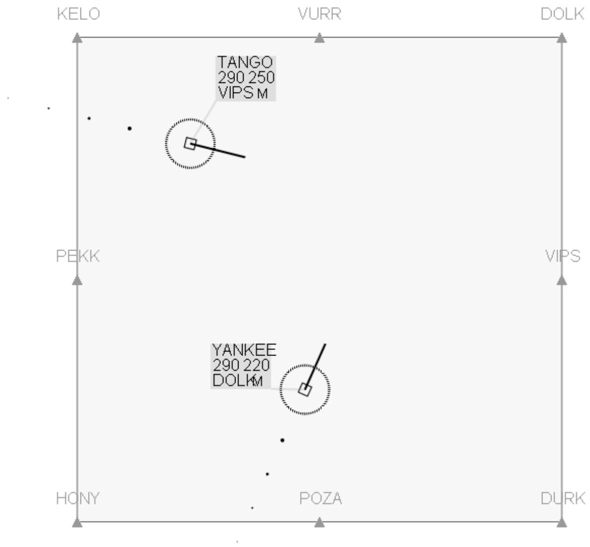


Figure 6.8: Question 23

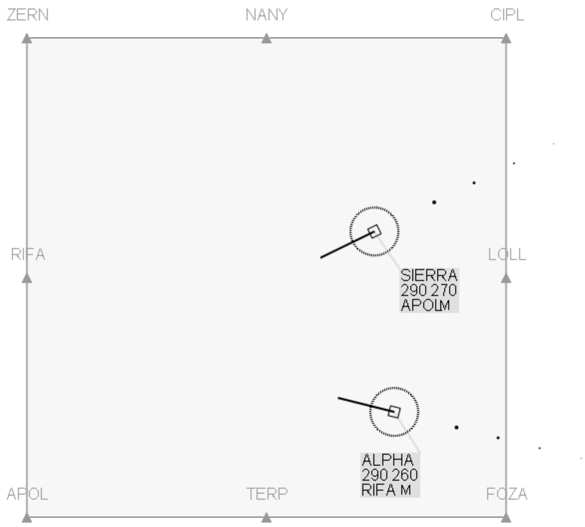


Figure 6.9: Question 24

Appendix N

Training Script CD&R Chapter - Control Group

Appendix M presented the training script that was developed to guide participants from the SSD group through the experiment. As explained in Part I, the CD&R chapter of the script was different for both groups due to the (un)availability of the SSD. As the rest of the script was the same for both groups, only the CD&R chapter for the control group is presented here to allow for comparison.

3 Conflict Detection & Resolution

As the mission statement of ATC is "to safely and efficiently organize and expedite the flow of air traffic from origin to destination and provide information and other support to pilots", your main task in the experiment will also be to manage traffic as safely as possible while also keeping efficiency in mind.

Due to several uncertainties in, for example, weather, turbulence effects and determining the location of an aircraft with radar systems, aircraft cannot fly in too close proximity of each other. To have aircraft fly at a safe distance from each other, a minimum separation distance has thus been defined. Aircraft should horizontally be separated by at least 5 NM and vertically by 1,000 ft. Each aircraft thus has a so-called Protected Zone (PZ) around it with a radius of 5 NM and a height of 2,000 ft. As this experiment focuses on horizontal separation only, you can disregard the vertical separation distance of 1,000 ft.

When an aircraft enters another aircraft's PZ, this is called a Loss of Separation (LoS), which is also depicted in Figure 3.1. A conflict arises when an actual LoS occurs but also when a potential LoS within a predetermined time window presents itself, thus when two or more aircraft follow trajectories that are set to cross in the near future. A conflict is thus not defined as a collision between two aircraft but rather as a (potential) LoS. This chapter is used to familiarize you with conflicts in the simulation environment.

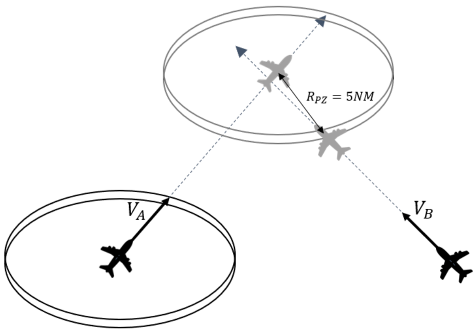


Figure 3.1: Schematic of a Conflict

SCENARIO 2

Simulation Environment

In the following scenario, a conflict is introduced in the simulation environment. Again, take your time to carefully read the following steps first before executing them.

1. On the PVD you see two aircraft, both on course to their respective exit waypoints and both surrounded by a circle. The circles around the projected aircraft positions are called the half-PZ circles and have a radius of 2.5 NM, hence, a LoS occurs when these circles overlap.

2. Select aircraft `VICTOR`.

3. In case two or more aircraft follow trajectories that are set to cross in the near future and will thus result in a LoS, a Short-Term Conflict Alert (STCA) will be given. When you will start the simulation in step 6, you will notice that the aircraft will soon turn orange. This is an STCA indicating a time to LoS of 120 seconds. If you wait, you will notice that the aircraft turns red. This is an STCA indicating a time to LoS of 60 seconds. The red STCA is also accompanied by an alarm sound. It should be noted that above all, safety is most important and PZ violations and STCA alerts are to be avoided at all times.

   Note that the times mentioned above are different in the simulation. As the simulations run at 3 times the actual speed, the STCAs indicate a time to LoS in the simulation of 40 and 20 seconds, respectively.

4. After waiting for the alerts and having experienced the orange and red STCA, you can resolve the conflict by giving a heading clearance to one or both aircraft, whichever solution you think is best.

5. The horizontal separation distance of 5 NM can be determined by means of the scale in the bottom right corner of the simulation environment, which shows the scale of 10 NM. It can furthermore be determined by means of the half-PZ circles.

6. You may now put on the headphones and press the start button on the top left corner of the simulator to start the scenario. This scenario lasts 3 minutes. When the scenario is finished, you are again asked to rate the perceived difficulty of the scenario.

7. Click on *Next Scenario* and continue with the next chapter.

Appendix O

Experiment Answer Sheet and Results

During the experiment, participants were asked several questions, of which they had to write down the answers on an answer sheet. The answer sheet is presented on the following pages, after which the answers (to questions 10, 15 and 27) of the participants are presented.

Participant number

Experiment Answer Sheet

Chapter 5

Question	Answer		
1	Conflict / No Conflict	Aircraft:	Left / Right
2	Conflict / No Conflict	Aircraft:	Left / Right
3	Conflict / No Conflict	Aircraft:	Left / Right
4	Conflict / No Conflict	Aircraft:	Left / Right
5	Conflict / No Conflict	Aircraft:	Left / Right
6	Conflict / No Conflict	Aircraft:	Left / Right

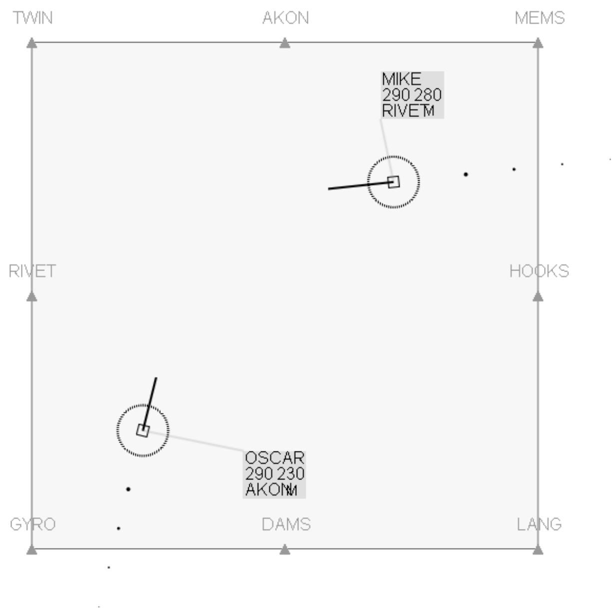


Figure 1: Question 1

1/12

Participant number

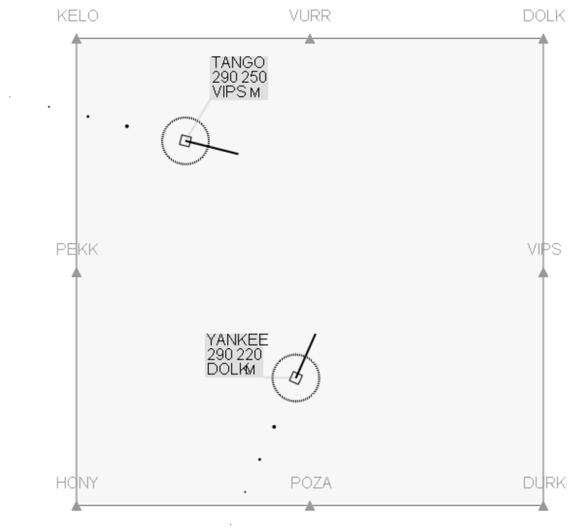


Figure 2: Question 2

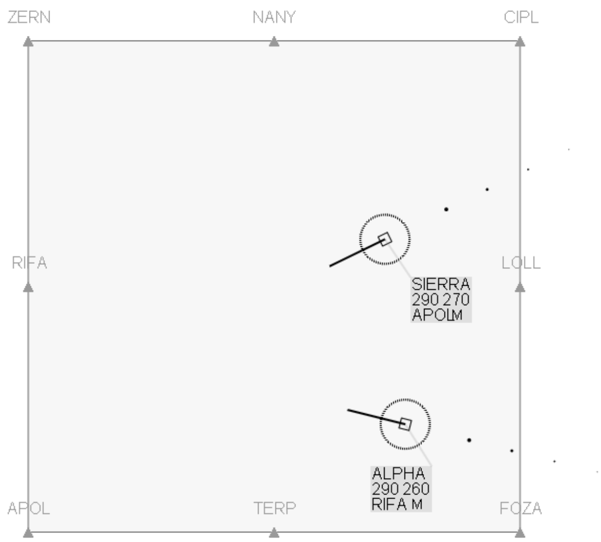


Figure 3: Question 3

Participant number

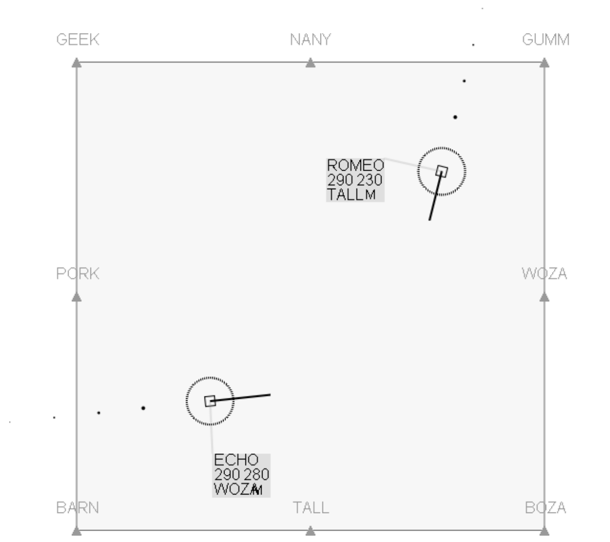


Figure 4: Question 4

Figure 5: Question 5

3/12

Participant number

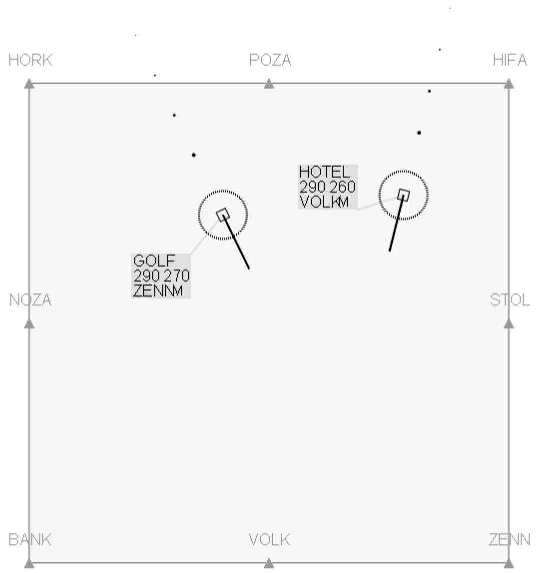


Figure 6: Question 6

Question 7: Describe your strategy for detecting conflicts in Questions 1 - 6:

Question 8: Describe your strategy for resolving conflicts in Questions 1-6 in case these were present:

4/12

Participant number

Question 9: Describe your strategy for detecting conflicts in the Training Phase scenarios:

Question 10: Describe your strategy for resolving conflicts in the Training Phase scenarios in case these were present:

Question	Answer		
11	Conflict / No Conflict	Aircraft:	Left / Right
12	Conflict / No Conflict	Aircraft:	Left / Right
13	Conflict / No Conflict	Aircraft:	Left / Right

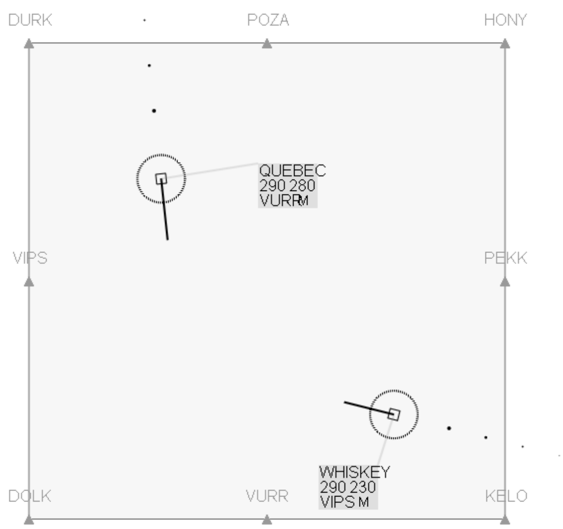


Figure 7: Question 11

Participant number

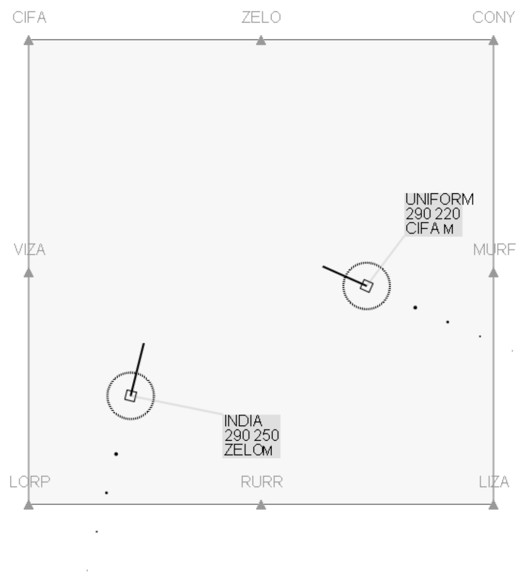


Figure 8: Question 12

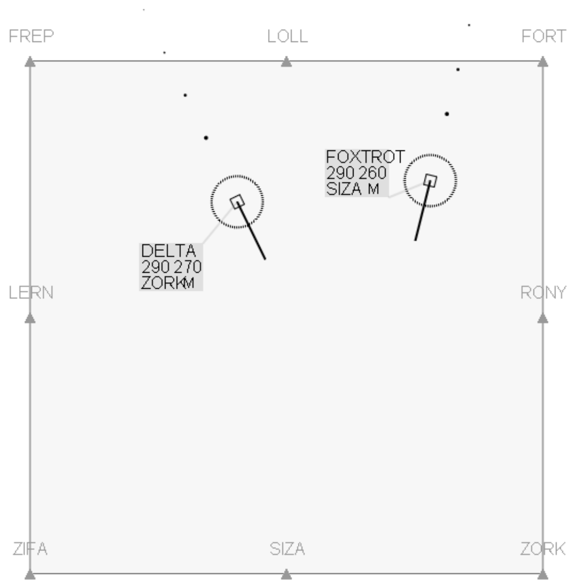


Figure 9: Question 13

6/12

Participant number

Chapter 6

Question 14: Describe your strategy for detecting conflicts in the Test Phase Scenarios:

Question 15: Describe your strategy for resolving conflicts in the Test Phase Scenarios in case these were present:

Question	Answer		
16	Conflict / No Conflict	Aircraft:	Left / Right
17	Conflict / No Conflict	Aircraft:	Left / Right
18	Conflict / No Conflict	Aircraft:	Left / Right
19	Conflict / No Conflict	Aircraft:	Left / Right
20	Conflict / No Conflict	Aircraft:	Left / Right
21	Conflict / No Conflict	Aircraft:	Left / Right
22	Conflict / No Conflict	Aircraft:	Left / Right
23	Conflict / No Conflict	Aircraft:	Left / Right
24	Conflict / No Conflict	Aircraft:	Left / Right

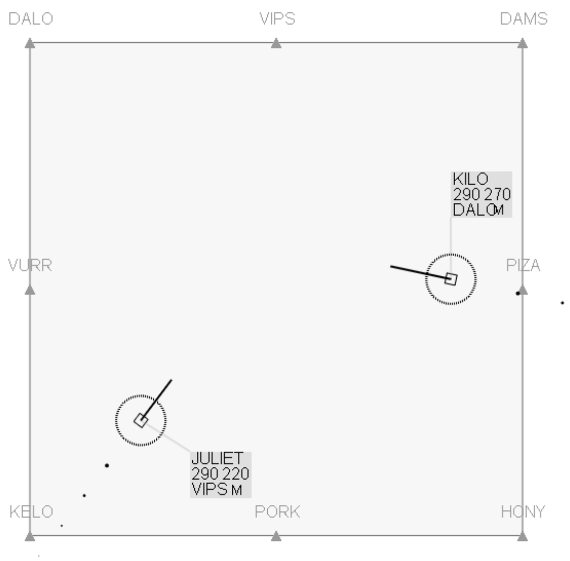


Figure 10: Question 16

7/12

Participant number

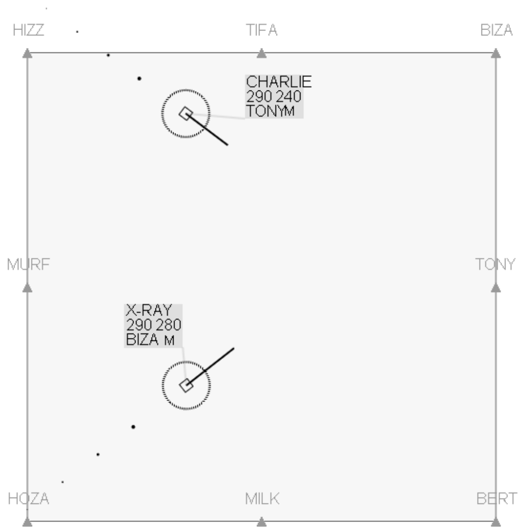


Figure 11: Question 17

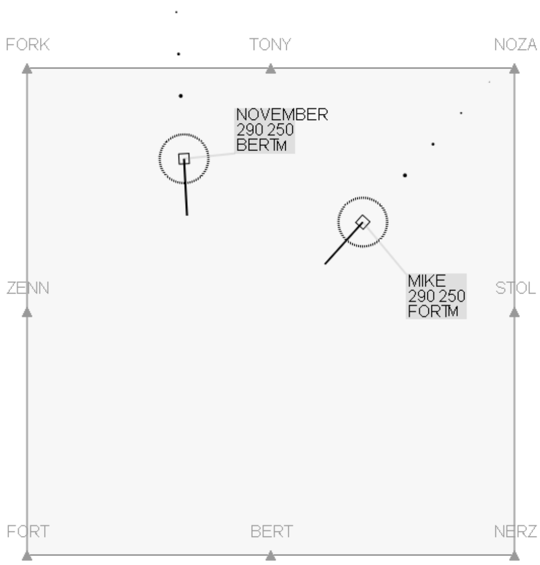


Figure 12: Question 18

8/12

Participant number

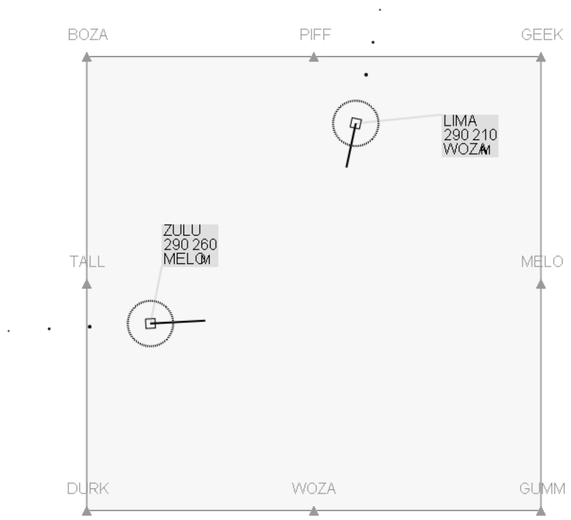


Figure 13: Question 19

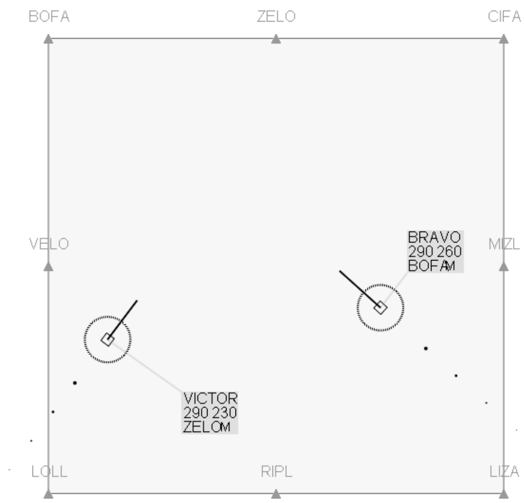


Figure 14: Question 20

9/12

Participant number

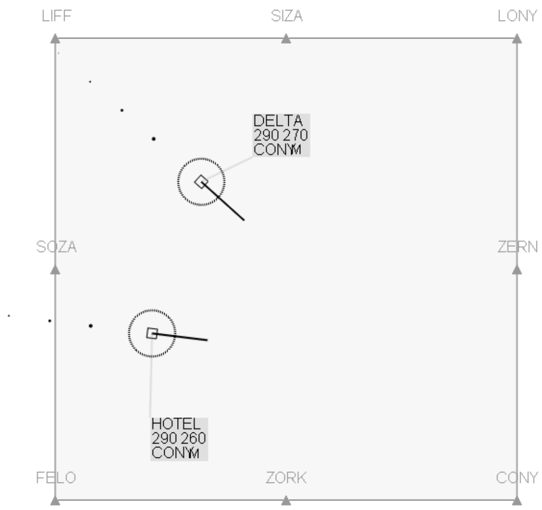


Figure 15: Question 21

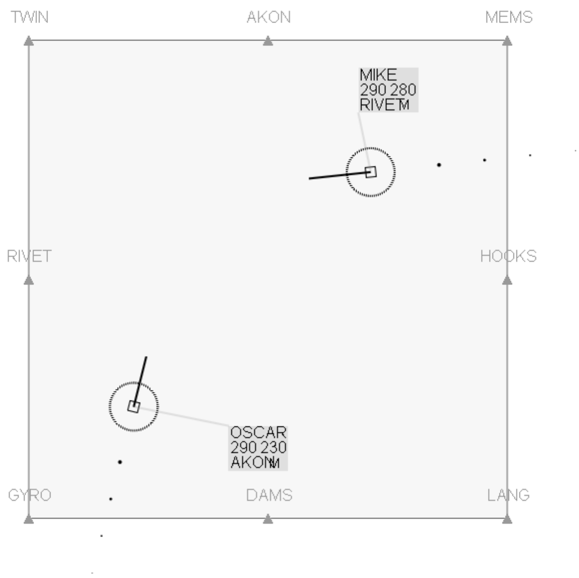


Figure 16: Question 22

10/12

Participant number

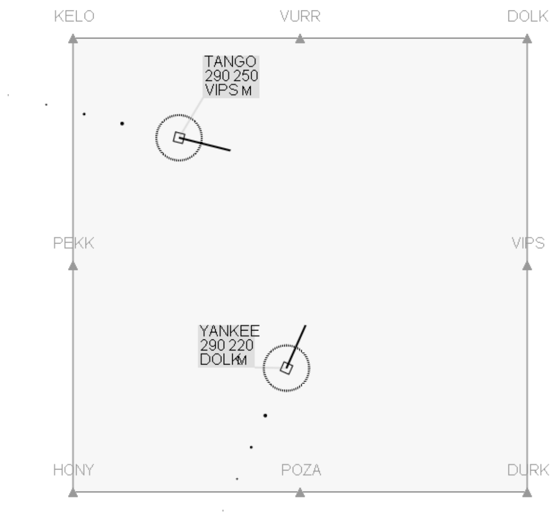


Figure 17: Question 23

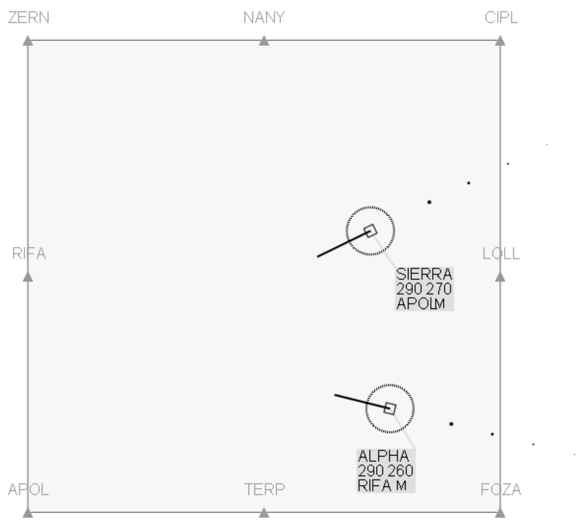


Figure 18: Question 24

11/12

Participant number

Question 25: Describe your strategy for detecting conflicts in Questions 16 - 24:

Question 26: Describe your strategy for resolving conflicts in Questions 16 – 24 in case these were present:

Question 27: What is in your opinion the main rule of thumb you learned during this experiment to resolve conflicts?

Question 28: Did you understand everything that was explained to you in the manual?

Question 29: Did you find parts of the experiment easy/difficult? If so, which parts and why?

Question 30: Did you feel bored at any point in the experiment? If so, during which part(s) and why?

Question 31: Is there anything else you would like to mention / state about the experiment and/or your performance?

12/12

Table O-1: Participant's Answers from the **SSD Group** regarding their Strategies in the Experiment

	Describe your strategy for resolving conflicts in the Training/Test Phase scenarios	
Participant	**Training Phase - Question 10**	**Test Phase - Question 15**
1	Check which aircraft are in conflict, then decide which aircraft changes heading (usually the one ith the minimum heading change)	Execute a heading change for an aircraft with lower speed. The heading change is sufficiently large to avoid conflicts, then apply DCT to all aircraft minimize miles.
2	First of all, I looked for the slower plane and if it was possible to change its heading. If the required 'delta heading' was big, I would go the fastest aircraft and I would try to change its heading. If also this 'delta heading' was still big, I would come back to the slower plane and I would change its heading.	I gave a heading to command the slower plane in all scenarios after giving a DCT command to the faster plane.
3	Act fast, just overshoot initially. Don't take too long deciding the best option. Moving the fast one is very effective and resolves issues quick	Overshooting a lot, get the fast one out the way and moving early
4	In case of symmetric problem, prefer to send one aircraft behind the other. Else, find the smallest comamnd in any direction that resolves the problem.	Same as before
5	First make sure there is no conflict anymore. Then let the faster aircraft go in front and look at the waypoints during a turn.	Most of the times trying to get the faster aircraft in front, also keeping in mind WP goals
6	Looking at triangles and decice the . . . action on these, also looking at velocity of aircraft and possibly 'give the right'.	Give the right when possible
7	Move the aircraft that has the smallest angle to go to larger non-red area	Similar to before but estimating the angles that are possible is harder so a larger safety margin in angles (before: move the aircraft that has the smallest angle change to go to the larger non-red area
8	Try to minimize commands to as few aircraft as possible. Use SSD to find out when to give commands	Try to let most aircraft fly directly to their COPX, and only alter aircraft with minimum commands such that they never fly in the opposite direction of COPX
9	Making the slow aircraft go behind faster planes. First make sure no conflict is emerging, then check where the planes actually need to go before sending them there. Wait until they pass behind other planes to give them their DCT.	Again, giving the fastest plane right of way, make others go behind. Getting one plane out of the way, and dealing with it later
10	Look at the shortest possible heading change or if that is not possible for getting to the waypoint, choose a heading that makes it easier to get there.	Smallest heading change or fastest route to waypoint
11	Try to steer slower aircraft behind faster aircraft. Try not to deviate from the shortest track too much.	Send one aircraft (generally the slower aircraft) to cross behind the faster aircraft
12	Redirect the slower aircraft behind the faster, this was usually the smallest HDG change. In one scenario speed was almost equal and in that case it was quicker to redirect the faster aircraft.	Identify the slowest aircraft and see if its reasonable to redirect it behind the other
13	Take decisive action, do not forget about 3rd aircraft. The faster you act, the faster everything is resolved.	Aircraft that can deviate the least from waypoint: radial heading (360, 45, 90,. . .), then send all aircraft to waypoints as it becomes visible that you can
14	1. First adjust only one trajectory. 2. Send other plane to exit point. 3. Adjust other plane to send finally to its exit point.	Adjust one plane, the one that seemed to have more flexibility to arrive at exit point (for example: longer trajectory/more distance to that point)

Table O-2: Participant's Answers from the **Control Group** regarding their Strategies in the Experiment

	Describe your strategy for resolving conflicts in the Training/Test Phase scenarios	
Participant	**Training Phase - Question 10**	**Test Phase - Question 15**
1	The faster aircraft again was given priority, however only by sending it straight to the COPX. The conflict, I resolved by changing the HDG of the slower aircraft	Also same, I did get a bit more 'risky' in my strategy as I learned that aircraft can go quite close behind each other. I did already insert a HDG (not EXQ) to resolve the possible new conflict, just in case!
2	Put faster aircraft behind slower aircraft. With 3 aircraft I tried to redirect one aircraft more than the other two.	Faster aircraft most of the times behind slower aircraft depending on trajectory. With 3 aircraft, I tried to change the heading of 1 aircraft drastically and the headings of the other two aircraft as little as possible.
3	As less time was present before conflict situations, the faster aircraft was chosen to divert faster. In these aircraft situations, the plane which is slowest was diverted	Divert the slower aircraft for minimum deviation. If the faster aircraft is further away and requires smaller change in heading, that option was chosen
4	Estimate which aircraft requires the smallest heading change to resolve the conflict. If there is no clear distinction between the required heading change for both aircraft, diver the aircraft which has the most available diversion space.	Find the solution which requires the smallest deviation from the desired flight path of all aircraft
5	Pretty much same as before, although sometimes I failed at giving commands correctly, needed quick action then. Plus I tend to send the slower airplane behind the quicker one, because slower one has tighter radius.	Like before but also, let quick aircraft go in front, let the least action needed airplane go to destination directly so not to worry about it more
6	Slowest aircraft around other aircraft(s) untill all conflicts have been resolved	Still the same: Looking which is the slowest (independent of amount of aircraft) and sending that one behind other aircraft
7	Change heading of aircraft in which the heading change is most in the direction of COPX	With multiple aircraft, speeds and distances to target get more important, often letting the fastest plane go in front
8	The aircraft with the lowest speed was often chosen. Furthermore, for these aircraft, the aircraft was chosen so the other two could (almost) directly to their exit points, so only one aircraft makes extra miles. Most of the times one can pass at the rear quite easily	Slowest aircraft is easiest to adjust heading. Most of the times, slower aircraft could pass at the rear of faster aircraft without too much deviation. Rather 1 aircraft deviating from heading than two (or three). Try to avoid flight paths crossing
9	Giving space to aircraft with lower velocity by sending higher velocity aircraft behind the one with a lower velocity (although sometimes it seemed like there was no velocity difference present).	I tried to always redirect the fastest aircraft behind the path of the slowest aircraft when a potential conflict occurred. I also took into account whether a redirection of the path brought the aircraft closer to the COPX or not
10	If i had enough time to think i would choose the path to lose minimum distance, so the plane would still get nearer. Otherwise the first option I could see.	Determining which aircraft would have least addition of distance to prevent collision
11	Depending on how far the aircraft is from the exit waypoint a command is given to aircraft that is far away from the exiting waypoint	Same as training
12	Avoid most times behind as the aircraft will sooner diverge, thus decreasing additional traveled miles. Try to resolve the conflict by a single command for simplicity	Try to resolve the conflict by 1 command. Try to resolve behind to avoid traffic flying parallel and leaving airspace at a non-desired location
13	Make sure PZ do not intersect. Leave space for HDG change. 50NM x 50NM	Change HDG, make a roundabout turn if more aircraft
14	The one with lowest velocity is redirected the least possible. If the high velocity aircraft seems easier to deflect, that is done	Depends on the situation, preferably only 1 aircraft is redirected the least amount of path

Table O-3: Participant's Answers to Question 27 from the answer sheet, regarding the Rule of Thumb

What is in your opinion the main rule of thumb you learned during this experiment to resolve conflicts?		
Participant	**SSD Group**	**Control Group**
1	Change the heading for one aircraft only, wait to ensure separation, then apply DCT to all aircraft	When suspecting a conflict, already resolve it. Only when you are sure there is none (in the future) look at getting the aircraft to the COPX
2	Slower plane has to cross the trajectory of the faster one only when the faster plane has already traveled it	Keep the extra time due to heading changes as small as possible
3	Go behind the fast one with the slow one	Divert slower aircraft away from the direction of faster aircraft
4	The slowest behind the fastest	Find the solution which requires the smallest deviation of the desired flight path for all aircraft involved. The slower aircraft gets a small heading change to pass behind the faster aircraft
5	Get the faster one in front	Let the quickest aircraft go in front
6	Give the right	Choose slowest to change heading, that's the most efficient way in terms of covered path
7	Make small adjustments to let one aircraft pass behind another. Generally the slower aircraft are easier to move away.	Depends on the situation, not a clear rule of thumb
8	Move the most conflicting causing aircraft out of the way	To quickly estimate future positions and try to minimize probability of PZ violation by ensuring flight paths do not cross
9	Getting the plane you order to move behind the other, not in front and preferably order the slow plane to do so	Always avoid conflicts by crossing behind an aircraft
10	Try to change heading of least amount of aircraft	Choose the aircraft that would still go towards the destination and not move from it
11	Change slow aircraft headings so that faster aircraft pass in front of them	It's not efficient if we use for more than 3 aircraft: Depending on how far the aircraft is from the exit waypoint a command is given to aircraft that is far away from the exiting waypoint
12	Slow aircraft behind fast	Try to resolve a conflict as soon as possible such that correction can be done when needed
13	Be fast, be decisive, steer aircraft towards empty space to avoid future conflicts	Check speed and prioritize: easier high speed difference, difficult same/low speed difference
14	Adjust one plane, away from the trajectory of the other plane	Wait and see

Appendix P

Audio Transcripts

Below, four example transcripts are shown to give the reader an impression of the verbal comments participants made while performing the scenarios. The tables each display a different situation (conflict yes/no, due to (in)correct execution of command or strategy).

Table P-1 shows a transcript of a participant executing the first scenario of the test phase. From the transcript it can be seen that the participant was assigned to the SSD group as he/she is referring to the removal of the SSD in the first sentence. This is thus an example of a transcript that could bias an evaluator during blind analysis of the data. It can furthermore be seen that the participant is referring to the correct strategy.

Table P-1: Example Audio Transcript of Scenario 1 of the Test Phase

Scenario 1 of Test Phase	
Time (s)	Verbal comment
00:03	Dit is echt een stuk vervelender
00:06	We gaan gewoon weer langzaam achterlangs sturen watn dat was een goede strategie
00:10	Ehh good strategy
00:12	KILO you're going behind 045 EXQ, aahh now I can't see where you are
00:22	I think this will go right
00:25	JULIET you will get a DCT command to VOLK
00:31	KILO has to wait until JULIET has passed
00:40	KILO I think by now can get a DCT command as well

While the participant from Table P-1 is giving the researchers a little bit more info about his/her strategy/thought patterns by not strictly adhering to the think-aloud protocol but rather using it as a guide, the participant from Table P-2 showed to strictly adhere to the think-aloud protocol by mentioning the very minimum. Although this makes the transcribing of audio recordings less laborious, it also makes it more difficult to correctly assess the participant's strategies and thought patterns as many thoughts were likely not mentioned out loud (e.g., the reason for sending the aircraft to the right).

Table P-3 and P-4 show two scenarios where a LoS occurred. Table P-3 shows that the LoS occurred due to an incomplete command (repetion of HDG 300), whereas Table P-4 shows that the LoS was a result of an incorrect command (extra correction to resolve the conflict).

Table P-2: Example Audio Transcript of Scenario 9 of the Test Phase

Scenario 9 of Test Phase	
Time (s)	Verbal comment
00:04	Ok INDIA is slowest aircraft
00:06	Think.. a conflict will happen
00:11	To prevent the conflict, sending it to the right with a HDG of 160
00:29	Ok conflict prevented
00:33	Waiting until they have passed each other
00:44	Sending PAPA to NOKL
00:49	And sending INDIA to HILL

Table P-3: Example Audio Transcript of Scenario 3 of the Test Phase

Scenario 3 of Test Phase	
Time (s)	Verbal comment
00:12	DCT TANGO naar waypoint
00:18	Change HDG SIERRA naar 270
00:29	Dat gaat niet goed
00:36	Oke dat ging niet goed
00:41	Change HDG 300
00:45	HDG 300, nou kom op
00:51	Ja nou wordt het boem
01:08	DCT SIERRA naar waypoint

Table P-4: Example Audio Transcript of Scenario 3 of the Training Phase

Scenario 3 of Training Phase	
Time (s)	Verbal comment
00:04	Ok checking for both what their destination is
00:07	They will collide with each other
00:10	Ehhh
00:17	I'm sending INDIA a little bit the left because that one's going faster so they won't collide with each other
00:25	So INDIA 25 HDG
00:41	Oh no it's not gonna work!
00:47	Oke, INDIA HDG 10 EXQ
01:00	Oh nee het gaat mis! Shit!
01:18	So INDIA DCT destination, ehh UNIFORM DCT destination and INDIA as well

Appendix Q

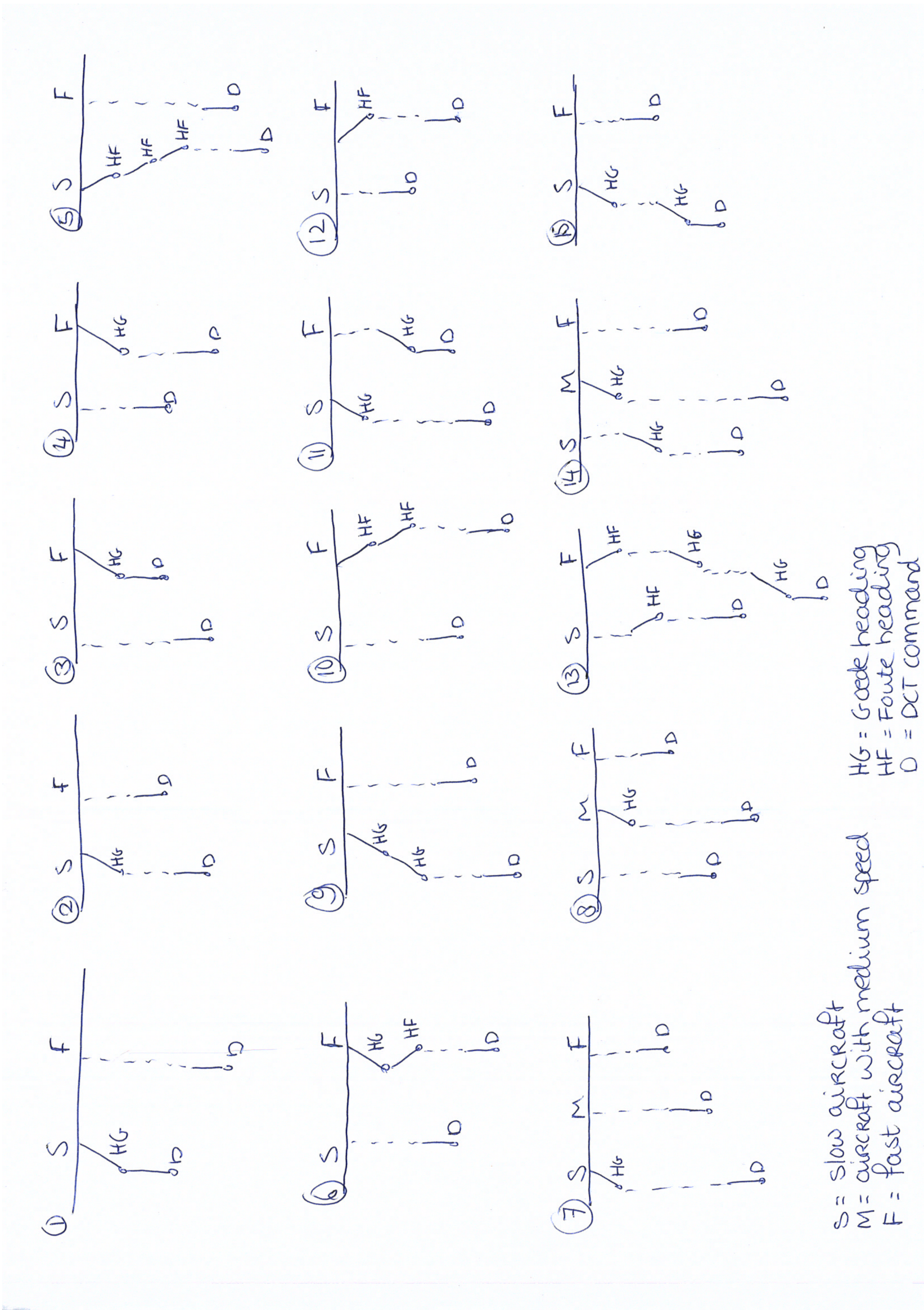
Strategy Classification

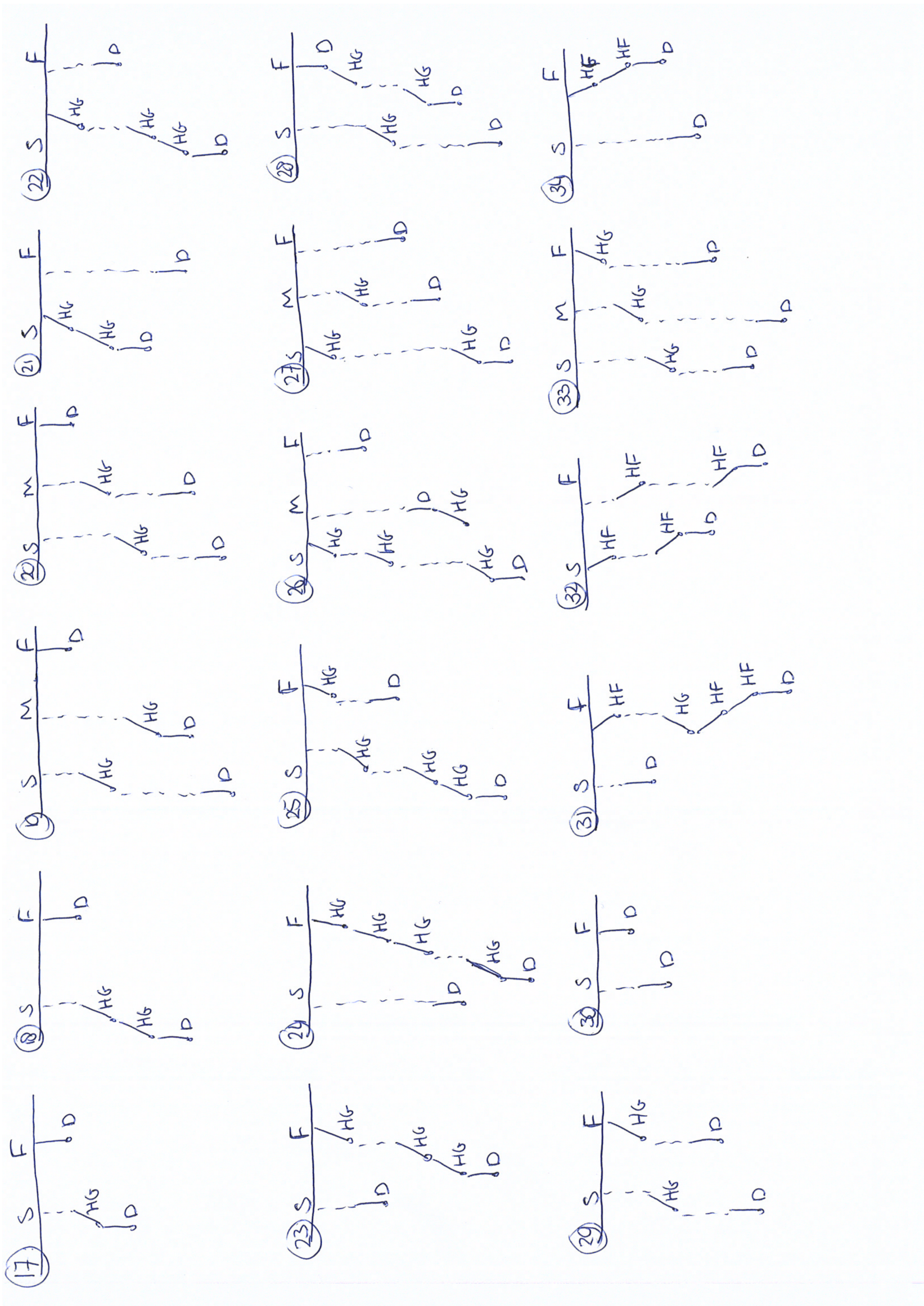
As discussed in Part I, to analyze participant strategies, they were first categorized. This categorization is discussed below.

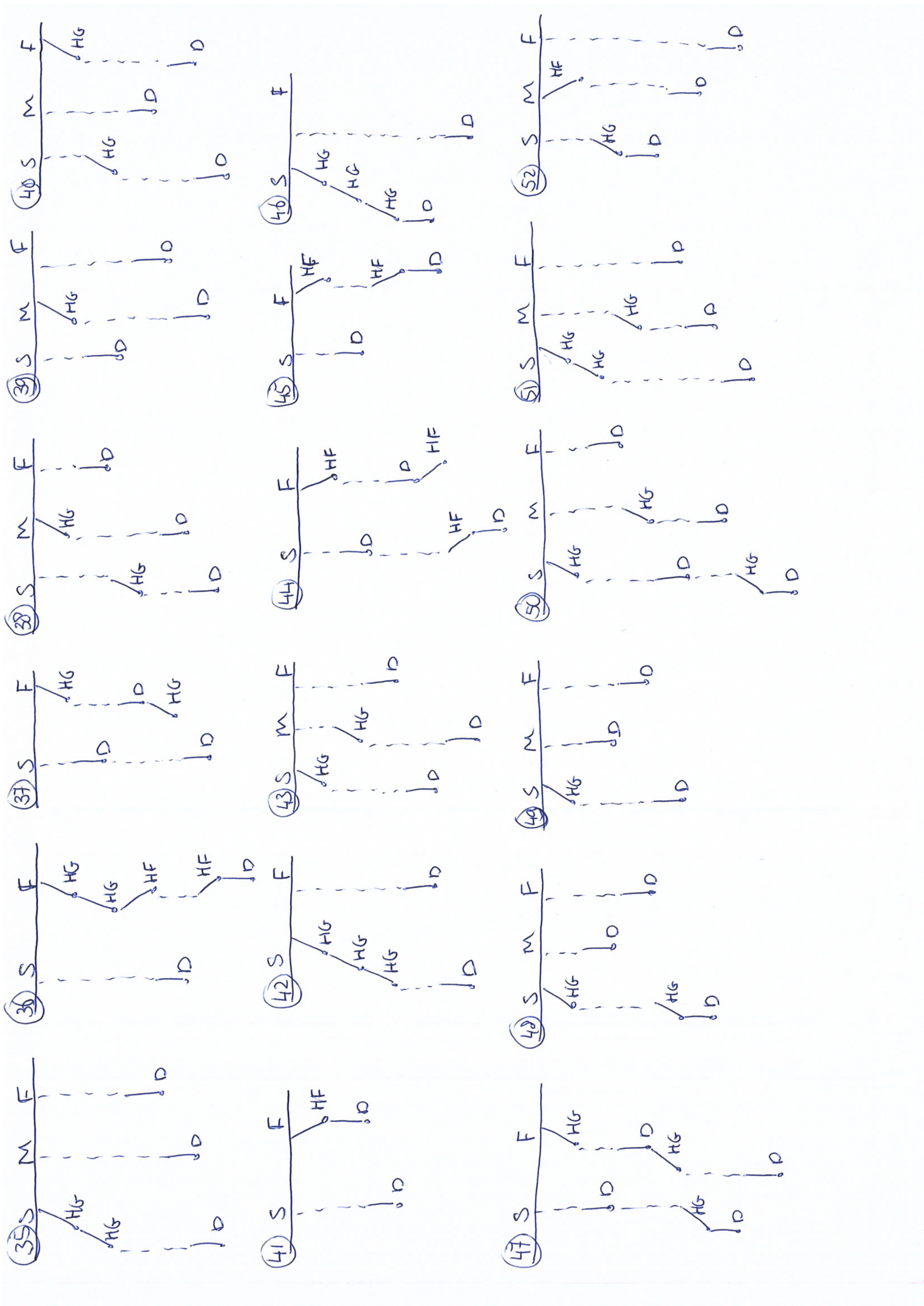
Each scenario performed by the participants was analyzed and a strategy was drawn based on the given commands for that scenario. Three types of commands were distinguished: a DCT command, a correct heading command and an incorrect heading command. A correct heading command would lead toward the optimal solution of sending the slower aircraft behind the faster aircraft, whereas an incorrect command would lead away from the optimal solution toward a sub-optimal solution. Within the strategy schematics, shown on the following pages, these commands were thus drawn in opposite directions. A DCT command was drawn as a vertical line.

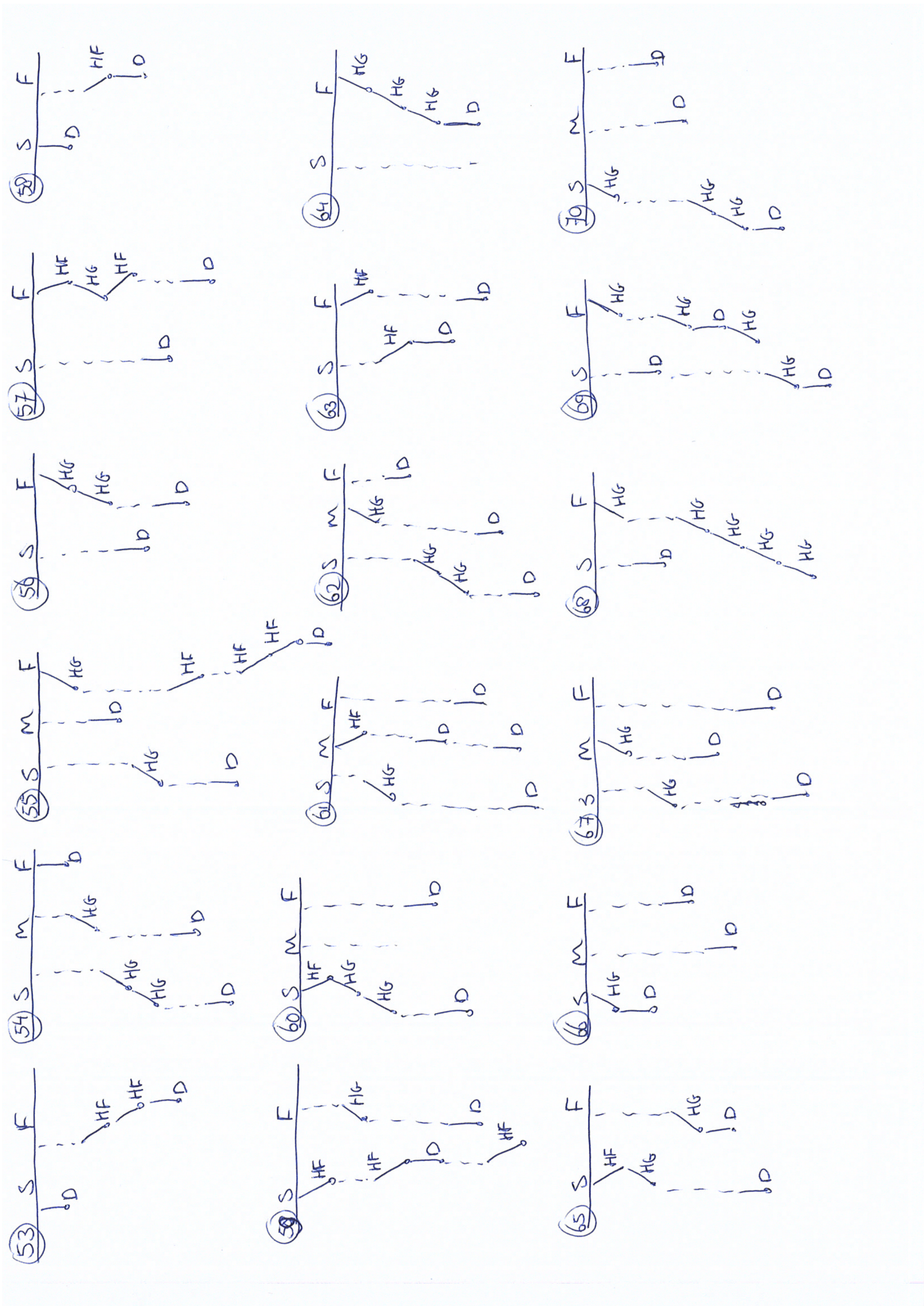
All strategies that were encountered throughout the analysis are shown on the following pages. As they were drawn in the order they were encountered, the two and three-aircraft scenarios are drawn in a mixed order. It can be seen that every strategy received a number. This number was assigned to the corresponding scenarios where this specific strategy was observed, to be able to later group the strategies and perform an analysis on the occurrence of different type of strategies per scenario category.

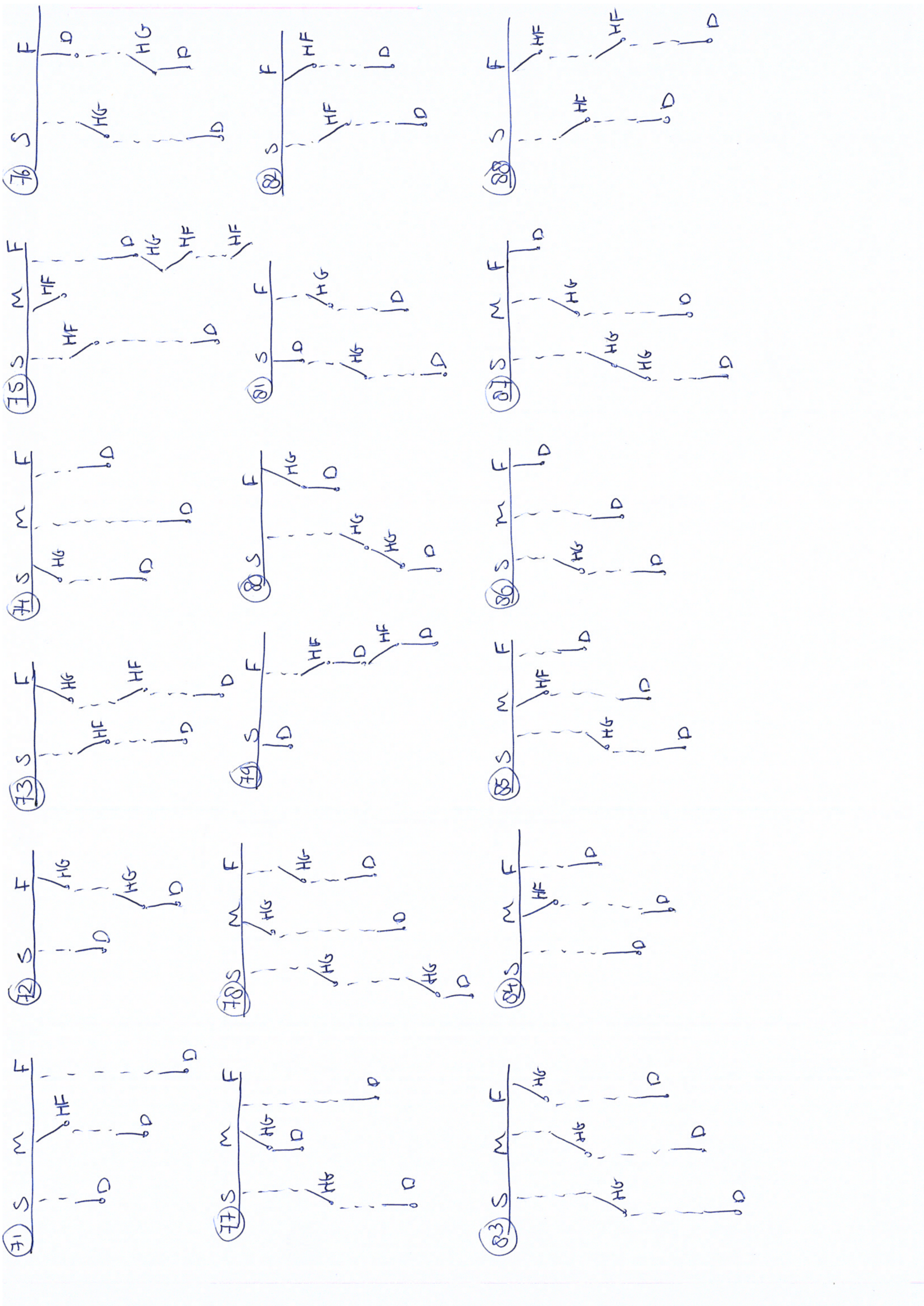
The strategies that tend to move to the left were categorized as optimal, whereas strategies that tend to move to the right were categorized as sub-optimal. The video recordings were used to verify whether a drawn strategy was indeed an optimal or a sub-optimal strategy in case this was not obvious from the drawings. After being categorized as optimal/sub-optimal, they could further be categorized based on the aircraft that received the first command, the number of aircraft involved in a solution and the number of corrections.

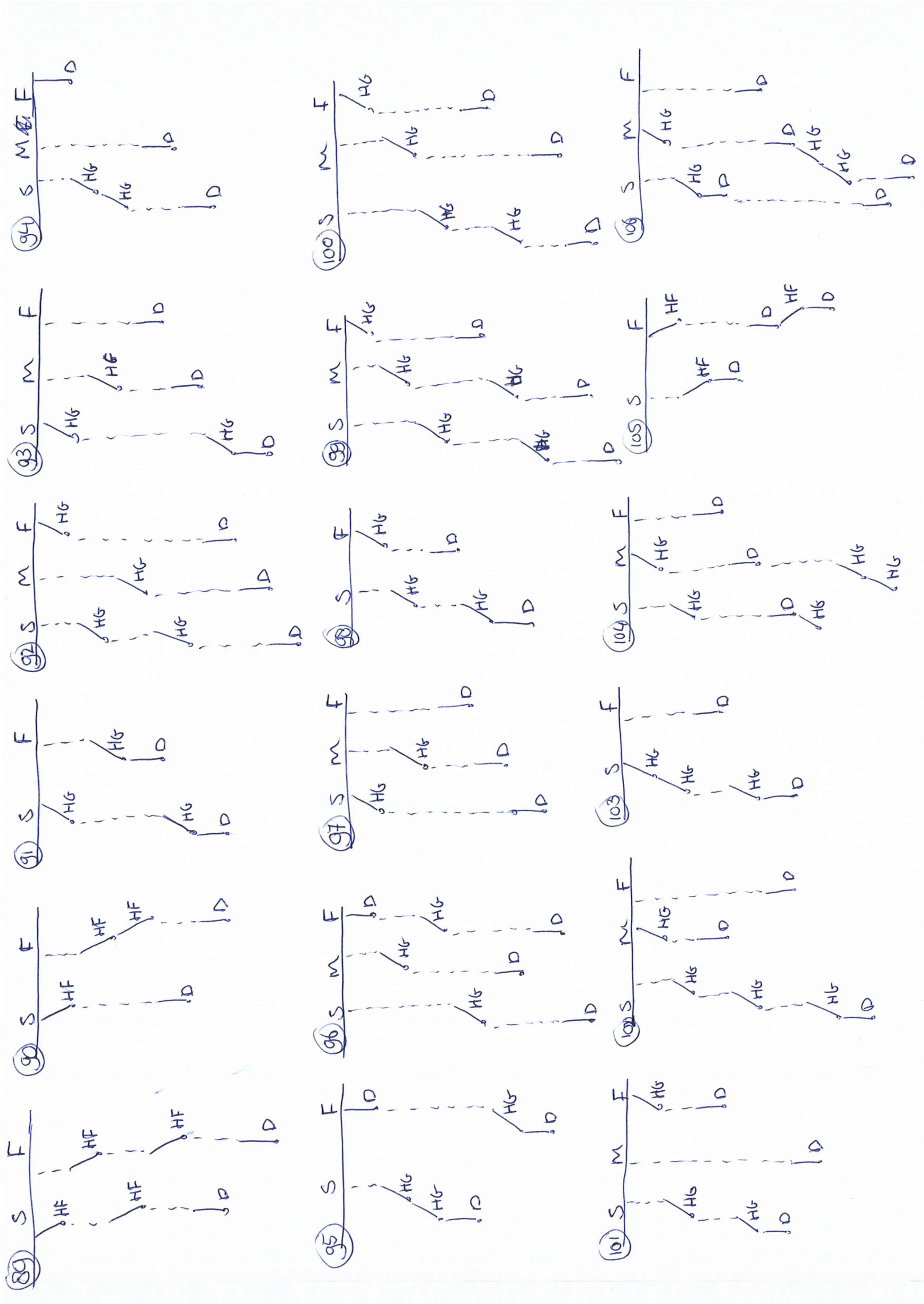

S = slow aircraft
M = aircraft with medium speed
F = fast aircraft
HG : Goede heading
HF : Foute heading
D = DCT command

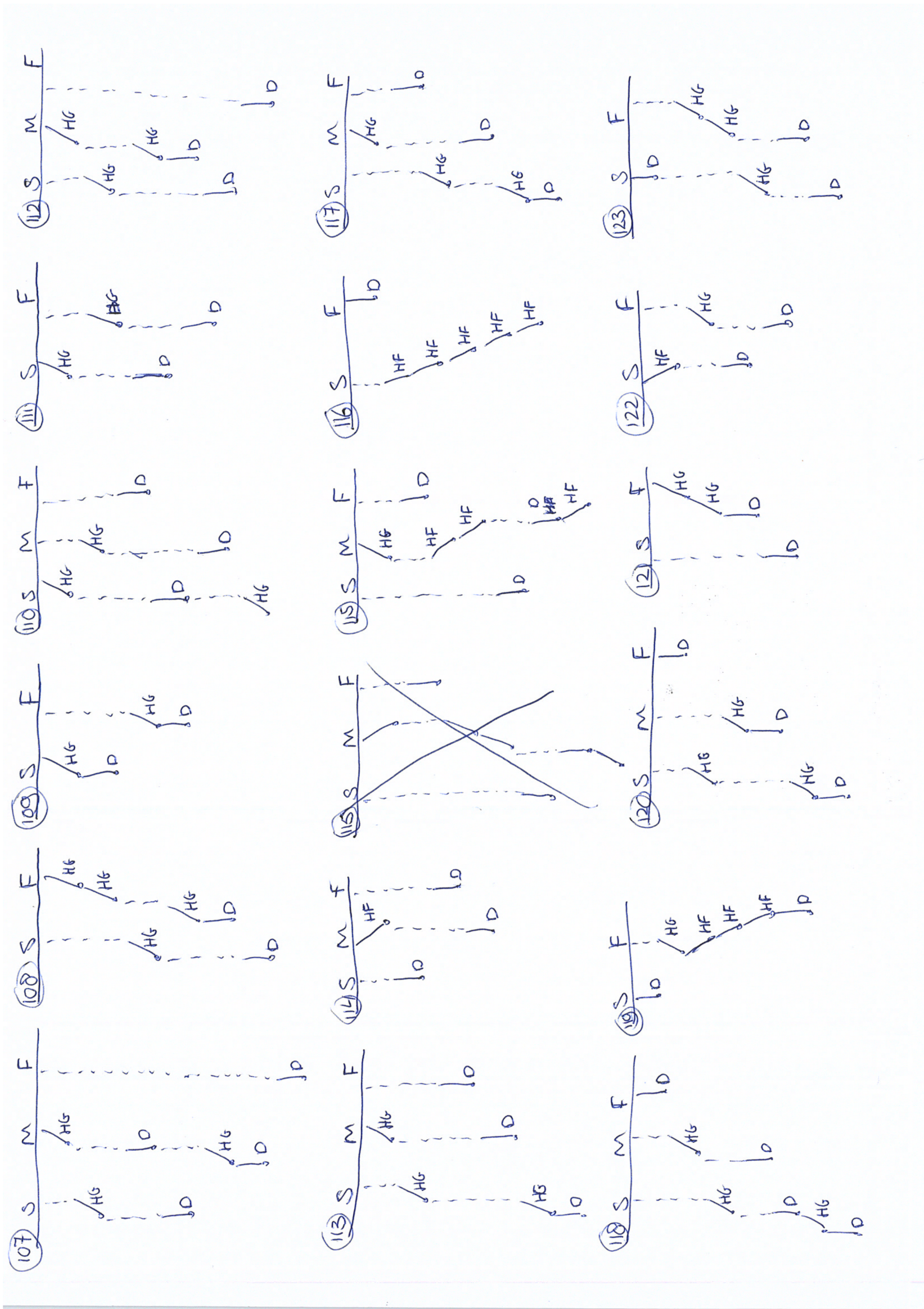












Appendix R

Additional Experiment Results

This appendix presents additional results of the experiment. The results that are shown here are the results for the number of STCAs, the track deviation and difficulty scores.

R-1 Short-Term Conflict Alerts

The number of STCAs, indicating a time to LoS within 40 seconds, were recorded per participant during the experiment. Figure R-1 shows the number of STCAs per scenario for per experiment phase, whereas Figure R-2 shows the number of STCAs per participant.

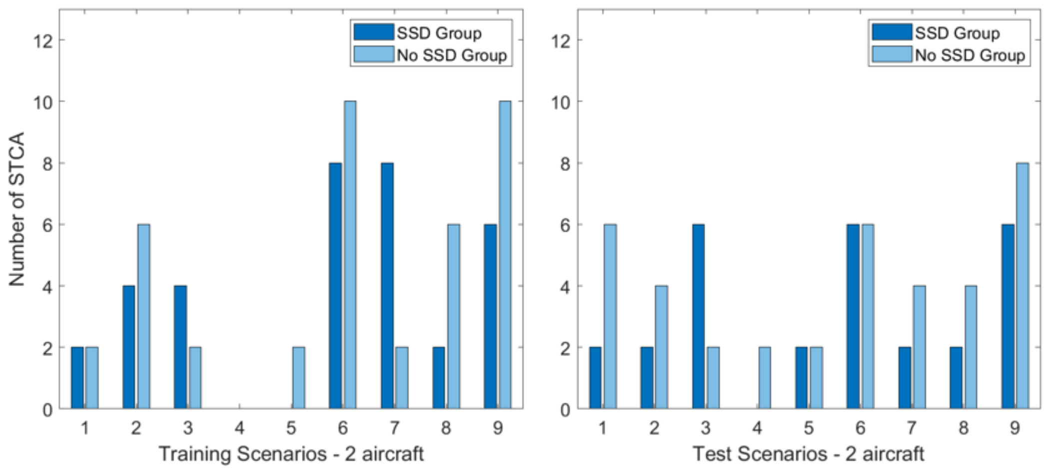
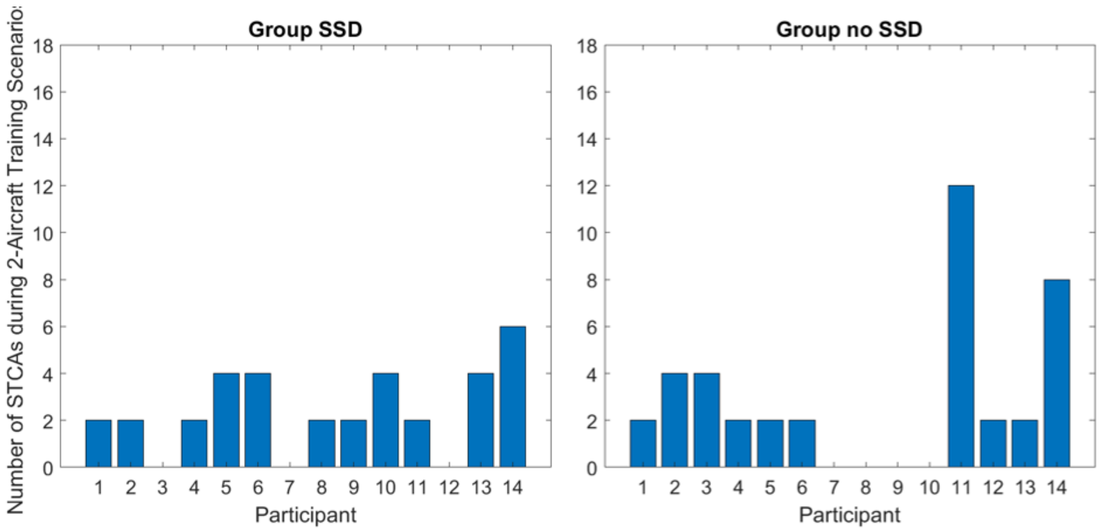
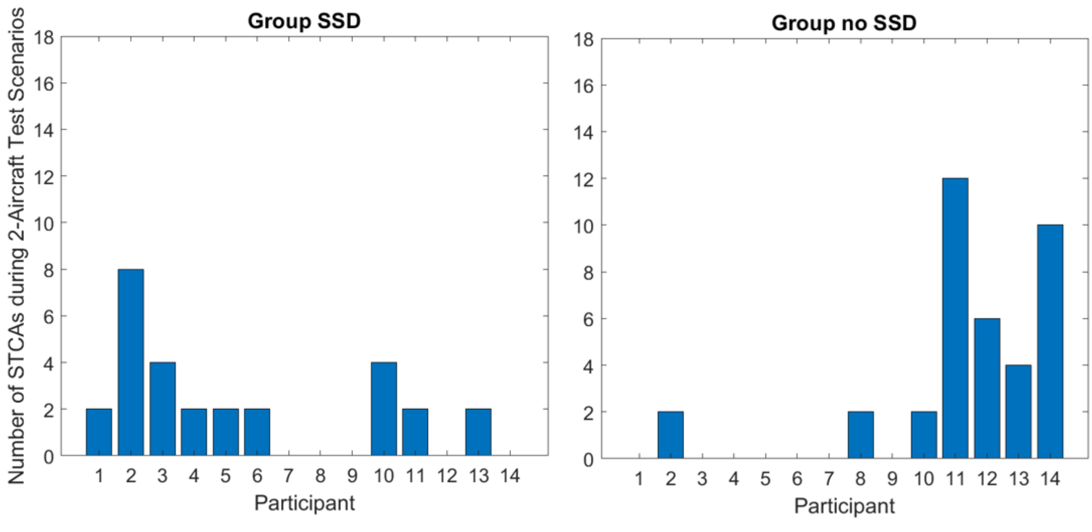


Figure R-1: Number of STCAs per Two-aircraft Scenario in the Training and Test Phase



(a) Training Phase

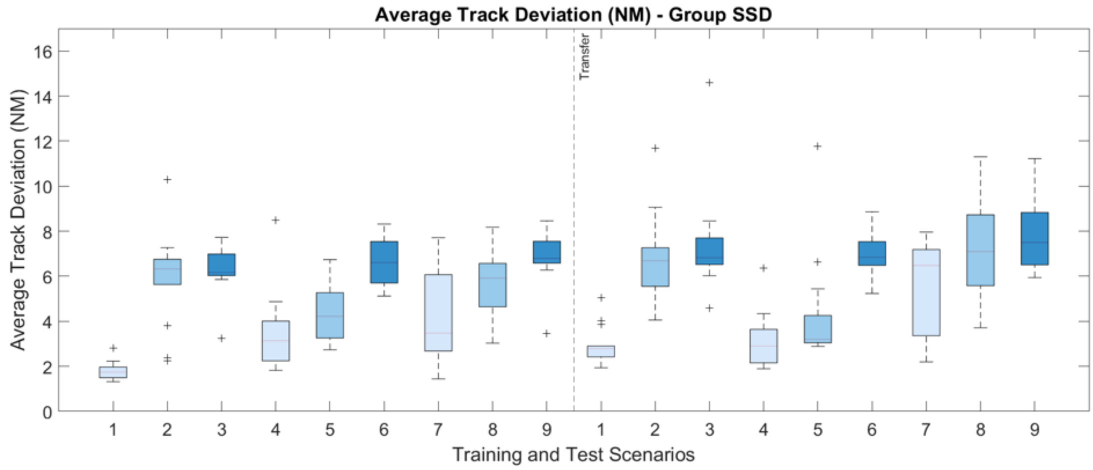


(b) Test Phase

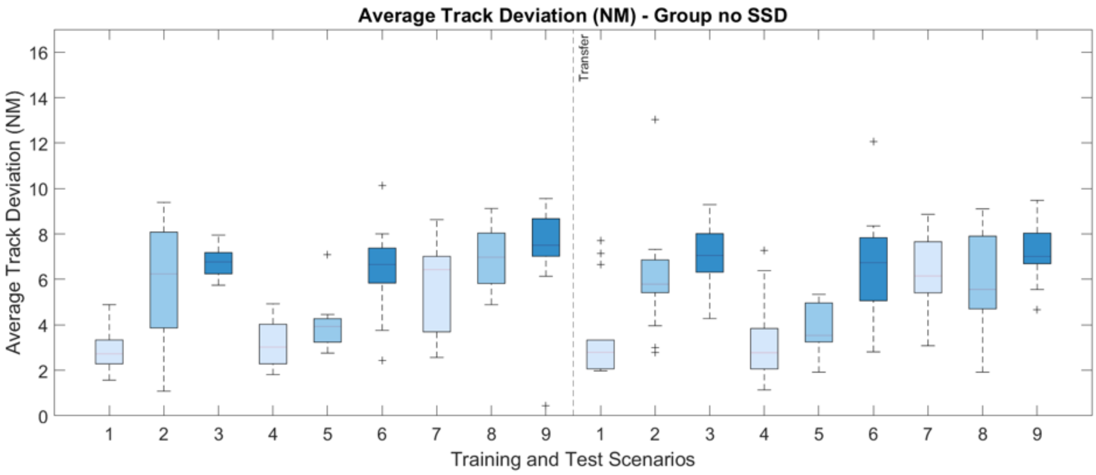
Figure R-2: Number of STCAs per Participant During the Two-aircraft Scenarios

R-2 Track Deviation

The track deviation was also discussed in Part I. The results presented here display the results in the original scenario order.



(a) SSD Group

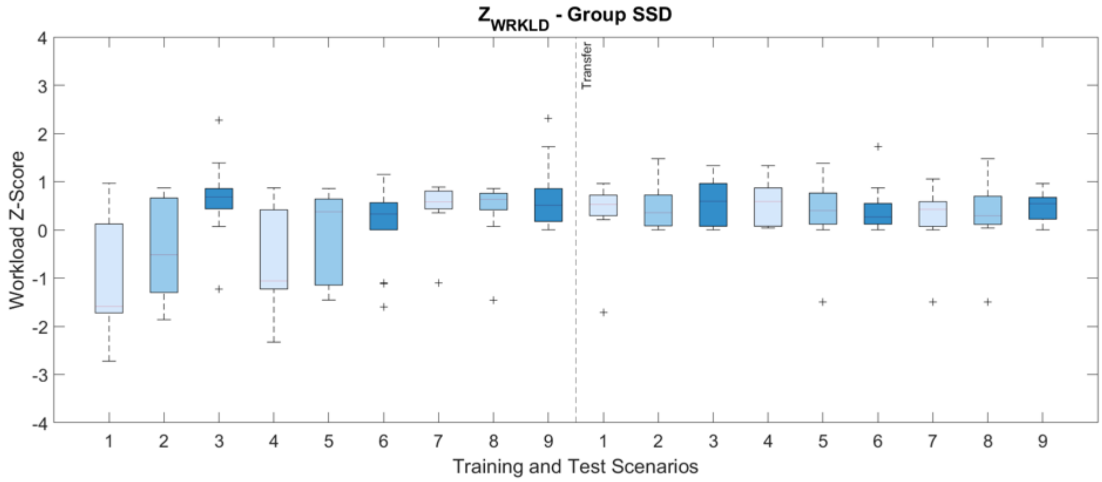


(b) Control Group

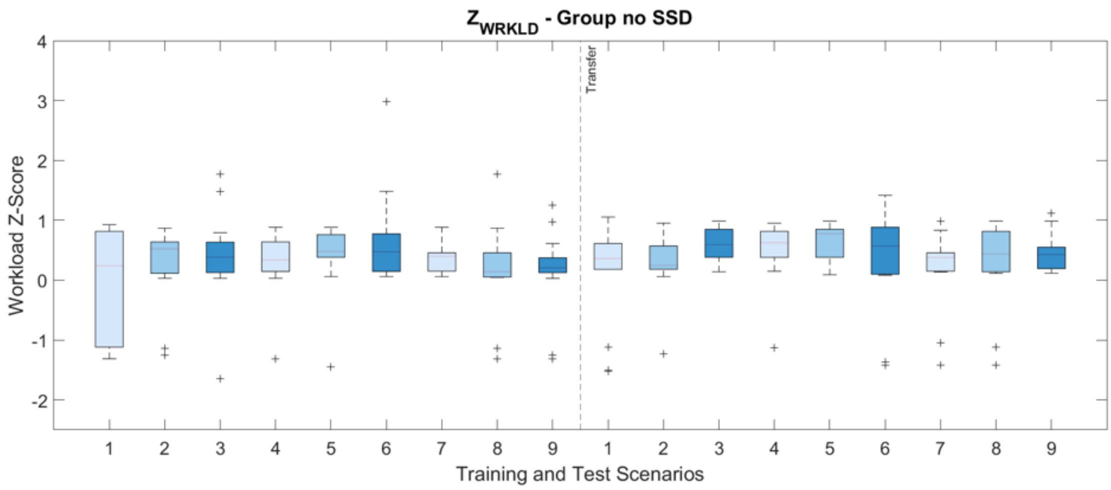
Figure R-3: Average Track Deviation (NM) Results, Displayed for Both Groups in the Original Scenario Order

R-3 Difficulty Scores

After every scenario, participants were asked to indicate the perceived difficulty level on a sliding bar, where sliding it left indicated an easy scenario and sliding it right indicated a more difficult scenario. The results have been normalized for each participant to account for participant subjectivity. The normalized Z-scores are shown per participant group in Figure R-4 and per scenario category in Figure R-5.



(a) SSD Group



(b) Control Group

Figure R-4: Z-Scores for the Difficulty Rating per Group, Displayed in the Original Scenario Order

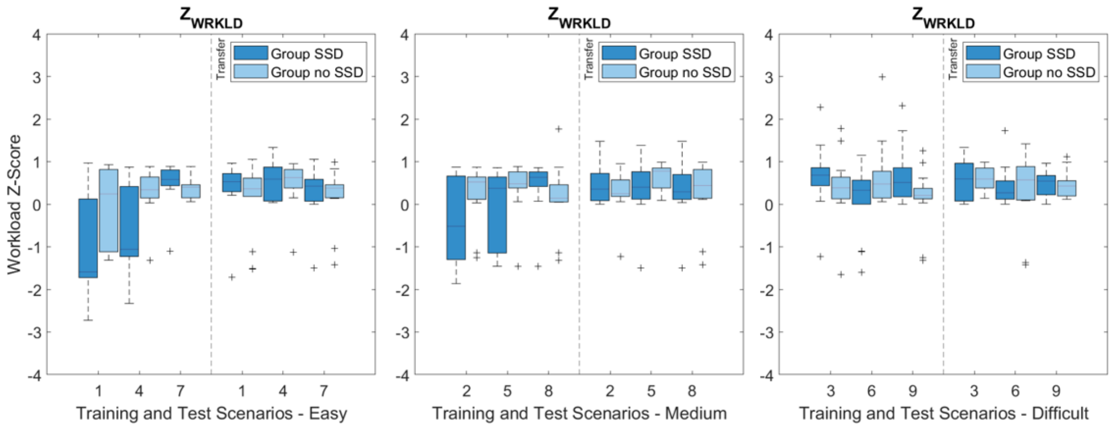


Figure R-5: Z-Scores for the Difficulty Rating, Displayed per Scenario Category

Appendix S

Statistical Results

This appendix presents the results of the statistical analysis that was performed in SPSS 25. Kolmogorev-Smirnov and Shapiro-Wilk tests were performed to check for normality. In each data set, several of the distributions appear to be non-normal distributions. Because of this, the relatively small sample size and the nature of the data (constant scenario order for all participants), it was decided to consider all data sets as non-parametric. The Kruskal-Wallis test was used to identify differences between groups, whereas the Wilcoxon Signed Rank Test and Friedman Test were used to identify differences between the training and test phase as well as differences between the three scenario categories (i.e., easy, medium, difficult), respectively.

The analysis was performed for the minimum separation distance, track deviation, number of correctly-answered still scenarios, strategy classification, input errors and number of LoS's.

S-1 Minimum Separation Distance

The minimum separation distance was evaluated by looking at the learning curve characteristics for all participants. These characteristics consist of three variables, namely two learning gradients and a delta value, as was explained in Part 1 of the report. Below, the results for the delta and learning gradient data are shown, respectively.

S-1-1 Delta

Table S-1: Test of Normality for Delta Values of the Minimum Separation Distance

Tests of Normality							
Category	Group	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Delta Easy	SSD	0.091	14	0.200*	0.976	14	0.948
	No SSD	0.154	14	0.200*	0.962	14	0.755
Delta Medium	SSD	0.195	14	0.154	0.864	14	0.035
	No SSD	0.130	14	0.200*	0.975	14	0.935
Delta Difficult	SSD	0.209	14	0.100	0.914	14	0.182
	No SSD	0.141	14	0.200*	0.937	14	0.381

*. This is a lower bound of the true significance.

Table S-2: Kruskal-Wallis Test for Delta Values of the Minimum Separation Distance

Test Statistics[a,b]			
	Delta Easy	Delta Medium	Delta Difficult
Kruskal-Wallis H	4.664	4.864	1.022
df	1	1	1
Asymp. Sig.	**0.031**	**0.027**	0.312

a. Kruskal Wallis Test
b. Grouping Variable: Participant Group

Table S-3: Friedman Test for Effects of Scenario Categories on Delta Values of the Minimum Separation Distance

Test Statistics[a]		
	SSD Group	No SSD Group
N	14	14
Chi-Square	1.857	1.000
df	2	2
Asymp. Sig.	0.395	0.607

a. Friedman Test

S-1-2 Learning Gradients

Table S-4: Test of Normality for Learning Gradients of the Minimum Separation Distance

Tests of Normality								
Category	Group	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Training Easy	SSD	0.148	14	0.200*	0.938	14	0.390
	No SSD	0.158	14	0.200*	0.930	14	0.304
Test Easy	SSD	0.125	14	0.200*	0.962	14	0.749
	No SSD	0.152	14	0.200*	0.974	14	0.923
Training Medium	SSD	0.281	14	0.004	0.896	14	0.099
	No SSD	0.119	14	0.200*	0.959	14	0.713
Test Medium	SSD	0.133	14	0.200*	0.961	14	0.735
	No SSD	0.117	14	0.200*	0.982	14	0.985
Training Difficult	SSD	0.219	14	0.067	0.849	14	0.022
	No SSD	0.158	14	0.200*	0.944	14	0.468
Test Difficult	SSD	0.132	14	0.200*	0.926	14	0.266
	No SSD	0.193	14	0.166	0.943	14	0.461

*. This is a lower bound of the true significance.

Table S-5: Kruskal Wallis Test for Learning Gradients of the Minimum Separation Distance per Scenario Category and Experiment Phase

Test Statistics[a,b]						
	Training Easy	Test Easy	Training Medium	Test Medium	Training Difficult	Test Difficult
Kruskal-Wallis H	0.135	0.019	0.357	0.103	0.019	0.008
df	1	1	1	1	1	1
Asymp. Sig.	0.713	0.890	0.550	0.748	0.890	0.927

a. Kruskal Wallis Test
b. Grouping Variable: Participant Group

Table S-6: Friedman Test for Effects of Scenario Categories on Learning Gradients of the Minimum Separation Distance

Test Statistics[a]				
	SSD Training	No SSD Training	SSD Test	No SSD Test
N	14	14	14	14
Chi-Square	0.429	1.857	0.143	0.571
df	2	2	2	2
Asymp. Sig.	0.807	0.395	0.931	0.751

a. Friedman Test

Table S-7: Wilcoxon Signed Rank Test for Differences between Training and Test Phases of the Learning Gradients of the Minimum Separation Distance per Scenario Category

Test Statistics[a]						
	SSD Easy	No SSD Easy	SSD Medium	No SSD Medium	SSD Difficult	No SSD Difficult
N	14	14	14	14	14	14
T	48.0	54.0	52.0	62.0	45.0	54.0
Z	-0.282	0.094	-0.031	0.596	-0.471	0.094
Asymp. Sig.	0.778	0.925	0.975	0.551	0.638	0.925

a. Wilcoxon Signed Rank Test

S-2 Track Deviation

The track deviation was evaluated by looking at the learning curve characteristics for all participants. Similar to the minimum separation distance, these characteristics consist of two learning gradients and a delta value, as was explained in Part 1 of the report. Below, the results for the delta data and the learning gradients are shown, respectively.

S-2-1 Delta

Table S-8: Test of Normality for Delta Values of the Track Deviation

Tests of Normality							
Category	Group	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Delta Easy	SSD	0.289	14	0.002	0.848	14	0.021
	No SSD	0.232	14	0.040	0.909	14	0.155
Delta Medium	SSD	0.144	14	0.200*	0.948	14	0.532
	No SSD	0.143	14	0.200*	0.913	14	0.175
Delta Difficult	SSD	0.222	14	0.061	0.816	14	0.008
	No SSD	0.164	14	0.200*	0.913	14	0.175

*. This is a lower bound of the true significance.

Table S-9: Kruskal Wallis Test for Delta Values of the Track Deviation

Test Statistics[a,b]			
	Delta Easy	Delta Medium	Delta Difficult
Kruskal-Wallis H	1.319	3.549	3.049
df	1	1	1
Asymp. Sig.	0.251	0.060	0.081

a. Kruskal Wallis Test

b. Grouping Variable: Participant Group

Table S-10: Friedman Test for Effects of Scenario Categories on Delta Values of the Track Deviation

Test Statistics[a]		
	SSD Group	No SSD Group
N	14	14
Chi-Square	5.286	7.000
df	2	2
Asymp. Sig.	0.071	**0.030**

a. Friedman Test

Table S-11: Post-Hoc Test for Effects of Scenario Categories on Delta Value of the Track Deviation

Test Statistics[a,b]			
No SSD: χ^2 = 7.000, p = 0.030			
	Delta Easy	Delta Medium	Delta Difficult
Delta Easy	x		
Delta Medium	0.113	x	
Delta Difficult	**0.042**	1.000	x

a. Post-Hoc Test
b. Group = No SSD

S-2-2 Learning Gradients

Table S-12: Test of Normality for Learning Gradients of the Track Deviation

Tests of Normality								
Category	Group	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Training Easy	SSD	0.171	14	0.200*	0.941	14	0.425
	No SSD	0.144	14	0.200*	0.931	14	0.316
Test Easy	SSD	0.189	14	0.191	0.909	14	0.153
	No SSD	0.196	14	0.148	0.870	14	0.042
Training Medium	SSD	0.093	14	0.200*	0.988	14	0.988
	No SSD	0.133	14	0.200*	0.931	14	0.318
Test Medium	SSD	0.148	14	0.200*	0.964	14	0.788
	No SSD	0.158	14	0.200*	0.944	14	0.467
Training Difficult	SSD	0.170	14	0.200*	0.943	14	0.462
	No SSD	0.338	14	0.000	0.655	14	0.000
Test Difficult	SSD	0.155	14	0.200*	0.956	14	0.654
	No SSD	0.207	14	0.106	0.894	14	0.092

*. This is a lower bound of the true significance.

Table S-13: Kruskal Wallis Test for Learning Gradients of the Track Deviation per Scenario Category and Experiment Phase

Test Statistics[a,b]						
	Training Easy	Test Easy	Training Medium	Test Medium	Training Difficult	Test Difficult
Kruskal-Wallis H	0.887	0.076	0.684	0.211	0.540	0.304
df	1	1	1	1	1	1
Asymp. Sig.	0.346	0.783	0.408	0.646	0.462	0.581

a. Kruskal Wallis Test
b. Grouping Variable: Participant Group

Table S-14: Friedman Test for Effects of Scenario Categories on Learning Gradients of the Track Deviation

Test Statistics[a]				
	SSD Training	No SSD Training	SSD Test	No SSD Test
N	14	14	14	14
Chi-Square	8.143	10.429	10.714	9.571
df	2	2	2	2
Asymp. Sig.	**0.017**	**0.005**	**0.005**	**0.008**

a. Friedman Test

Table S-15: Post-Hoc Test for Effects of Scenario Categories on Learning Gradients of the Track Deviation

Test Statistics[a,b]			
SSD Training: $\chi^2 = 8.143$, p = 0.017			
	Gradient Easy	Gradient Medium	Gradient Difficult
Gradient Easy	x		
Gradient Medium	**0.014**	x	
Gradient Difficult	0.267	0.771	x

a. Post-Hoc Test
b. SSD, Training

Table S-16: Post-Hoc Wilcoxon Signed Rank Test for Effects of Scenario Categories on Learning Gradients of the Track Deviation

Test Statistics[a,b]			
No SSD Training: $\chi^2 = 10.429$, p = 0.005			
	Gradient Easy	Gradient Medium	Gradient Difficult
Gradient Easy	x		
Gradient Medium	0.176	x	
Gradient Difficult	**0.004**	0.558	x

a. Post-Hoc Test
b. No SSD, Training

Table S-17: Post-Hoc Rank Test for Effects of Scenario Categories on Learning Gradients of the Track Deviation

Test Statistics[a,b]			
SSD Test: $\chi^2 = 10.714$, p = 0.005			
	Gradient Easy	Gradient Medium	Gradient Difficult
Gradient Easy	x		
Gradient Medium	**0.014**	x	
Gradient Difficult	**0.014**	1.000	x

a. Post-Hoc Test
b. SSD, Test

Table S-18: Post-Hoc Test for Effects of Scenario Categories on Learning Gradients of the Track Deviation

Test Statistics[a,b]			
No SSD Test: χ^2 = 9.571, p = 0.008			
	Gradient Easy	Gradient Medium	Gradient Difficult
Gradient Easy	x		
Gradient Medium	**0.007**	x	
Gradient Difficult	0.113	1.000	x

a. Post-Hoc Test
b. No SSD, Test

Table S-19: Wilcoxon Signed Rank Test for Differences between Training and Test Phases of the Learning Gradients of the Track Deviation per Scenario Category

Test Statistics[a]						
	SSD Easy	No SSD Easy	SSD Medium	No SSD Medium	SSD Difficult	No SSD Difficult
N	14	14	14	14	14	14
T	59.0	56.0	55.0	32.0	42.0	27.0
Z	0.408	0.220	0.157	-1.287	-0.659	-1.601
Asymp. Sig.	0.683	0.826	0.875	0.198	0.510	0.109

a. Wilcoxon Signed Rank Test

S-3 Still Scenarios

During the course of the experiment, participants had to answer several still scenarios. In case participants identified a conflict and sent one of two aircraft in the correct direction, this was counted as a correct answer. Per participant, the number of correct answers was counted to be able to perform statistical tests on this data.

Table S-20: Test of Normality for Number of Correctly-Answered Still Scenarios

Tests of Normality							
Category	Group	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Training	SSD	0.164	14	0.200*	0.920	14	0.216
	No SSD	0.156	14	0.200*	0.936	14	0.365
Test	SSD	0.241	14	0.027	0.935	14	0.360
	No SSD	0.218	14	0.069	0.929	14	0.295

*. This is a lower bound of the true significance.

Table S-21: Kruskal Wallis Test for Number of Correctly-Answered Still Scenarios per Experiment Phase

Test Statistics[a,b]		
	Before Transfer	After Transfer
Kruskal-Wallis H	2.568	1.299
df	1	1
Asymp. Sig.	0.109	0.254

a. Kruskal Wallis Test
b. Grouping Variable: Participant Group

Table S-22: Wilcoxon Signed Rank Test for Difference in Correctly-Answered Questions Before and After Transfer Manipulation

Test Statistics[a]		
	SSD Group	No SSD Group
N	14	14
T	23.0	25.5
Z	-0.456	-1.072
Asymp. Sig.	0.642	0.284

a. Wilcoxon Signed Rank Test

S-4 Strategies

Participant strategies were classified based on being optimal (i.e., vector slow aircraft behind fast aircraft) or sub-optimal. A score of 1 was assigned in case a participant used an optimal strategy in a scenario. After summing the scores, a statistical analysis could be performed. The results of this analysis are shown below.

Table S-23: Test of Normality for Number of Optimal Strategies

Tests of Normality							
Category	Group	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Easy - Training	SSD	0.369	14	0.000	0.639	14	0.000
	No SSD	0.443	14	0.000	0.576	14	0.000
Easy Test	SSD	0.285	14	0.003	0.771	14	0.002
	No SSD	0.357	14	0.000	0.735	14	0.001
Medium - Training	SSD	0.372	14	0.000	0.681	14	0.000
	No SSD	0.349	14	0.000	0.724	14	0.001
Medium - Test	SSD	0.263	14	0.010	0.812	14	0.007
	No SSD	0.417	14	0.000	0.618	14	0.000
Difficult - Training	SSD	0.335	14	0.000	0.751	14	0.001
	No SSD	0.350	14	0.000	0.731	14	0.001
Difficult - Test	SSD	0.369	14	0.000	0.639	14	0.000
	No SSD	0.277	14	0.005	0.708	14	0.000

*. This is a lower bound of the true significance.

Table S-24: Kruskal Wallis Test for Number of Optimal Strategies in Two-Aircraft Scenarios per Experiment Phase

Test Statistics^{a,b}						
	Training Easy	Test Easy	Training Medium	Test Medium	Training Difficult	Test Difficult
Kruskal-Wallis H	2.250	2.525	0.135	3.122	0.113	0.275
df	1	1	1	1	1	1
Asymp. Sig.	0.134	0.112	0.713	0.077	0.737	0.600

a. Kruskal Wallis Test

b. Grouping Variable: Participant Group

Table S-25: Kruskal Wallis Test for Number of Optimal Strategies in Three-Aircraft Scenarios per Experiment Phase

Test Statistics^{a,b}		
	Training	Test
Kruskal-Wallis H	0.380	3.363
df	1	1
Asymp. Sig.	0.538	0.067

a. Kruskal Wallis Test

b. Grouping Variable: Participant Group

Table S-26: Friedman Test for Difference in Number of Optimal Strategies between Scenario Categories during the Training Phase

Test Statistics[a]		
	SSD Group	No SSD Group
N	14	14
Chi-Square	0.452	0.941
df	2	2
Asymp. Sig.	0.798	0.625

a. Friedman Test

Table S-27: Friedman Test for Difference in Number of Optimal Strategies between Scenario Categories during the Test Phase

Test Statistics[a]		
	SSD Group	No SSD Group
N	14	14
Chi-Square	3.556	7.316
df	2	2
Asymp. Sig.	0.169	**0.026**

a. Friedman Test

Table S-28: Post-Hoc Test for Effects of Scenario Categories on Number of Optimal Strategies

Test Statistics[a,b]			
No SSD: $\chi^2 = 7.316$, p = 0.026			
	Count Easy	Count Medium	Count Difficult
Count Easy	x		
Count Medium	0.089	x	
Count Difficult	0.392	1.000	x

a. Post-Hoc Test
b. No SSD, Test

Table S-29: Wilcoxon Signed Rank Test for Differences in Number of Optimal Strategies between the Training and Test Phase per Scenario Category

Test Statistics[a]						
	SSD Easy	No SSD Easy	SSD Medium	No SSD Medium	SSD Difficult	No SSD Difficult
N	14	14	14	14	14	14
T	3.0	0.0	10.5	10.0	21.0	7.0
Z	-1.342	-2.000	-1.469	0.707	1.265	-0.816
Asymp. Sig.	0.180	**0.046**	0.142	0.480	0.206	0.414

a. Wilcoxon Signed Rank Test

Table S-30: Wilcoxon Signed Rank Test for Differences in Number of Optimal Strategies between the Training and Test Phase for Three-Aircraft Scenarios

Test Statistics[a]		
	SSD	No SSD
N	14	14
T	5.0	10.0
Z	-0.707	1.857
Asymp. Sig.	0.480	0.063

a. Wilcoxon Signed Rank Test

S-5 Input Errors

As explained in Part I, an input error was defined as an incomplete command. Per participant, the number of input errors over all scenarios was counted to be able to perform statistical tests on the data. The results are shown below.

Table S-31: Test of Normality for Number of Input Errors

Tests of Normality							
Category	Group	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Training	SSD	0.411	14	0.000	0.607	14	0.000
	No SSD	0.332	14	0.000	0.705	14	0.000
Test	SSD	0.245	14	0.023	0.935	14	0.009
	No SSD	0.418	14	0.000	0.397	14	0.000

*. This is a lower bound of the true significance.

Table S-32: Kruskal Wallis Test for Number of Input Errors per Experiment Phase

Test Statistics[a,b]		
	Training	Test
Kruskal-Wallis H	1.378	2.381
df	1	1
Asymp. Sig.	0.240	0.123

a. Kruskal Wallis Test
b. Grouping Variable: Participant Group

Table S-33: Wilcoxon Signed Rank Test for Difference in Number of Input Errors during the Training and Test Phase

Test Statistics[a]		
	SSD Group	No SSD Group
N	14	14
T	21.0	2.5
Z	-0.179	-1.983
Asymp. Sig.	0.858	**0.047**

a. Wilcoxon Signed Rank Test

Table S-34: Kruskal Wallis Test for Corrected Number of Input Errors per Experiment Phase

Test Statistics[a,b]		
	Training	Test
Kruskal-Wallis H	0.361	4.665
df	1	1
Asymp. Sig.	0.548	**0.031**

a. Kruskal Wallis Test
b. Grouping Variable: Participant Group

Table S-35: Wilcoxon Signed Rank Test for Difference in Corrected Number of Input Errors during the Training and Test Phase

Test Statistics[a]		
	SSD Group	No SSD Group
N	14	14
T	21.0	3.0
Z	-0.179	-1.930
Asymp. Sig.	0.858	0.054

a. Wilcoxon Signed Rank Test

S-6 Losses of Separation

Per participant, the number of LoS's over all scenarios was counted to be able to perform statistical tests on the data. The results are shown below.

Table S-36: Test of Normality for Number of Losses of Seoaration

Tests of Normality								
Category	Group	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Training	SSD	0.347	14	0.000	0.735	14	0.001
	No SSD	0.478	14	0.000	0.516	14	0.000
Test	SSD	0.372	14	0.000	0.681	14	0.000
	No SSD	0.388	14	0.000	0.684	14	0.000

*. This is a lower bound of the true significance.

Table S-37: Kruskal Wallis Test for Number of Losses of Separation per Experiment Phase

Test Statistics[a,b]		
	Training	Test
Kruskal-Wallis H	1.808	0.003
df	1	1
Asymp. Sig.	0.179	0.957

a. Kruskal Wallis Test
b. Grouping Variable: Participant Group

Table S-38: Wilcoxon Signed Rank Test for Difference in Losses of Separation during the Training and Test Phase

Test Statistics[a]		
	SSD Group	No SSD Group
N	14	14
T	22.5	25.5
Z	0.000	1.100
Asymp. Sig.	1.000	0.271

a. Wilcoxon Signed Rank Test