Design, evaluation and acceptance of a visual support tool for Air Traffic Control

Stephanie Wiechers

December 2021

TU Delft
LVNL
knowledge & development centre
Mainport Schiphol

The cover page photo is a radar screen image taken at LVNL. Credit: Ferdinand Dijkstra.

Delft University of Technology

In partial fulfillment of the
MSc Aerospace Engineering & MSc Science Communication

MSc Thesis

Design, Evaluation and Exploring Acceptance of a Visual Support Tool for Air Traffic Control

Author	S.M.G. Wiechers	Student no. 4381726
Aerospace Engineering	Prof. Dr. Ir. M. Mulder	AE C&S
Thesis Committee	Dr. Ir. C. Borst	AE C&S
	Dr. Ir. M. M. van Paassen	AE C&S
	Dr. Ir. M.C.A. van der Sanden	CDI
	Ir. G. de Rooij	AE C&S
	Dr. Ir. M.A. Mitici	AE ATO
	Ir. F. Dijkstra	[External] LVNL
Science Communication	Dr. Ir. M.C.A. van der Sanden	CDI
Thesis Committee	Drs. Ir. C. Wehrmann	CDI
	Dr. Ir. C. Borst	AE C&S

December 10, 2021

Preface

This report concerns the research conducted in the light of obtaining two masters at the Delft University of Technology. The project was conducted with the objective of researching the topic from two different angles and writing two theses: one for the MSc Aerospace Engineering, specialization Control and Simulation, and one for the MSc Communication Design for Interaction. The thesis project was started on September 1st, 2020 and will be defended on December 22nd, 2021. I would like to thank several people who have contributed to this project. First of all, my supervisors from TU Delft, Dr. Ir. Clark Borst, Prof. Dr. Ir. Max Mulder, Dr. Ir. Maarten van der Sanden, Ir. Gijs de Rooij, Drs. Ir. Caroline Wehrmann and Dr. Ir. René van Paassen, who have guided me though the process of writing a thesis, providing me with useful insights, feedback, cheerfulness during meetings and motivational talks at the right moments. Second, I would like to thank Ferdinand Dijkstra, for being the project supervisor from LVNL and providing me with data, information and practical knowledge. I am also very grateful for his constant support and enthusiasm, and for the opportunity to work together. Finally, I would like to thank all experiment participants from LVNL, and a special thank you to Jorien Dijkstra and Jonah Bekkers, who not only participated in the experiment but also gave me extremely valuable insights into the world of ACC and holding support.

Stephanie Wiechers
December 10, 2021

Contents

I

Thesis Paper

Design and Evaluation of a Support Tool for Planning Adherence While Holding Inbound Air Traffic

Author: S.M.G. Wiechers
Supervisors: C. Borst, M. Mulder, G. de Rooij, M.M. van Paassen and F. Dijkstra

Abstract – **Just before final approach, the Initial Approach Fix (IAF) can be crossed within a four-minute window around the Expected Approach Time (EAT) under current standard industry practice. As Air Traffic Control is moving toward a more strategic approach, a higher level of adherence is required. This paper introduces a tool that gives Area Control (holding stack management) insights in the present situation and in the impact of a turn to IAF command on the EAT adherence, aiming to decrease delivery window size and increase predictability. It consists of two main elements, namely a countdown timer (Delta-T) and a visual representation of the problem's margins and boundaries (EAT for Control Operation Location dots) in terms of possible turn-to-IAF locations, combined with a support element (declutter feature). Using the tool, the Air Traffic Controller can make well-sustained choices on the moment a turn-to-IAF command is given while staying the active controller. Ten professional area controllers assessed the proposed tool in a proof-of-concept experiment, from which an improved adherence of 37% to EAT followed. Additional findings suggested improved workload and a more predictable control strategy. All participants indicated a changed control strategy while using the tool and supported the implementation of the tool elements in their systems.**

Index terms– **Holding Stack, Holding Support, Air Traffic Control, Area Control, Display, Tool, Expected Approach Time**

I. Introduction

The final stage of flight, before arriving at the destination airport, is when the aircraft passes its Initial Approach Fix (IAF) and starts the approach phase. It has already descended and slowed down significantly at that moment, and is ready to start the landing process. However, in some cases there is no capacity for landing yet. In this case, the aircraft needs to be delayed, either by vectoring or by flying a holding pattern at the location of the IAF.

Standard practice at Dutch Air Traffic Control (LVNL), operating at Amsterdam Schiphol Airport, currently avoids holding under regular conditions. When, however, extreme conditions (adverse weather, emergency, delays of +7 minutes) dictate holding as the only option left to absorb delays, the systems offer little support to their operators (Personal communications with area controllers, 2020). This results in low predictability and large deviations from the planning during an already extreme scenario. Current standard practice allows for a four-minute window, from two minutes earlier to two minutes later, in which the IAF can be crossed around the Expected Approach Time (EAT), a condition that is currently not always met in extreme situations.

The goal of an Area Controller who is responsible for managing a holding stack is to empty the stack from the bottom. Aircraft are to exit the stack at the IAF. A planning is made regarding the EAT of each aircraft by a planner (there is one dedicated person responsible for the planning). The EAT is the exact moment an aircraft is to cross the IAF before continuing approach to the airport. Adhering to this planning means higher predictability in the Terminal Control Area (TMA), less need for vectoring or other means of delaying aircraft in the TMA, a traffic flow that has been optimized for runway capacity, all leading to less workload for Approach Control (APP). Hence, there are two main tasks when managing a holding stack: lowering aircraft, where a maximum of one aircraft can be present per flight level, and ensuring aircraft turn toward the IAF at the right moment to comply with the planning. For the second task, there are many factors and conditions that cause a high level of complexity. These, in turn, lead to unreliable estimations of timing, causing low EAT adherence. From analyzing real-world holding data (see LVNL (2019)), it was found that the main complicating factors are wind, aircraft type and pilot response delay. These factors cause the lap time of holding loops to vary considerably.

The need for a decision support tool is imminent: it can give accurate estimates of time where current displays require the Air Traffic Controller (ATCo) to make an estimation based on history dots, and provide a trigger as to which aircraft require the controller's attention. Requirements for a decision support tool are that it should allow the controller to validate whether their intended control action and its timing will have the intended result, in order to improve safety, workload and accuracy. The task of lowering aircraft is straightforward, using the current system's vertical view, and there is no indication of demand for additional support. The new tool discussed in this paper will focus on providing support to comply with planning, i.e., the time-location management of aircraft leaving the stack.

Earlier research in the field of holding support tooling by Mac an Bhaird (2020) was geared at aligning aircraft at higher levels in the stack. However, LVNL radar data show that the duration of a holding loop is unpredictable, meaning that the predicted EAT adherence error will increase again when lowering in the stack (LVNL, 2019). Other studies conducted at LVNL have focused on other flight phases and/or control tasks, e.g., turn-to-Instrument Landing System (see Dirkzwager (2019)) and strategic conflict handling (see Bakker (2019); Ottenhoff (2020)).

This study introduces a first concept for such a decision support tool that can be implemented as a *non-critical* (not critical for safety) augmentation to Air Traffic Control (ATC) systems in a modular and adaptive manner, meaning that it will introduce several features that can be added to the displays available at to ATC and, in the final stage of implementation, be turned on or off as the ATCo wishes to use them. The new tool is in the form of a partial solution space that hinges on an extended leg control strategy, combined with a trigger, namely a countdown timer, and an additional support element to reduce screen clutter. The main drivers of the design are increased performance, workload and solution predictability.

The concept is evaluated using an extreme yet realistic holding scenario at one of Schiphol Airport's holding fixes. The scenario involves a large amount of traffic and relatively strong winds, as to represent a situation that could strongly benefit from a decision support tool. The display design is implemented in and tested using a medium-fidelity ATC simulator called SectorX, which was developed at Delft University of Technology.

A detailed background of the problem is provided in Section II. This is followed by the functioning of the prediction algorithm in Section III that serves as the backbone of the proposed support tool concept, which is introduced in Section IV. The setup of evaluation of the concept is explained in Section V, and the subsequent results are presented in Section VI. This paper concludes with the implica-

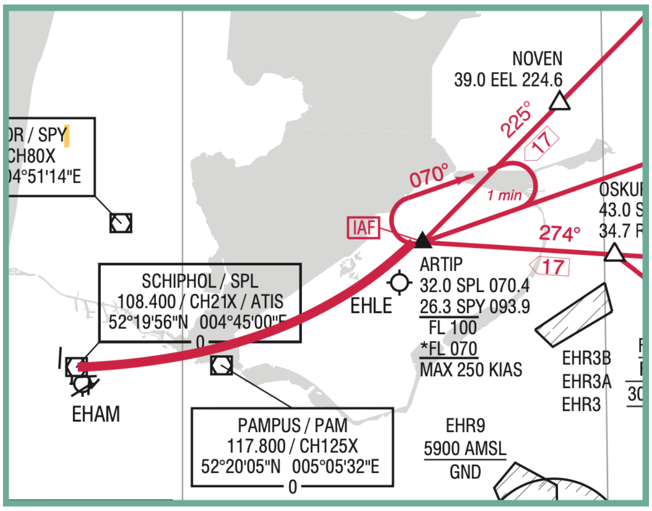


Figure 1: Overview of airspace: holding stack location ARTIP and Schiphol Airport (bottom left). Aircraft approach via ARTIP from the North-East (top right) (LVNL).

tions on ATC, a discussion on real-world conditions and suggestions for display and experiment design improvements in Section VII and the final conclusion in Section VIII.

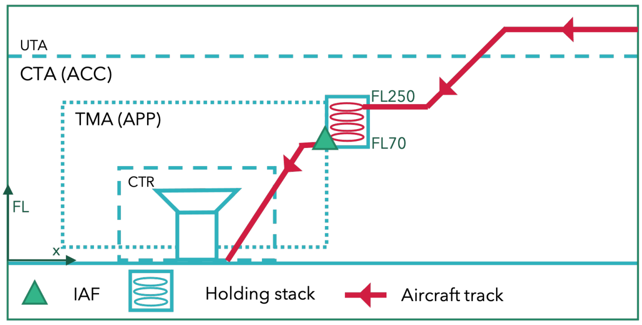


Figure 2: Overview of aircraft trajectory

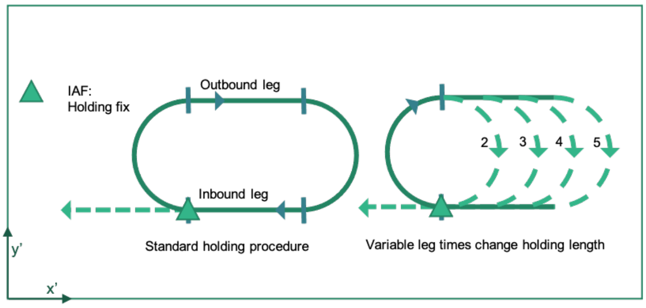


Figure 3: Theoretical holding pattern

II. Theoretical Background

Depending on traffic density, noise abatement constraints, and various other conditions, different airports have varying levels of required EAT adherence accuracy. The specific context used for this research is LVNL at Schiphol Airport. First, the specifics and challenges of holding patterns are explained. Then, the holding control task is explained and the current tools available in holding and common control strategies are described. Finally, requirements for new tooling are discussed.

A. Holding Pattern Control Task and Complexities

When an aircraft is inbound to land at Schiphol Airport, it is first guided by the responsible area controller toward its IAF. The dense air traffic and lack of support tools make it difficult for the Area Control Center (ACC) to achieve a high accuracy in EAT adherence, which currently results in a higher workload in the TMA as approach controllers have to match the incoming traffic with the landing capacity by, e.g., vectoring. It is noted that standard practice varies per airport: at NATS (London ATC), holding patterns are used to create a more comprehensible traffic situation, whereas currently, LVNL refrains from using them as a standard practice because of limited support from the system and the resulting limited predictability (Personal communications with Knowledge Development Center (KDC), 2020–2021, Personal communications with area controllers, 2020).

Figure 1 shows a part of the airspace around Schiphol Airport and a schematic overview of the routes that are flown. IAF ARTIP can be seen, and a schematic holding loop is depicted using ARTIP as its holding fix; the outbound leg of this holding pattern is flown at heading 70. The route from ARTIP to Schiphol is also depicted schematically. In Figure 2 an overview is given of the parts of the airspace an aircraft crosses and the location of the holding is shown. The IAF is shown at the point where an aircraft exits the holding and continues into the TMA, however, it should be noted that the IAF is altitude-independent.

The formal geometry of a holding pattern, as shown in Figure 3, consists of a holding fix, two legs and two turns that are identified as inbound (approaching the IAF) and outbound (flying away from the IAF). Under different conditions, for example, due to wind or shorter leg times, the precise shape of the pattern that is flown will vary. It is standard practice to fly a holding pattern with right-hand turns at most airports, including Schiphol. An aircraft enters the holding at the top of the holding stack (around Flight Level (FL) 200-250). It starts flying holding loops which have a standard time of four or five minutes: one minute for each leg below FL140, 1.5 minutes for each leg above FL140 and rate 1 turns[1]. By altering leg length, the standard loop duration can be controlled, where the minimum loop duration is determined by the turns (a rate 1 turn means 2 minutes for a 360° turn) and total loop duration by the length of the legs, as indicated by the numbers (standard loop duration in minutes) in Figure 3. A pilot has a free choice of speed in holding (usually an Indicated Airspeed (IAS) of 180–230kts chosen by traffic inbound at Schiphol (Personal communications with KDC, 2020–2021)) and therefore both leg length and turn geometry need to vary in order to meet the standard holding criteria. As the assigned holding stack controller empties out the stack from the bottom, they let the aircraft in the stack descend to lower flight levels. Aircraft leave the holding pattern at the IAF between FL70 and FL100, after which they enter the TMA as can also be seen in Figure 2.

As in real life holding speeds vary and winds are nonzero, the actual geometry, duration and size of holding patterns vary as well, as visualized in Figure 4. Analysis of data shows large deviations of holding loop duration, where loops of five to seven minutes in total are very common (see LVNL (2019)). This can be accounted for by deviations in both turn times and leg times. While flying a holding pattern, the maximum bank angle is constrained at 25° (International Civil Aviation Organization (ICAO) regulation), which inhibits flying rate one turns in almost all cases at Schiphol due to the high IAS flown close to sea level. A pilot is free to choose an IAS in holding (ICAO regulation); this, in combination with wind conditions cause the timing of flight legs to vary in practice. In current practice, a standard four-minute holding lap time is used as a rule of thumb to make time estimates, which does not align with the lap times that are flown in reality. The uncertainty regarding the lap times and the mismatch between this rule and reality is one of the main reasons a high level of EAT adherence is difficult and time-consuming to manage. Because of the uncertainty and variance in lap times, control actions have to be taken in the last loop for improved EAT adherence.

[1]See `https://www.skybrary.aero/index.php/Holding_Pattern`, accessed on 2020-10-03.

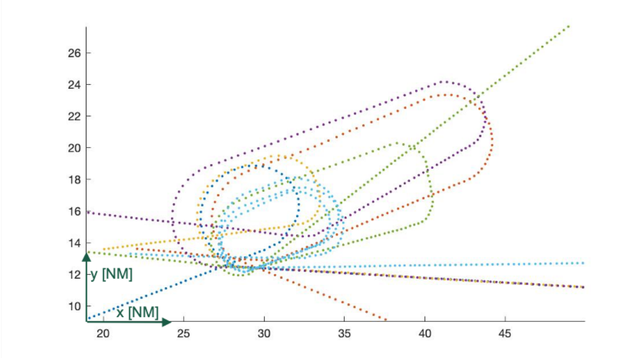


Figure 4: Plot of actual holding patterns, colored per aircraft (LVNL, 2019)

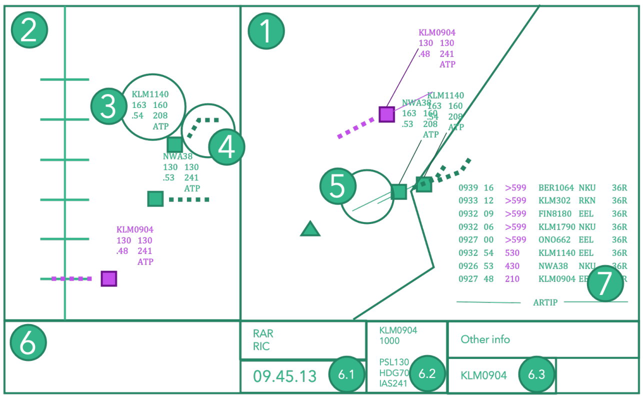


Figure 5: Schematic overview of current display. Colors have been changed for clarity.

Both on paper and in practice, the timing of one holding loop can be influenced by altering the leg times, while turns have a fixed duration due to bank angle constraints (see variable leg times in Figure 3 with indications of loop time in minutes for different leg times under the assumption of rate 1 turns). This is especially relevant in the final stage of the holding, when the EAT is nearing: then the ATCo can decide to actively influence the pattern by changing the length of the outbound leg, by giving a *turn to IAF* command. Current EAT adherence is required to have a two-minute accuracy, which is not met in some extreme cases (Personal communications with area controllers, 2020).

B. *Current Tools, Control Task and Strategies*

This section describes the current tools available to a holding stack controller, the specifics of their task and the strategies currently employed for time-location management when leaving the stack.

Current tools When multiple aircraft are flying a holding pattern above each other, a holding stack is present. In this case, an ATCo is appointed as a dedicated holding stack controller. Only when this is the case, the vertical view is enabled and the controller uses the EAT that is presented in the aircraft label as a primary source of planning.

The current display used in holding situations consists of two elements where the aircraft are seen. Figure 5 shows a schematic overview of the current display. The Plan View Display (PVD) (1) gives a top view or radar screen image, while the Vertical View (VV) (2) shows the traffic situation from the side, giving an insight into the vertical positions of the aircraft and their vertical separation. To gain information about the individual aircraft as well as the situation, the ATCo currently has access to several tools. In Figure 5 the elements are numbered; these numbers in the following paragraphs reference to these elements.

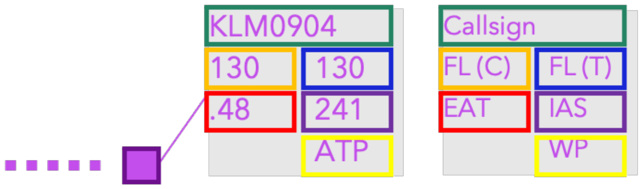


Figure 6: Schematic overview of label.

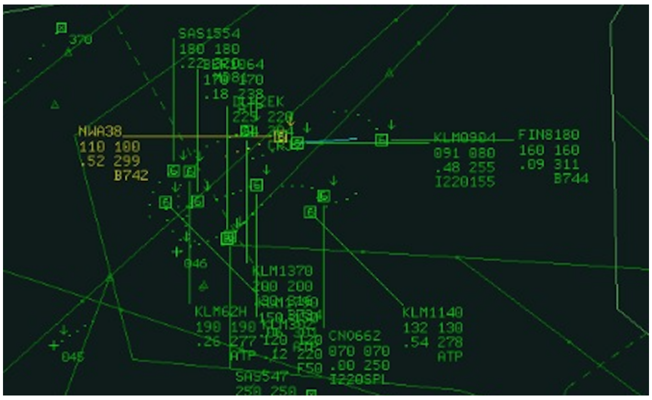


Figure 7: Example of screen clutter in PVD during holding

First of all, each aircraft has a label (3) in which the following elements are present, see Figure 6: aircraft ID (callsign), current flight level (FL (C)), target flight level (FL (T)), EAT, IAS, target heading or waypoint (WP), aircraft type (at location of WP if no waypoint or target heading are present). Each aircraft is trailed by five history dots (4), spaced two radar updates ($2 * 5s$) apart. This means the oldest history dot is the aircraft location exactly one minute ago. An optional feature that can also be used to estimate the aircraft's future location is the speed vector (5), which extrapolates current heading and speed to one minute into the future; the speed vector can be enabled for all aircraft or for the selected aircraft only.

Second, when an aircraft is selected, its color changes from green to yellow and a ribbon (6) in the bottom of the screen shows additional information on the selected aircraft; (6.1) is current time; (6.2) shows callsign, flight level selected by the pilot, heading and IAS; (6.3) shows callsign again and addtional information such as aircraft type. Only one aircraft can be selected at a time, which then changes color in every display element (PVD, VV, stack list (see below)). Aircraft can be selected from all displays.

Finally, the stack list (7) shows predicted time over IAF, EAT, predicted EAT adherence error, aircraft ID, last waypoint and runway. In the current systems, predicted time over IAF is not updated after the IAF has been crossed for the first time: the time listed in the stack list is actually the moment the aircraft started holding, i.e., the first time the IAF was crossed and the holding stack was entered. This is contrary to non-holding situations, where the EAT in the stack list is used as a primary means of planning and the predicted time over and predicted EAT adherence error are the main drivers for deciding whether to induce delay or compensate time.

As mentioned, history dots and speed vectors are used to estimate future location. In general, however, a dedicated holding stack controller turns off speed vectors for all, or all but the lowest, aircraft. This is because speed vectors are considered to clutter the screen while holding as they do not provide more information than the history dots due to the fixed speed (chosen by the pilot). An example of the screen clutter in the PVD during holding is shown in Figure 7. An advantage of the history dots is that they also provide insight into to altitude history and descent rate of the aircraft in the VV.

Control task The control task of a holding stack controller consists of two main elements. These are lowering aircraft in the stack and ensuring they leave the stack again (crossing the IAF and continuing into the TMA) at the planned time. The task description is based on Personal communications with area controllers (2020).

Since flying at higher altitudes is more fuel-efficient, pilots prefer holding at higher altitudes. However, an aircraft needs to enter the TMA between FL70–FL100 and the holding stack only extends up to FL250. Therefore, the first task of the controller is to lower aircraft in the stack to both ensure aircraft are low enough to be able to continue into the TMA and to create space at the higher levels in the holding stack such that additional traffic can enter the holding stack. Safety regulations dictate that only after a level has been cleared (the aircraft has descended to the next flight level), a command can be given to another aircraft to lower to the cleared level. For this task, the VV is mainly used.

The second task is aimed at meeting the EAT. For this, the ATCo also uses the PVD to see the radar location of the aircraft. In order to cross the IAF at the EAT, the controller alters the final loop geometry, resulting in a longer or shorter lap time. A longer lap time can either be achieved via an extended leg (see Figure 3) or by giving an intermediate heading command, essentially altering turn geometry. A shorter lap can be achieved by turning the aircraft back to the IAF before the end of the outbound leg (see Figure 3).

EAT adherence strategy The APP planner provides an EAT planning, which comprises of the exact moment in time a pilot is to cross the IAF before continuing to Schiphol. The ATCo currently has the freedom to ensure the pilot crosses this point within a four-minute window around the planned EAT, meaning maximum two minutes earlier or later than planned. Two different strategies are employed here by ATCos, depending on the individual controller.

The first approach to EAT adherence hinges on making worst-case estimations on the timing and then planning to be two minutes too late (-2:00). Then, faster reaction or flight time means the IAF is crossed earlier than expected which is perfectly within the given four-minute window. It is noted that the small time margins mean the window will not be crossed on the other (+2:00) side through a faster reaction or flight time.

The other approach to EAT adherence is the exact opposite, namely to use perfect-case estimations on the timing, and aim at two minutes too early (+2:00). Crossing the IAF two minutes too early (+2:00) means that there are two minutes to be compensated for by the ATCo, meaning there is a positive amount of time *remaining*. Then, if anything goes worse than expected, the IAF is crossed later than planned which again fits the four-minute window.

From observation in the simulator at LVNL, where an ATCo was observed who employed the +2:00 minute approach, it was found that, in fact, the deviation from +2:00 minutes (too early) from EAT is relatively small and rarely gets below +1:00 minute from EAT. This implies that the EAT accuracy can be improved by providing the ATCos with better tooling, to enable them to validate their own estimates. The reason for flying at two minutes margin is that this is seen as standard practice by ATCos, and they have indicated that additional support is a prerequisite for reducing the size of the margin by which the EAT needs to be met (Personal communications with area controllers, 2020).

The current strategies used to alter holding loop timings are visualized in Figure 8. There are three options altering loop time and one option that does not impact the EAT but the time an aircraft arrives at Schiphol. The first is (1) shortening the final holding loop by turning to the IAF early. The second option, (2) alters the loop pattern geometry, which is done by giving an intermediate heading (some heading between the out- and inbound heading) and then turning to the IAF later. This strategy is used with the idea that the aircraft has then already partially turned toward the IAF, such that it is not flying further away, and with the idea that it helps with the timing at which the aircraft arrives at the IAF (Personal communications with area controllers, 2020). Third, (3) extends the outbound leg, meaning the pilot is instructed to continue at the outbound leg heading and then gets a turn-to-IAF command.

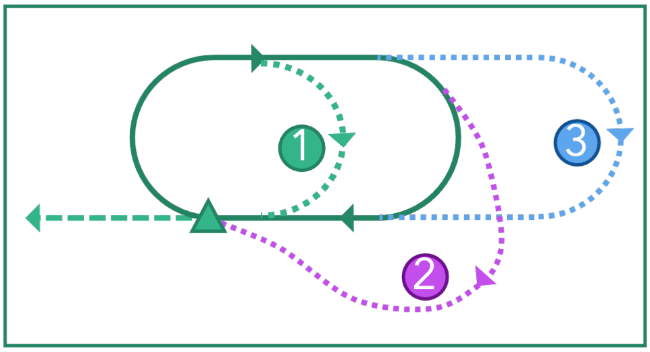


Figure 8: Visualization of strategies changing the timing of a holding loop

As an addition to the strategies altering loop times, a controller also has the option to send the aircraft directly to Schiphol, using a so-called direct-to-SPL command. This means that the pilot is instructed not to cross the IAF again, but continue straight to Schiphol. If the aircraft is expected to cross the IAF later than planned, this can be done to compensate for some of the delay. A combination of these strategies is possible, e.g., first extend the outbound leg and then fly an intermediate heading.

C. *Requirements on New Tooling*

There are a several standard practices and control strategies currently used in holding. The design of the proposed visual support tool was not constrained by current practices, shortcomings and system limitations: the design process of the tool was approached from a requirements point of view and assumed that (1) control strategies and standard practices can change under a new system and (2) current technological limitations can be overcome by implementing and adding the required features to the system. However, the standard workflow at LVNL and the manner in which people work was an important factor since all experiment participants shared this background.

Based on the objectives of ATC, namely safety, reliability, economic and environmentally sustainable operations, a support system should have several characteristics. First, it should allow ATCos to identify risks, and therefore the system should give insight into the real-world situation rather than only present a solution. Second, the most important driver in ATC are people; therefore, a system should always keep its end-user (ACC) in mind, and should be designed in such a way that it triggers people to engage with it. Especially in the domain of ATC, it is known that controller acceptance of new technologies and new technological support systems is generally on the low side (see Bekier et al. (2012)). This means that technology acceptance is a critical factor in the success of improving EAT adherence. Finally, higher EAT adherence means more predictability and opportunities for efficient flight in the TMA, improving economic and environmental operations due to the possibility for optimized flight paths.

A final requirement is that the tool is not safety-critical. Because of this reason, the tool does not include guidance on separation. Additionally, aircraft are already vertically separated in a holding stack which means lateral separation is not an issue.

III. Prediction Algorithm

This section discusses the prediction algorithms that determine the integrated information presented to the controller in the new tool. To assist the ATCo in reaching a higher EAT adherence, an approach is taken showing the margins and boundaries on the actions they can take. For larger predictability, an extended leg strategy is supported, where the ATCo's actions are putting an aircraft on an extended

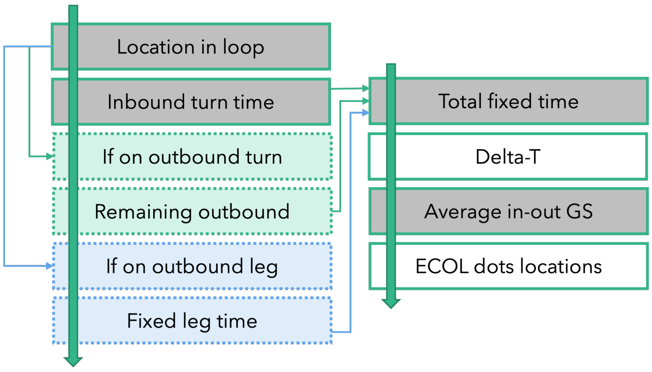


Figure 9: Flowchart algorithm

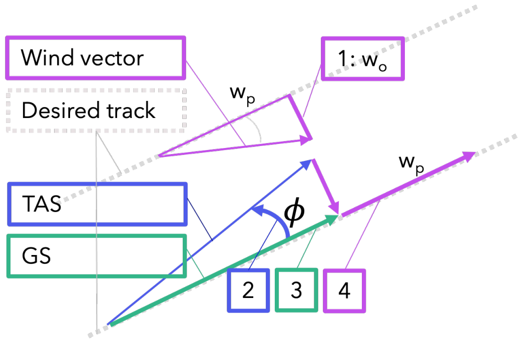


Figure 10: Graphical representation of ground speed prediction

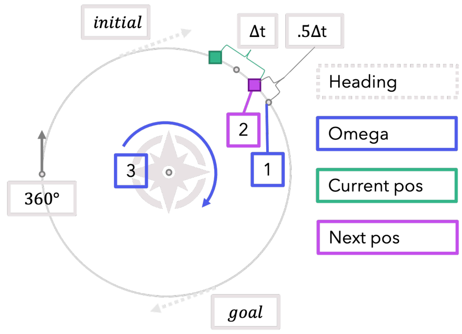


Figure 11: Graphical representation of turn time prediction

leg and giving a turn-to-IAF command. Showing the margins and boundaries means that the interface and tool merely visualize data in an integrated, logical manner such that the controller can interact with it. Essential to the problem at hand is the predicted EAT adherence based on the location at which the pilot starts to turn toward the IAF which is determined by two things: the moment the instruction is given by the ATCo and when it is followed up by the pilot. By giving a prediction about the EAT adherence, an analysis step is automated using the aircraft performance (speed, performance characteristics, descent path) and contextual factors (altitude, wind), relating these to constraints (planned EAT). Automating this analysis step comes at the benefit of speed (faster calculation) and accuracy (calculation over estimation).

To make that prediction, the holding loop is split into multiple components. The leg and turn times are calculated separately. For each heading, the predicted ground speed is based on the IAS and the wind field; the algorithm makes use of KNMI medium-detailed weather data which is the same as currently used in LVNL systems, but can easily be adapted to facilitate using more detailed wind fields. During the turn time calculation, ground speed determines angular velocity which is numerically integrated for total turn time.

A. *Algorithm functioning*

The algorithm flow is illustrated in Figure 9. The gray boxes are intermediate steps, the green and blue boxes are optional steps, the white boxes are the end product. First, the location in the holding loop is determined, which can be outbound turn, outbound leg,

Figure 10 visually supports the explanation of the steps in the ground speed prediction algorithm; the numbers in the figure correspond to the steps below. To prevent clutter, the desired track is depicted twice. It works as follows:

1. Wind component orthogonal to desired track $w_o = sin$(H(wind)-H(desired) where H() is the heading;
2. Compute angle ϕ between TAS and GS. Assumption TAS $\gg$ wind speed yields $\phi = \arcsin\left(\frac{\text{orthogonal wind}}{\text{TAS}}\right)$;
3. Along-track component of the TAS = $\cos(\phi) \cdot$ TAS;
4. Full ground speed vector = along-track TAS (green) + along-track wind w_p (purple).

The turn time algorithm, as supported by Figure 11 hinges on calculating the angular velocity based on the ground speed prediction at future locations. A numerical integration is done where $\Delta t = 5s$ (equal to one radar update), ω (number 3: omega in Figure 11) is the (predicted) angular ground speed when flying a turn, gravitational constant g =9.80665m/s^2 and V is the (predicted) ground speed.

1. Predict heading at $1.5\Delta t$ from the current moment (i.e., predicted moment) by adding $1.5\omega\Delta t$ to the heading at the current predicted position (green);
2. Predict the next heading (purple) by adding $\omega\Delta t$ to the current predicted heading;
3. Predict the next omega, by taking the heading computed in step [1] and using $\omega = \frac{g\tan(\phi)}{V}$ and the ground speed;
4. Add Δt to the turn time prediction.

These steps are continued until the difference between the next predicted heading and the desired heading is smaller than $\omega\Delta t$. The last step takes the difference between the two headings and divides them by the last predicted ω, yielding time, and adds this time to the turn time prediction.

B. *Wind*

"Wind is one of the most influential inputs in the aircraft trajectory predictions process" (Magaña, 2016, p. 58) and is in reality one of the most difficult factors to account for. The need for including wind in trajectory predictions is further substantiated by Bakker (2019) and has been indicated as the essential factor by LVNL (Personal communications with area controllers, 2020). Reynolds et al. (2013, p. 1) state: *"accurate wind information is of fundamental importance to some of the critical future air traffic concepts."* This is especially valid for the research at hand. In the specific case of holding patterns, wind works in the opposite direction on the in- and outbound legs, leading to a significant change between in- and outbound ground speed. It cannot be expected of a human to memorize complete wind fields at different altitudes that change over location and time, emphasizing the potential benefit of an integrated tool that automatically takes wind effects into account.

Two types of weather forecasts are currently used at LVNL in several support systems, where every hour a new dataset is provided with a 10-minute interval prediction for the first three hours and a 1-hour interval prediction for the following four hours (Personal communications with KDC, 2020–2021). The most detailed forecasts include information about the wind vectors at various heights and locations, but also about other weather conditions such as temperature and prediction of rain, thunderstorms, and humidity. These files contain a 4-D grid such that at every point information on these factors is present. The other forecast type is simpler and more widely used. This type contains wind and temperature predictions per flight level, which do not vary throughout the ten-minute prediction interval or over the, in this case, span of the holding area. Current LVNL prediction algorithms make use of the simple wind data. However, as this research is aimed at improving the systems, this is not proposed as a reason for not using the more detailed forecasts. Still, simple wind data will be used over full weather fields for two reasons.

First, using full weather fields increases model complexity as integration over each point in the weather grid is required. Since the duration of a holding leg is in the order of one minute and the spatial domain on which holding loops are flown is limited, while the accuracy increase is very small (order of one second). Therefore, the benefits of higher accuracy gained by using the full wind do not outweigh the increased complexity and computational power required.

Second, the update frequency of the prediction is also of influence on the level of weather prediction accuracy and the level of trajectory prediction accuracy that can be obtained (Reynolds et al., 2013, 2015). Main drivers in the accuracy of a trajectory prediction influenced by wind have been identified to be the magnitude and forecast latency (Robert and Smedt, 2013). Based on this, considering update frequency of the predictions is hourly, the impact of changing to full weather fields is outweighed by that of the prediction update frequency .

Finally, an uncertainty between the predicted and actual wind (field or vector) always remains, which can be modeled using a nominal wind value from the prediction combined with a stochastic variable (Casado, 2012). The influence of such wind uncertainties on trajectory predictions has been evaluated in (Lee et al., 2009). This influence has been shown to be very small when the forecast time and elapsed (flight) time are of the levels that are used for the holding loops in this research. Therefore, it will be assumed that the uncertainties in wind field prediction lead to a negligible trajectory uncertainty in holding loops.

C. Validation

The validation of the algorithm was done using real-world radar and simplified wind data. The data used were collected at August 10, 2019 by LVNL (LVNL, 2019); in the morning from 7:00AM to 9:00AM, wind conditions were extreme, causing LVNL to decide a holding stack needed to be installed. Heading of the wind over time and at different flight levels was between 228-237 degrees; intensity of the field was between 37-44 kts. From this dataset, seven holding loops were isolated to use as validation data. For each loop, the time the aircraft crossed the IAF at the end of the loop was registered and stored in a list of imaginary EAT data. Then, the prediction algorithm was run for three aircraft locations per loop. This was done for aircraft locations in the outbound turn at heading 30, on the verge of outbound turn and outbound leg and further down the outbound leg at 30% of its length; the EAT was set to the actual time over IAF and then the predicted turn-to-IAF location was compared with the actual turn-to-IAF location.

After validating the prediction algorithm on multiple holding loops, using different aircraft locations, it was found that overall the prediction error is very small. In twelve cases, the expected impact on EAT adherence error was smaller than one second, in five cases the expected impact was larger than one but smaller than two seconds. For each case, the expected impact was bounded by five seconds. The error was expressed in the distance between the actual location where the aircraft started the inbound turn and the predicted location where the aircraft should start the turn. In order to put this into perspective, the resulting distance was divided by the distance between the two final radar updates on the outbound leg (location where turn started and the radar update before that). This fraction is then multiplied by the radar update frequency (5s) to get an estimate of the error in time between the predicted and actual turn-to-IAF location. Since any additional distance flown on the outbound leg also needs to be covered on the inbound leg, the error estimate was doubled to give an idea of how large the impact of the error is on the EAT adherence.

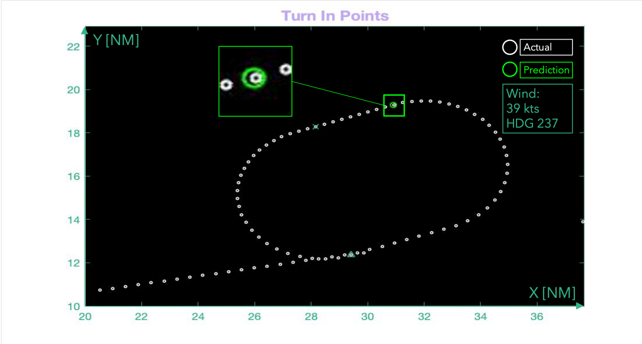


Figure 12: Validation of prediction algorithm

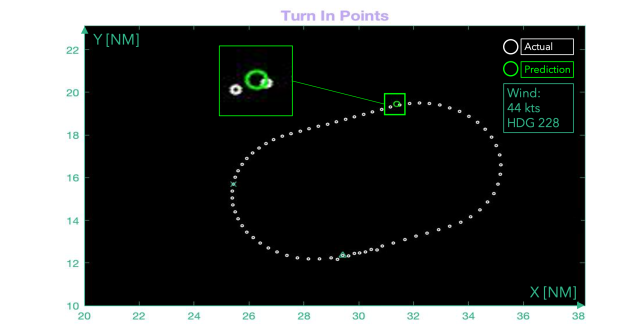


Figure 13: Validation of prediction algorithm

In Figures 12 and 13 two examples of the validation of the algorithm using real-life data are shown. The EAT is set to the actual time over IAF, the open green dot represents the turn-to-IAF location at which the EAT adherence error is predicted to be zero, the highlighted feature shows the predicted turn-to-IAF location and the actual turn in point. Figure 12 shows that the predicted turn-to-IAF location and actual turning point used to reach the IAF at the set EAT are almost exactly at the same position under the following conditions: aircraft on outbound leg, wind intensity 39 kts, wind heading 237. Figure 13 shows that the predicted turn-to-IAF location and actual turning point used to reach the IAF at the set EAT are slightly off under the following conditions: aircraft on outbound leg, wind intensity 44 kts, wind heading 228. The accuracy is still acceptable, as the distance between predicted location and actual location is 0.21 radar update which corresponds to one second in time, or an impact in the order of two seconds on the EAT adherence.

IV. Proposed Concept

The different components of the tool are outlined in this section. The components are: prediction updates without control action in stack list, difference between expected approach time (planned) and expected time over IAF (Delta-T), visualization of EAT adherence for Control Operation Locations (ECOL dots) and a feature reducing screen clutter in the PVD. The explanation of the tool components is followed by a section on the way ATCos are expected to use the tool.

A. Stack List Update

In the current systems, the next predicted time the aircraft crosses the IAF (predicted time over) is not updated. The time shown instead of the predicted time over in the current stack list is actually the time the holding stack was entered (first moment the IAF was crossed). To comply with the EAT adherence times presented in the stack list in all other (non-holding) flight situations, the predicted time over IAF is the moment the aircraft is predicted to reach the IAF again at the end of the present holding loop. In other words, it represents the predicted EAT adherence without performing any

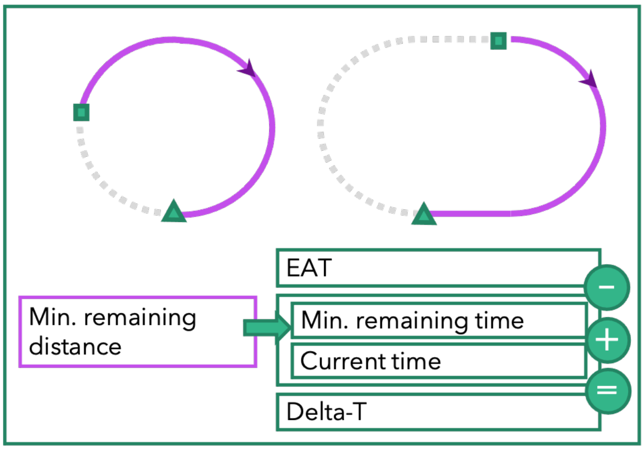


Figure 14: Visualization of delta-T calculation

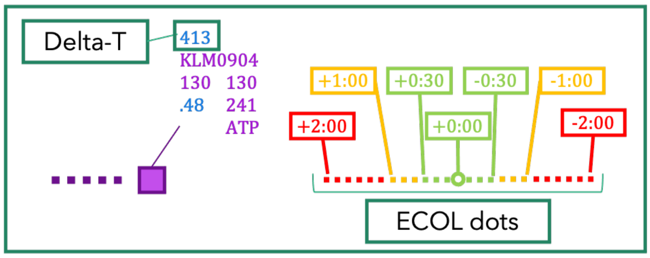


Figure 15: Visualization of delta-T and ECOL dots. Turn in early (first red dot) means +2:00 minutes too early (to be compensated for) at the IAF.

control actions, giving the controller an idea of how far in the future a control action is required.

B. Delta-T

The first innovative system addition brought by the tool is the prediction of EAT adherence error when a turn-to-IAF command would be given *now* based on the current aircraft location, see Figure 14 for an illustration of the calculation. First, the minimum remaining distance is found, which is dependent on the aircraft position. Based on this, the minimum amount of time required to reach the IAF is calculated. This is added to the current time to determine at what moment in time the IAF can be crossed. The difference between the EAT and the first possible IAF crossing moment is the delta-T, where a positive delta-T means the IAF can be crossed before the EAT.

The EAT adherence error is visualized as a timer at the top line (line zero) of the aircraft label, both in the PVD and VV, see Figure 15. This location has been chosen as it does not take up the space of any important piece of information and because additional information is often presented in line zero (Personal communications with NLR, 2021). The reason for depicting it as a clock or timer is that the EAT adherence error is depicted in a similar manner during different flight phases and is therefore in accordance with the mental model of the controller.

The delta-T shows the difference between EAT (planned) and predicted time over IAF. In this case, a positive time means that the AC will cross the IAF too early (e.g., 1027 means that the aircraft would be 10 minutes and 27 seconds too early if turning toward the IAF and starting approach at this moment). Another way to put this is that there is a positive amount of time to be compensated for. On the other hand, a negative time indicates that the aircraft will cross the IAF later than planned.

The delta-T is always shown in the label for all aircraft because this gives an overview of how far in the future a control action is required to meet the EAT. The delta-T serves an additional purpose: it provides a trigger for action while showing the boundaries to the problem. It does this by showing the remaining time up to which an action needs to be performed instead of only giving a trigger when the action needs to be done immediately, such that the ATCo is able to anticipate on the moment a control action will be required.

C. ECOL Dots

The other main aspect of the tool is the visualization of the EAT adherence of Control Operation Locations (ECOL dots). These dots visualize the predicted EAT adherence error at future possible turn-to-IAF aircraft locations. They are placed on the extended, extrapolated outbound leg (see Figure 3 for a reference to the outbound leg) and show the locations where the EAT adherence error will be between +120 and -120 seconds with 10-second intervals, see Figures 15 and 16.

The dots can be used for two purposes. The first is that the location of the dots relative to the aircraft and general holding pattern geometry can be used as a decision aid when the ATCo has to choose between putting the aircraft on an extended outbound leg or to fly another holding loop. The relative location of the dots to the aircraft gives an estimate of extended leg length, supporting the controller into deciding whether to fly an additional holding loop or continuing outbound heading.

The second and main use is to determine the optimal turn-to-IAF location and therefore control action (command) location. The middle, open green dot represents the point where the inbound turn should be started for a predicted EAT adherence error of 0s. The surrounding dots give an indication of the sensitivity of the solution and the margins. Since pilot reaction time plays a large role regarding the actual turn-in location, an ATCo can use the surrounding dots to estimate how much earlier a command should be given for the best EAT adherence. If, for any reason, it is not possible to give an aircraft a turn-to-IAF command at the point optimal for EAT adherence, the ECOL dots give the ATCo the tools to know the predicted EAT adherence error at other locations on the outbound or extended leg, too, and therefore allow for bounding the EAT adherence to different levels when it is impossible to steer at 0 seconds error.

D. Declutter Feature

Screen clutter in the PVD is already an issue while holding (see Figure 7 for a photo of the real-world PVD during a holding situation). Since the tool proposes to add even more elements to the PVD (delta-T in line zero and ECOL dots), screen clutter would be increased even further while using the tool. This would make it difficult to use the ECOL dots, as their projection is tangled with aircraft locations and labels. To solve this, all aircraft except for the selected one are faded in the PVD. The other aircraft are still visible, but much darker, making the selected aircraft, its label, and the ECOL dots clearly distinguishable and legible. In Figure 16 the feature is shown, where colors have been changed for readability.

E. Expected Strategy

In current holding situations, ATCos estimate the future positions of aircraft by (1) using a standard four-minute lap time and (2) extrapolation of history dots and speed vectors, in such a way that they gain a rough estimate of the future location. As discussed in Section II, holding lap times vary and wind causes in- and outbound ground speeds to vary significantly, making these estimates unreliable. In the current situation, there is full freedom in control strategies and controllers use very different approaches to changing holding loop times, such as intermediate headings, alternate headings, extended legs and shorter legs (Personal communications with area controllers, 2020). With the use of the tool, support is offered when flying an extended outbound leg. It is therefore expected that with the use of the tool controllers show a more predictable control strat-

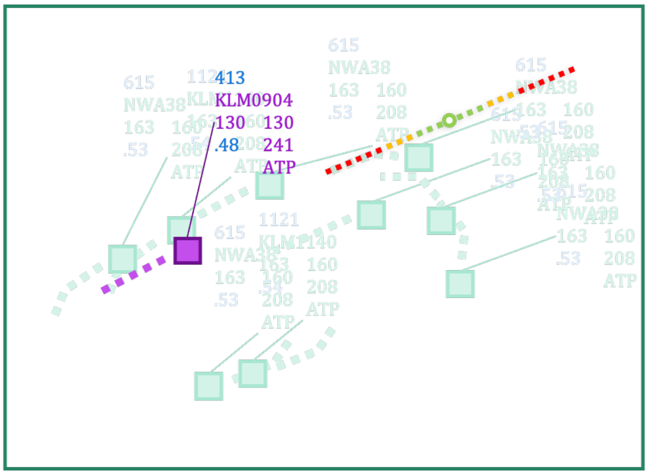


Figure 16: Visualization of PVD including tool (delta-T, ECOL dots) with declutter feature. Colors have been changed for clarity.

egy, namely one where aircraft are put on an extended outbound leg (i.e., they are told to keep flying the outbound leg heading) after which the ATCo gives a turn-to-IAF command. It is expected that with the presence of the tool, controllers will focus on this strategy and inform themselves using the ECOL dots and/or delta-T to decide when to give the turn-to-IAF command as the tool gives them insight into the expected impact of their control actions. The steps taken when using the tool are thus (1) based on tool, determine whether the current loop could be the final loop, (2) if so, lock on heading and (3) give turn-to-IAF command when EAT is expected to be met.

V. Evaluation

The main objective of the tool is to assist the area controller in the turn-to-IAF control task, with the aim of improving EAT adherence. A proof-of-concept experiment was conducted to assess the performance of this tool. Ten professional area controllers from LVNL were tasked with managing a holding stack in a simulated environment, both with and without tooling.

A. Participants, Instructions and Procedure

The control task that was tested in the experiment was the alignment of the time the IAF was crossed with the planned EAT through giving turn-in commands to aircraft that are in their final holding loop, with the explicit goal of getting an EAT adherence of $\pm 30s$ while aiming as close to zero as possible. Part of the task was to lower the aircraft through the stack, ensuring a realistic representation of workload. The other part was that aircraft needed to be put on heading 70 (extended leg) and to give turn-in commands towards ARTIP. Commands were given via the command panel. Not part of the task were: radio R/T with pilots, giving alternate headings, using different control strategies than varying the turn-in moment on the outbound or extended leg.

The study was performed in a mixed setup (two groups of five, per group within-subject setup), where all participants experienced two similar scenarios and all participants did one of the scenarios with tool and the other without. The groups were distinguished by the order of the measurement runs, with and without tool. The scenarios were different but comparable in conditions and traffic density, meaning they are expected to have a similar level of difficulty and workload. The order of the scenarios and presence of the tool was varied. No difference between the two scenarios is expected. That means that the order of presence/no presence of the tool is expected to be the only variable. The result of this is that the two groups are distinguished by the order of the tool being present only. In total two measurement runs were recorded per participant.

Before the measurement, each participant was briefed about the aim of the experiment, and told to focus on minimizing the EAT adherence error in each scenario, an explanation of the tool, and did a training run with the display. The training time is relatively short since the baseline display closely resembles the real LVNL interfaces that the participants work with on a daily basis. An overview of the experiment procedure is given in Table 1.

Table 1: Overview of experiment procedure

Task	Time
Survey	5 min
Briefing, explanation	20 min
Training	10 min
Survey	5 min
Group 1 (A1–A5): no tool Group 2 (B1–B5): tool	20 min
Survey	5 min
Group 1 (A1–A5): tool Group 2 (B1–B5): no tool	20 min
Survey	5 min
Semi-structured interview	20 min

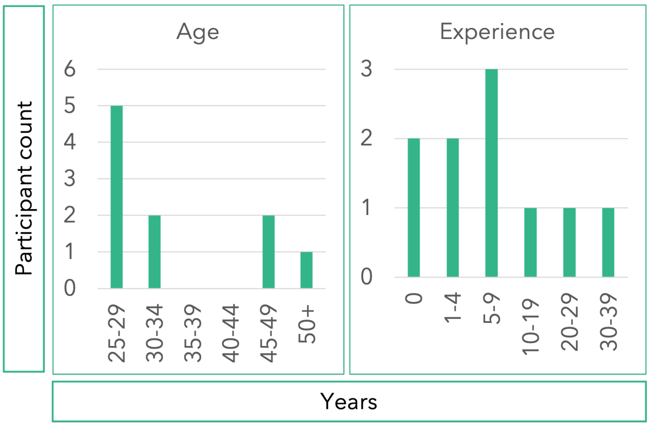


Figure 17: Age and experience distribution of participants

All participants were professional area controllers from LVNL, of whom eight are fully licensed professionals and two are in the final stage of their education, meaning that they currently only work under supervision. The distribution of age and experience (grouped) is shown in Figure 17. It can be seen that half of the participants fall within the youngest age group (25-29 years); seven out of the ten participants have less than ten years experience. At 57, retirement is mandatory which is the reason 50+ is the last age group.

B. Training Specification

During the training, the participants were presented with a low traffic density holding scenario, and had the same task (manage a holding stack with the focus on EAT adherence) as they had during the measurement runs. The training scenario was built in such a way that the most logical strategies to follow were either turning in early (before the theoretical inbound turn location of the aircraft, see Figure 8, Strategy 1) as well as extending the outbound leg and turning to IAF later (see Figure 8, Strategy 3).

C. Independent Variables

There were two independent variables, namely (1) the presence of the support tool (consisting of delta-T for turn in now in label, delta-T for holding loop in stack list, ECOL dots, and declutter in plan view) and (2) the order of the with-without tool scenarios which is a between-participants variable (Group 1 and Group 2). It was varied per participant whether the tool was present in the first or second scenario. The rest of the interface was kept constant as described under control variables.

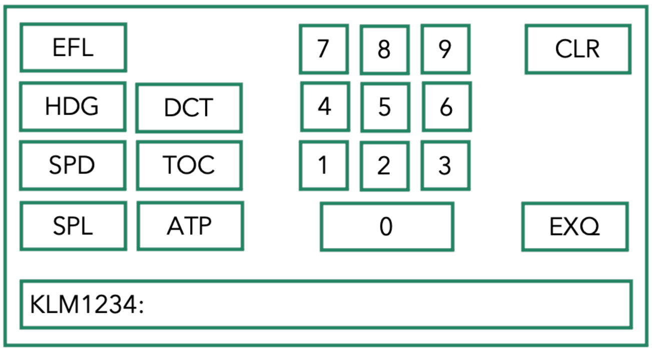


Figure 18: Command interface layout

An equal number of participants started with Scenario 1 (S1) or Scenario 2 (S2); half of each of these groups started with tooling ("yes", Group 2) and half of them without ("no", Group 1). After that, the participants were presented with the other scenario and condition than the one they started with.

D. *Control Variables*

The control variables are:

- Number of aircraft (14 in stack from the beginning);
- Scenario duration (20 minutes);
- Pilot reaction time (modeled via a Gaussian distribution, representing real-life pilot reaction times. Mean is $15s$, as this is the most common pilot reaction time, minimum is $5s$, maximum is $25s$ (numbers based on Personal communications with KDC (2020–2021));
- Wind field (intensity = 50 kts, angle difference with outbound leg = 20°), exact direction differs per scenario (S1 50°, S2 270° as to prevent bias, wind fields uniform and altitude independent);
- Available control actions and control panel layout (speed, heading, altitude, turn to IAF/SPL), see Figure 18;
- Possibility to sort stack list (using "space" on the keyboard);
- Simulation layout (radar display, presence of vertical view, aircraft with labels and history dots, clock);
- IAF (ARTIP);
- Traffic characteristics (the set of aircraft with EAT $<$ scenario duration contained: 1x A321, 1x B737, 2x B738, 1x B77W, 2x E190, 3x E75L; traffic at higher levels contained the same aircraft types; modeled via Base of Aircraft Data (BADA), see Nuic et al. (2010); order is varied; all callsigns are different per scenario. A new aircraft enters the stack at FL 240-260 every $200 \pm 100\sigma$ seconds, with their IAS varying between 180–230 kts to ensure a realistic scenario. The distribution of these variables is similar over both scenarios).
- Modeling choices made in the simulator are the standardized performance of aircraft per type according to BADA 3.1 (Nuic et al., 2010), a standardized rate of descent varying per aircraft category, using International Standard Atmosphere parameters and a uniform wind model to calculate ground speed.
- Holding patterns are modeled according to ICAO standards, such that either rate-one turns are flown or slower as turns are allowed a maximum 25° bank angle. Legs are $90s >$ FL140 and $60s \leq$ FL140

E. *Dependent Measures*

In the experiment, the effect on EAT adherence was assessed with the aim of showing a proof-of-concept. From time and location of aircraft the dependent measure is the EAT adherence error. The second effect that was assessed is the workload. For this, the corresponding measures are engagement with different screen components and corresponding actions (e.g. selection, organize elements) and perceived workload. Finally, it was assessed whether control strategy became more predictable while using the tool. This was done using the ground tracks, control action timing and type (heading 70, alternate heading, turn-to-IAF) and (if present) the ECOL dots. The following dependent measures were taken:

1. EAT adherence;
2. Label drag counts;
3. Stack list sort events;
4. Aircraft selection per display component (vertical view, plan view, stack list);
5. Subjective workload using a 5-point Likert scale;
6. Timing of control action;
7. Type of control action;
8. Ground tracks;
9. ECOL dot location and activation.

Besides the quantitative data that were recorded during the experiment, participants were asked to complete a survey and answer several interview questions. In the survey, questions were asked about the ease of use, perceived workload and perceived EAT adherence which were all presented on a Likert scale. Usefulness and satisfaction were evaluated using the scales from Laan et al. (1997). During the interview, participants were asked to describe whether they changed their strategy upon using the tool and how, and whether they saw any improvements for the tool.

The measurements were taken at the following moments: (1) before starting the experiment, participants were asked to indicate how they see the need for a holding support system and what they think is the current level of EAT adherence. (2) After the briefing and training, they answered a set of questions on ease of use, workload, perceived usefulness (qualitative) and perceived EAT adherence (in seconds). This set of questions was (3) also answered after both measurement runs. The dependent measures listed above were obtained after the two measurement runs. The reason for measuring immediately after each experiment component is that this prevents other parts of the experiment from coloring the opinion and that in this way, it was possible to track any possible change of opinion throughout the experiment.

F. *Hypotheses*

The first hypothesis is that there will be no significant difference in performance between scenarios S1 and S2. In terms of dependent measures, this means no significant difference will be found within the group/presence of tool combinations between the scenarios in terms of EAT adherence scores.

The second hypothesis is that performance will improve when working with the tool, with respect to the scenario without using the tool. In terms of dependent measures, this means that the EAT adherence error decreases.

The third hypothesis is that the workload decreases when using the tool. This means, in terms of dependent measures, that (1) aircraft selection decreases and (2) people indicate a lower subjective workload. It is also expected that (3) label drag events increase when workload decreases because an ATCos is expected to start playing with the labels and that (4) stack list sort events decrease when the ATCo does not require the stack list for gaining information about the traffic situation.

The fourth hypothesis is that the control strategy changes when using the tool, in favor of putting aircraft on extended legs (Strategy 3 in Figure 8), more often (higher predictability) and for longer distances. This means that the ground track pattern of the final loop changes. In terms of dependent measures, that the final loop outbound leg length will increase with the use of the tool.

The fifth hypothesis is that group B (tool first) will show behavior triggered with the tool, even when the tool is removed in the second measurement run. In terms of dependent measures, this means (1) a

Figure 19: Photo of experiment setup

higher frequency of using an extended leg strategy and (2) less stack list sort events.

G. *Apparatus*

The experiment is performed at the iLABs facility at LVNL on 60" screens (approximating the size of the real radar screens used at LVNL). In the room, a normal working environment as ATCos are used to is simulated, meaning that they have the freedom to talk and chat as it will increase their comfort with the situation and better approach their regular working environment. The setup is shown in Figure 19. Participants are using a mouse instead of a track ball to control the screen.

The simulation environment is the medium-fidelity simulator SectorX, an ATC simulator developed at Delft University of Technology written in Java. The package was adapted to closely resemble the specific systems that ACC at LVNL uses, to reduce training time and prevent confounds in the results attributed to learning and an unfamiliar control environment.

Control actions could be given via a control panel (see Figure 18) which closely resembled the control panel ATCos normally have at their disposal. A difference is that in a real-life situation, the actual commands are given to the pilot via voice R/T instead of via the control panel. Generally, voice R/T is associated with a higher workload.

Another difference is that in SectorX, ATCos are free to drag labels to any position, whereas in the actual systems, each label is positioned at a fixed distance and always has one of eight fixed angular positions (0°, 45°, 90°, etc.)

VI. Results

Following the experiment procedures described in Section V, the obtained results regarding performance, workload and strategy are presented in this section. Since the size of the total participant group was 10 people which is a relatively small size regarding statistical significance, and accordingly it is not possible to determine a normal distribution in every case, the statistical test performed in order to draw conclusions is the non-parametric Wilcoxon signed-rank test unless indicated differently. Advantages of using this test compared to the more common Analysis of Variance (ANOVA) test is that there is no need for the data to follow a normal distribution. For each test, an α-value (threshold) of 0.05 is used. Since the participant group was split into Group 1 (participants A1–A5) and 2 (participants B1–B5), in some cases only five data points are available per group sample (e.g., average EAT adherence per participant, perceived EAT adherence, stack list sort events per participant) which means that in these cases, it was never possible to find statistically significant results as p (test result from Wilcoxon signed-rank) will always be larger than 0.05. Where possible, the Z-statistic is given; if this was not possible because of small sample size the T-statistic is given.

Data are generally shown per participant group and in chronological order, unless stated otherwise. Group 1 consists of all participants who had no tool in the first measurement run (A1–A5) and did have access to the tool in the second measurement run. The participants in Group 2 (B1–B5) did the first measurement run with the tool and the second measurement run without tool.

A. *Difference Scenarios*

The EAT adherence errors were compared between the two scenarios within the groups using a Wilcoxon rank-sum test to assess whether significant differences in results could be found. The performance scores were compared within the groups under the same condition for the presence of the tool, while it was tested whether the different scenarios showed different results. It was found that the performance results varied non-significant (Group 1 no tool Z=0.3952, $p = .7604$, with tool Z=1.0431, $p = .2550$; Group 2 with tool Z=-.6674, $p = .5434$, no tool Z=.8912, $p = .3400$). Since the scenarios were similar and showed no significant performance differences the scenarios are not distinguished as an independent variable in the further presented results.

B. *Performance*

Both groups consist of different people, meaning that their baseline holding performance is different. The baseline level of holding performance is different for both groups. Comparing overall EAT adherence between Group 1 and Group 2 yields Z=-2.8781, $p = .0040$, no tool scenario yields Z=-2.1537, $p = .0313$, with tool scenario yields Z=-1.9689, $p = .0490$). Therefore, comparing the results between the groups does not give any indications about the performance of the tool and statistical tests will only be performed to test the effects of the display per group.

The main measure for performance is the EAT adherence. In Figures 20 and 21 histograms of the EAT adherence of each group can be seen. The scores are divided in buckets of 30s. It can be seen in Figure 20 that for Group 1, the spread of the EAT adherence error is decreases when using the tool, and that the highest counts are found in the buckets [-30 to 0]s and [0 to 30]s while there is only one score exceeding 60s. In the case without the tool, the spread is wider and deviations are larger from the EAT adherence. In Figure 21 the EAT adherence scores for Group 2 can be found with and without tool. Here, too, the spread of EAT adherence error scores is larger without the use of the tool. In the case of Group 2, the scores without the tool have less resemblance with a normal distribution, and a dip can actually be seen for scores in the bucket [0 to 30]s.

In Figure 22 all data points for the absolute EAT adherence error are visualized. The reason for showing the absolute EAT adherence error instead of the actual (positive or negative) values is that while looking at data for the full group, the boxplots show averages close to zero in all cases as the positive and negative scores cancel each other out. The absolute data helps to get a better insight in the size of the errors regardless of a participants personal preference of aiming to be earlier or later than the actual EAT. The boxplots are shown in chronological order, while the colors indicate the presence of tooling in the scenario (green for no tool, pink for with tool). All data points are visualized next to the boxes; the outliers are also visualized above the whiskers in red. It can be seen that in both groups, the level of EAT adherence improves when using the tool. A Wilcoxon signed-rank test returned a significant improvement of EAT adherence for both groups (Group 1: Z=2.1369, $p = .0311$; Group 2: Z=-2.1577, $p = .0188$).

In the data in this plot, several clusters appear to be visible as some EAT adherence scores have not been obtained. The clusters are not very outspoken at any score above $45s$. Standard practice is currently to steer toward $+120s$, which is not a point at which clusters are visible. Another interesting observation is related to relatively high scores per participant, where one person has 50% or more scores of more than $60s$ deviation from EAT. It was found that in Group 1, no tool, there was one participant who scored above$60s$ deviation for 60% of the time. In Group 2, the overall scores are higher both with and without the tool. For this group, in the first run (with tool) there were participants with 71% and 50% scores higher than $60s$, while in the scenario without tool there was one participant who did this 71% of the time. However, the participants with high scores in Group 2 were not the same over the scenarios.

In Figure 23 the mean of the absolute EAT adherence per participant is visualized. Analogous to Figure 22 where every separate data point is included, a trend can also be seen in both groups that the mean EAT adherence per participant decreases. Since the number of data points per group per scenario is only 5 in this case (five participants per group), the result is always non-significant. It can be seen that for all participants except B3, the mean absolute EAT adherence error has decreased while using the tool. For B3, the mean absolute error has increased; individual EAT adherence results for this participant show a constant high value for EAT deviation (i.e. low adherence: 43% and 50% of scores above $60s$ with and without tool, respectively), while during the interview this participant stated that "after one aircraft had a delay, it was needed to delay some other aircraft in order to maintain separation", indicating that this participant was already thinking about creating horizontal separation for approach even though this was not explicitly asked for.

In Figure 24 the expected EAT adherence from survey questions can be found. Participants had to indicate the level of EAT adherence on a 5-point Likert scale (30s, 1min, 1.5mins, 2mins, 2.5+mins). For the with-tool case, the scores are the average of the scores for full tool, delta-T and ECOL dots. 'Current' refers to the level of EAT adherence that the participants believe is met within the current systems. 'Training' refers to the level of EAT adherence the participants believe can be obtained within the experiment, both with and without using the tool, after going through the training phase. Finally, the results under 'Measurement' show the level of EAT adherence participants believe to have met with and without using the tool during the measurement runs. From this figure, it can be observed that the EAT adherence error is expected to decrease upon using the tool, both after it has been introduced in training (Z(1)=, $p = .0020$) as well as after the participants have worked with it in the measurement runs (Z(1)=, $p = .0469$).

In Figure 25 the perceived EAT adherence (same as expected EAT adherence in Figure 24) and the actually obtained EAT adherence (mean per participant, same as in Figure 23) can be found. The perceived EAT adherence scores were obtained using a 5-point Likert scale (30s, 1min, 1.5mins, 2mins, 2.5+mins); in the with-tool case the scores are the average of the scores for full tool, delta-T and ECOL dots. Since the survey questions about EAT adherence are about the bound within which the error should fall 95% of the time, in line with the current way of the allowed EAT adherence error margin, the resulting values from the measurement runs are taken such that they can be compared with these. Therefore, the value for actual EAT adherence represents the 95% bounded absolute EAT adherence error in this figure. It should be noted that in this figure, both for training as well as for measurement the scores without tool are presented first (left, green) and the scores with tool second (right, pink). Several scores overlap; it can be seen that overall, the participants rate the level of EAT adherence they think they obtained ('P') lower than the level of EAT adherence they actually achieved ('A'). This also becomes clear from the interview data, where many participants indicated that they think they scored very low on the EAT adherence (e.g. *"I don't think my performance does right to your tool, but if I've trained more on this I do think my performance will become better"*, *"I don't think the statistics will honor the potential of this project"*).

An observation that can be made when looking at the data on an individual basis, A4 and B5 rated their expected performance with and without tool at the same level. In contrast to the other participants, these participants had a better perceived performance than actual performance. For A1, this was also the case in the scenario

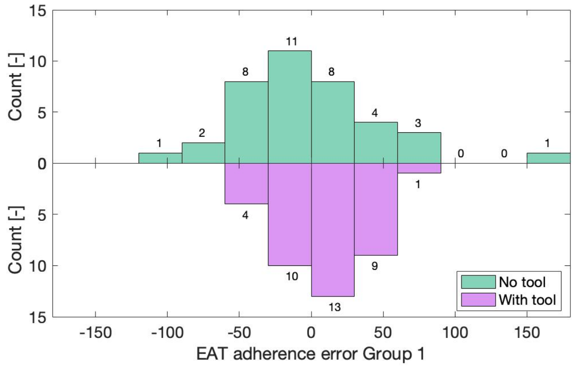


Figure 20: Histogram of Group 1

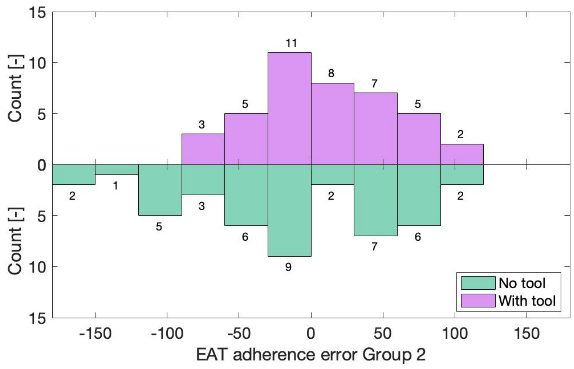


Figure 21: Histogram of Group 2

without the tool.

C. *Workload*

Workload considers both physical as well as subjective workload.

Physical workload In Figure 26 the number of label drag events is presented, where one event is the picking up and moving of a label. In both groups, it can be seen that the number of times labels were dragged (moved to another location) decreases when the tool is present. Relatively large outliers are seen at participants A3 and B5; for both participants the amount of events is large without and with the tool, but in the latter case (with tool) the amount of events is even higher. The perceived workload of participant A3 and was

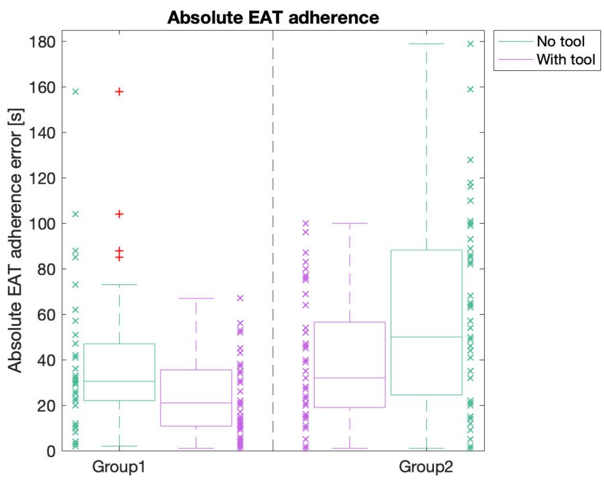


Figure 22: Absolute EAT adherence

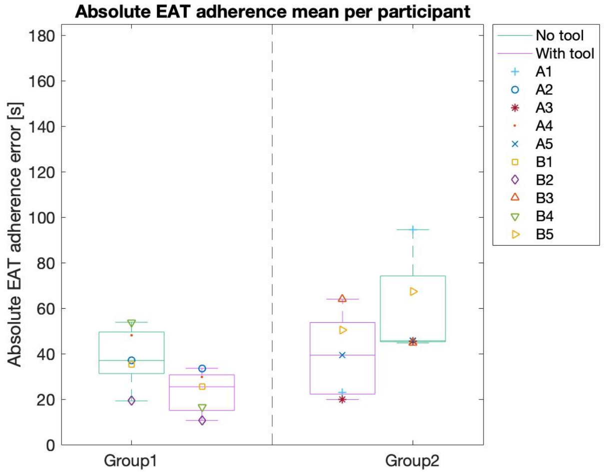


Figure 23: Absolute mean EAT adherence per participant

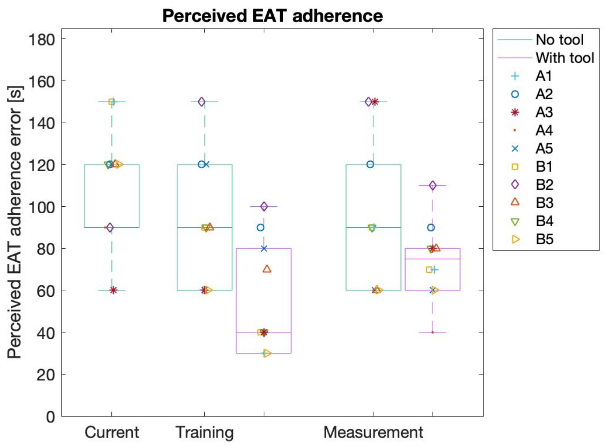


Figure 24: Perceived EAT adherence per participant based on survey

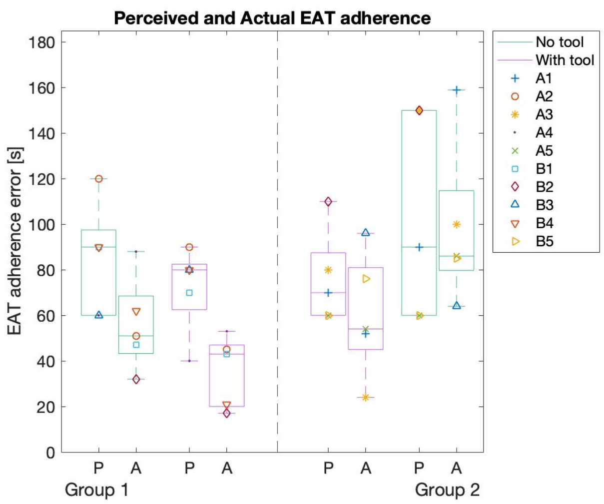


Figure 25: Perceived and actual 95% bounded absolute EAT adherence

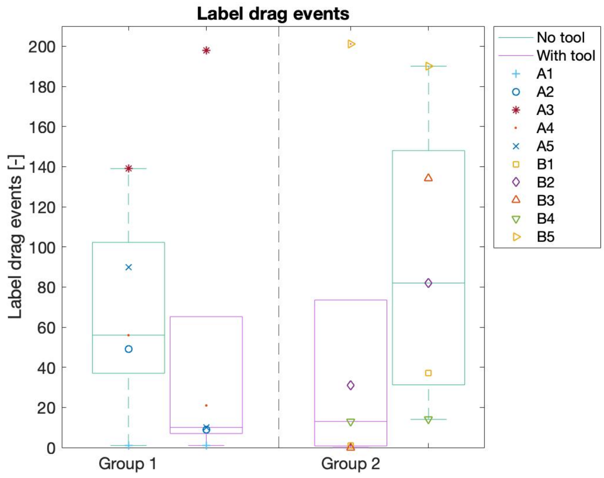


Figure 26: Label drag events

high (no tool) and neutral (tool). For participant B5 the perceived workload was low (tool) and very high (no tool); participant B5 was the only participant to indicate a very high perceived workload.

In Figure 27 the stack list sort events are found. In both groups, it can be observed that the number of stack list sort events decreases upon using the tool. From interview data, it was found that participants indicated they used the stack list less with the presence of the tool, e.g., *"I've barely looked at the stack list"*.

In Figure 28 the number of times aircraft were selected by participants per display component (vertical view, plan view, stack list) can be found. It can be seen that for both groups, the number of times an aircraft is selected in the vertical view decreases with the presence of the tool. For Group 1, nothing can be said about the number of times aircraft are selected in the plan view, but the spread between the participants in selection of aircraft is larger without the tool than with the tool. For Group 2, there is an indication of a slight decrease with the presence of the tool of selecting aircraft via the plan view. In both cases, the number of aircraft selected via the vertical view is by far the largest, followed by the aircraft selected via the plan view. In both groups, the data show that some participants barely use the plan view to select aircraft, both in the scenario with and without tool. In both groups, the aircraft are rarely selected via the stack list.

Subjective workload considers the indication participants have given on their workload.

In Figure 29, the perceived workload of the participants is shown. This was measured using a five-point Likert, ranging from very high to very low. It can be seen that without the tool, one participant rated the workload as very high, five as high, three as neutral and one as low. With the tool, one participant rated the workload as high, four as neutral and five as low.

Interview data give the following quotes regarding workload:

> *"when you have such a tool, you could make a better planning on at what times you would send [the aircraft] through and experience more calmness in that"*

D. Strategy

The total count of the use of each strategy can be found in Figures 30 and 31. In Figures 32 and 33, the final outbound leg duration can be found. It shows that for both groups, Strategy 1 (turn immediately after outbound turn) has been employed more often in the first run. For Group 1, a sharp decrease in the use of Strategy 1 is seen after the tool has been introduced (24 instances without tool followed by 11 instances with tool). For Group 2, Strategy 1 has been used 27 times in the first run, with the tool,

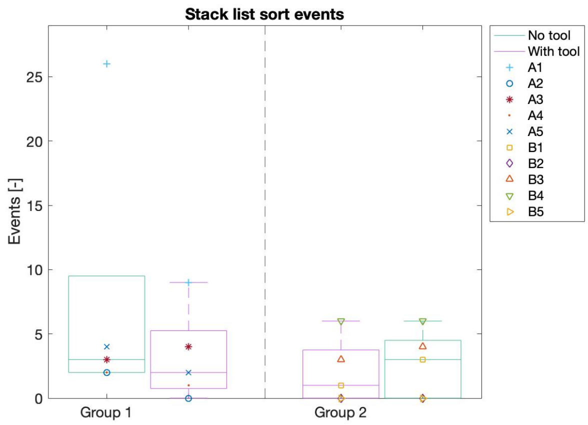


Figure 27: Stack list sort events

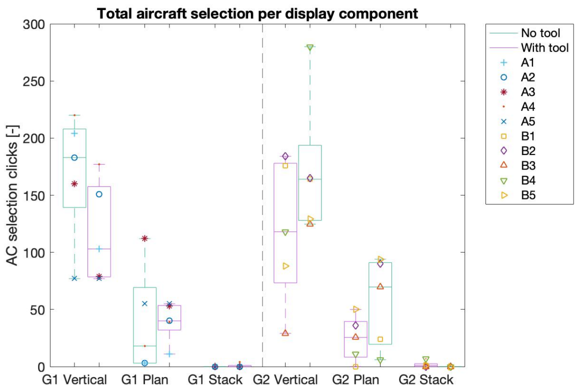


Figure 28: Aircraft selection per display component

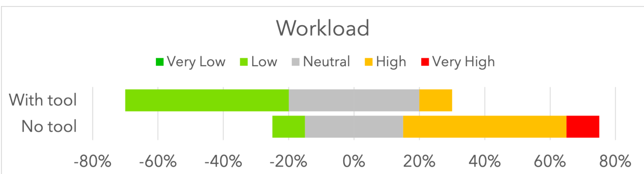


Figure 29: Perceived workload

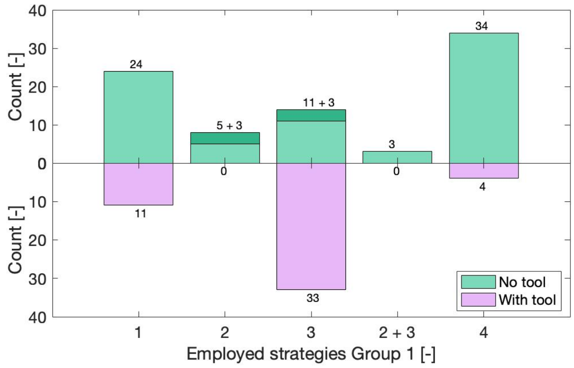


Figure 30: Count of strategy use, Group 1. In light green use of indicated strategy, in dark green the use of Strategy 2 and Strategy 3 that is also counted in Strategy 2 + 3.

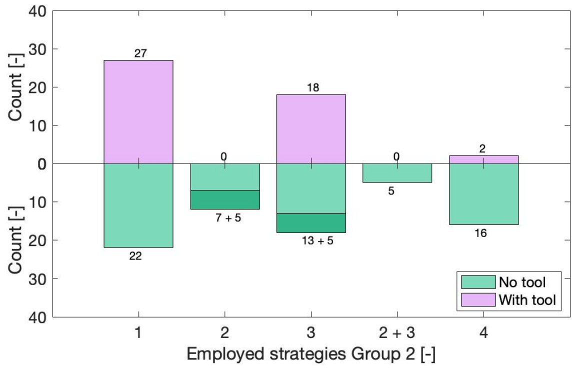


Figure 31: Count of strategy use, Group 2. In light green use of indicated strategy, in dark green the use of Strategy 2 and Strategy 3 that is also counted in Strategy 2 + 3.

and 22 times without the tool. For Group 1, the total amount of extended legs (Strategy 3, outbound leg time > 60s) was 14 and 33 without and with tool, respectively. For Group 2, this amounted to 18 times, both with and without tool. It can be seen that in both groups, both with and without tool, there were some cases of leg times $> 180s$, which results in an estimated total loop time of > 8 minutes. This occurred four times without and five times with the tool for Group 1 and four times with and six times without the tool for Group 2.

In Figures 34 to 39 examples of ground tracks that occurred during the experiment are visualized. It can be seen that during the holding loop, the wind has some influence on the ground track pattern, as the holding flown in the simulation is (virtual) waypoint-

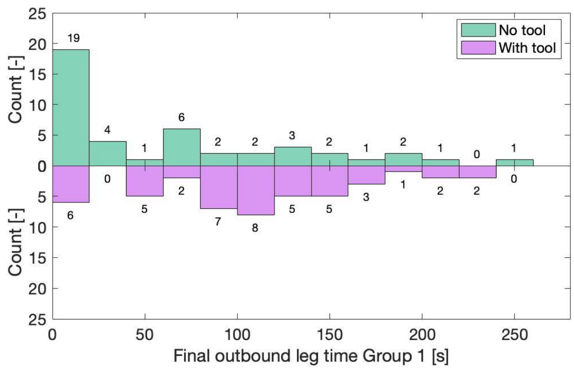


Figure 32: Duration of final outbound leg [s], Group 1

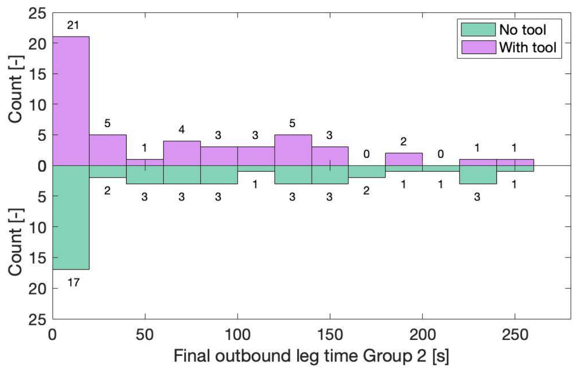


Figure 33: Duration of final outbound leg [s], Group 2

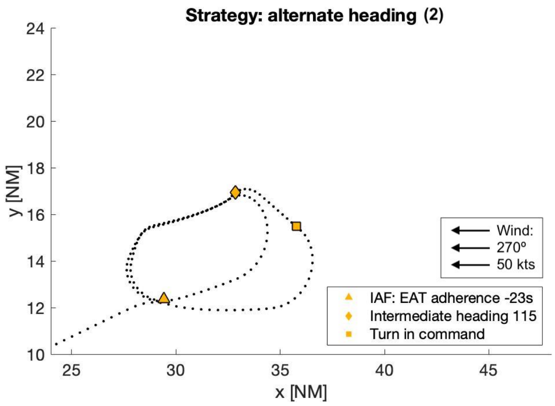


Figure 34: Ground tracks with tool; strategy alternate heading

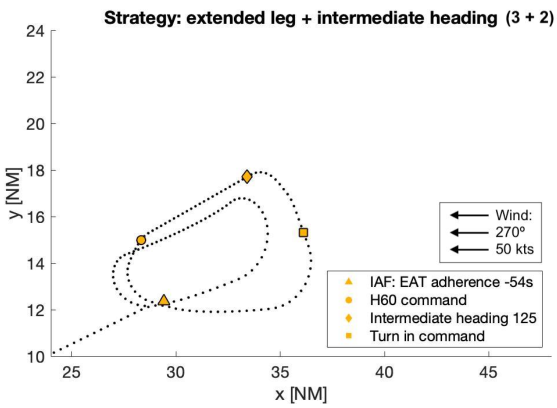


Figure 35: Ground tracks no tool; strategy extended leg + intermediate heading

based instead of track-based. When a controller locks the aircraft on heading 70 (H70), the FMS of the simulated aircraft will take the wind into account to keep this heading, such that it does not drift.

In each figure, several important locations are marked when applicable. These are: IAF location, extended leg command (e.g., H70), intermediate headings, and the turn-to-IAF (turn in) command. In the legend of these figures, the EAT adherence error is also shown.

"Depending on the stage of the flight, I switched from one part to another part of the tool. I used the delta-T to determine where the aircraft is [in the hold], how far it still needs to go, and then I combine the delta-T with the ECOL dots to decide where in my loop I am and at what moment I can start turning"

Example Ground Tracks No Tool When the tool was not present, four different strategies could be distinguished. These were: alternate heading strategy (Figure 8, Strategy 2), a combination of extended leg and intermediate heading (Figure 8, Strategy 2+3), extended leg (Figure 8, Strategy 3) and turn in early (Figure 8, Strategy 1). A decision was made to give a direct-to-Schiphol command (Strategy 4) 34 times in Group 1 without tool and 16 times in Group 2 without tool. Participant B2 did not use the direct-to-Schiphol option.

In Figure 34 the ground tracks show an example of the strategy where the aircraft is put on an alternate intermediate heading (Strategy 2). In this case, the heading chosen by the controller is H115. Analyzing the data shows that this heading is chosen randomly: when this strategy is applied, intermediate headings of 115, 125, etc. are present. After giving the alternate heading command, the controller lets the aircraft continue on this heading for a while before giving the turn-to-IAF command. In the interview data it was found that many controllers mentioned *"intermediate headings"*. The EAT adherence error in this example is -23s, which is one of the lowest EAT adherence errors present in the data for using this strategy.

In Figure 35 the ground tracks show an example of the strategy where the controller first puts the aircraft on an extended leg, but then later also decided to put the aircraft on an intermediate heading (Combination of Strategy 3 + Strategy 2). In this case, the extended leg the controller chose was heading 60, and the intermediate heading 125. The EAT adherence error in this case is -54s. During the interviews, one person said the following about estimating when to turn to IAF: *"now it is more guesstimating, a little to the left and a little to the right"*, while other participants have said similar things about having to estimate distance just with the tools they currently had. One of the things that also became clear during the interviews is that they use history dots to estimate the distance that will be covered in a certain amount of time, which means that with strong winds, the estimate will always be off.

In Figure 36 the ground tracks show an example of the strategy where the controller puts the aircraft on an extended leg (Strategy 3). The specific case of this example occurred during the second measurement run, meaning that the controller had already worked with the tool in the first measurement run where an extended heading command always needed to be H70. It can be seen that the aircraft was locked on heading 70, until the controller estimated that the EAT would be met and gave a turn-to-SPL command (Strategy 4). Generally, a turn-to-SPL command is given when someone believes time needs to be made up for (negative EAT adherence error). E.g., from interview data: *"when you give an escape [i.e., direct-to-command] to Schiphol"*. It can also be seen that the aircraft actually "crossed" the IAF 41s earlier than the EAT. In this scenario, the ground speed was higher in the inbound leg than it was during the outbound leg, which can also be seen from the spacing of the dots in in- and outbound leg. Therefore, an estimate on ground speeds made during the outbound leg results in a lower estimate than the actual ground speeds during the inbound leg if wind is not taken into account.

Example Ground Tracks With Tool With the presence of the tool no more than three strategies could be distinguished: an extended leg strategy (Figure 8, Strategy 3), a turn-to-IAF early strategy (Figure 8, Strategy 1), and direct-to-SPL commands (Strategy 4). Strategy 2 did not occur at all with the tool present, while a decrease of 82% and 83% in use of Strategy 4 was seen for Group 1 and Group 2, respectively. For each participant, the extended leg strategy has been used under the presence of the tool at least once, however, not exclusively. A direct-to-SPL command was given six times in total, of which four times in Group 1 and two times in Group 2. Upon analyzing the ground track data, it can be seen that in each of the cases where Strategy 4 is employed, the aircraft has already passed the ideal turn-to-IAF location.

In Figure 37 an example of the strategy of extending the outbound leg (Strategy 3) is visualized, showing the final two loops.

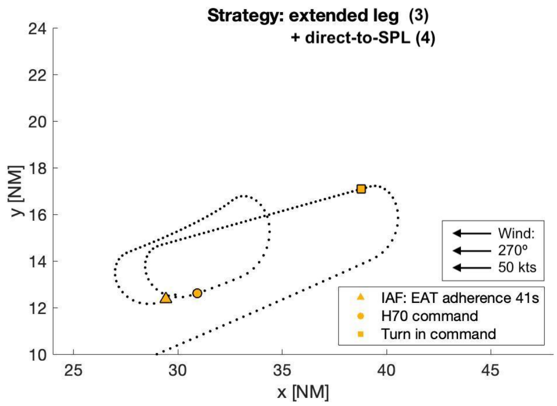


Figure 36: Ground tracks no tool; strategy extended leg

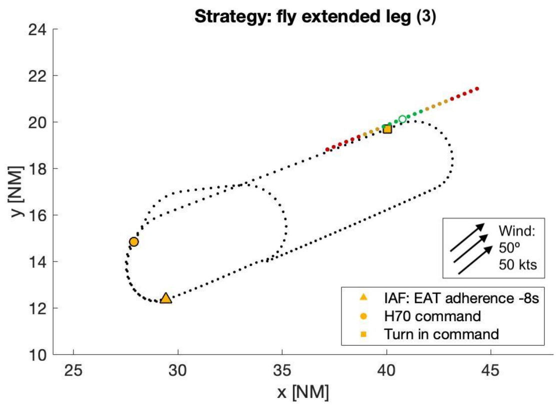


Figure 37: Ground tracks with tool; strategy extended leg

The smallest loop is the first one in time, the largest loop is the last. Here, a heading 70 command is given during the outbound turn. The turn-to-IAF (turn in) command was given slightly before the indicated "perfect" or EAT $\pm 0s$ location. The resulting EAT adherence error from this flight is 8s.

In Figures 38 and 39 an example of the ground tracks where the outbound leg has not been extended and visualized, showing the final two loops. The largest loop is the first one in time, the smallest loop is the last. It can be seen in Figure 38 that the ECOL dots appear quite far away from the standard holding loop tracks during the second-to-final loop, meaning that in the case the controller would have chosen to already lock the aircraft on heading 70, the total holding loop would have been more than doubled in size. (see ghost track in gray, Strategy 3). If the controller would have decided to do this, the additional loop would not have been required. In this case, as shown in the figure, the controller decided to fly an additional loop. In Figure 39 the result of this can be seen. After continuing to fly the regular holding track, instead of extending the outbound leg, the ECOL dots show that in order to meet the EAT, the turn in location is almost immediately after the outbound turn. From data, it becomes clear that the controller selected the aircraft and immediately after that gave the turn-to-IAF command (Strategy 1), indicating that after seeing the ECOL dots and the optimal turn-in location, the controller decided to give the control input.

It can be seen when analyzing the ground tracks of the scenarios with the tool, that two types of strategy are used: either lock the aircraft on heading 70 and turn in when the turn in location is approaching (Strategy 3), or finish the loop early (turn-to-IAF before the end of the standard holding leg, Strategy 1). It can also be seen that in some cases, controllers turn in slightly later than on the ideal location, which already became clear from the EAT adherence data in Section VI.B. When analyzing the ground tracks, it becomes clear that this indeed happens at the moments where separation between two aircraft (see interview results in Section VI.B) is low.

During the interview, participants were asked about their strategy. The following quotes present some of the participant outlooks: *"I fully focused on the dots"*, *"I think the statistics will show that the dots work less good because of system performance [the experiment setup is different than the radar screen participants normally work on] and accustomization"*,

VII. Discussion

Upon analyzing the results, it can be concluded that the trends from data are in line with the hypotheses. For both groups, performance increases with the use of tooling; workload reduces; the control strategy becomes both simpler and more predictable, and a clear difference in results between the groups was visible. Even though some results were statistically significant, in many cases it was not possible to draw any statistically significant conclusions due to the small group size. Accepting the most important hypothesis, namely

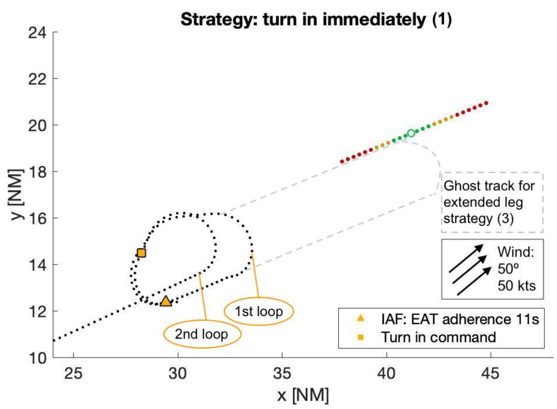


Figure 38: Ground tracks with tool; strategy turn in immediately/missed extended leg; initial dot location

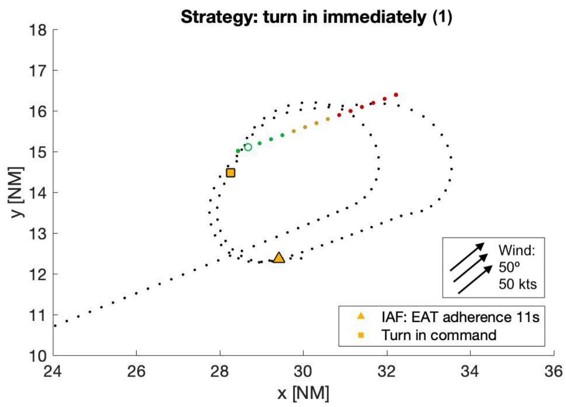


Figure 39: Ground tracks with tool; strategy turn in immediately/missed extended leg; additional loop dot location

that EAT adherence would improve under the tool, could, however, be substantiated with statistically significant data. To further test the other hypotheses, additional testing should be done. Since it has been found that the groups differ between each other, participants in a new test should also be split into two groups, which means that additional participants are required to be able to find statistically significant results. More training could possibly also mitigate the effects between the groups, as the training scenario duration was very short (10 minutes), or having participants test the tool in the simulator facility at LVNL (AAA sim) instead of in SectorX to increase the fidelity level.

A. *Implications on ATC and Aircraft*

The most interesting finding, besides a significantly relevant improved EAT adherence, is that most controllers actually perform much better than they think. One of the reasons a holding support system is required is that ATCos have indicated that it is not possible to improve on the EAT adherence without support and that they really require the current two-minute margin. In contrast, our data show that ATCos are too conservative in their estimation of their own skill in this area. A reason for this could be that holding is not common practice at Schiphol, and therefore people do not have much real-life experience on the topic. Another reason could be that the experiment tools are different from the ones ATCos work with on a day-to-day basis, and the fidelity of the simulator is lower.

The main cause for improved EAT adherence can be explained using Figure 36. The wind conditions are such that the ground speed on the outbound leg is lower than on the inbound leg. When using history dots to estimate the EAT adherence that will be achieved (the distance between oldest history dot and current location is one minute) under these type of wind conditions, the estimated inbound velocity will be slower than it is in reality. Using this estimation method under the conditions in Figure 36 would have resulted in a negative delta-T (i.e., the IAF would be crossed too late), which potentially explains the controllers decision to give an "escape" command (immediately to Schiphol) as this compensates for some delay (adds positive delta-T). However, the realized EAT adherence was +0:41s, meaning the aircraft crossed 41s too early. If, in this case, the controller would have had access to the ECOL dots to validate a decision, it would have immediately become clear that the optimal turn-to-IAF command location had not yet been crossed. A final note here is that the mentioned effect becomes larger with longer extended legs, which may be a reason for the current practice of refraining from making very long extended legs, explaining the frequent occurrence of alternate heading and mixed strategies without the tool.

In the broader context of ATC applications and the functioning of the entire organization, higher EAT adherence means the possibility to fly fixed arrival routes and more predictability in the TMA. This, in turn, creates possibilities for routes that have been optimized for noise abatement constraints and fuel efficiency.

More predictability in strategy also means two other things. First of all, it is more efficient and makes it easier to learn the task as only one strategy is required. The negative effect that is paired with that is that the task requires (and stimulates) less creativity in designing a solution, which can cause people to pay less attention and stop critically challenging what they see but rely on the system instead. Even though the tool is not safety-critical, full reliance on the advisory system is seen as problematic and should be avoided due to the criticality of the ATCo's job; participants noticed that they started to build their control strategy on the tool and that they wanted to rely on the information presented. From this it follows that the impact of using the tool on situation awareness should be assessed more extensively since it is an essential factor in a real-world situation.

Finally, a strategy of flying longer extended legs improves passenger and crew comfort: flying less turns over additional loops means less turns, which is more comfortable for passengers. When the EAT is met with higher accuracy, this reduces the need for vectoring in the TMA, further reducing the amount of turns that need to be flown.

B. *Implications of Real-World Conditions on Tool Functioning*

During the experiment, several things were different from real-world conditions. The most important things are that it is known that workload significantly increases when radio R/T communication is used for giving commands. Before implementation of the display it is essential to test the impact on workload of both using the display and having to do radio R/T communication simultaneously. It should be noted here that a more predictable strategy and less control actions also means less R/T communication.

In a broader context, a lower workload means a holding stack manager could potentially take up other tasks if also possible in terms of safety restrictions. Lower workload could also mean longer shifts, freeing up capacity in terms of available controllers or allow for a tighter planning, improving overall airspace capacity. In this case, it is difficult to draw conclusions from the results on workload due to the aforementioned radio R/T communication factor, the fact that moving labels can be both done to create a better overview or to fill the time when workload is actually low.

Another implication of real-world conditions versus the simulated experiment environment is that wind influence is more complex. The uniformity assumption probably causes a small deviation since the deviation in wind field intensity between the flight levels at which holding takes place shows variations of maximum 3 kts between levels. This implies that the variation of wind speed over the span of the holding pattern on one level will be in that order or even smaller than 3 kts. However, since wind is one of the defining factors in EAT adherence error and critical to the correct estimation of turn-to-IAF location, wind prediction should always be handled with much care and the prediction errors caused by wind field simplifications should be tracked in the future to ensure that the wind data continues to meet the standards required for the desired accuracy of flight time prediction.

Finally, the proposed tool is focused on EAT adherence and has therefore not taken separation criteria into account. This was received positively by the participants, who indicated to believe there is more power in a tool that specifically focuses onto supporting for one element (in this case, EAT adherence) over attempting to do multiple things at the same time (i.e., also providing separation support). The fact that separation is not included in the tool requires controllers to maintain separation criteria actively, instead of requiring them to trust the tool to do this. The participant feedback implied that they had a more positive attitude toward use of the tool because of this focus and because of it not being safety-critical.

C. *Display Improvements*

After the experiment, several participants indicated that they disliked having to select aircraft to see the ECOL dots, one of them suggesting *"I would actually want to be able to see the [ECOL] dots at any time, so that I can use them as a trigger and create an overview of when I need to pay attention to an aircraft"*. This is interesting because other participants indicated they mainly used the delta-T, both as a means to generate an overview of when action would be required but also to determine the turn-to-IAF command timing, one of them noting *"It is nice to have a time that counts down to the moment where you have to turn in"*.

It suggests that some people have a more visual way of working, while others have a more textual way of working. One person indicated *"I prefer the visual thing, those ECOL dots"*, another *"the delta-T in the label, that does not contribute much for me"*, contrary to people who said *"So [...] I didn't look at the dots as much and [looked at] the time in the label much more"* and *"that delta-T is what will make the holding more efficient"*. To meet the needs of both groups, in an improved version of the tool it should be possible to not only have the delta-T in line zero of the labels always activated, but also to project the ECOL dots in the VV at all times for all aircraft. For the PVD, it is suggested to keep only showing the ECOL dots for the selected aircraft and decluttering the PVD using the fade feature. Finally, it is suggested that all features can be turned on- and off individually, such that using the features is voluntary.

An additional improvement to the tool is a trigger that shows when it will not be possible to fly another full holding loop. This can best be explained using Figures 38 and 39: the final loop needed to be terminated almost immediately after the outbound leg. This could have easily been prevented by giving a prediction on whether it will be possible to fly another full loop. It is suggested to add an addition trigger to the ECOL dots that shows the predicted optimal turn-to-IAF location in the next holding loop. Based on this, the controller can decide whether or not to fly an additional loop. The predicted full loop times are best taken from adaptive previous loop data to ensure the most accurate prediction; if this is not possible due to computational power or other constraints, it is recommended to not use the standard lap time of four minutes but to make a separate prediction of the turn times and add this to the standard leg times.

Finally, modularity should be key when implementing the tool in real-life. Where most participants were very enthusiastic about the decluttering feature, one person indicated the were *"not sure whether fading is the solution"*. In a similar fashion, many participants said something in line with this statement: *"if it were up to me, you could implement this [the delta-T] tomorrow and those ECOL dots the day after"*. As people expressed preferences for slightly different ways of working, involving stronger use of different parts of the tool, it is recommended to safeguard autonomy in that sense that the ATCo is allowed to activate and de-activate tool components separately.

D. Experiment Design Recommendations

The main recommendation is that participants appreciated it a lot that they felt part of a proof-of-concept experiment and that not everything had been fully decided for them yet. One participant stated *"people are just afraid of talking to operations [the ACC group]"*, another said *"every time they fully develop a tool, and then ask for our opinions afterward. They are afraid to show us something that is not completely finished, and because of that we can't give feedback during the process. And what happens then is that someone has put a lot of effort into a tool on which the feedback is that it is wrong in the basis."* This last comment implies that this can induce a vicious circle between operations and developers, making it even harder to ask for feedback the next time. The recommendation is therefore to keep in a continuous dialogue with operations, and do this by being present on the floor and talking to people, possibly showing mockups of the innovation.

Another recommendation is that for performing an experiment in this niche group, a successful way to gain participants for an experiment has been to have people (in this case two) involved from the beginning onward in the process. These people have, in turn, enthused their colleagues to participate in the experiment as they themselves were also enthusiastic about the project. Even though this has not been tested, this seemed to follow from the process and way people talked about participating in the experiment.

The two-group setup is recommended when the goal and time allocation constraints of the experiment are similar. The reason for this is that it should be made certain that the order of presence of having the tool or not should not matter for the EAT adherence improvement. Since the impact of using the tool on EAT adherence error was the most important data type from the experiment, it is important to verify that the EAT adherence difference is not induced by the order of the presence of the tool.

Regarding the specifics of the simulation, two recommendation are made: the first on changing the pilot delay and the second concerns the holding pattern simulation setup. In the current setup, the pilot delay was always set to a value between 5–25s. During the experiment, it became clear that pilots have different reaction times for different commands. For increased realism, it is recommended to set the pilot reaction time for altitude and heading (angle) commands to 3–15s and the pilot reaction time for direct to IAF or airport commands to 5–25s. For designing this in a future setup, it is recommended to ask even more specific questions on reaction times of all possible commands instead of asking for pilot reaction times in general.

The second recommendation on the specifics of the simulation is to change the way a holding pattern is simulated. Since it was found in historic data that leg lengths vary, it is recommended to do research on the distribution of leg lengths in real life per aircraft type and airline. In the current setup, leg times were modeled according to ICAO standards, but for a more realistic representation of reality holding leg times should vary. Since some airlines do fly holding patterns based on leg time, more research is needed before this can be implemented. It is expected that with a larger variation in holding leg times, estimation of holding loop times is harder and the four-minute rule-of-thumb is further off (since total loop times vary more). It is expected that in this situation, the effect of using the tool becomes even larger.

VIII. Conclusions

A visual support tool to aid ACC controllers in improving the EAT adherence during holding was designed and evaluated in an initial proof-of-concept experiment. The display supported in determining the optimal turn-to-IAF location on the final outbound leg or extended leg in combination with a countdown timer that shows the boundaries and margins on the problem as when to give a turn-to-IAF command. Because of the way the display was designed, this nudged the controller to employ a specific control strategy, namely to put the aircraft on an extended outbound leg instead of using intermediate headings altering the inbound turn geometry. The tool was tested in an experiment at LVNL, where ATCos performed a holding management scenario both with and without the tool. The results of the experiment are promising, showing trends of improved EAT adherence (EAT adherence improved by 45% from 50 seconds to 35 seconds in Group 1 and 29% from 43 seconds to 31 seconds in Group 2), improved workload, and a more predictable strategy. Before implementation of the tool, several factors should first be investigated: the impact of real-life factors, namely using radio R/T on workload while using the tool and actual pilot delay on the EAT adherence, and the impact of a display iteration, namely always showing the ECOL dots in the VV on the EAT adherence. It is also important to assess the impact of using the tool on situation awareness, which should also be done in a future research.

Acknowledgements

The author would like to thank Ferdinand Dijkstra, LVNL and Ferway for providing historic datasets on holding patterns, and LVNL and specifically Jorien Dijkstra and Jonah Bekkers from ACC, without whom it would not have been possible to truly understand the core of the problem while holding at LVNL nor achieve the accuracy, simulation realism and task complexity that was aimed for during the case study. The author also wants to thank all experiment participants for their time, enthousiasm and valuable feedback.

References

E. S. Bakker. Design and evaluation of a visual interface for an en route air traffic control merging task. Master's thesis, TU Delft, 2019. Unpublished.

M. Bekier, B. Molesworth, and A. Williamson. Tipping point: The narrow path between automation acceptance and rejection in air traffic management. *Safety Science - SAF SCI*, 50, 02 2012. doi: 10.1016/j.ssci.2011.08.059.

E. Casado. *Identification and Initial Characterization of Sources of Uncertainty Affecting the Performance of Future Trajectory Management Automation Systems*. PhD thesis, Toulouse, FRA, 2012.

M. Dirkzwager. Design and evaluation of a visual interface for separation support in time-based approach air traffic control. Master's thesis, TU Delft, 2019. Unpublished.

J. Laan, A. Heino, and D. Waard. A simple procedure for the assessment of acceptance of advance transport telematics. *Transportation Research Part C: Emerging Technologies*, 5:1–10, 02 1997. doi: 10.1016/S0968-090X(96)00025-3.

A. Lee, S. Weygandt, B. Schwartz, and J. Murphy. Performance of trajectory models with wind uncertainty. In *AIAA modeling and simulation technologies conference*, page 5834, 2009. URL `https://arc.aiaa.org/doi/pdf/10.2514/6.2009-5834`.

LVNL. EHAM Chart from LVNL. URL `https://www.lvnl.nl/eaip/2019-08-01-AIRAC/html/eAIP/EH-AD-2.EHAM-en-GB.html`. Accessed 2020-11-02.

LVNL. Luchtverkeersleiding Nederland, dataset containing flight data around ARTIP, including holding data, 2019.

L. Mac an Bhaird. Design evaluation of holding stack management tool. Master's thesis, TU Delft, 2020. Unpublished.

C. Magaña. *Trajectory prediction uncertainty modelling for Air Traffic Management*. PhD thesis, 2016. URL `http://theses.gla.ac.uk/7700/1/2016CasadoPhD.pdf`.

A. Nuic, D. Poles, and V. Mouillet. Bada: An advanced aircraft performance model for present and future atm systems. *International Journal of Adaptive Control and Signal Processing*, 24:850 – 866, 10 2010. doi: 10.1002/acs.1176.

M. M. Ottenhoff. Effects of wind and trajectory uncertainty in a 4d trajectory management interface. Master's thesis, TU Delft, 2020. Unpublished.

Personal communications with area controllers. Interviews with area controllers at LVNL, 2020.

Personal communications with KDC. Interviews with F. Dijkstra at KDC, LVNL, 2020–2021.

Personal communications with NLR. Interviews with M. van Apeldoorn, NLR, 2021.

T. Reynolds, Y. Glina, S. Troxel, and M. McPartland. Wind information requirements for NextGen applications phase 1: 4D-Trajectory Based Operations (4D-TBO). Project report, Massachusetts Institute of Technology, Lincoln Laboratory, 2013.

T. Reynolds, M. McPartland, T. Teller, and S. Troxel. Exploring wind information requirements for four dimensional trajectory-based operations. Eleventh USA/Europe Air Traffic Management Research and Development Seminar (ATM2015), 2015.

É. Robert and D. D. Smedt. Comparison of operational wind forecasts with recorded flight data. Navigation and CNS Research Unit, EUROCONTROL, at: Tenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2013), 2013.

II

Literature Review and Preliminary Research as Graded Under AE4020

DELFT UNIVERSITY OF TECHNOLOGY

IN PARTIAL FULFILLMENT OF THE MSC
AEROSPACE ENGINEERING & SCIENCE COMMUNICATION

CONFIDENTIAL

Midterm report

Design and Evaluation of a Visual Support Tool and Exploring the Emotional Relation Between Air Traffic Controller and Interface Innovation

Author:
Stephanie Wiechers 4381726

June 11, 2021

Contents

List of Figures

List of Tables

1

Introduction

Global welfare has been rising to levels that our ancestors could not have imagined. A lot has changed since the Wright brothers first accomplished heavier-than-air flight. Once a luxury only pertained by the lucky few, rising welfare standards around the world and technological developments that cut the cost of flying have made travel by aircraft a common good. In the meantime, globalization created multinationals with offices around the world. Until the Covid-19 crisis hit in the beginning of 2020, global air traffic volumes had continuously been on the rise. A steep drop in air travel was the result, as tourists currently have to stay closer to home and online meetings have become the norm. What the future brings is uncertain to all, however the one thing that is certain is that even in this deep crisis, air travel has not come to a complete stop and is starting to show signs of recovery. Therefore, doing research on making air travel and airspace control safer and more efficient is still a relevant topic.

The final stage of flight, before arriving at the destination airport, is when the aircraft passes its Initial Approach Fix (IAF) and starts the approach phase. It has already lowered and slowed down significantly at that moment, and is ready to start the landing process. However, in some cases there is no capacity for landing yet. This can be for various reasons, from planning, procedures or capacity management, to extreme weather conditions or an emergency at the destination airport. When an aircraft has to wait in the sky, it starts flying a holding pattern at the location of the IAF.

Standard practice at Schiphol Airport currently refrains from holding under regular conditions. When, however, extreme conditions (weather, emergency) dictate holding as the only option left to absorb delays, the systems offer little support to their operators. This results in low predictability and large deviations from the planning. One of the impediments that comes with enhancing flight time predictability through 4D trajectory management is an increased workload for Air Traffic Controllers [Zeghal and Dowling]. The need for a 4D decision support tool becomes imminent for two reasons: it allows to estimate time where current displays are designed to convey aircraft location only, and it allows the controller to validate its decisions, improving safety, workload and accuracy.

1.1. Problem statement

The airspace around Schiphol airport is divided into multiple sectors. Before making a final approach, the aircraft passes the Initial Approach Fix. It is possible that at peak moments a holding pattern is installed at these locations, essentially delaying the arrival time of an aircraft. In the current operational environment these delays are more often generated using other strategies, such as vectoring. Only when weather conditions are extreme and delays increase beyond a certain level (over 7 minutes), aircraft are required to hold. Upon passing the IAF, the final approach starts, where margins are small and there is little room for error. Often, the approach controller still has to correct for aircraft as appropriate separation is required upon landing, both creating detours and flying variable routes as well as varying speed to line all incoming traffic up for safe landing. If these corrections could be prevented, predictability and safety would increase while noise pollution would go down as a result of flying fixed arrival routes.

In a general attempt to improve predictability, safety and controller workload, benefiting the efficient use of resources in airspace, landing capacity and fuels, Luchtverkeersleiding Nederland (LVNL, Air Traffic Control at Schiphol Airport) is moving from a tactical planning based on three-dimensional waypoints towards planning and controlling aircraft in four dimensions. Part of that is ensuring aircraft can adhere to their Expected Approach Times (EATs) with higher accuracy; current accuracy goal is an error bound of 2 minutes on both sides. It has been found that often there is a bias on steering toward either +2 minutes or -2 minutes per IAF, resulting in a 4-minute difference between IAFs. This research aims to reduce the error to 30 seconds on each side, resulting in a 1-minute window or a factor 4 reduction.

The problem that is to be solved by this research is:

In the current practice the traffic entering the Terminal Control Area (TMA) becomes more unpredictable at the least favorable moments: peak moments and in extreme weather conditions, when holding stacks are installed. This is because traffic is delivered with a relative difference of up to 4 minutes between the IAFs. The main reason for this problem is that there is no support offered to holding stack controllers, meaning they are not able to validate their decisions and need to prioritize safety over accuracy. The secondary problem that needs to be solved is the known resistance toward new support systems within the organization.

The main objective of this research that follows is:

To assist the area controller in adhering to a desired EAT at IAF by means of designing a tool using EID that offers support in the final phase of holding (final outbound holding leg), that is readily accepted by its end users.

Even though the problem at hand might seem quite simple at first glance, the amount of variables in researching how the current interfaces can better support area controllers in doing their job is endless. Therefore the scope of the research is actively focused on certain topics, from which sub-goals follow. One of the limitations is that the concept will be build around one IAF, assuming one fixed target runway and with that target height, as these variables do not have a large impact on the functioning of a tool or managing a holding stack according to interviews with KDC [2021].

The first objective is to gain knowledge about current LVNL systems, procedures and practices, specifically in the context of holding patterns as that is the situation in which the tool should operate. This is important as the outcome of the research should form a basis for LVNL to improve EAT adherence within the present interfaces and systems.

Secondly, the principles of Work Domain Analysis (WDA) and Ecological Interface Design (EID) will be used for the tool, where the conscious choice is made to avoid exploring other display design options as to limit the scope of the research. In essence this boils down to the interface showing the implications of certain control actions, allowing the controller to stay in charge of the situation. The associated research goal is to gain knowledge about EID and investigate in what way its principles can be applied to contribute to the research objective. An implied result of creating higher situation awareness through EID is creating an improved notion of workload. In order to bound the scope of the research, the set of considered solutions in the visualized solution space shall have the aim of being both comprehensive as well as limited to that what can be reasonably considered logical practice.

The third sub-goal is to identify uncertainties deemed relevant by ATCos. These are then to be further investigated such that they can be included in the prediction algorithm forming the basis of the solution space to be shown on the display. Other uncertainties can safely be excluded from further investigation, narrowing the scope of the research to an attainable amount.

The fourth sub-goal is to evaluate the performance of the proposed concept. Therefore, the proposed display is to be assessed in an experiment with the foreseen end users, both in terms of performance as well as in terms of user experience. Additionally, from this the essence of the proposed tool is to be derived, as to identify an aspect that can be more easily implemented in current LVNL systems.

Finally, at the basis of improving accuracy with a tool lies the willingness of end-users to use the support that is offered and the effectiveness by which it is used. Creating an interface is not enough: the people will be the ones responsible for the change. Therefore, the last sub-goal is to find out what is required for a behavioral change to occur within the present context.

1.2. Research questions

This research is twofold: it will aim at exploring how to provide the air traffic controller with an accurate insight in the current situation and the impact of certain decisions, as well as how a tool that gives these insights can be used to trigger a behavioral change. In order to reach the aforementioned objectives, several different (sub-)questions need to be answered first, taking an integral approach to the research and considering the end-user to be the most important driving factor from the beginning onward.

1. What information would support an area controller in ensuring an aircraft can adhere to a desired EAT upon passing IAF?
   (a) What are current holding procedures at LVNL?
   (b) What are current best practices in controlling a holding stack (at LVNL)?
   (c) What are the physical as well as regulation-based (Netherlands) constraints in the holding procedure?
   (d) Are there single realistic historical scenarios that can be used to represent the situation accurately enough from controller perspective? (As to limit the scope and let depth prevail over breadth; if this is the case, choose a scenario.)
2. How is trajectory prediction data best presented using the principles of Ecological Interface Design?
   (a) What are the principles of Ecological Interface Design?
   (b) What does the current state-of-the art (display) look like?
   (c) To what level of integration should parameters be presented, ranging from just the visual outcome to only the building blocks?
   (d) What would be the ideal-world display look like (disregarding current system limitations)?
   (e) How to best present a version of the ideal-world information system within the current system, using minimum alterations to the current presentation?
3. What are the uncertainties software and ATCos have to cope with in case of trajectory predictions, and what level of detail is needed in modeling them?
   (a) To what level can wind field predictions be simplified to ensure the wind component of the prediction error for minimum time until IAF is within a standard deviation of $\sigma = 2s$?
   (b) How should the influence of pilot reaction time be modeled and accounted for?
   (c) How to predict flight leg and turn times, keeping calculations simple and total error within $\sigma = 6s$?
4. How does the proposed support system influence controller performance?
   (a) What are relevant confounding factors in the experiment and can these be mitigated?
   (b) What independent variables should be measured to evaluate EAT adherence?
   (c) How should subjective questions about the tool be asked to prevent bias?
   (d) How should subjective questions be qualified and assessed to draw a conclusion from them?
5. How can a behavioral change be triggered in ATCos at ATC, by means of or stimulated by a novel graphic support interface (technological innovation)?
   (a) What is the role of learning when a technological innovation is introduced?
   (b) What is the framework of teamwork, collaboration and interaction within LVNL?
   (c) What is an area controller's mental model regarding current operations and how does this influence innovation processes within LVNL?
   (d) What role does emotion play in determining how people work with (new) systems?
   (e) Is it possible to determine favorable circumstances for innovation?

The research questions above can be used to reach the sub-goals in the way that they (1) ensure the foundations for the project are present, (2) make sure the framework in which the final objective can be met is laid out, (3) provide the necessary insights and limitations to the scope, in such a way that the initial concept can be created which is then (4) evaluated and from which final conclusions can be drawn, and (5) lays the foundations for actual change.

As other research has already been (successfully) performed on visual support tools for ATC using EID to improve situational awareness, workload, efficiency and performance in other phases of flight, this research is seen to be feasible as long as the scope is limited to encompass an amount of topics and parts that can be researched in the period envisioned for the research. In other words, in consideration with LVNL and TU Delft a set of realistic assumptions and essential details is composed to ensure that the research outcome is useful and complete enough, as well as delivered timely. The long-standing contribution that this research will contribute to the body of knowledge are insights in how an EID-based visual support system can improve the performance of the controller, offering further scientific backbone to the concept of using EID for supervisory control tasks in ATC, and insights on how the Technology Acceptance Model (TAM) can be extended with information and participation in the context of a technology *development* process.

1.3. Report structure

This report is structured as follows. First, a description of the work domain of air traffic management in general and LVNL in specific is given. A detailed explanation is given on holding patterns, both in theory and in practice, followed by the current available tooling and common control strategies. In Chapter 3, the methods, assumptions (based on literature review) and algorithms used for trajectory prediction are described. Chapter 4 discusses the current layout and its shortcomings, gives display design principles and a rationale, and explains the proposed interface concepts. Then Chapter 5 shifts the focus from the technical requirements of the problem towards the people needed to materialize upon the EAT adherence improvement tooling. It presents several background concepts that are relevant to introducing a technological innovation and explores how these theories can be used to create favorable circumstances for promoting change while designing an innovation. Finally, Chapter 6 describes the planned future work, both in terms of testing the tool and concepts proposed throughout the report as well as the steps that need to be taken to bring this research to a successful end.

2
Work Domain Analysis

The work domain of Air Traffic Control (ATC) has become more complex every year with increasing volumes until the Covid-19 crisis, resulting in a crowded airspace and high workloads. In this chapter, the way ATC works is explained, including the various roles people have and the way the airspace is divided. The focus here will be on LVNL and Schiphol Airport, since the research conducted aims to propose a display that LVNL can implement. The research aim, to increase Expected Approach Time (EAT) accuracy at the Initial Approach Fix (IAF), is a result of the extremely tight planning of Schiphol Airport caused by the high traffic density and noise constraints. As the conditions are different at other airports around the world, their plannings are also different and often do not require such high accuracy. This makes it less relevant to look into the practices at other airports and their Air Traffic Control.

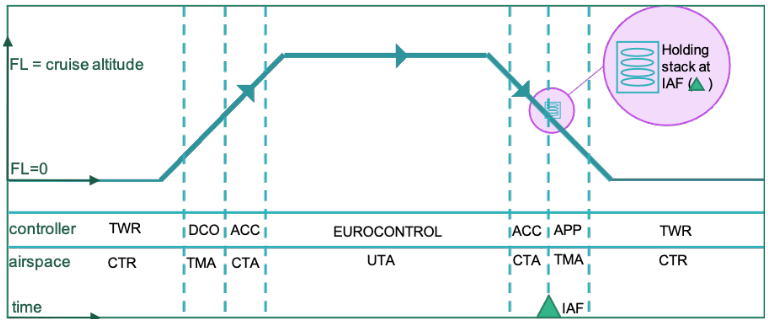
After explaining the airspace structure, the chapter starts zooming in more and more, first by narrowing down to holding patterns, then by zooming in on the objectives of air traffic control, and finally by describing in detail the functioning of LVNL and specifically the way they manage holding patterns. These sections are provided to give a complete background on the subject, in such a way that the different types of intended readers will all be knowledgeable of the theoretical framework of holding stack management after reading this chapter.

2.1. Airspace Structure and Control Flow

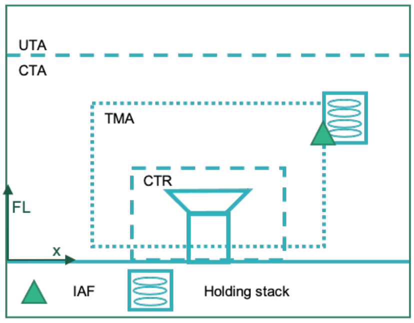
The global airspace is divided into multiple sections, based on both their height expressed in Flight Level (FL, the aircraft's altitude at ISA pressure per 100ft) and their lateral position, considering points of interest on the ground such as distance to airports and military terrain. Per country the exact limits of each boundary differ; the following numbers refer to the division in the Netherlands. First there is the Upper Control Area, which considers the airspace above FL245 and is, in the case of the Netherlands, controlled by Maastricht Upper Area Control Center, which is part of EuroControl. The airspace below is divided into Flight Information Regions, where in this case the EHAA FIR is considered, spanning the Netherlands and a piece of the North Sea. Area controllers from LVNL are responsible for this part of the airspace; it is sub-divided into five smaller regions to ensure a manageable workload. Within the FIR, surrounding any airport is the Terminal Control Area (TCA), responsible for guiding the aircraft through the approach phase, until they are handed over to the Tower (TWR) which is the control zone on the ground, responsible for taxi, landing and departure. In Figure 2.1a the different parts of airspace are shown in relation to the different stages of flight (not to scale), and in Figure 2.1b the division of the airspace around an airport is schematically shown[1].

Analogous to the different parts the airspace is divided into, different controllers are responsible for the air traffic at different locations and with that, different phases of flight. Looking at it from the perspective of a pilot going from A to B, the aircraft starts of at an airport where the Tower (TWR) is responsible. After take-off, the aircraft enters the Terminal Maneuvering Area and is controlled by a Departure Controller (DCO) during the first phases of climb. When it crosses the boundary of the TMA, it enters the CTA and an area controller is responsible for ensuring separation between aircraft in these zones. As the aircraft continues to climb towards its cruising altitude, it will exit the CTA and enter the UTA where in the case of the Netherlands and Europe,

[1]Source: interviews with LVNL, Knowledge Development Center (KDC) and the aeronautical information packages from LVNL, see `https://www.lvnl.nl/eaip/2021-05-06-AIRAC/html/index-en-GB.html`



(a) Airspace division per stage of flight



(b) Airspace division around an airport

Figure 2.1: Schematic overviews of airspace division, not to scale

Eurocontrol takes over the responsibilities. When an aircraft is approaching its destination, the same stages are passed in the reversed order, and the same divisions between controllers are made [Borst, 2019]. However, there are two differences: first of all, a holding stack may be present at the Initial Approach Fix (IAF) where the aircraft has to wait before entering the TMA. In that case, an area controller will be assigned to this holding stack, regulating traffic. Around Schiphol Airport the area controller will feed the aircraft into the TMA and therefore determines at what moment they pass the IAF, while around Heathrow, the Approach Controllers (APP) pull aircraft from a holding stack when they have enough capacity for landing. This is immediately the second difference: upon departure a DCO is responsible while for approach the role is called APP. However, at Schiphol these tasks are not separated; generally there will be four approach/departure controllers working at the same moment, managing the traffic in the TMA, plus one planner[2].

2.2. Around the IAF: Holding Patterns

When an aircraft is inbound to land at Schiphol, it is first guided by the responsible area controller toward its Initial Approach Fix (IAF). The dense air traffic and lack of support tools make it difficult for the Area Control Center (ACC) to achieve a high accuracy in Expected Approach Time (EAT) adherence, which currently results in a higher workload in the Terminal Control Area (TMA) as approach controllers have to match the incoming traffic with the landing capacity by e.g. vectoring. It is noted that standard practice varies per airport: at NATS (London ATC), holding patterns are used and create a more comprehensible traffic situation; currently LVNL refrains from them as a standard practice because of limited support and limited predictability[2]. This section will explain how holding patterns work and how they can be used to influence the moment an aircraft passes the IAF.

There are three holding stack locations around Amsterdam Schiphol Airport, seeFigure 2.2. The red lines represent the routes by which aircraft fly toward the holding, plus an indication of the holding pattern geometry. The zoom panel shows the holding at IAF ARTIP, as well as a possible route to Schiphol. The theoretical geometry of a holding pattern, as shown in Figure 2.3a, consists of a holding fix, two legs and two turns. Under different conditions, for example due to wind or shorter leg times, the precise shape of the pattern that is flown will vary. It is standard practice to fly a holding pattern with right-hand turns at most airports, including Schiphol. An aircraft enters the holding at the top of the holding stack. It starts flying holding loops which have a standard time of four or five minutes: one minute for each leg below FL140, 1.5 minutes for each leg above FL140 and rate 1 turns [SKYbrary, a]; the standard IAS flown at holdings around Schiphol is 220kts. However, each pilot is allowed to choose at what speed she flies a holding pattern and therefore not only leg time, but also turn time varies. As the assigned holding stack controller empties out the stack from the bottom, she lets the aircraft in the stack descend to lower flight levels. Aircraft leave the holding pattern at the Initial Approach Fix (IAF) between FL70 and FL100, where they enter the TMA.

As in real life holding speeds vary and winds are nonzero, the actual geometry, duration and size of holding patterns vary as well, as visualized in Figure 2.3b. Both in theory and in practice, the timing of one holding loop can be influenced by altering the leg times, while turns have a fixed duration due to bank angle constraints. This is especially relevant in the final stage of the holding, when the EAT is nearing: then the ATCo can decide to actively influence the pattern by changing the length of the outbound leg, by giving a *turn to IAF* command. Current EAT adherence is required to have a 2 minute accuracy, which is not met in some extreme

[2]Source: interviews with LVNL and F. Dijkstra, KDC

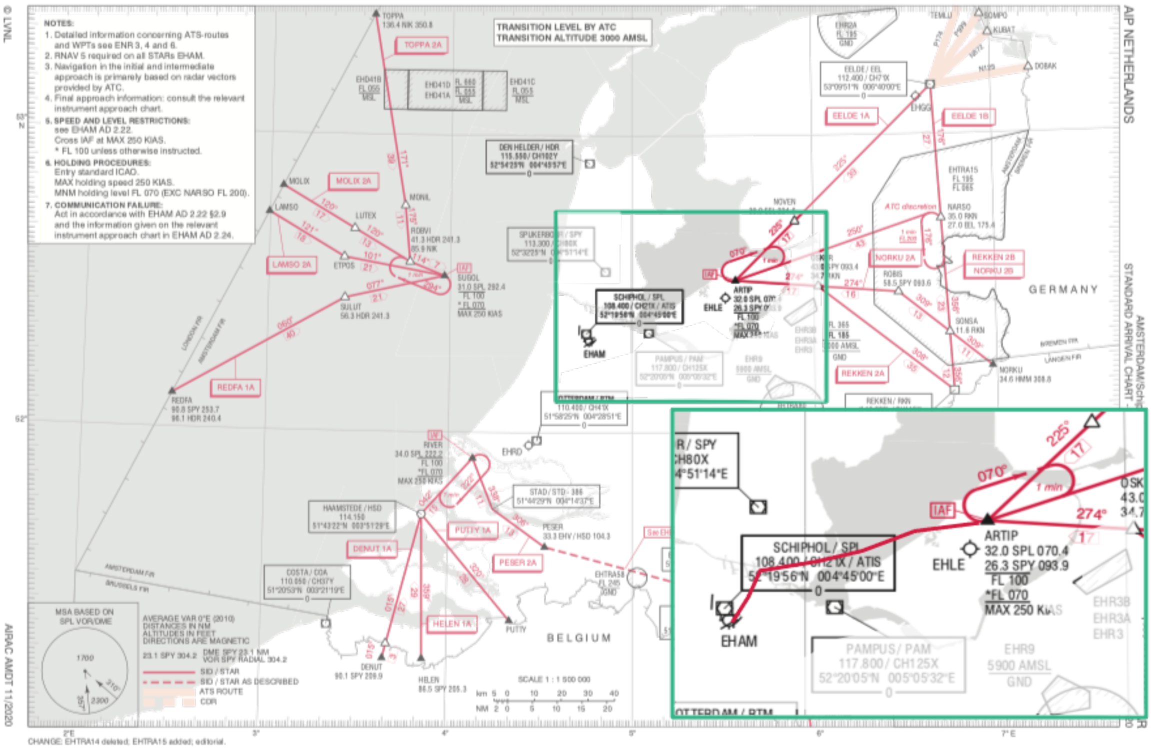


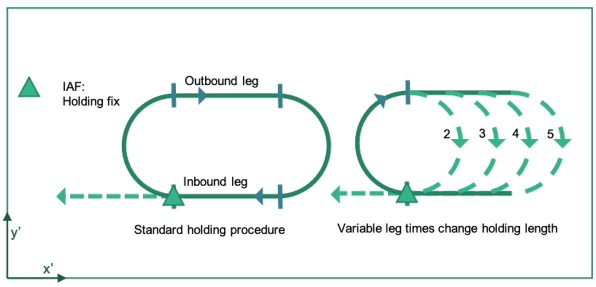
Figure 2.2: Holding stacks around Amsterdam Schiphol Airport LVNL

cases[2]. From the perspective of Approach Control in the TMA, a higher level of adherence to EAT is desired most, followed by a target velocity (preferred TMA entry velocity is 250kts) and target flight level (preferred flight level is determined by IAF/runway combination)[2]. All of these things result from the short time-span aircraft spend in the constrained space of the TMA, creating limited room for deviations and flexibility.

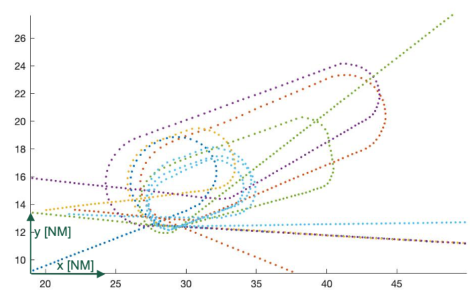
2.3. Objectives of Air Traffic Control

The International Civil Aviation Association (ICAO) defines its vision for Air Traffic Management (ATM) systems as:

> *To achieve an interoperable global air traffic management system, for all users during all phases of flight, that meets agreed levels of* safety*, provides for optimum* economic operations*, is* environmentally sustainable *and meets* national security requirements *[ICAO].*

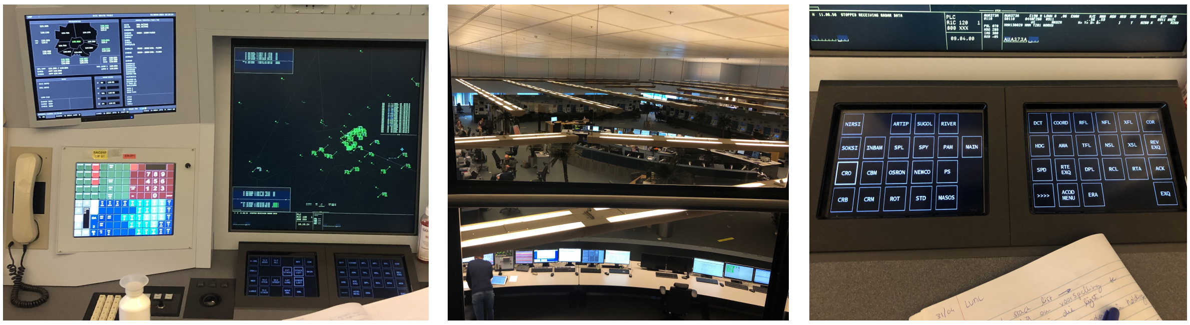


(a) Theoretical holding pattern



(b) Plot of actual holding patterns [LVNL, 2019]

Figure 2.3: Theoretical and practical holding patterns



(a) ATCo workplace (b) The “zaal”: work domain of ATCos (c) Command panel

Figure 2.4: Work domain of the ATCo[4]

At LVNL these pillars are translated into their corporate strategic goal: “becoming the world’s best air traffic control organization in terms of safety, people and delivery reliability [LVNL, 2020a]”. These objectives are actively pursued with initiatives on improving safety through encouraging employees to identify risks and develop safety and security management systems, revise departure and arrival routes to reduce noise and emission effects for the people living around Schiphol, and implement time-based separation to increase capacity and allow for higher delivery reliability [LVNL, 2020a].

It is in line with these objectives to steer toward a higher EAT accuracy. Based on the above objectives, a support system should have several characteristics. First, it should allow air traffic controllers to identify risks, and therefore it is important that the system gives insight into the real-world situation rather than only present a solution. Second, the most important driver in ATC are people: therefore a system should always keep its end-user (Area Control) in mind, and should be designed in such a way that it triggers people to engage with it. Especially in the domain of ATC, it is known that controller acceptance is generally on the low side (see Bekier et al. [2012]). That means technology acceptance is a critical factor in the success of improving EAT adherence.

The Future of Holding at LVNL At Schiphol airport, the goal is *not* to turn holding into a standard practice. The main reason for this is efficiency, as holding costs additional fuel[3]. Simultaneously to this research, other research projects are executed that have the aim to improve EAT adherence in other (non-holding) situations. Together, an overall higher EAT adherence will allow for flying fixed arrival routes in the TMA, which is the main goal of LVNL[3]. These routes will allow for more efficient flight trajectories in the TMA and lower noise pollution, which is one of the main drivers for the limitations on traffic at Schiphol airport.

2.4. Work Environment

In the current situation, the Air Traffic Controller has its own workspace in the so-called “zaal” (i.e. room), shown in Figure 2.4b. The different teams, controlling different parts of the airspace all have their own physical location. On the left is approach control, sitting in a circle such that it is easy to speak to everyone else working at APP at that moment as it is such a small space where they have to manage the traffic. In the middle at the straight desks are the military controllers. Closest to the photographer is a planner workbench, just like the oval workbenches to the right of the military controllers. Finally, in the back of the room and to the right are the ACC desks. These are not positioned in a circle but next to each other, as for ACC it is more common to only need to work together with the people controlling the airspace right next to theirs.

Each ATCo has their own workplace, which is flexible and dependent on the part of the airspace they are managing at that moment. The amount of ATCos working at a moment in time is dependent on the occupancy rate of the airspace. At peak hours, sectors become smaller and more controllers are needed; when holding stacks are installed due to e.g. weather conditions or an emergency, separate holding stack controllers are assigned who then also get their own desk. The layout of the radar screen varies, depending on the type of activity: in a regular situation, so without holding, the radar screen only shows the top view (like in Figures 2.4a and 2.6a) and the stack list, see Figure 2.5a. When a couple of aircraft enter the holding, but the ATCo still manages those aircraft next to the other traffic in its sector, the ATCo often chooses to use

[3]Source: interviews with F. Dijkstra, KDC

the extended stack list shown in Figure 2.5b. Finally, if a dedicated holding stack is installed, the ATCo will have access to the vertical view.

The vertical view is only available to a dedicated holding stack controller. According to ATCos at LVNL, it is not possible to manage both the sector traffic and a complete holding stack. As holding occurs in extreme situations, it requires a lot of communication with the pilots, meaning that the ATCo has to explain to all pilots entering the holding what is going on and why they have to hold. This is done for safety reasons, such that the pilots themselves can decide to either hold or deviate to another airport (e.g. considering a limited amount of fuel taken aboard for holding). Besides the workload, which could in part be improved by better support systems but not relieved completely as communication with aircraft is still of vital importance, the second argument given has to do with the space on the screen. Managing a larger sector requires quite a lot of space on the radar screen, such that the vertical view makes it more difficult to manage the traffic as part of the traffic entering the sector is seen much later, inducing additional workload. One final comment that is made regarding these considerations is that the information was obtained from interviews with LVNL and therefore contains a bias toward the limitations of the current situation.

Besides the radar screen which is explained below, the ATCo has several tools that can be used for giving commands and communicating with other Air Traffic Control Centers. Using the phone EUROCONTROL can be contacted. Many other features exist, but are not relevant to flying holdings and will therefore not be discussed here. The most relevant feature outside the radar screen is the command panel, as shown in Figure 2.4c. This is used to give all commands, for example target FL, target velocity (SPD), waypoints, but also to enable different views on the radar screen. Besides entering these commands in the command panel, the ATCo also gives the command to the pilot via radio.

2.4.1. Features of the Radar Screen

Stack lists The layout of the stack list is normally as follows, from left to right: expected IAF crossing - EAT (planned) - EAT inaccuracy - aircraft ID or flight number - waypoint - runway. When the ATCo enables the extended stack list, the current and cleared altitude, in flight levels, are also presented in the list, in that order. Besides the addition of FLs, the major difference between both stack lists can be found in the sorting order. In the regular case, the aircraft are sorted on EAT (planned). In the case of the extended stack list, the order is based on the FL (current).

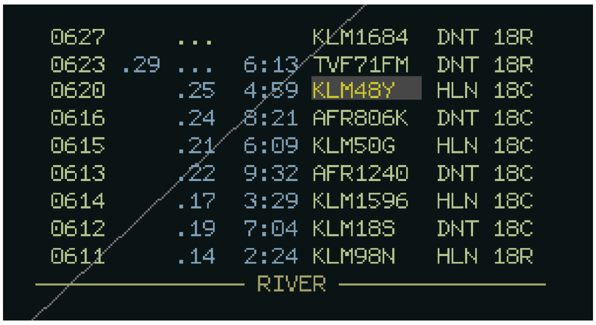
The sorting of the lists is such that the first aircraft to continue to Schiphol (SPL) is on the bottom. That means the first EAT or the lowest FL is on the bottom. The lists do not automatically re-sort; on the desk there is a button which sorts the list again when clicked. From observing ATCos it became clear that this sorting is something that they do routinely and seemingly without actively thinking about it: re-sorting the list is therefore rule-based behavior rather than knowledge-based.

Finally, there is one very important thing that the reader should note here. The predicted IAF crossing times presented in the second column of the stack list stop updating after the IAF has been crossed. In other words: as soon as the aircraft enters a holding pattern, the prediction times are not updated anymore.

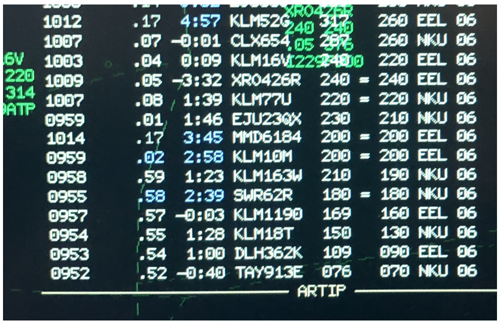
Vertical view When a dedicated holding stack controller is present, the vertical view can be made active. As seen in Figure 2.6a, the screen becomes very cluttered as the holding fills up. In fact, at the moment the still is taken, there are 17 aircraft present in the holding space. This makes it extremely difficult if not impossible to distinguish the aircraft in the top view: since aircraft are separated by height and fly a similar track, it makes more sense to look at their positions from the side. A caption of the vertical view at the same moment is shown in Figure 2.6b. This obviously gives a better overview of the situation than the top view, improving the controller's situational awareness.

History dots and speed vectors Based on radar updates (approximately 5 seconds), the ATCo has additional tools to get a better idea on the past and future trajectory of the aircraft. In Figure 2.6 one can distinguish (if looking closely) five dots behind the aircraft. These represent the last five radar positions, and can be used to get an idea on how fast the aircraft is going and whether it is e.g. descending. The ATCo also has the option to enable a speed vector, which is a line from the aircraft toward the predicted location in five radar updates based on current heading and velocity.

Label Another feature from which the ATCo gets a lot of information is the aircraft label. Its layout is given below. The EAT is the amount of minutes past the closest hour, meaning that if it is currently 8:53 .54 implies

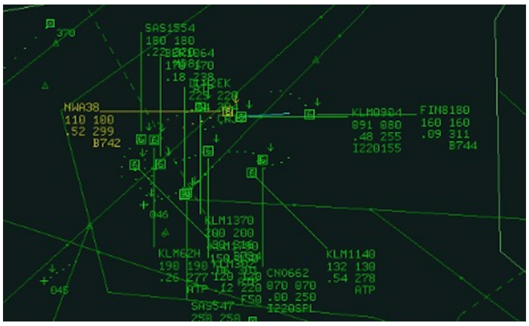


(a) Stack list at RIVER

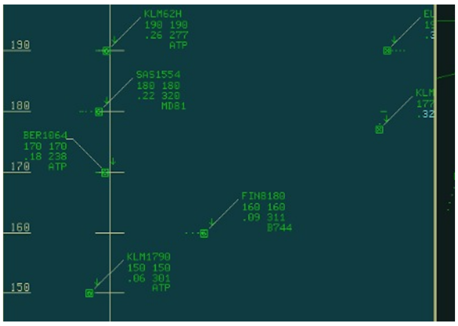


(b) Extended stack list at ARTIP

Figure 2.5: Normal and extended stack list



(a) Top view



(b) Vertical view

Figure 2.6: Radar screen close-ups when holding stack is present

that the planned EAT is in one minute, and .01 implies it is at 9:01. Finally, the bottom right entry either shows the next waypoint (e.g. ATP for ARTIP or SPL for Schiphol), or the aircraft type (e.g. B737). It is possible to move the aircraft labels and make them readable again in the top view when the holding stack is full and they are overlapping like in Figure 2.6a.

Aircraft or flight ID		`KLM1790`	
FL (current)	FL (cleared)	`132`	`130`
EAT	Speed (kts)	`.54`	`278`
	WP or type		`ATP`

2.4.2. Holding Stack Control Task Strategies

There is a couple of standard practices and control strategies currently used in holding. While the design of the visual support tool will not be constrained by current practices, shortcomings and limitations, it is good to be aware of the standard workflow at LVNL to gain a better understanding of the way people work.

Stack list versus vertical view The stack list and extended stack list are used as a primary measure on planning when aircraft pass the IAF. The ATCo gets an overview on who needs to pass first, EAT adherence error (not updated in holding), and in the case of the extended stack list whether the pilot has already lowered enough to continue to approach. When a dedicated holding stack controller is installed, the vertical view is enabled and the controller uses the EAT that is presented in the aircraft label as a primary source of planning.

History dots and speed vector Speed vectors are not used by all ATCos. In general, when someone is working as a dedicated holding stack controller and the vertical view is present, speed vectors are turned off as they are considered to clutter the screen at that moment. This can be further explained by the speed constraints present in holding, meaning that the different aircraft will not have an extremely large variation in speed - and an aircraft does not vary its own speed significantly during holding. For this reason, the speed vector does not give more information than the history dots, in fact, it gives less information. That is the case as the history dots also provide insights in the altitude history of the aircraft, providing insight in both speed and descent rate.

EAT accuracy The ACC planner provides an EAT planning, which comprises the exact moments in time a pilot is to cross the IAF before continuing to Schiphol. The ATCo has the freedom to ensure the pilot crosses this point within a four-minute window around the planned EAT, meaning maximum two minutes earlier or later than planned. Two different strategies are employed here by ATCos, depending on the person.

The first hinges on making worst-case estimations on the timing and then planning to be two minutes too late (-2:00). Then, if anything goes better than expected, the IAF is crossed earlier than expected which is perfectly within the four-minute window given.

The other strategy is the exact opposite, namely to use perfect-case estimations on the timing, and aim at two minutes too early (+2:00)[5]. Then, if anything goes worse than expected, the IAF is crossed later than planned which again fits in the four-minute window.

From observation in the simulator at LVNL, it was found that in fact the deviation from +2:00 minutes (too early) from EAT is relatively small and rarely gets below +1:00 minute from EAT. This implies that the EAT accuracy can be improved by providing the ATCo with better tooling, to enable them to validate their own estimates, as well as by exploring how the tool can trigger a behavioral change as to change the aim from +2:00 minutes from EAT to 0:00 minutes or exactly at EAT. The reason for flying at two minutes margin is that this is seen as standard practice by ATCos, and they do not wish to refrain from keeping this safety measure without additional support.[6]

[5]Cross the IAF two minutes too early (+2:00) means that there are two minutes to be compensated for by the ATCo, meaning there is a positive amount of time *remaining*

[6]Source: interviews with LVNL (2021)

3

Flight Time Prediction Methods

This chapter will discuss the different parts that form the basis of the trajectory prediction algorithm. Mondoloni and Rozen [2020] argues for the ambiguity that is present due to the uncertainty on decisions made in the future. Because of this, the formulas used in trajectory (and corresponding time) prediction may be well substantiated, while an uncertainty regarding the timing of certain control actions remains. In this chapter, both the formulas and corresponding assumptions are discussed, as well as other uncertainties and assumptions that are inherent to the model. First, the influence of wind and wind prediction accuracy on trajectory prediction is discussed, including the functioning of the algorithm used for estimating the ground speed under different headings. Then, the approach for turn time prediction is given, which is one of the most vital parts in predicting remaining holding loop duration. Another essential part in this prediction is the estimation of fixed leg time that is induced when the pilot is on the outbound leg. Next, the assumptions regarding pilot reaction time are presented.

3.1. Prediction Components

In Figure 3.1, the different parts of a holding pattern are outlined. These are the parts in which the loops are split up for predicting the time it will take to fly the total remaining loop. (1) and (3) are the turns, which means a turn time prediction is required. This is explained in Section 3.3. (2) and (4) are the out- and inbound legs, for which the time it takes to fly a leg has to be predicted. This is explained in Section 3.4. Knowing the time it takes to fly a turn and being able to predict the flight leg time, it is possible to make a prediction of the time it takes to reach the IAF. This concerns two things: first, the trajectory that is taken in order to reach the IAF as soon as possible. This is outlined in Figure 3.1, where S1 represents the scenario where the aircraft is already on the outbound leg, which requires a certain amount of flying "back" over the inbound leg after a turn has been made. S2 represents the case where the aircraft is still in the outbound turn, which requires the outbound turn to be finished first before the pilot can start the inbound turn. Finally, to resemble the mental model of the ATCo as closely as possible, the prediction will make use of a delay in pilot reaction time, as explained in Section 3.5.

The full prediction algorithm has been tested using Matlab. This was done under multiple wind fields, varying in intensity as well as heading, and setting pilot reaction time in the prediction equal to zero. In Figure 3.2 a plot of the predicted turn-in locations versus real-life data is shown. The white dots represent a plot of one (final) holding loop that was flown in reality. Only the final loop was plotted, explaining the lack of incoming tracks. The wind field vector at that time could be best estimated by a heading of 252 and speed of 19.2 kts. A fictitious EAT was added which coincides with the moment the aircraft passed the IAF in reality. The yellow and green dots represent the predicted turn-in locations to meet reach the IAF at a certain predicted time, where the open green dot (see zoom in figure) represents the EAT and each neighboring dot introduces a 10-second margin. Since all white dots are positioned one radar update (5 seconds) from each other, and the open green dot *almost* coincides with the actual turn-in location. Therefore, the objectives in RQ 3(a) and 3(c) are met using this prediction algorithm while testing the aforementioned conditions. Running multiple tests with various (realistic) scenarios all showed similar results.

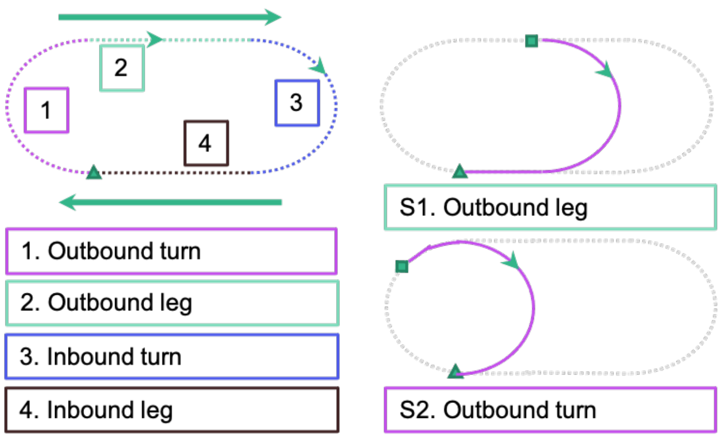


Figure 3.1: Different parts of a holding loop and possible trajectory alterations

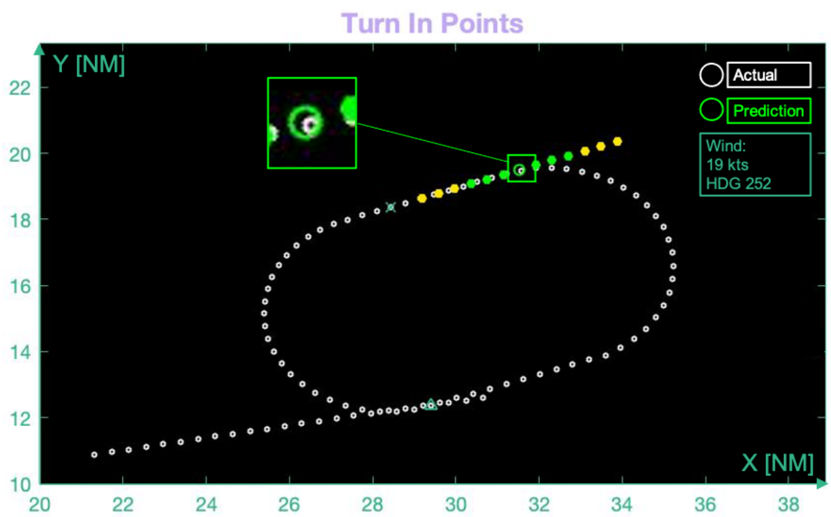


Figure 3.2: Plot of algorithm prediction versus real turn-in moment

3.2. Influence of Wind

Wind is a highly influential component in trajectory predictions [Magaña and Juan, 2016] and is in reality one of the most difficult things to estimate. The need for including wind in trajectory predictions is further substantiated by E.S.Bakker [2019] and has been indicated as the essential factor by LVNL[1]. Reynolds et al. [2013] state:

> *"accurate wind information is of fundamental importance to some of the critical future air traffic concepts"*

This is especially valid for the research at hand. In the specific case of holding patterns, the influence of wind on the in- and outbound legs works in opposite directions leading to a significant change between in- and outbound ground speed. Even though an ATCo familiarizes herself with present wind fields before controlling traffic[1], it is impossible to memorize complete wind fields at different altitudes that change over location and time.

Detailed weather forecasts from KNMI are currently used at LVNL in several support systems, where every hour a new dataset is provided with a 10-minute interval prediction for the first three hours and a 1-hour interval prediction for the following four. These forecasts include detailed information about the wind vectors at various heights and locations, but also about other weather conditions such as temperature and prediction of rain, thunderstorms, humidity. Taking the full weather fields into account will lead to a higher accuracy in leg and turn time prediction but will also increase model complexity as integration over each point in the weather grid is required. Since the duration of a holding leg is in the order of one minute and the spatial domain on which holding loops are flown is limited, the benefits of higher accuracy by using the full wind may not outweigh the increased complexity and computational power required. Additionally, even while using a highly detailed grid, the update frequency of the prediction should be considered regarding the level of weather prediction accuracy and the level of required trajectory prediction accuracy [Reynolds et al., 2013, 2015]. Main drivers in the accuracy of a trajectory prediction influenced by wind have been identified to be the magnitude and forecast latency [Robert, 2013]. As a result, this research will make use of a constant wind vector in its trajectory predictions.

Finally, an uncertainty between the predicted and actual wind (field or vector) remains, which can be modeled using a nominal wind value from the prediction combined with a stochastic variable [Casado et al., 2012]. The influence of such wind uncertainties on trajectory predictions has been evaluated in [Lee et al., 2009]; it has been shown to be very small when the forecast time and elapsed (flight) time are of the levels that are used for the holding tracks in this research. From this it will be assumed that the uncertainties in wind field prediction lead to a negligible trajectory uncertainty in holding loops. Additionally, considering the goal of the research at hand, adding a stochastic variable does not contribute to the quality of the predictions made and is therefore excluded from the scope.

Even though the current ground speed and TAS are known, the ground speeds when flying a turn vary (as in practice there is always some wind), just like the ground speeds on the in- and outbound legs. For predicting the ground speeds throughout the entire holding loop, using the wind vector, TAS, and heading, the following algorithm is used (see Figure 3.3):

1. Compute wind component orthogonal to desired track based on wind vector and desired heading. Taking the sine of the difference between wind heading and desired heading, the size of the orthogonal component can be determined.

2. As the size of the ground speed vector plus the size of the orthogonal-to-track component of the wind speed are known, the angle between TAS and GS can be computed. Here, it is assumed that TAS $\gg$ wind speed such that the angle can be approximated by $\arcsin\left(\frac{\text{orthogonal wind}}{\text{TAS}}\right)$.

3. Knowing the angle ϕ between TAS and desired track, the along-track component of the TAS is computed using $\cos(\phi) \cdot \text{TAS}$.

4. Finally, the along-track wind component is added, resulting in the full ground speed vector. It is noted here that in Figure 3.3, the wind component is still indicated in purple, but the *full* ground speed consists of the green GS vector and the purple wind vector added.

[1] Source: interviews with LVNL, 2019

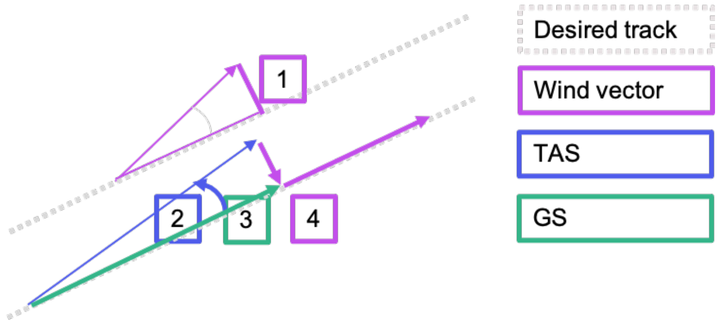


Figure 3.3: Graphical representation of ground speed prediction

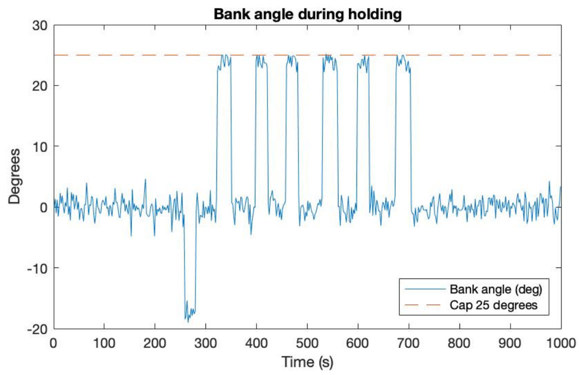


Figure 3.4: Bank angle fluctuation when flying a holding loop

3.3. Turn Time Prediction

When flying a turn while a wind field is present, not only the direction but also the magnitude of the ground speed vector will change continuously. In order to make an accurate estimation of turn times, two factors are essential: a prediction of the turn anatomy and knowledge about the ground speed vector at different locations in the turn. The pilot controls indicated airspeed and heading; standard civil practice is to fly a turn at rate 1 (3 degrees/second), or slower if the bank angle would exceed 25° otherwise [SKYbrary, b].

Since Schiphol airport and its surroundings are approximately at sea level, a standard holding (indicated) airspeed of 220 kts would result in a standard holding true airspeed of 220 kts. Upon investigation of flight data when holding stacks were present at ARTIP [LVNL, 2019] it can be seen that often, pilots choose to fly even faster. The result of this is that rate 1 turns cannot be flown as the required bank angle exceeds 25°. That means a standard turn in a holding pattern around Schiphol generally exceeds the standard of one minute. Further analysis of the dataset also shows continuous fluctuations in bank angle, implying the Flight Management System (FMS) flies a fixed track while varying bank, yet always capped at 25° as shown in Figure 3.4. This was verified by KLM pilots[2] and the KLM FMS pilot manual[KLM, 2019]. Based on this, the simulated prediction around Schiphol will assume that turns are always flown at 25° bank.

For the prediction of turn times a simple approximation will be used. Centrifugal force F_c is induced by the lift vector under bank angle ϕ. The size of the total lift vector can be determined by using the gravity vector, see Equation (3.2). Combining the two yields the formula for calculating turn rate in Equation (3.3), where V is the predicted ground speed at that location and heading, g the gravitational constant on earth, R the turn radius and ϕ the bank angle which is set to 25° in the simulation.

$$F_c = m\omega^2 R = m\omega V \tag{3.1}$$

$$F_g = mg \tag{3.2}$$

$$mg\tan(\phi) = m\omega V \quad \Rightarrow \quad \omega = \frac{g\tan(\phi)}{V} \tag{3.3}$$

[2]Source: interview with KLM pilots, February 2021

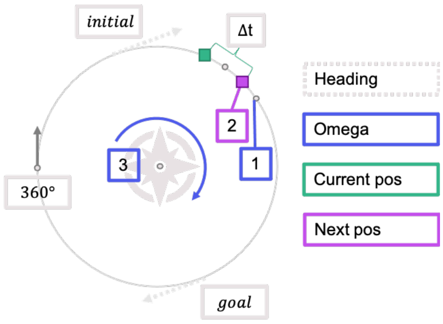


Figure 3.5: Graphical representation of turn time prediction

Based on the turn rate ω, the (remaining) turn time is estimated using the initial heading, which can be the current heading when on the outbound turn or the outbound leg heading when predicting the inbound turn time, and the desired heading, which is the heading of the leg (in- or outbound) following the turn at hand. The algorithm is visualized in Figure 3.5. The green square represents the current predicted location, the purple square represents the next predicted location which is $\omega\Delta t$ away, and the gray dots represent the locations that are $0.5\omega\Delta t$ and $1.5\omega\Delta t$ away from the current predicted location. In gray in the background is the visualization of the ground track, which is not used in the algorithm but included for visual purposes here. Before starting the algorithm, the turn rate is estimated based on the ground speed at the initial heading as explained before. Turn time prediction starts at 0s.

1. Predict the heading at $1.5\Delta t$ from the current moment (i.e. predicted moment) by adding $1.5\omega\Delta t$ to the heading at the current predicted position (green).
2. Predict the next heading (purple) by adding $\omega\Delta t$ to the current predicted heading.
3. Predict the next omega, by taking the heading computed in step [1] and using Equation (3.3) and the ground speed algorithm described in Section 3.3.
4. Add Δt to the turn time prediction.

These steps are continued until the difference between next predicted heading and desired heading is smaller than the time step. Since turn rates in holding are approximately $2.5° - 3°/s$, this may lead to some overshoot where the difference becomes negative. The last step takes the difference between the two headings and divides them by the last predicted ω, and adds this to the turn time prediction.

3.4. Flight Leg Prediction

When the aircraft is already flying on the outbound leg, there is a fixed amount of inbound leg time left after it has turned in again before reaching the IAF. In order to estimate the remaining leg time, the position relative to the IAF is determined. Then, the axes are transformed such that the positive x-axis lies along the outbound leg and the positive y-axis travels from the IAF to the outbound leg. The x'-coordinate of the aircraft in the transformed axes determines the remaining inbound leg length, as visualized in Figure 3.6. Using the inbound heading to predict the inbound ground speed, the fixed inbound leg time is found.

3.5. Pilot Reaction Time

Through interviews with air traffic controllers, it has become apparent that pilot reaction time is an important variable when managing a real-life holding stack. Pilot reaction time is a measure of the time between giving the turn-to-IAF or heading command and the aircraft actually commencing of the turn. When asking ATCos, they indicate that on average, pilot reaction time lies between 15-20s[3]. While estimating the turn-in moment, the ATCo recognizes that a 20 second delay in reaction time yields an approximate delay of 40 seconds (depending on wind conditions) in IAF-crossing since the additional outbound leg time results in additional inbound leg time. Therefore, one of the essential tasks while managing a holding stack in particular and upon managing air traffic in general is to make realistic estimates on pilot reaction time. This really is

[3]Source: interviews with LVNL, 2021

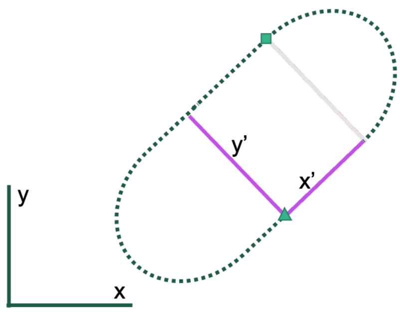


Figure 3.6: Relative position to IAF under transformed axes

a social communication task, where interviewees from LVNL have indicated that the set of factors they use to make the estimate is *"extremely large"* and based on experience combined with many subtle signs they pick up while communicating with pilots.

This communication is done using a radio, similar to a phone, and therefore the input signals are transmitted via voice primarily. The perceived awareness and swiftness of reaction of the pilot is determined based on how active they sound, on how well their English is, and on how easy and fast they copied previous commands. Other factors that the ATCo takes into account is the airline's familiarity with Schiphol (e.g. a KLM pilot will be very familiar with procedures while an Air China pilot will be less acquainted with the standard holding operations at Schiphol) and the origin (since a pilot may be more energetic after two hours of flight than after fourteen hours of flight).

In this research, the predicted pilot reaction time will be assumed equal to a constant fifteen seconds for the following reasons. First of all, fifteen seconds is taken as it is in line with the mental model of the ATCo. A constant is chosen over a stochastic variable as the effectiveness and impact on EAT adherence of a visual support tool in holding is to be measured. Introducing a stochastic variable for pilot reaction time can possibly confound the results, as it is likely to stochastically influence the realized EAT adherence in the experiment. The final reason is that it is not possible to adequately replicate the subtle and unconscious signals that are picked up by the ATCo through voice without using spoken feedback, making the introduction of a stochastic pilot reaction time additionally confounding to the results. To further limit pilot reaction time as a confounding factor, the participants are briefed about the pilot reaction time that is fixed to fifteen seconds to prevent any further confounds in EAT adherence due to unknown pilot reaction time.

To ensure a realistic scenario, the pilot reaction time in the experiment will consist of a Gamma-distribution. The mode is 15s, and it is slanted such that pilot reaction times of $< 15s$ are less common than pilot reaction times of $> 15s$. For further detail on the exact layout used in the experiment, see Chapter 6.

4

Interface Design

This chapter concerns the interface design considerations and resulting concepts. It starts by outlining the layout, features, and shortcomings of the current display. After that, some principles on display design are given that are based on literature along with considerations based on previous research in the area (of visual support systems in air traffic control). Then, the different parts of the interface and their improvements are discussed, followed by a section on the proposed interface design. This is split into two scenarios: one with a dedicated holding stack controller present, thus making use of the vertical view, and one where an ATCo is responsible for all traffic in the sector while there are also a couple of aircraft in the holding under his responsibility.

4.1. Current Display Layout and Shortcomings

The current display, as described in Section 2.4.1, is schematically represented in Figure 4.1. When the airspace gets crowded and aircraft are flying on top of each other, the vertical view brings the solution. Since it gives an overview of the situation from the side, it makes clear how the aircraft are separated and what an aircraft's position in the stack or in the vertical plane is. One of the shortcomings of the current operating display is that this vertical view is only available when a dedicated holding stack controller is present. The reason for this is found in situational awareness and safety requirements[1], as an ATCo who also has to manage traffic entering or leaving the sector is required to have a continuous overview of the situation in the entire sector, ensuring horizontal separation. Making the vertical view available at all times could, in theory, distract from looking at the top view and therefore reduce situational awareness in the x,y-plane. Even though the ATCo is already used to switching views, as in the bottom of the screen (not shown in Figure 4.1) there is a ribbon with additional information on the aircraft as well as space to enter commands, permanently allowing the vertical view will not be proposed here. The reason for this is that situational awareness decreases, as the ATCo will have less overview of the traffic entering the sector at the place where the radar screen has been replaced by the vertical view. In Section 4.4.2, an improvement for such a situation is proposed.

Another obvious possibility for improvement is found in the stack list. In the old stack list, the expected IAF crossing is not updated when an aircraft starts the holding. In Figure 4.2 a schematic example of both the old and a proposed (regular) stack list are presented. It should be noted here that the colors are not representative for the colors currently used in real LVNL systems. The changes made in the proposed stack list are highlighted in gray for clarity. In the first column, the *next* predicted IAF crossing is updated. The reason for using the *next* IAF crossing moment and not the *last* IAF crossing moment is that this way, a quick overview of what the remaining time an aircraft must stay in the holding is given. An implication of this new functionality is that during holding, the sorting of the stack list should not be done based on predicted IAF crossing but on planned EAT, as that will give a more realistic and ordered overview of where the ATCo will have to put its focus and attention as to where the most imminent control action may be necessary.

The second proposed change is an implication of the first. As the times are updated, it becomes clear how much error there still is in EAT adherence. As the aircraft are compensating time in the holding and this error is reducing, there will be one point where the aircraft enters the final holding loop. As this time enters below a certain threshold, the ATCo has the option to extend or shorten the final holding loop to ensure the EAT is

[1] Source: interviews with F. Dijkstra, KCD (2021)

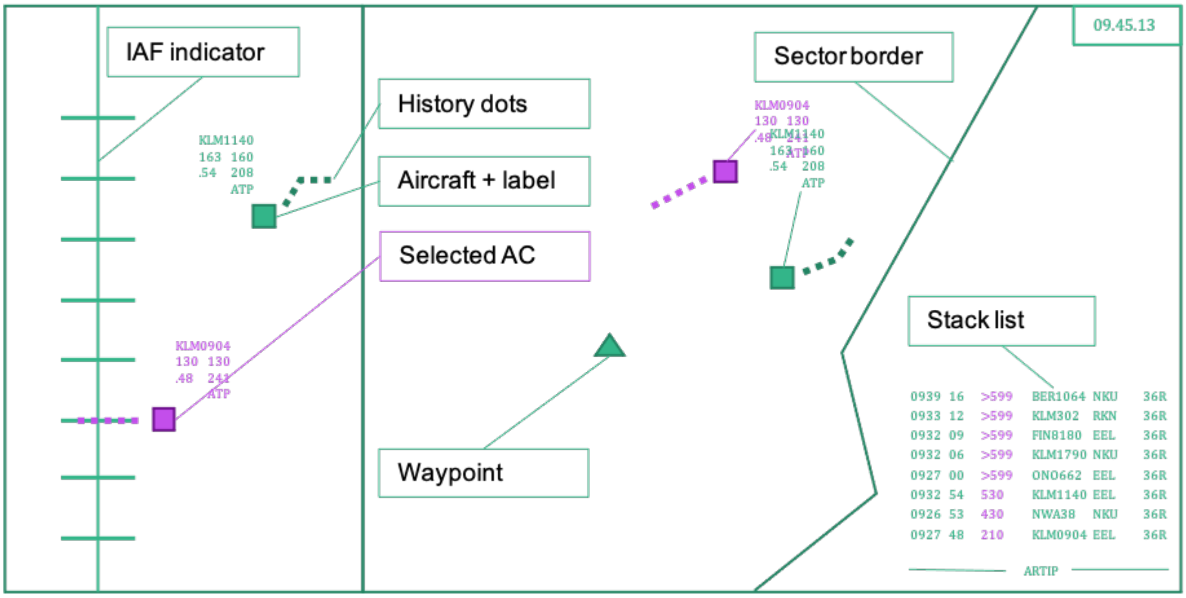
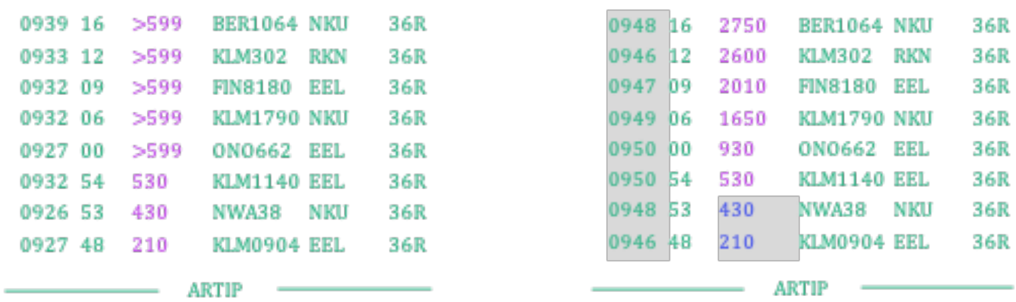


Figure 4.1: Schematic overview of current layout



(a) Current stack list layout

(b) Proposed stack list layout

Figure 4.2: Overview of stack lists, current and proposed at time 09:45:13

met with higher accuracy. In the new, proposed stack list, these times will change color to notify the ATCo that there are control options available by means of which they can improve the EAT accuracy. This is done in order to give a quick indication (in one glance) that does not require any reading of where action can improve EAT adherence.

4.2. Display Design Foundations

As the principles of the current design are laid out, this section will spend some attention to previous research on both the principles guiding good interface design and on previous design concepts in ATC and holding stack management.

4.2.1. Ecological Interface Design

Rasmussen [1987] proposed a taxonomy where human performance is modeled in terms of three levels that follow from the level of familiarity with the task at hand: skill- rule- and knowledge-based (SRK) behavior. This distinction in behavior is useful in support system design as “humans are not [...] input-output devices but goal-oriented creatures [Rasmussen, 1987]” leading to a need for better understanding of the human operator before creating any support tool.

Vicente and Rasmussen [1988],[1992] present a theoretical framework where all three types of human behavior are considered, which they call Ecological Interface Design (EID). The idea is to allow the human operator to perform a control action at the lowest and therefore fastest and easiest level possible while keeping the flexibility to provide support for unfamiliar and unanticipated situations. Translated to the context of

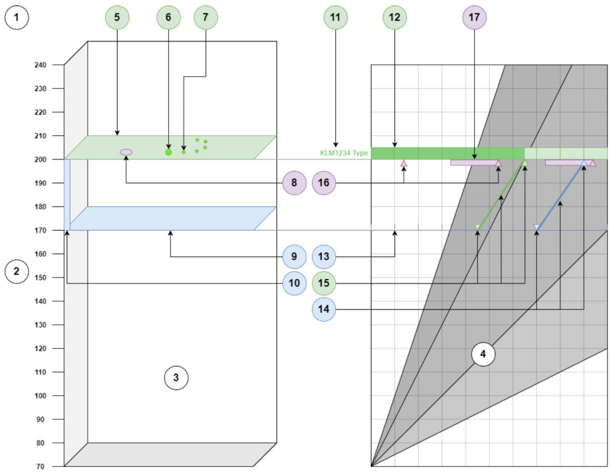


Figure 4.3: Holding display concept by Mac an Bhaird et al. [2020]

the visual support display to be proposed this translates into not only giving Air Traffic Controllers the tools to quickly make decisions but also showing them the work domain, which gives an advantage in cases of increased complexity. In such a way, critiques to full automation [Wickens et al., 1998] are overcome through actively keeping the human in the loop. This will be done by computing the predicted impact of a control action and presenting it to the controller such that she is able to quickly validate the decisions made and improving them when deemed necessary by the controller. The result is a display that does not restrict the controller to a set of proposed options but rather shows her the boundaries and possibilities within the domain.

4.2.2. Previous Design Concepts

Previous research by Mac an Bhaird et al. [2020] proposes a holding stack management display with a main focus on lowering aircraft in the stack, see Figure 4.3. Here, the visualization is not focused on the turn-in moment in the final loop but rather on creating the possibility for alignment early on in the holding process. The pink triangle (16) represents the predicted IAF crossing moments, while the green triangle (15) represents the EAT. The pink bars (17) stretching to the left of the IAF crossings represent the variability in time that can be created by turning in earlier. When both triangles are aligned and the aircraft flies a theoretical holding loop that lasts exactly four minutes, the EAT accuracy should theoretically be met with $\pm 0s$.

Regarding the research at hand, some considerations are made on what parts of the concept proposed by Mac an Bhaird et al. [2020] are relevant in the present scope. The major relevant part is the EAT adherence concept presented here. It will be argued why a different set of assumptions and goals leads to different priorities in designing a tool that supports EAT adherence. Pre-aligning an aircraft is only effective if the predicted holding loop times are exact. If not, additional work is induced as multiple turn-in commands need to be given. Since pilots are free to adjust their speed [SKYbrary, a] and wind conditions change over time and altitude, accurate predictions can not be made in practice. It also implies that each aircraft in the stack needs to be managed continuously. The proposed tooling has the objective of lowering workload in the management of a holding stack, eventually creating possibilities for ATCos to manage more traffic simultaneously in a safe and efficient manner. Therefore, the concept is designed such that it assists in determining the final turn-in moment while during the rest of the time an aircraft is in the holding stack it requires minimal attention and thus makes use of ACC capacity in the most efficient manner.

Research by Dirkzwager et al. [2019] proposes a display concept, see Figure 4.4. The research does not concern holding patterns, but the tool visualizes the feasible and infeasible moments where an aircraft can turn toward the runway in terms of safety and separation regulations. Its principle is simple: the future aircraft trajectory is extrapolated and a green dot represents a predicted feasible location to turn, while a red dot a predicted infeasible location to turn. It also gives the possibility to change the heading at which the aircraft is flying towards the runway and continuously updates the feasibility of moments accordingly, while visualizing the track that the aircraft would fly. The relevance for the current research is found in the visualization of turn-in moments.

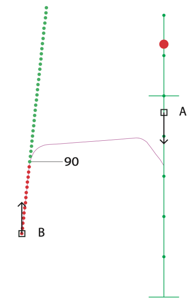


Figure 4.4: Turn-in concept by Dirkzwager et al. [2019]

4.2.3. Considerations on Controller Acceptance

As will be discussed further in Chapter 5, technical aspects are not the only shaping factors in the success of a technological innovation. The people who work with it play a crucial role. The framework presented regarding behavioral change here is meant to explore the perceived usefulness of a novel technology within the context of the Technology Acceptance Model [Davis et al., 1989]. A link between acceptance and the learning and emotions that are experienced upon first engagement is explored. To make things more concrete, that means that a constructivist learning process should be logged as to provide insights for further improvement of the tool (giving insight into controller expectations) and the design of the tool should be focused on triggering positive emotions. This underlines the importance of innovating at the intersection of current operations (full alignment with mental model) and radical change (maximum improvement). As any human controller is a human being with thoughts, priorities and emotions, these factors are critical in designing for controller acceptance. The preliminary interface concepts that are discussed in Section 4.3 are therefore created with the controller's mental model in mind.

4.3. Preliminary Interface Concepts

Based on the nature of the problem, two types of information are relevant to assist the controller: the EAT adherence error and the implications of taking a certain control action. For both of these, a prediction can be made. A case is made for the exact information that is most relevant to present, as well as the principles that should guide the presentation. An important factor that determines the effectiveness by which the information presented is actually understood and processed is the mental model of the receiver. The display design layout should connect to the mental model, meaning in this case that the visualization should follow several guiding principles regarding pictorial realism, the moving part, visual adherence, as well as have a close resemblance to the design of LVNL screens.

4.3.1. EAT Adherence Error: Time

When the aircraft are in holding, the most common practice to influence the timing of the last loop and therefore the timing at IAF is by first giving a command to stay on the outbound leg until told differently, followed by a "turn to IAF"-command. When this command is given, the pilot starts the inbound turn. Another way in which this is sometimes done[2] is by giving an (intermediate) heading command, e.g. by giving the heading the aircraft would normally have halfway during the turn, followed by the inbound heading or turn to IAF command. This approach is taken if the ATCo expects a pilot to react late to the turn to IAF command, and results in similar turns as in the first case.

In order to show the EAT adherence and give the ATCo an idea of the time that remains before any control action needs to be taken, the choice is made to show the EAT adherence error if the turn to IAF command were to be given *right now*. The reason for this is that ATCos have indicated[2] that they prefer information to be presented as integrated as possible. Earlier considerations on representing the time were a straight depiction of the time it would take the IAF when the turn in command was given right now, or the absolute time at which the IAF would be reached. Both of these were considered less desirable as they required additional calculations from the ATCo without providing a better insight in the situation.

The relative error is presented as follows: negative time means arriving at the IAF too late such that a positive time indicates crossing earlier than the EAT, or in other words: positive time means there is still a

[2]Source: interviews with ATCos from LVNL, 2020

(a) In label

(b) As a separate clock

Figure 4.5: Visualizations of EAT adherence error in label or as a separate clock

positive amount of time to be compensated for to arrive on time. The presentation of EAT adherence error in the stack list is considered here to ensure a consistent representation. Therefore, the same colors will be used for an error that can be compensated for in the current turn and for an error which requires flying an additional loop in both lists, as indicated in Figure 4.2b. Not only the depiction, but also the location of portraying a time is considered. Considered are a presentation in the stack list only, in the aircraft label, or as a clock close to the aircraft.

The advantage of using stack list only is that it prevents additional screen clutter, but the disadvantage is that it means the controller will not have a complete overview when looking at the aircraft itself. The consideration made between a timer in the label versus somewhere close to the aircraft is that in the label, the screen will look more ordered and less cluttered, while a separate clock will provide more clarity and space to show the EAT adherence error. As to give the best situational awareness to the holding stack controller without overly cluttering the screen, the display concept will put the EAT adherence error in the labels of the aircraft, see Figure 4.5a when either the vertical view or the extended stack list have been enabled. Besides this option, the ATCo should be able to manually turn the additional information on or off at any moment in time, for example when only one aircraft is flying a holding and therefore no additional vertical separation tools are in use. In the case of a separate clock, see Figure 4.5b, there are two options: either have it activated for all aircraft or only for the selected aircraft. From this follows a trade-off between better situational awareness and added screen clutter. Since putting the time in the label provides the best solution on those two factors, it is preferred over a separate clock.

4.3.2. Visualization of Control Action

To give the ATCo additional tools that allow them to further gain insights in the situation and verify their intended course of action, not only a clock will be present in the tool. Deciding on what to visualize begins with an assessment of the control options available to the ATCo. These are heading, speed, altitude, and turn to IAF commands. The timing of a holding loop and its alteration is for the largest part influenced by heading and turn to IAF commands; altitude has a very little effect due to changes in ground speed and the velocity bounds in holding are small, leaving little room for variation. This leaves two scenarios: either a command is given that determines the moment in which the holding loop is ended, or a command is given that alters the holding loop geometry. In current practice, the latter occasionally happens to compensate for a small deviation of time. However, with the proper tooling it could become more clear for the ATCo what the implications of the different timing of a turn-in command are, eliminating the need for such a deviation. It is for this reason that the control action to be focused on is primarily the turn-in command or a heading which also leads to the pilot turning toward the IAF. Here, the prediction algorithm will assume a turn-in command will be given, which is the most common practice. The reason for not allowing both options is that it is impossible to know what the next heading command will be, making a prediction less accurate, and that in the end, the aircraft will need to reach the inbound leg, leading to a similar turn as when a turn to IAF command would have been given. For further explanation, the reader is referred to Section 3.3.

Considering the control action, the display will visualize what the impact is of taking this control action at various moments in time, where time is visualized as the extrapolation of the aircraft's track or predicted holding pattern. Since the goal of the display is to present a solution space, thus ensuring the ATCo remains in control and keeps their autonomy, see Section 5.5.1, the display will not present only the location of where the optimal predicted turn-in moment is. For various possible turn-in locations along the track the aircraft flies, the algorithm will compute what the deviation from the EAT will be if turned in at that location. Doing this, dots will be colored green ($\pm 30s$), yellow ($\pm 60s$) or red (larger deviation or not feasible due to separation constraints). These dots are located at the predicted radar update locations, ensuring the interface adheres to the actual functioning of the radar screen and aircraft locations that will be made visible in the future.

An additional consideration is the visualization of the predicted holding loop track. From interviews with

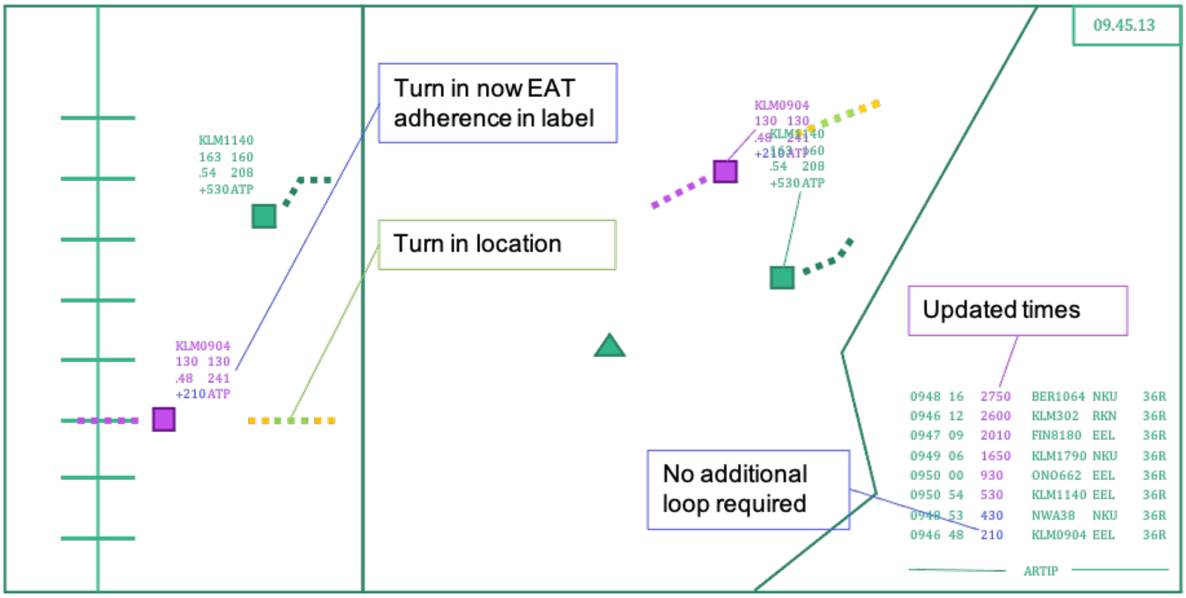


Figure 4.6: Schematic overview of proposed layout

ATCos[3] it was found that they desire a tool which is as simple as possible, giving a quick insight without requiring additional integration of variables. For this reason, the focus of the tool will be on giving insight on what happens upon turning in now or later along the outbound leg. The tool will not give visualizations of the trajectory, neither will it give a full overview on what happens if the ATCo decides to give different/additional headings. The reasons for this are to prevent screen clutter and since additional heading commands are generally given by ATCos to influence pilot reaction time. However, the concepts proposed in this research were not limited by the desire for a simple tool. If the concepts appear too complex upon evaluation, they can be simplified upon implementation, ensuring that this research can still make a valuable contribution to the body of knowledge on ATC tooling.

Since aircraft are displayed on radar screens, their locations do not update continuously but per radar update. Using dots that are separated from each other by one (predicted) radar update corresponds will therefore give the representation that aligns best with an ATCo's mental model.

4.3.3. Vertical View Layout

Mac an Bhaird et al. proposes a three-dimensional vertical view display, where aircraft are displayed in their holding planes as to give a full view on their position in the holding. The way in which a display is used is guiding for determining the suitability of such a three-dimensional view. If the aim of the display is to give any user a high situational awareness and overview of the location of the aircraft, the three-dimensional view is highly suitable. However, the user for the display proposed in the current research is an ATC professional. In this context, an overview of any situation can be gained quicker when all unnecessary information is removed. Since the vertical view in two dimensions, combined with history dots and the fact that the aircraft is holding give an experienced ATCo enough information to know location and trajectory, the current two-dimensional vertical view layout combined with a top-view radar screen will be used. For clarity and better situation awareness, a line representing the IAF and bars for flight levels is present on the vertical view.

4.4. Proposed Interface Design

In Figure 4.6, a schematic overview of the proposed layout is given. It contains the improvements proposed above. The changes have been highlighted: EAT adherence error is visualized in the label, the possible turn in locations and their predicted EAT error bound (green $\pm 30s$, yellow $\pm 60s$), updated stack list and highlighting of EAT errors that can be compensated through adjusting the current loop. The dots indicating the EAT adherence error at possible turn-in locations will be referred to as ecology dots.

[3] Source: interviews with LVNL (2021)

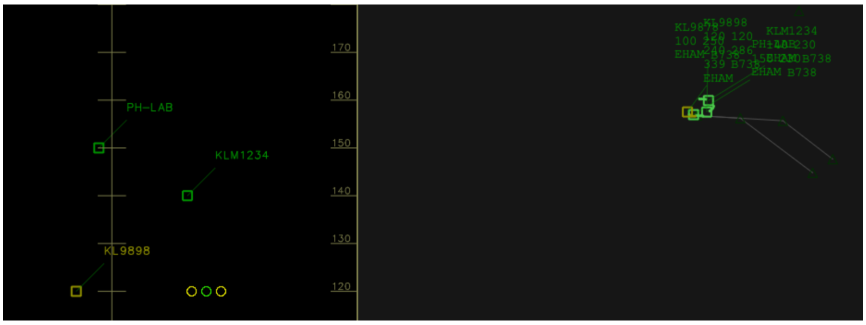
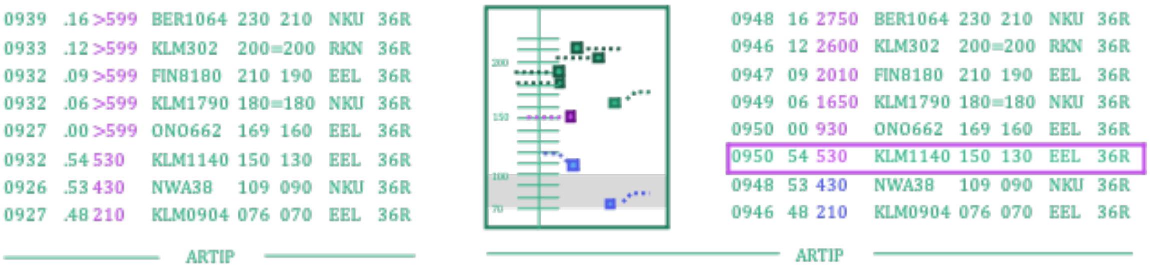


Figure 4.7: Current running interface in SectorX



(a) Proposed extended stack list

(b) Current extended stack list

Figure 4.8: Current and proposed extended stack list

4.4.1. Dedicated Holding Stack Controller

When a dedicated holding stack controller is present, the following elements will become available: EAT adherence in label, updated stack list, and ecology dots. Even though the top view is very cluttered while a holding stack is present (see Figures 2.6a and 4.7), the extended label and ecology dots will be available in both views. When selected, the aircraft and ecology dots will highlight, making it possible to distill it in the top view when desired. It will be possible to click and select aircraft by selecting the aircraft themselves, their label, and by selecting them from the stack list. When hovering over any of these locations, the corresponding locations will also be highlighted temporary just like the ecology dots; when clicking, the highlight and ecology dots will stay on until the aircraft is deselected. When hovering over the ecology dots, the expected EAT adherence upon turning in at that point is shown in the aircraft label EAT adherence location.

In Figure 4.7, a still of the current running interface in SectorX is shown. It is noted here that not all elements are running yet. The future work that needs to be done in SectorX before performing an experiment is discussed in Section 6.7.

4.4.2. Holding Stack Combined with Other Traffic

When managing a couple of aircraft that are in a holding, the current tooling only provides for the extended stack list. Several additional improvements are proposed to the extended stack list, as visualized in Figure 4.8. A consideration has been made whether to allow an ATCo who manages both a sector as well as aircraft in a holding pattern a vertical view or not. Prioritizing safety, maintaining overview, and situation awareness, the use of a vertical view gives a larger negative than positive impact on these factors as the vertical view then takes up a large part of the radar screen.

The first improvements proposed are the already-mentioned improvements to the stack list, possibility to click and select aircraft from anywhere, EAT adherence error in the aircraft label and ecology dots in the top view. Then, the extended stack list will be enlarged with a schematic overview of the vertical view, called the mini stack view. This overview only contains the aircraft locations plus history dots, as to give the controller a quick overview of the situation. When the aircraft enter their final holding loop, in other words, the EAT adherence error can be reduced to zero by giving a control input in the current loop, the aircraft change color (blue, in the example in Figure 2.5b). The selected aircraft is highlighted in both radar screen, stack list and mini stack view. In the mini stack view, the appropriate flight levels for leaving the holding are also indicated to give a quick overview of the current situation to the controller.

5
Behavioral Change

In this chapter, the theoretical foundations of the subject in the context of communication are explored. The aim of the chapter is two-sided: to gain insight into the different aspects relevant to the research question, and to determine where the gap in scientific knowledge lies. Different angles that each cover a potential factor that influences whether a behavioral change takes effect or not are discussed. This research will aim to generate preliminary insights on how the Technology Acceptance Model (TAM) can be extended with information and participation in the context of a technology *development* process. As this research is exploratory, promising results can then be used for further research with larger subject groups.

5.1. Communication Problem

Based on conversations with Air Traffic Controllers, an initial research question is posed. The following factors helped shape the question: first of all, the development of a new tool means that people who want to use it will have to engage with it and learn how to work with it. Every person will have a different mental model, and therefore people's willingness to work with new technology may vary. Finally, the design of the tool can invoke emotions upon engagement. It is important to realize the impact of different design choices, as people with different disciplinary backgrounds are likely to engage and react in different ways. Research by Westin et al. [2015] suggests that the problem-solving style of a support system impacts controller acceptance, where strategic conformance is proposed as the key driver for this acceptance. Here, strategic conformance refers to the degree to which the automation's style of problem-solving matches that of the controller, both in process as well as in solution.

The objective of the research is to make sure EAT adherence goes down from 2 minutes error to 30 seconds error, by means of designing, testing and evaluating a graphic support interface for ACC. Where the previous chapters have discussed the technical solution to this problem and the foundations required for the design of a tool, this chapter will focus on the social aspects toward introducing a technological innovation. A driving factor for this part of the research is that in order for any innovation to valorize, a behavioral change is needed in the people using it.

Since the design of this tool is an example of a multidisciplinary engagement between Air Traffic Control, Aerospace Engineering and Communication Design for Interaction, the collaboration between these parties and the extent to which they can learn from each other is a prerequisite for success. The relevance of the research can be found in evaluating the strengths and weaknesses associated with introducing a visual support tool within a domain of high complexity, high responsibilities and many involved individuals. The following sections will each pay attention to the theoretical foundations of learning, collaboration, mental models, the impact of emotions and an assessment of favorable circumstances for change from literature, in that order.

5.2. Learning

Learning is important for multiple reasons: a new tool requires people to learn to work with it, learning can be the goal as well as the means to an innovation process, and based on observations at LVNL, understanding and solving complex problems is a motivator for people - so learning how to use a new product could actually boost energy and creativity. Designing a tool in such a way that a learning process is supported is not a trivial task. Two factors should be considered: what knowledge needs to be gained and the organization of the environment to do so in an efficient way Paquette et al. [2006]. The process of transferring information is often supported by some image, such as a flowchart or an icon, as it can enhance the understanding of the receiving party. This can speed up the learning process, as the content or meaning of the information is picked up faster by the receiver. Different approaches to learning contexts exist, as discussed below. In order to support the interaction design with the graphic support tool in the experiment, the process of discovery learning is seen as the most accurate fit since subjects will perform the experiment by themselves and they are known to already have extensive knowledge about the domain of holding control, making it likely for them to build any new knowledge on top of their present mental model (Section 5.4) on holding control.

5.2.1. Passive versus Constructive Learning

As opposed to constructive learning, where the learner builds the new knowledge themselves, other ideas about knowledge are that it can be innate (as proposed by Descartes) or passively absorbed. Innate knowledge is not relevant in the context of this research, but for the latter it should be argued why it will not be taken along as a learning method, by first outlining some principles. In the process of passive learning, the information should be passively received by the learner, implying that engagement with the world is not necessary. In a constructivist approach, it is argued that prior knowledge influences the meaning an individual will distill from information presented Phillips [1995]. As this research will be done with highly skilled Air Traffic Control professionals in mind, who are known for being critical toward newly introduced technologies LVNL [2020b], Teperi and Leppänen [2010], it is highly unlikely that their learning process will take place by means of just assuming newly presented information as knowledge (passive learning). A more likely course of action would be for these professionals to engage with information presented and distill new knowledge from that process (constructive learning). For this reason, the next sections will focus on different approaches that can be taken toward constructive learning.

5.2.2. Collaborative Learning

Collaborative learning allows for a combination of social learning and a constructivist approach, where subjects create knowledge themselves. Important drivers to choose this form of learning are the social and emotional development that is experienced by the learner, one's view on the nature of knowledge, the emphasis on teamwork, and the interactive nature. The process can be seen as learning through an active process, collaborating with others with the aim to produce a shared meaning after interaction and negotiation [van der Linden et al., 2000].

5.2.3. Discovery Learning

De Jong and Van Joolingen [1998] describe discovery learning as a "self-directed and constructivist form of learning", where the learner is expected to actively seek knowledge. From a task or situation, knowledge about important characteristics can be inferred according to De Jong and Van Joolingen [1998]. This is contrary to a learning situation where a teacher presents the student with the knowledge to be learned.

The research conducted by De Jong and Van Joolingen [1998] is additionally interesting, as they performed an experiment where subjects were to learn upon engaging with a computer simulation, as is also the case in this research. Potential chances and problems in designing a simulation-based discovery learning situation are: lack of hypotheses generation and adaptation in the learner and the amount in which people tend to stick to their original ideas as well as confirmation bias; the chance to design an experiment with the focus on an hypothesis as it increases the further usefulness of the experiment outcome for the research in this context; and finally guiding the subjects through the experiment in a planned manner such that they can follow a successful learning route [De Jong and Van Joolingen, 1998]. Since the objective is to determine how a behavioral change can be triggered upon introducing a support tool, in the context of which an experiment is performed evaluating the reaction professionals have upon engagement, this form of learning is relevant: confirmation bias may play a big role upon introducing new tools at LVNL and the experiment should be designed with behavioral change in mind, both in terms of how this change is to be measured and in terms of how new insights gained can actually lead to this change.

5.3. Teamwork and Interaction

Within Air Traffic Management in the Netherlands, the different parts of the airspace are controlled by different teams. Three of these teams work (physically) at LVNL: approach control at Schiphol (APP), area control around Schiphol (ACC) and military airspace control. These teams do not work together as the domains they control do not overlap, however they do impact each other as their airspaces neighbor each other. This especially concerns the former two, where the accuracy in which ACC can adhere to the planning impacts the workload of APP. Next to this form of interaction, the domain of ACC is sub-divided into multiple sectors. The number of divisions depends on the amount of traffic present above the Netherlands at that moment, where each ATCo professional controls a different part of the airspace. This requires close collaboration and a focus on clear communication within the ACC team. Finally, not only do these people interact with each other in the conduct of their work, they also interact with the tools at hand (support systems such as radar screens and prediction tools) and with in- and outbound traffic, instructing and informing the pilots. The success of the operation not only relies on the work of the individual, but also largely on collaboration and communication [Teperi and Leppänen, 2010]. The same applies to the succes of a support tool: people will need to engage with it, collaborate (with each other and the tool) and the communication process between human and machine should successfully convey the right and intended messages.

5.3.1. Human Interaction

The interaction within the team of ACC is collaborative in nature. The competence of each ATCo is determined by the amount in which they are able to share knowledge, communicate, and solve problems in a collaborative way. Through in-team interaction, ATCos learn from their more experienced colleagues [Teperi and Leppänen, 2010]. The interaction with APP and pilots is informative and/or instructive in nature. Literature suggests that collaborative transdisciplinary training involving different airport users is currently missing [Prince and Salas, 1999, Schroeder et al., 2006]; this is confirmed by current practice at LVNL where simulation training is performed without the use of *real* pilots. Collaboration and communication within the teams works fluidly, and there seems to be a strong in-group feeling within the own team [LVNL, 2020b]. From both observation and conversations with LVNL, it was found each ACC individual has a strong sense of responsibility and is highly professional. However, interviews with the KDC have also made clear that a friendly rivalry with other stakeholders within ATM exists, which in extreme cases have led to people insinuating that other teams would be doing less than they were, and therefore *the others* should be responsible for solving a problem. For the research at hand, it is important to understand the complex social network found within LVNL such that the setup entices support, rather than resistance, under experiment participants.

5.3.2. Interaction with Support Systems

Even though the use of support systems is increasingly widespread in ATM, the human controller remains the central decision-maker in the field [ICAO]. The paradox here is that the organizational structure and culture can create conflict in learning to work with these systems [Teperi and Leppänen, 2010]. This is partially supported by field research at LVNL, yet partially negated: each individual within the organization has a different willingness to interact with *new* support systems. A general trend to be distilled from observation is that younger people are more willing to engage with new technologies. Research on truly cooperative tools with a focus on interaction between human and support system indicates that essential factors are trust (in the system) and controller autonomy (to conceive a problem-resolving strategy). One factor that led to a more positive experience with a support system was identified to be the level to which the support system would actually enhance collaboration between controllers [Guiost et al., 2006]. The goal of a tool should be that people learn from it together, in a constructivist and discovery manner. From this follows the hypothesis that a tool may be more effective in triggering behavioral change if the tool lets the ATCos discover what they can do themselves through using it and the ATCos can learn about possibilities they can exert using their own capabilities. In other words: if the tool can give insights in the EAT adherence possibilities and capabilities the ATCos personally have, it may be a strong motivator to steer towards a higher EAT adherence.

Finally, an important part of interaction is learning from the tool *together*. Upon introducing the tool, the contact should be in such a way that ATCos get the feeling that this research is aimed at helping them and initiated to be on their side, instead of imposing anything on them. If the project comes to life and people start talking about it, working with the tool becomes something more in-group people do. The majority of the people may be on board with the project when doing this right. But while dealing with humans, it may happen that some people will remain unconvinced or unwilling to cooperate. The focus will be on the motivated part of the group, especially as this research is exploratory.

5.4. Mental Model

For innovation to take place within a social network, it is important that it takes place in exactly the right moment: it should strike the right balance between being conservative and being revolutionary. In other words, when an innovation builds upon known systems, making it incremental, it is more likely to be accepted. To make sure a new technology does not estrange people, it is vital to know what their mental model looks like. In Chapter 2 a detailed description of the current ACC work domain and LVNL support systems is given. In this section, the focus will be on describing the mental model of ATCos regarding the *way* in which their job should be done, based on observations and semi-structured interviews. It should be noted here that interviews were held with two different ATCos, as it is difficult to find people that are both motivated to help with external research and also have the time to do so; the representation given is therefore largely biased by the way these people view their job. Interviews were also held with Ferdinand Dijkstra from the Knowledge Development Centre, who knows many ATCos personally and has been performing research for LVNL for a long time.

To become an air traffic controller, one has to pass a strict selection and will have to follow a long education and trainings. In the context of the responsibility that comes with the job, it is easily seen why. Another factor that may be shaping in how ATCos (in general) view the work they do, is the organizational culture. ATCos take a lot of pride in their job, and value autonomy in their work. Autonomy can be found in many different aspects: the freedom ATCos have to plan their own shifts, the way in which they solve the problems that daily challenge them and the freedom to work on additional projects.

Culture and attitude toward innovation has two sides: the one hand there is a group of people who believe that they do not need tools, as they are highly skilled and have a resistance toward technological innovations since in the end, the complexity of the work requires people to make the decisions for safety. On the other hand, there are people who believe that technological innovations can be used to their advantage, and that building support systems is actually a way in which they can either validate their own decisions, promoting safety, or steer towards higher accuracy in planning adherence, for example because a tool can free up workload and mental capacity while retaining safety. A tool can in that sense make an ATCo experience its work in a more fun and satisfactory manner: if a tool allows you to do a better job, that promotes the pride you take in delivering aircraft in an even more accurate manner.

The ATCos that have been interviewed fall in the latter category while discussions with Mr. Dijkstra give more substance to the first viewpoint. One of the two ATCos has actually taken the initiative to request and start a research project on holding support tools, while the other has volunteered to help and give feedback on the current project. It should be obvious here that they have a large willingness to innovate, learn and work with new technologies. However, two people do not constitute a culture. The truth must be somewhere in the middle, but it is important to note here that an organizational culture can have a large influence on how a team handles an innovation, regardless of personal opinions, as illustrated in the case study by Vakola [2012], see Section 5.5.1.

Remarks that are often seen in previous studies done regarding support tools for LVNL is that the interface does not match with the LVNL interface E.S.Bakker [2019], Ottenhoff et al. [2020], Dirkzwager et al. [2019]. The result in an extreme case would be that feedback only comes on already known shortcomings regarding the match with people's mental model, instead of feedback that can be used to improve the proposed tool. Another reason to put an emphasis on visually adhering to the ATCos mental model is that using their own visual language will trigger a different emotion than a strange visual language. The power to promote change is to be found in what is already known.

5.5. Favorable Circumstances for Innovation

In this section, some examples from literature regarding the circumstances in which behavioral changes upon introducing a new technology could or could not distill are evaluated, following by a critique on how these relate to the problem at hand. Openness to change is defined by Miller et al. [1994] as the "willingness to support the change and the positive affect about the potential consequences of the change". The other factor, technology, should be interpreted as computer or digital technology, such as software or an application. Within this context, behavioral change is conceptualized as the result of openness to change, where it is assumed that an openness to change in combination with the new technology will lead to a behavioral change.

5.5.1. Case: Resistance to Change in the Public Sector

This case, as described by Vakola [2012] concerns a medium-sized company in the public sector, with an employee base that has an average age of 48 years, 35% higher educated, and “characterized by bureaucracy, predictability, stability and control”. In the case study, the top management has decided to invest in a new technology, which is to be implemented by an external company. The emotions linked with new technologies within the company are that employees do not trust it and a strong sense of “that is not how we do things here”. Resistance to change was identified as the main issue that blocked the program, split up into four categories: *people were afraid their performance would be tracked* (and turn out lower than that of colleagues), fear and stress about incompetence to work with the new technology, *the union resisted the change*, and there was a lack of trust in the management since *many initiatives were left unimplemented in the past.* There was a minority of employees who actually were open to the innovation and willing to (learn to) work with the new system. In the end, the management did push through to implement the change, but it was costly, time consuming, and key users would indicate various flaws and mistakes in the system and its implementation. In the long term, the system has not made a valuable contribution to the organization and was taken out of operation [Vakola, 2012].

Link to LVNL Since ATCos can only perform their job until they are 57 years, the average age is higher for the case; average education levels at LVNL are relatively high; and having a focus on safety and predictability, the characterization of the organization presents some resemblance. The emotion linked with technological change is also recognized by people who have been working with LVNL for longer; here, too, a group who is actually advocating for the innovations exists. If it is possible to take every key user on in the process of developing and implementing a technological innovation, the problems that are seen in the case may be prevented. Having people collaborate can lead to both a more efficient innovation as well as implementation process, where they can give input along the way. In the end, this leads to a better system for the users.

The factors emphasized above are the ones that are recognized by professionals who have worked with LVNL before. The first two factors, anxiety about performance tracking and resistance to change by the union of workers, have been present for longer. As in the past ten to five years, more innovation projects have started while the organization is sluggish in terms of innovation implementation and development. Reasons for this are regulations and safety issues that all require innovations to pass through long bureaucratic processes first. The result of this is that many projects have turned into floating or broken promises, but have not materialized in actual improvements. Therefore, in the last five years a sentiment of distrust regarding the realization of innovations has started to emerge at LVNL.

5.5.2. Identifying Predictors of Openness to Change

During an extensive reorganization at the U.S. Department of Housing and Urban Development, Wanberg and Banas [2000] conducted research on the relation between the level up to which employees were open to change based on context-specific variables and individual-specific variables. It was found that personal resilience was a strong predictor for acceptance and openness to change, determined by three factors: self-esteem, perceived control, and optimism. These factors are personal and not influenced by the workplace directly or on a short term. However, the study also found that three context-specific variables would also impact the level of change acceptance, being: information received, participation in the process, and self-efficacy or perceived competence. On the other hand, low levels of change acceptance could be predicted by (low) job satisfaction, workplace irritations, and people’s intention to quit [Wanberg and Banas, 2000].

Link to LVNL Even though in this context, the change is organizational and not technological, the way people cope with a change in their situation may still be representative. However, this study is conducted within a different context and therefore caution should be taken regarding its validity in the present context. From observations, people at LVNL seem to have high job satisfaction, take pride in the work they do and enjoy the autonomy they have. As for that, a lower change acceptance over the ACC workforce is probable to result from the moment their autonomy is put at risk, as this threatens people’s job satisfaction. Factors that in this case are most likely to influence acceptance are the information presented about the change and the ability of people to contribute to the innovation. Since ATCos are used to working with complex technology, are high-educated, and have a maximum age of 57 (being relatively young), perceived competence to work with new technology is unlikely to be a driving factor for resistance to change [Wanberg and Banas, 2000].

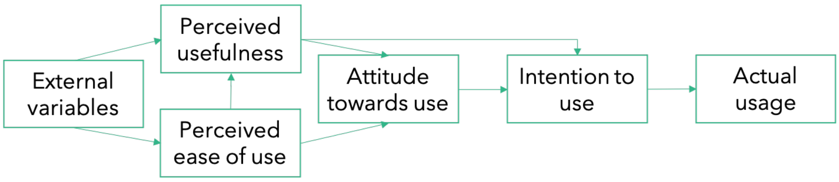


Figure 5.1: Technology Acceptance Model, adapted from Davis et al. [1989]

5.5.3. Acceptance of Technology

For innovation to happen trough introducing a new technology, the computer system will first need to be accepted. To explain and improve user acceptance, Davis et al. [1989] has proposed the *technology acceptance model* (TAM) in which they describe and predict people's intention to work with a computer system. The factor that mainly determines this is perceived usefulness of the system, and to a much lesser extent ease of use. Where its evaluation was initially done using a word processing tool, the concept is widely used and seen as a solid way to represent technology acceptance. The TAM layout is shown in Figure 5.1. One interesting factor that should be noted here is that the tool presented in the experiment is used on a voluntary basis, which explains the strong focus on intention of using as a measure for technology acceptance.

Research by Westin et al. [2015] discusses how the TAM can be used in the context of ATC decision making tools. First of all, they find that the TAM has been mainly applied to the acquisition and analysis of information, but not so much toward the actual decision-making process that follows. This is a critical factor for improving EAT adherence through the use of a support tool, as is the subject of the current research. What is more, Westin et al. explain how a higher conformance of a system to the human's problem-solving style can be used to overcome initial controller acceptance issues in expert user groups. However, they also argue that the highest level of conformance is only possible on an individual basis as each controller will have a unique problem-solving style.

Link to LVNL Potential resistance and potential acceptance in the case of ATCos can be linked to the TAM considering the way they view their job. The job comes with high responsibility, uncertainty and requires creative and non-standard solutions continuously. Ironies of automation, as introduced by Bainbridge [1983], are that a computer or automation system can deliver standard solutions while the system operates faulty upon an unexpected situation. It also becomes harder for the human controller to spot these errors. Considering these ironies and the TAM, a natural and logical response to a digital support system would be that the system can only get in the way of safe operations and is not useful: an explanation on why technologies are so often not accepted in the world of ATC. Currently, LVNL is implementing iLABs, which can be seen as a Living Lab where ATCos can first-hand experience new technologies. Since it is still under construction at the time of this research, nothing is yet to be said about it being a possible solution to this problem.

The precautions mentioned above are taken into account into the proposed design of the tool as explained in Sections 4.3 and 4.4, making sure the technology in fact makes it easier for the human controller to operate in uncommon situations and by all means staying away from full automation, keeping a focus on controller autonomy and the ever critical human factor in ATC. Based on the above, the TAM is seen as a promising framework for explaining initial controller acceptance. The current research will explore the potential link between acceptance and initial emotions upon the first engagement with the technology assuming the framework of the TAM.

5.6. The Role of Emotion

As mentioned above, the willingness of controllers to work with support systems is partially determined by trust in the technology (see Guiost et al. [2006] and Section 5.5. Emotions toward a system, technology or group can play a large role in the way people perform. Unidentified bias can also serve as a confounding factor when performing research. It is therefore important to identify the role emotions play in the context of the research, as well as what emotions are relevant in the present situation.

5.6.1. Emotion and the Mental Model

The process of decision making in the domain of air traffic control is one entailing complexity, risk and ambiguity. It is not an easy process to get though the selection for being an air traffic controller, and each person who passes is highly capable of doing so and working under stressful conditions. From different interviews and observations at LVNL, as well as with people who have worked with ATCos for a longer time, it has been concluded that people at LVNL take a lot of pride in their job and value their autonomy. Any innovation should therefore acknowledge the complexity of their job and inherently respect their skill and autonomy.

Another factor that is important regarding the work domain is recognition. The emotion one experiences upon engaging with a visual support tool can in a large part be triggered by the design of the tool in combination with the subject's mental model (Section 5.4). Based on previous experiences, expectations and knowledge, different emotions are triggered. In the specific case of air traffic controllers at LVNL, it was already found that some have a large resistance against working with new technologies, while some are excited to be part of innovation and are enthusiastic to be part of a development process [LVNL, 2020b]. This is a normal process that happens for any technological innovation, and refers to early and late adopters.

5.6.2. Trust

An important driver for people to engage with a support system is trust according to Guiost et al. [2006], also identified as one of the attributes describing teamwork as a basis for human-machine interaction by Degani et al. [2017]. Trust can be seen as a two-sided prerequisite: for any larger system or hyper object, not only is trust in the technological system vital, but also trust in any other person in the system; this is in line with human interaction with systems that is *similar* the that with other humans [Degani et al., 2017]. In other words: for an ATCo, it is essential that both the technological support system as well as one's colleagues can be trusted to be fail-safe: upon any error, failure or mishap, the system or colleague will communicate what happened.

5.6.3. Conveying Emotion

In a regular ACC (holding or other) task, communication is done through speech. It is therefore possible for both controller as well as pilot to convey emotion through the structure of their voice [Gill, 2008]. An interesting aspect is that holding situations mainly occur through extreme weather or anomalies and problems at Schiphol. Since voice communication always conveys more than just the words, it can be used to send out additional messages from the side of ACC, e.g. that the situation is under control and that there is nothing to worry about, and from the side of the pilot, who may e.g. be worried, paying little attention, or in a hurry. These aspects will not come into play during a holding control experiment where commands are given through the computer, which may influence the exact timing of commands given.

5.7. Research Framework

Combining the context-specific factors of information and participation introduced by Wanberg and Banas [2000] on change with the reason for technology rejection as presented in the TAM [Davis et al., 1989], the following framework is presented regarding introduction of new technologies at LVNL, which also serves as an explanation for the different opinions people have regarding these systems. It should be noted that this applies the literature presented above and theories to the case at hand, and will serve as a principle guideline and starting point for the research.

Upon introducing a novel technology, ATCos will have a bias on its positive or negative effects. This could be caused by misinformation on how useful the system will be to them, leading to an increased or lowered openness to change. In order to change the perceived usefulness of a novel technology, people should be educated on the actual shortcomings and capacities. As ATCos are intelligent and highly knowledgeable people, this learning process is suggested to be done using a constructivist approach. In the context of the research, where a tool is proposed and evaluated, the constructivist approach to learning will have the additional advantage of engagement, during which feedback on the tool can be gathered.

6
Research Outlook

This chapter discusses the research outlook, which comprises the approach that will be taken to gain further insight into the research question plus the expected outcomes, and the planned future work that is necessary to execute the desired approach.

6.1. Methodology

The overall methodological approach for investigating the research problem will involve a mixed-methods approach combining a semi-structured interview, a survey and a human-in-the-loop simulation experiment. The research will be exploratory, in the sense that it explores the possible impact of a holding support tool on EAT adherence, it aims to get insights in what parts of the tool are suitable for further exploration to possibly implement in LVNL systems and it explores how a behavioral change can be triggered by a technological innovation, laying the groundwork for further (larger participant group) evaluation of such a theory.

First, the participants will receive a briefing where the goal and approach of the research will be explained. They will also be asked to sign a consent form (which is still to be created and shall inform them about the processing of their data and ask permission for logging and recording their answers). The full session will last two hours, consisting of the aforementioned briefing (15 minutes), a training scenario (10 minutes), two holding stack simulation experiments (25 minutes each), a survey (15 minutes) and a semi-structured interview (25 minutes), plus five minutes for task-switching etc. The experiment is conducted in ATC simulation environment SectorX, and will quantitatively log the performance of the participant. The survey will evaluate how the participants experienced the support tool and the experiment, what their stance is toward innovation, and how they view current operations and the need for change (of both the system/innovation as well whether they believe it is necessary to improve EAT adherence at all). Finally, short semi-structured interviews will be held to explore suggestions and ideas for future improvement of LVNL's holding support systems and to explore ATCos' stance on improving EAT adherence and how they would expect a behavioral change to take place within their workforce.

6.2. Subject Group

The envisioned participants are air traffic controllers (ACC) or student ATCos from LVNL because they are the target group whom is to use the tool. People from this group have a lot of experience and therefore specific strategies for control that a layman does not have, influencing the performance and how the tool is used. Another reason is that they are the only group who can give feedback on what they would like to see implemented into the LVNL systems. Finally, they are also the ones who know the specific environment in which the communication challenge lies, making them an essential subject group to generate a specific communication plan that triggers behavioral change in this specific context. Since from previous research it is known that gathering participants from this group is very hard[1], a challenge with a leaderboard is introduced to create more enthusiasm under ATCos.

[1] Source: private conversations with E.S. Bakker (2019) and M.M. Ottenhoff (2020), who have performed earlier research and experiments with this group

Table 6.1: Variables to be measured

Measure	Reasoning	Log
All traffic data	Playback	AC, variables, time
Controller inputs		
IAF crossing	Compare with EAT	AC, time
Turn to IAF	Expected control action	AC, time
Turn to IAF then SPL	Expected control action	AC, time
Turn to SPL	Expected control action	AC, time
Heading	Unexpected control action	AC, time, value
Speed	Unexpected control action	AC, time, value
Altitude	Relative position in stack	AC, time, value
Clicks: stack list	AC selection	AC, variable type
Clicks: ecology dots	AC selection	AC, dot location + type
Clicks: aircraft	AC selection	AC, label or AC selection
Hover: stack list	AC highlight	AC, variable type
Hover: ecology dots	AC highlight	AC, dot location + type
Hover: aircraft	AC highlight	AC, label or AC selection
Thoughts	Qualitative evaluation	Log what is being said

6.3. Experiment

The control task that is to be tested in the experiment is the alignment of the time the IAF is crossed with the planned EAT through giving turn-in commands to aircraft that are in their final holding loop. This is to be done in a similar manner as in LVNL systems, using simulated environment (SectorX). This implies differences in screen size, command options, and some other minor details. Apart from these minor differences, the interface is very similar to the LVNL operating systems. An advantage is that ATCos will require little training and will therefore be able to focus on the tool that is to be tested.

Part of the task is the lowering of aircraft through the stack, ensuring a realistic representation of workload. The experiment will be used for two things: to assess the effectiveness of the tool and its impact on EAT adherence, and to let ATCos experience the effect this support can have on their work. The last thing is meant to ensure people become part of the development of the tool and with that setting a first step toward triggering a behavioral change. The support tool will be most helpful in more extreme situations where it is more difficult to make an estimation, which influences the conditions of the scenarios.

Link to Research Question Performing an experiment to measure EAT adherence and get controllers to familiarize with and experience the tool, links to the following research questions. First of all, the interface is based on information gained while answering RQ 1-3 (concerning required information for support, trajectory prediction foundations and display design principles). Then, the influence on controller performance is measured, which answers RQ 4 (how does the proposed support system influence controller performance).

6.3.1. Goal

The quantitative goal of the experiment is to measure the influence of the support tool on EAT adherence. In order to do this, several variables will be logged during the experiment. An overview is given in Table 6.1, indicating what is to be measured, why, and what exactly needs to be logged. AC comprises the following aircraft parameters: ID, current FL, cleared FL, TAS, heading, AC type, predicted EAT adherence error. These are also the variables that are in the aircraft label during the experiment in SectorX.

The qualitative goal is to evaluate what aspects the controllers found pleasant to work with and what can be adjusted or what they'd like to see implemented in the display. These things are discussed in the survey and semi-structured interviews. The controllers are also asked to think out loud during the experiment In order to give this input, it is necessary that they have had a first-hand experience with the tool first.

6.3.2. Scenarios

The scenarios will involve strong wind conditions so that the tool will have the most influence as it is realistic: holding normally occurs only in extreme situations, for example when weather conditions are extreme and landing capacity is low. Additionally, the scenarios will have a completely filled holding stack since the vertical view and dedicated holding stack controller will only be present when many aircraft are in the holding. The stack will be full from the beginning of the experiment, as the goal is to measure the impact of the tool under realistic workload conditions.

6.3.3. Operationalization: Variables

This section discusses the variables of the experiment in more detail.

Independent Variables There are two independent variables, namely wind and the presence of the support tool (predicted EAT adherence for turn in now and ecology dots, for reference see Figure 4.6). It is logged per participant whether this was present in the first or second scenario. The rest of the interface is kept constant as described under control variables.

The type of wind fields (intensity, direction) and amount of variation over time that will be the same for each participant. The exact direction of the wind field will differ per scenario as to prevent bias in the second scenario. To limit the confounds created by the wind fields, some further research is to be done on the most suitable angles.

Dependent Variables The dependent variables are EAT adherence error and an indication of workload from the amount of control actions. These variables will be logged by the computer; the full list of logged variables can be seen in Table 6.1. Perceived workload is assessed through a survey question, as discussed under Section 6.4.

Control Variables The control variables are: number of aircraft, scenario duration, pilot reaction time, available control actions and control panel layout (speed, heading, altitude, turn to IAF/SPL), simulation layout (radar display, presence of vertical view, aircraft with labels and history dots, clock), IAF. The IAF is ARTIP. The scenarios will involve a set of different aircraft that is kept constant over the different scenarios, but the order will not be exactly the same. A new aircraft enters the stack at FL 240-260 every $4 \pm 1.5\sigma$ minutes, with their airspeed varying between 220-330 kts to ensure a realistic scenario. The distribution of these variables will be similar over both scenarios.

Pilot reaction time will be modeled via a fixed gamma-distributed array that contains the same numbers in a different order over each scenario. The distribution is chosen such that it represents real-life pilot reaction times. Its modulus will be $15s$, as this is the most common pilot reaction time, and the minimum will be $10s$ while the maximum will be $25s$[2].

6.4. Survey

The aim of the survey is to evaluate how the participants experienced the support tool and its specific components.

Operationalization It will make use of a digital questionnaire where the participants need to give a rating, plus in some cases room for additional comments or feedback. The rating will be done on an ordinal scale with five levels. It is considered to use the acceptance scale by van der Laan et al. [1997]. Further research will be done on the exact type of scale when creating the survey. Room for comments is also present in the survey to provide more depth to the answers. The combination is preferred as the use of a scale allows for more straightforward comparison of participants and gives a quick and clutter-free idea of their opinions and standpoints.

Besides the full survey, the participants will be asked to score their perceived workload after both scenarios. There will also be a question in the survey where they are asked to rate the relative workload of the scenarios (with and without tool) to each other. Other questions will concern screen clutter, the usefulness of the different aspects of the tool, how the tool was experienced in general. The final question in the survey will concern the participants stance toward innovation in the systems, and facilitate the transition towards the semi-structured interview on this topic.

[2] Numbers based on interviews with F. Dijkstra, KDC (2021)

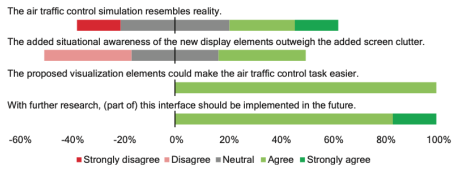


Figure 6.1: Example of visualization of survey results by Ottenhoff et al. [2020]

Link to Research Question The survey is aimed at gaining knowledge about how the controller experienced the interface, and therefore answers RQ 2 and RQ 4(c): the presentation of the trajectory prediction and how to evaluate the tool in a subjective manner.

Sampling The data gathered from the survey will be visualized using sampling methods analogous to Ottenhoff et al. [2020]. The reason for this is that this type of scale does not imply that the ordinal results can be compared quantitatively, as they are subjective, but it does give the reader a quick and visual insight on the outcome of the survey. The full results of the survey, including all feedback, will be presented in the appendix of the report, while the most relevant results and quotes will be included in the research article.

6.5. Semi-structured interview

At the end of the session, a semi-structured interview will be held. The nominal time for this is set to 25 minutes, but if a participant wishes to elaborate further and provide more insights, extra time is reserved. The aim is twofold: first of all, it is to explore further suggestions and ideas for improving the holding support interfaces and get a more in-depth view on how ATCos perceive current systems and the proposed system. The second goal is to explore people's emotions related to innovation as well as EAT adherence, and to get a better understanding on the possibilities and limitations when aiming to procure a behavioral change within LVNL.

Operationalization Questions will be asked about what level of EAT adherence is important and why, how they feel about research projects such as the one at hand, and whether they are happy with the level to which they are involved.

Link to Research Question The semi-structured interview will be used to answer RQ 5 and specifically to explore the emotions that form the basis for people's stance toward the project. These emotions will then be linked to the theoretical framework, with the goal of using emotions as a predictor for the success of behavioral change under the implementation of new technologies.

Sampling The sampling of the semi-structured interview will be done by open coding, after which axial coding will be done. The reason for this is that open coding allows for breaking up the data into categories without introducing bias from the start. Since each participant can have different emotions and responses, the goal is not to identify one central emotion, but rather to explore the relation between a participant's responses and be able to later link it to the theories of behavioral change. The full transcripts will be included in the appendix, while some of the most relevant quotes and the most relevant findings will be presented in the article.

6.6. Expected Results

The research at hand is focused on the following objective: to make sure EAT adherence goes down from 2 minutes error to 30 seconds error or less, by means of designing, testing and evaluating a graphic support interface for ACC. This report introduces a visual support system that has the aim of doing so, and discusses the theoretical foundations of triggering the behavioral change that is required to valorize that system. By doing so, it gives (preliminary) answers to the sub-questions:

1. *What information would support an ATCo in ensuring an aircraft can adhere to a desired EAT upon passing the IAF?* This has been answered in Chapter 2 and have been implemented in the proposed display layout.

2. *How is trajectory prediction data best presented using the principles of EID?* This has been researched in Chapter 4 and its findings have been implemented in the proposed display design.

3. *What are the uncertainties software and ATCos have to cope with in case of trajectory predictions, and what level of detail is needed in modeling them?* The trajectory prediction methods and corresponding assumptions are discussed in Chapter 3 and form the basis of the prediction algorithm used.

4. *How does the proposed support system influence controller performance?* The foundations for answering this question have been laid in this chapter, and will be measured in the continued research.

5. *How can a behavioral change be triggered in ATCos, by means of or stimulated by a novel graphic support interface (technological innovation)?* The foundations for answering this question have been laid in Chapter 5 (theoretical framework) and in this chapter (research plan).

Expected results are a measure of EAT adherence improvement, qualitative feedback for the further improvement of the tool, and a preliminary theory on how emotion can be seen as a predictor for behavioral change under the implementation of a new technology. The relevance of the research can be found in these results, as it adds to the body of knowledge on using EID for ATC tooling and specifically on holding patterns, and it also introduces preliminary insights on the relation between emotion and behavioral change. It is noted that due to the exploratory character of this study, any relation that will be found must be subjected to further research in order to verify its validity.

6.7. Planning and Future Work

Before performing the experiment, several things need to be finished first. These are outlined below.

Programming SectorX The programming steps that need to be taken in SectorX are shown in Table 6.2.

Table 6.2: Programming steps that need to be taken in SectorX

Task	Expected time
Add full aircraft labels to stack list	1/2 day
Show ecology dots (predicted EAT at possible turn-in locations) in top view SectorX	1 day
Verify ecology dots location	3 days
Program stack list	5 days
Create training scenarios	2 days
Create experiment scenarios in SectorX	5 days
Create on-off option for tool	1 day
Create data logger	3 days

Social Research The steps that need to be taken for the social research are shown in Table 6.3.

Table 6.3: Steps that need to be taken for social research

Task	Expected time
Create survey	3 days
Create semi-structured interview protocol	3 days

Pilot After the above-mentioned work has been completed, a pilot will be performed on 2–3 peers to test whether all scenarios work accordingly and to overcome any problems in the experiment, survey and interview setup. Findings from the pilot will be used to further improve the setup. It is noted that the pilot is not performed on professional ATCos and therefore is not fully representative, but this is still the preferred method due to the limited amount of ATCos that are available as a participant in order to prevent confounds due to a learning effect.

Experiment Preparations After the full preparations have been performed, participants will be emailed with an invitation for the experiment. The invitation setup can be found in Appendix A.

Final Thesis After the experiment has been performed, the results are processed and analyzed. Then, the final thesis paper is to be written and the presentation to be planned, made, and given.

Bibliography

L. Bainbridge. Ironies of automation. In *Analysis, design and evaluation of man–machine systems*, pages 129–135. Elsevier, 1983.

M. Bekier, B. Molesworth, and A. Williamson. Tipping point: The narrow path between automation acceptance and rejection in air traffic management. *Safety Science - SAF SCI*, 50, 02 2012. doi: 10.1016/j.ssci.2011.08.059.

C. Borst. Course slides from “AE4302: Avionics and Operations”, 2019.

E. Casado, C. Goodchild, and M. Vilaplana. Identification and initial characterization of sources of uncertainty affecting the performance of future trajectory management automation systems. ATACCS ’12, page 170–175, Toulouse, FRA, 2012. IRIT Press. ISBN 9782917490204.

F. Davis, R. Bagozzi, and P. Warshaw. User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35:982–1003, 08 1989. doi: 10.1287/mnsc.35.8.982.

T. De Jong and W. R. Van Joolingen. Scientific discovery learning with computer simulations of conceptual domains. *Review of educational research*, 68(2):179–201, 1998.

A. Degani, C. Goldman, O. Deutsch, and O. Tsimhoni. On human–machine relations. *Cognition, Technology Work*, 19, 09 2017. doi: 10.1007/s10111-017-0417-3.

M. Dirkzwager, C. Borst, M. Mulder, and M. M. Van Paassen. Design and evaluation of a visual interface for separation support in time-based approach air traffic control. Master’s thesis, TU Delft, 2019.

E.S.Bakker. Design and evaluation of a visual interface for an en route air traffic control merging task. Master’s thesis, TU Delft, 2019.

S. Gill. *Cognition, Communication and Interaction: Transdisciplinary Perspectives on Interactive Technology*. 01 2008. ISBN 978-1-84628-926-2. doi: 10.1007/978-1-84628-927-9.

B. Guiost, S. Debernard, T. Poulain, and P. Millot. Task allocation in air traffic control involving a common workspace and a cooperative support system. *IFAC Proceedings Volumes*, 39(4):90–96, 2006. ISSN 1474-6670. doi: https://doi.org/10.3182/20060522-3-FR-2904.00015. URL https://www.sciencedirect.com/science/article/pii/S1474667015330342. 9th IFAC Symposium on Automated Systems Based on Human Skill and Knowledge.

ICAO. International civil aviation organization, global air traffic management operational concept. URL https://www.icao.int/Meetings/anconf12/Document%20Archive/9854_cons_en%5B1%5D.pdf.

KDC. Interviews with the knowledge development centre, f. dijkstra, mainport Schiphol, september 2020 - april 2021, 2021.

KLM. B777 flight crew operation manual, 02 2019. Flight, Management, Navigation – FMC Descent and Approach.

A. Lee, S. Weygandt, B. Schwartz, and J. Murphy. Performance of trajectory models with wind uncertainty. In *AIAA modeling and simulation technologies conference*, page 5834, 2009. URL https://arc.aiaa.org/doi/pdf/10.2514/6.2009-5834.

LVNL. Luchtverkeersleiding Nederland, amsterdam/schiphol standard arrival chart - instrument. URL https://www.lvnl.nl/eaip/2020-09-24-AIRAC/html/index-en-GB.html.

LVNL. Luchtverkeersleiding Nederland, dataset containing flight data around ARTIP, including holding data, 2019.

LVNL. Luchtverkeersleiding Nederland, corporate vision and strategy, 2020a. URL `https://en.lvnl.nl/about-lvnl/corporate-vision-strategy`.

LVNL. Luchtverkeersleiding Nederland, exploratory and explanatory interviews with air traffic controllers from LVNL, j. bekkers and j. dijkstra, september - april 2021, 2020b.

L. Mac an Bhaird, C. Borst, and M. Mulder. Design evaluation of holding stack management tool. Master's thesis, TU Delft, 2020.

C. Magaña and E. Juan. Trajectory prediction uncertainty modelling for air traffic management. 2016. URL `http://theses.gla.ac.uk/7700/1/2016CasadoPhD.pdf`.

V. D. Miller, J. R. Johnson, and J. Grau. Antecedents to willingness to participate in a planned organizational change. 1994.

S. Mondoloni and N. Rozen. Aircraft trajectory prediction and synchronization for air traffic management applications. *Progress in Aerospace Sciences*, 119, 09 2020. doi: 10.1016/j.paerosci.2020.100640.

M. M. Ottenhoff, C. Borst, and M. Mulder. Effects of wind and trajectory uncertainty in a 4d trajectory management interface. Master's thesis, TU Delft, 2020.

G. Paquette, M. Léonard, K. Lundgren-Cayrol, S. Mihaila, and D. Gareau. Learning design based on graphical knowledge-modelling. *Educational Technology Society*, 9:97–112, 01 2006.

D. C. Phillips. The good, the bad, and the ugly: The many faces of constructivism. *Educational researcher*, 24 (7):5–12, 1995.

C. Prince and E. Salas. Team processes and their training in aviation. in: Handbook of aviation human factors. 1999.

J. Rasmussen. Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *Systems, Man and Cybernetics, IEEE Transactions on*, SMC-13, 01 1987. doi: 10.1109/TSMC.1983.6313160.

T. Reynolds, Y. Glina, S. Troxel, and M. McPartland. Wind information requirements for nextgen applications phase 1: 4d-trajectory based operations (4d-tbo). 2013.

T. Reynolds, M. McPartland, T. Teller, and S. Troxel. Exploring wind information requirements for four dimensional trajectory-based operations *. 2015.

É. Robert. Comparison of operational wind forecasts with recorded flight data. 2013.

D. J. Schroeder, L. Bailey, J. Pounds, and C. Manning. A human factors review of the operational error literature. 2006.

SKYbrary. Holding pattern, date accessed: 2020-10-12, a. URL `https://www.skybrary.aero/index.php/Holding_Pattern`.

SKYbrary. Rate of turn, date accessed: 2020-11-16, b. URL `https://www.skybrary.aero/index.php/Rate_of_Turn`.

A.-M. Teperi and A. Leppänen. Learning at air navigation services after initial training. *Journal of Workplace Learning*, 22:335–359, 08 2010. doi: 10.1108/13665621011063469.

M. Vakola. *Resistance to change: Technology implementation in the public sector*, pages 112–117. 01 2012. doi: 10.4135/9781483387444.n12.

J. van der Laan, A. Heino, and D. De Waard. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research. Part C: Emerging Technologies*, 5(1):1–10, Feb. 1997. ISSN 0968-090X.

J. van der Linden, G. Erkens, H. Schmidt, and P. Renshaw. Collaborative learning. In *New learning: Three ways to learn in a new balance*, chapter 3. Springer, 2000.

K. Vicente and J. Rasmussen. On applying the skills, rules, knowledge framework to interface design. *Proc. Ann. Meet. Human Fact. Ergon. Soc.*, 32:254–258, 10 1988. doi: 10.1177/154193128803200501.

K. J. Vicente and J. Rasmussen. Ecological interface design: theoretical foundations. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(4):589–606, 1992.

C. Wanberg and J. Banas. Predictors and outcomes of openness to changes in reorganizing workplace. *The Journal of applied psychology*, 85:132–42, 03 2000. doi: 10.1037/0021-9010.85.1.132.

C. Westin, C. Borst, and B. Hilburn. Strategic conformance: Overcoming acceptance issues of decision aiding automation? *IEEE Transactions on Human-Machine Systems*, 46:41–52, 11 2015. doi: 10.1109/THMS.2015.2482480.

C. Wickens, A. Mavor, R. Parasuraman, and J. Mcgee. The future of air traffic control: Human operators and automation. 01 1998.

K. Zeghal and F. Dowling. 4d trajectory management: an initial controller perspective. URL `https://www.eurocontrol.int/eec/public/standard_page/EEC_News_2008_1_4DTM.html`.

III

Book of Appendices

A

Experiment Documents

A.1. Experiment Invitation

The experiment invitation e-mail text is presented in the frame below, and the invitation in Figure A.1.

Beste ACC's,

Bij deze wil ik jullie graag uitnodigen mee te doen aan een experiment om de vertical view/stack te verbeteren voor tijdens het holden - dus op het moment dat er een dedicated stack controller is.

Wie vraagt dit eigenlijk aan jullie? Leuk dat je de tijd neemt om tot zo ver te lezen! Ik ben Stephanie Wiechers, 26 jaar, en studeer op dit moment communicatie en lucht- en ruimtevaart in Delft. Oorspronkelijk kom ik uit Breda, maar woon nu al een tijdje in Rotterdam, en vind het leuk om in mijn vrije tijd te tennissen en te pottenbakken. Sinds oktober ben ik bezig mijn master afstudeeronderzoek te doen dat gaat over holding stacks en dan specifiek de invloed van het indraaimoment op de EAT. Hierbij heb ik al veel samengewerkt met Jonah en Jorien, en ben ook al eens mee gaan kijken in de sim om te zien hoe het er in het "echt" aan toe gaat. Maar goed, dé manier om er achter te komen wat jullie belangrijk vinden is natuurlijk door het jullie in het echt te vragen. Vandaar deze oproep.

Het verdere idee is dat over een paar jaar een aantal nieuwe hulpmiddelen in de vertical view beschikbaar zijn + eventueel in de radarschermen om ondersteuning te bieden bij het indraaien. Met zo'n tool wordt het voor jullie makkelijker in te schatten wanneer je kan indraaien, en ook overzichtelijker - dat is in ieder geval de bedoeling. Aangezien de ontwikkeling nu nog in een vrij vroeg stadium is, is dit hét moment om dingen aan te passen en precies te maken zoals jullie ze graag zouden zien! Het enige nadeel is wel dat de uitvoering dus niet meteen morgen in de systemen gaat zitten.

In het kort Op basis van wat ik allemaal heb gezien en gehoord in de afgelopen 8 maanden, heb ik een eerste idee voor een tool in elkaar gezet. We zullen aan de slag gaan met een gesimuleerd holding scenario in een iets versimpelde interface, waarbij jullie zelf kunnen ervaren of het prettig werkt, wat nuttig is, wat juist niet, of dat er onderdelen zijn die jullie graag in de systemen zouden terugzien of dat het juist helemaal anders moet.

Na het draaien van het holding scenario is er ook nog genoeg tijd ingecalculeerd waar ik graag van jullie wil horen hoe je het hebt ervaren. Dit is dan ook meteen een oproep om eventueel van te voren na te denken over wat er nu nog minder goed werkt in de praktijk met holden en op welke punten jullie graag verbetering zouden willen zien.

Wanneer Tussen XXX t/m XXX. In het totaal (scenario + feedback) gaat het ongeveer 2 uur duren, omdat ik jullie er niet te veel mee wil belasten en van zo veel mogelijk feedback wil krijgen!

Qua tijden is het super flexibel: bijvoorbeeld voor/na een shift als je toch op LVNL bent. Als je thuis aan de slag bent en het liever vanaf daar wilt doen dan ga ik er alles aan doen om dat te regelen. Dus laat me vooral weten wat voor jou goed uitkomt en dan gaan we dat plannen. Als je het nog niet helemaal weet maar wel graag mee doet kunnen we ook samen naar een moment zoeken.

Extra extra Last but not least: voor wie het leuk vindt is er een klassement. Wie het scherpst op de EAT kan sturen met de tool wint. Meedoen is geheel vrijwillig, maar ik kan wel verklappen dat de winnaar een taart krijgt!

Ik hoop van jullie te horen!

Groetjes, Stephanie

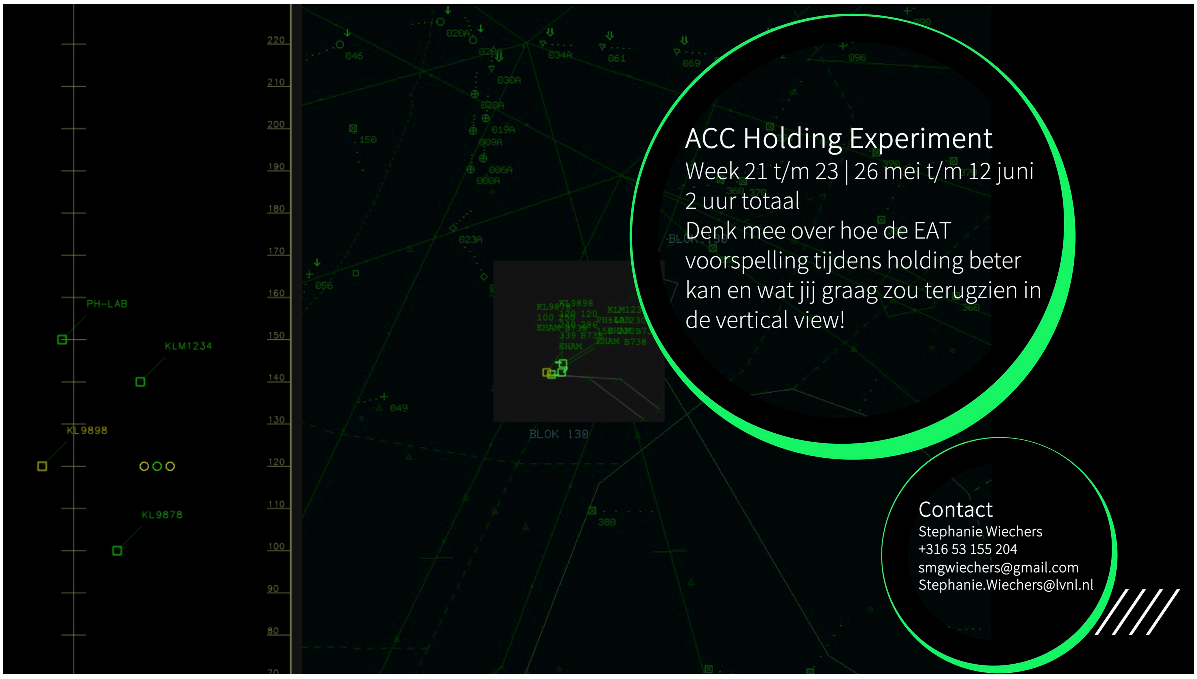


Figure A.1: Experiment invitation

A.2. Experiment Briefing

The experiment briefing and tool explanation was done in person. The accompanying slides can be found in this section. Participants were encouraged to ask questions and it was verified during the briefing whether all information was clear to them. After the briefing, participants started the training scenario.

Experiment Holding Support LVNL

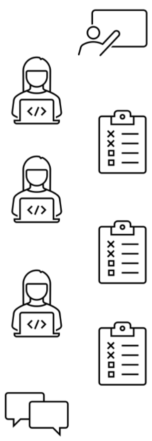
Augustus 2021

TU Delft LVNL

1

Planning

- Introductie & uitleg
- Oefenscenario
- Kort vragenmoment
- Scenario 1: met tooling
- Kort feedbackmoment
- Scenario 2: zonder tooling
- Vragenlijst
- Uitgebreide feedback



TU Delft LVNL

2

Doel van het experiment

- We gaan een tool testen voor holding
- Taken tijdens het experiment
  - Stack management @ ARTIP
  - Commands: dalen in stack, heading 70/extended leg, uit holding
  - Geen onderdeel: praten met piloot, vluchten in holding stack inweven, alternatieve headings geven dan 70
- EAT adherence op ±0s krijgen!
- **Tijd** is focus (meer dan sequence/level/separatie)
- Tijdens alle delen van het experiment mag je zeggen wat in je opkomt (bv. interessante situatie want…)

TU Delft LVNL

3

Doel van het experiment

Exploratief onderzoek
- Kan tooling ondersteunen in hogere EAT adherence?
- Werkt de tooling prettig?
- Hoe zien jullie de link tussen onderzoek naar, implementatie en gebruik van nieuwe tooling?
- Alle mogelijke feedback verzamelen!

Functionaliteit || Implementatie & Gebruik

TU Delft LVNL

4

De interface en de tool

- De tool maakt uiteindelijk gewoon data inzichtelijk zodat jullie die beter kunnen gebruiken om zelf beslissingen te nemen.
- Om te kunnen vergelijken en ervaren of het prettiger werkt met of zonder tooling gaan we twee scenario's runnen: een met en een zonder.

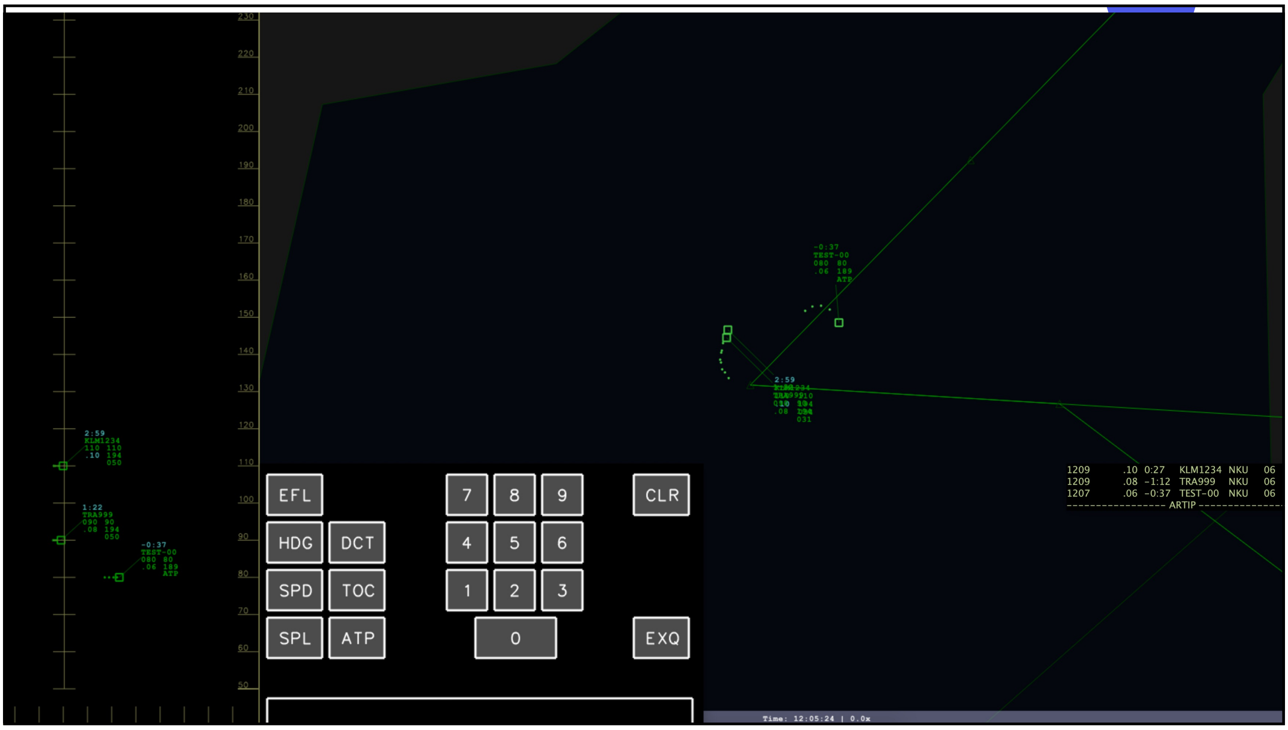
TU Delft LVNL

5

Basics interface

TU Delft LVNL

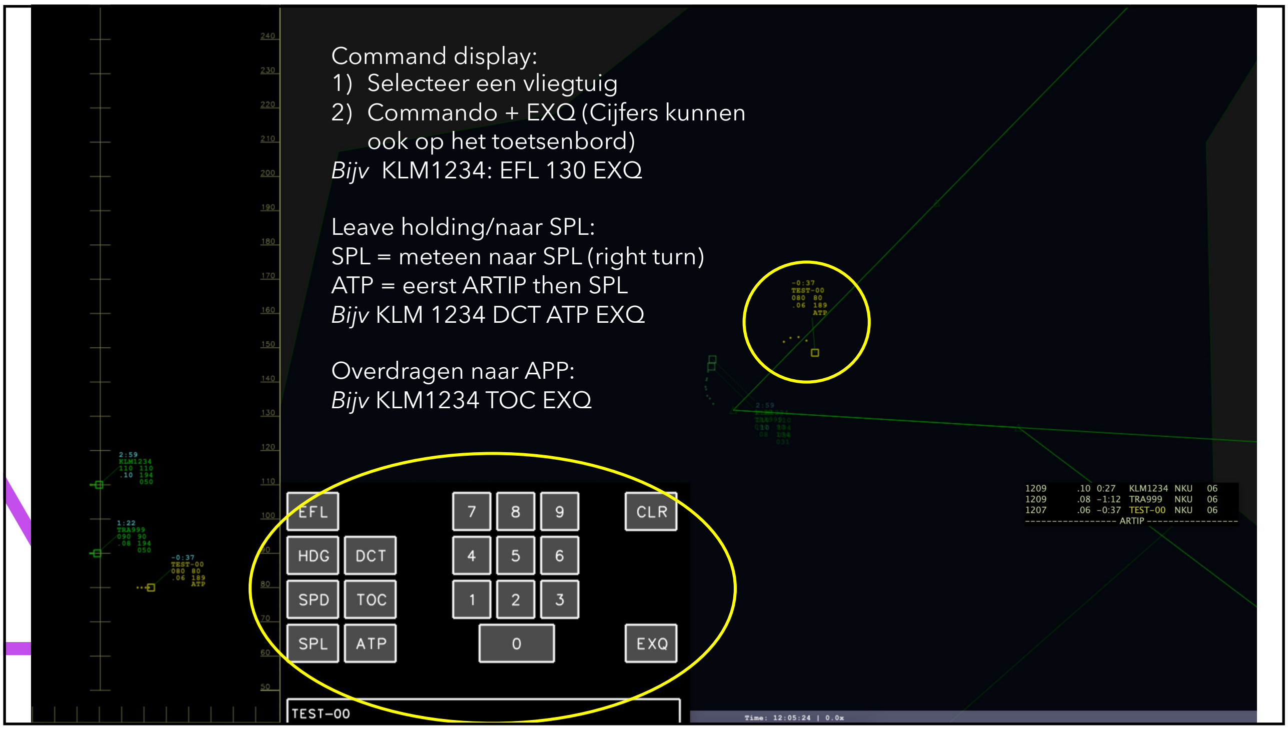
6


EFL
HDG
DCT
SPD
TOC
SPL
ATP
7
8
9
4
5
6
1
2
3
0
CLR
EXQ
1209 .10 0:27 KLM1234 NKU 06
1209 .08 -1:12 TRA999 NKU 06
1207 .06 -0:37 TEST-00 NKU 06
ARTIP

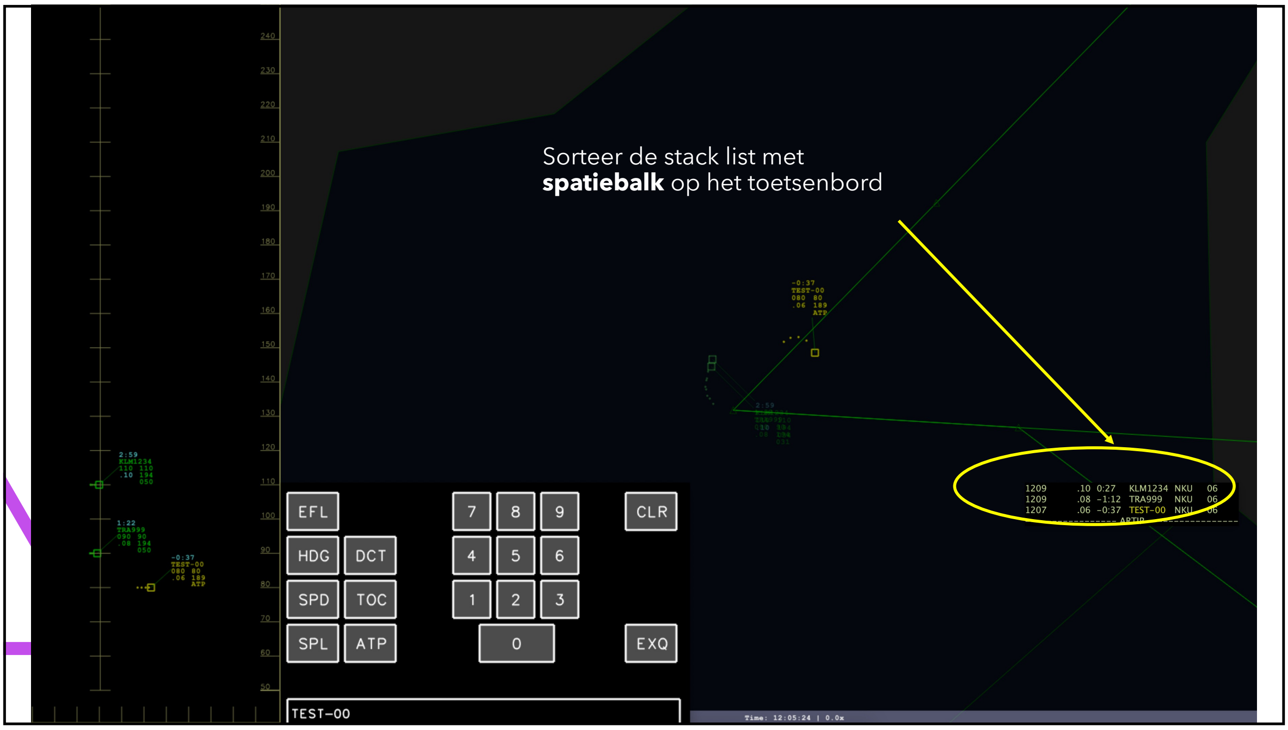
7


Selecteer een vliegtuig door er op te klikken op al deze plekken
- Het geselecteerde vliegtuig is geel
- Basic interface: alles blijft fel
- Tool: niet-geselecteerde vliegtuigen faden in top view
Symbool & label
Symbool & label
AC ID
EFL
HDG
DCT
SPD
TOC
SPL
ATP
7
8
9
4
5
6
1
2
3
0
CLR
EXQ
TEST-00

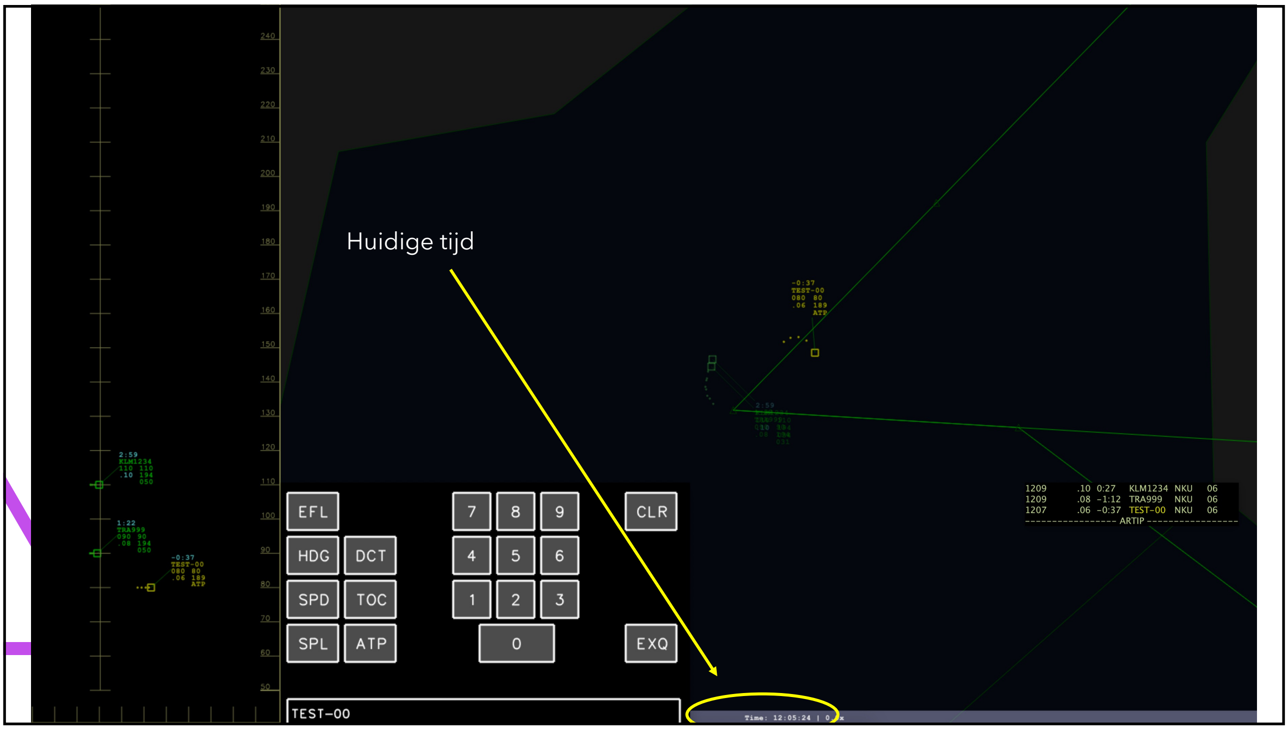
8


Command display:
1) Selecteer een vliegtuig
2) Commando + EXQ (Cijfers kunnen ook op het toetsenbord)
Bijv KLM1234: EFL 130 EXQ
Leave holding/naar SPL:
SPL = meteen naar SPL (right turn)
ATP = eerst ARTIP then SPL
Bijv KLM 1234 DCT ATP EXQ
Overdragen naar APP:
Bijv KLM1234 TOC EXQ
EFL
HDG
DCT
SPD
TOC
SPL
ATP
7
8
9
4
5
6
1
2
3
0
CLR
EXQ
TEST-00

9


Sorteer de stack list met **spatiebalk** op het toetsenbord
EFL
HDG
DCT
SPD
TOC
SPL
ATP
7
8
9
4
5
6
1
2
3
0
CLR
EXQ
TEST-00

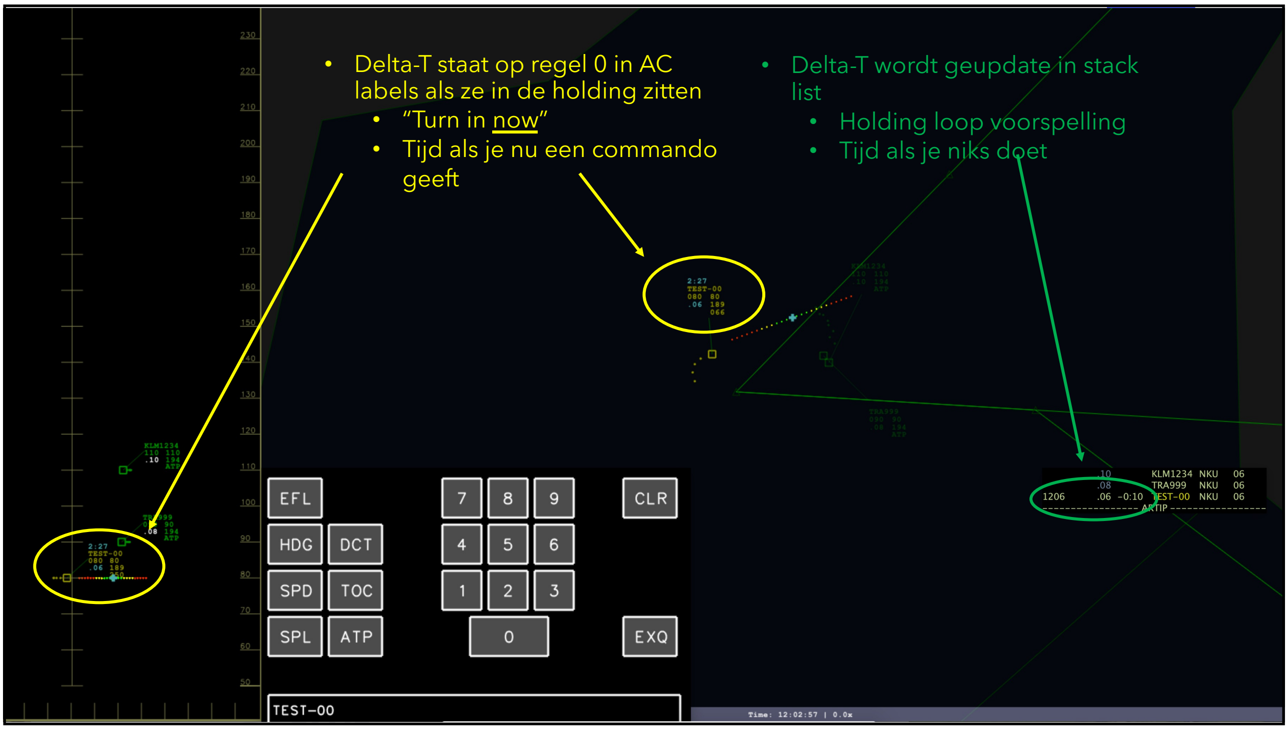
10



11

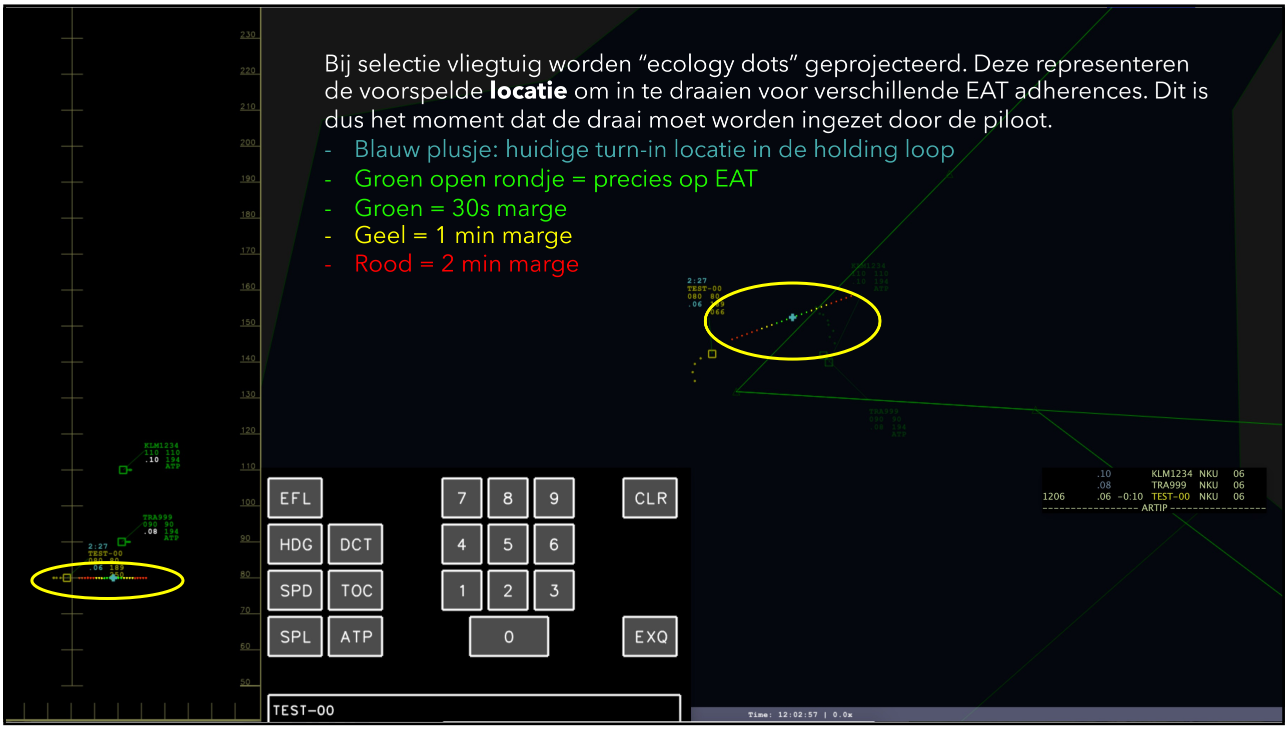
Features tool

12



• Delta-T staat op regel 0 in AC labels als ze in de holding zitten
• "Turn in now"
• Tijd als je nu een commando geeft
• Delta-T wordt geupdate in stack list
• Holding loop voorspelling
• Tijd als je niks doet
EFL
HDG DCT
SPD TOC
SPL ATP
7 8 9
4 5 6
1 2 3
0
CLR
EXQ
TEST-00

13



Bij selectie vliegtuig worden "ecology dots" geprojecteerd. Deze representeren de voorspelde **locatie** om in te draaien voor verschillende EAT adherences. Dit is dus het moment dat de draai moet worden ingezet door de piloot.
- Blauw plusje: huidige turn-in locatie in de holding loop
- Groen open rondje = precies op EAT
- Groen = 30s marge
- Geel = 1 min marge
- Rood = 2 min marge
EFL
HDG DCT
SPD TOC
SPL ATP
7 8 9
4 5 6
1 2 3
0
CLR
EXQ
TEST-00

14

Opmerkingen en verdere uitleg

- Er is pilot delay van ~15s voor een realistisch verloop. De tooling houdt hier rekening mee.
- Er zijn een hoop dingen net anders dan de LVNL tooling. Het uiterlijk hiervan is geen onderdeel van het onderzoek, ik heb wel heel erg m'n best gedaan om dingen zo realistisch mogelijk te maken maar er zijn wat dingen anders. Ik hoop dat het lukt om hier in je feedback doorheen te kijken en te focussen op de tooling.
  - Bv niet-doorzichtige stack list, vrij groot command panel, soms zullen labels misschien net afwijken, scherm is natuurlijk een stuk kleiner, alle vliegtuigen komen vanaf NKU in de stack list

TU Delft LVNL

15

BELANGRIJK

- In de eerste paar seconden is het belangrijk dat je nog nergens op klikt. Dat heeft te maken met de simulatiesoftware, die vastloopt als hij probeert berekeningen te maken zonder dat de vliegtuigen al hebben gevlogen. Na een paar seconden is er genoeg data en is er niks aan de hand!
- In het eerste rondje duurt het even voor het programma herkent dat alles in de holding zit. Daarom heeft hij als de simulatie net start nog niet voor alles een voorspelling. Dit zal in het echt natuurlijk niet zo zijn!

TU Delft LVNL

16

Tijden stack list / Delta-T

- Stack list: delta-T wanneer je geen actie onderneemt (blijf holding vliegen)
- Label delta-T: wat als je **nu** indraai **commando** geeft

TUDelft LVNL

17

Verloop experiment

Focus punten feedback

- Nuttig voor holding support?
- EAT adherence met/zonder tooling
- Onderdelen: delta-T en ecology dots
- Zou je de tool willen gebruiken in het echt?
- Verbeteringen en mogelijke extra features

TUDelft LVNL

18

Leuk dat je meedoet en veel plezier!

19

A.3. Survey and Interview Questions

The full survey and interview questions and protocol can be found below. The scales used to evaluate perceived usefulness and satisfaction are taken from van der Laan et. al. (1997).

Survey and Interview Questions

Design and Evaluation of a Visual Support Tool and Exploring the Emotional Relation Between Air Traffic Controller and Interface Innovation

By Stephanie Wiechers

Introductie

Dit document vormt de basis van de subjectieve vragen die betrekking hebben op het experiment wat we gaan uitvoeren, holding support (zowel in het algemeen als specifiek over het concept), en is daarnaast bedoeld om op in het kader van een wetenschappelijk onderzoek informatie te verzamelen.

Vertrouwelijkheid

Alle gegevens die je hier invult zullen vertrouwelijk worden gebruikt. De enige mensen die er toegang tot hebben ben ik plus, als het nodig is, mijn directe afstudeerbegeleiders. In alle gevallen zal ik ervoor zorgen dat dit geanonimiseerd wordt verwerkt. Alles wat je invult wordt vertrouwelijk behandeld en op een anonieme manier in mijn afstudeerverslag verwerkt (denk aan “P1”). Verder zal ik eventuele quotes altijd eerst aan je voorleggen voor ik ze verwerk. Op de volgende pagina word je gevraagd een “informed consent [Engels]” in te vullen, waarmee ik toestemming vraag je gegevens te verwerken.

Opzet

Er zal steeds als je moet wachten tot na een bepaald onderdeel een lege pagina zijn met “omslaan na …. [onderdeel X]”. Ik zal dit ook aangeven tijdens het experiment.

Timing

Er zijn best veel survey vragen. Deze lijken voor een groot deel vrij veel op elkaar. Je hoeft hier niet super lang over na te denken; alleen als de vraag niet duidelijk is dan is het handig om me om verheldering te vragen! Een aantal van de vragen zijn op nét een andere manier gesteld om zo altijd een helder en duidelijk beeld van jullie mening te krijgen. Om zo veel mogelijk tijd over te houden voor feedback en een gesprek aan het einde wil ik dus vragen om niet al te veel na te denken bij het invullen van de survey vragen maar gewoon je eerste gedachte op te schrijven.

Informed consent

Please tick the appropriate boxes	**Yes**	**No**
Taking part in the study		
I have read and understood the study information dated __/08/2021, 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.	☐	☐
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.	☐	☐
I understand that taking part in the study involves a simulation experiment combined with survey questions and an audio-recorded semi-structured interview, both of which will be destroyed after completing the research.	☐	☐
Risks associated with participating in the study		
I understand that taking part in the study involves the following risks: potential mental discomfort through reflective insights.	☐	☐
Use of the information in the study		
I understand that information I provide will be used to draft up a report.	☐	☐
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.	☐	☐
I agree that my information can be quoted anonymously in research outputs.	☐	☐
I agree to joint copyright of the written information during the workshop to the researcher.	☐	☐
Future use and reuse of the information		
I give permission for the anonymized audio transcripts and anonymized interview data that I provide to be archived as long as the research lasts.	☐	☐
I give permission for the anonymized survey data and interview transcripts to be archived such that they can be used for future holding support researches.	☐	☐

Signatures

_______________________ _______________________ _________

Name of participant Signature Date

I have presented 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 Signature Date

Algemene vragen

Naam ____

Leeftijd ____

Geslacht ______________________ Links/rechtshandig ______________

Email ____

Telefoon ____

Aantal jaren ervaring als verkeersleider ________________________________

Posities gehad binnen luchtverkeersleiding ______________________________

____

____

____

Ervaring met eerdere innovatie-onderzoeken ______________________________

____

____

____

Holding support

Vind je dat er een holding support systeem moet komen?

Zeker niet ☐ ☐ ☐ ☐ ☐ Heel graag

Waarom? ____

____

____

Wat voor verwachtingen zou je van zo'n systeem hebben? ____________________

____

____

____

Hoe scherp denk je dat EAT adherence in holding nu is? Hij valt 95% van de tijd binnen…

o	o	o	o	o
>2.5 mins	2 mins	1 min	30 s	10 s

Toelichting? Vind je hier iets van? ________________________________

____

____

Uitleg experiment verloop

Training scenario

Gebruiksgemak & werklast

A1. Op basis van de uitleg en training lijkt het gebruiken van de tool me …

Heel moeilijk						Vanzelfsprekend

A2. Op basis van de uitleg en training verwacht ik dat de werklast **met tool** ten opzichte van de werklast zonder tool als volgt verandert:

Neemt sterk af						Neemt sterk toe

Effectiviteit

A3. Deze tool lijkt me:

Nuttig						Nutteloos
Aangenaam						Onaangenaam
Slecht						Goed
Prettig						Vervelend
Effectief						Overbodig
Irritant						Aantrekkelijk
Behulpzaam						Waardeloos
Ongewenst						Gewenst
Maakt me alert						Slaapverwekkend

EAT adherence

Bij de volgende vragen gaat het er om dat hij 95% van de tijd binnen… valt

A4. Welke EAT adherence denk je dat je gaat halen in de simulatie **zonder** de tool?

o	o	o	o	o
>2.5 mins	2 mins	1.5 min	1 min	30 s

A5. Welke EAT adherence denk je dat je gaat halen in de simulatie met het gebruik van de tool?

o	o	o	o	o
>2.5 mins	2 mins	1.5 min	1 min	30 s

A6. Welke EAT adherence denk je dat je zou halen als **alleen de delta-T in de stack list** zou updaten?

o	o	o	o	o
>2.5 mins	2 mins	1.5 min	1 min	30 s

A7. Welke EAT adherence denk je dat je zou halen met **alleen de delta-T in de labels?**

o	o	o	o	o
>2.5 mins	2 mins	1.5 min	1 min	30 s

A8. Welke EAT adherence denk je dat je zou halen met **alleen de ecology dots?**

o	o	o	o	o
>2.5 mins	2 mins	1.5 min	1 min	30 s

Scenario 1

Gebruiksgemak & werklast

B1. Ik vond het werken met de interface in scenario 1…

Heel moeilijk						Vanzelfsprekend

B2. Ik vond de werklast van scenario 1…

Heel erg zwaar						Heel erg licht

B3. Was het realistisch?

Effectiviteit

B4. De mogelijkheden die ik in scenario 1 had om een overzicht van de verkeerssituatie te krijgen en de EAT adherence zo dicht mogelijk naar 0 te krijgen waren…

Nuttig						Nutteloos
Aangenaam						Onaangenaam
Slecht						Goed
Prettig						Vervelend
Effectief						Overbodig
Irritant						Aantrekkelijk
Behulpzaam						Waardeloos
Ongewenst						Gewenst
Maakt me alert						Slaapverwekkend

EAT adherence

B5. Welke EAT adherence denk je dat je hebt gehaald in scenario 1?

o	o	o	o	o
>2.5 mins	2 mins	1.5 min	1 min	30 s

Scenario 2

Gebruiksgemak & werklast

C1. Ik vond het werken met de interface in scenario 2…

Heel moeilijk						Vanzelfsprekend

C2. Ik vond de werklast van scenario 2…

Heel erg zwaar						Heel erg licht
Veel lichter dan S1						Veel zwaarder dan S1

C3. Was het realistisch?

Effectiviteit

C4. De **delta-T** update in de **stack list** was … voor het overzicht van de verkeerssituatie en EAT adherence

Nuttig						Nutteloos
Aangenaam						Onaangenaam
Slecht						Goed
Prettig						Vervelend
Effectief						Overbodig
Irritant						Aantrekkelijk
Behulpzaam						Waardeloos
Ongewenst						Gewenst
Maakt me alert						Slaapverwekkend

C5. De **delta-T** in de **labels** was … voor het overzicht van de verkeerssituatie en EAT adherence

Nuttig						Nutteloos
Aangenaam						Onaangenaam
Slecht						Goed
Prettig						Vervelend
Effectief						Overbodig
Irritant						Aantrekkelijk
Behulpzaam						Waardeloos
Ongewenst						Gewenst
Maakt me alert						Slaapverwekkend

C6.De **ecology dots** waren … voor het overzicht van de verkeerssituatie en EAT adherence

Nuttig						Nutteloos
Aangenaam						Onaangenaam
Slecht						Goed
Prettig						Vervelend
Effectief						Overbodig
Irritant						Aantrekkelijk
Behulpzaam						Waardeloos
Ongewenst						Gewenst
Maakt me alert						Slaapverwekkend

EAT adherence

C7. Welke EAT adherence denk je dat je hebt gehaald in scenario 2?

o	o	o	o	o
>2.5 mins	2 mins	1.5 min	1 min	30 s

C8. Welke EAT adherence denk je dat je zou halen als **alleen de delta-T in de stack list** zou updaten?

o	o	o	o	o
>2.5 mins	2 mins	1.5 min	1 min	30 s

C9. Welke EAT adherence denk je dat je zou halen met **alleen de delta-T in de labels?**

o	o	o	o	o
>2.5 mins	2 mins	1.5 min	1 min	30 s

C10. Welke EAT adherence denk je dat je zou halen met **alleen de ecology dots?**

o	o	o	o	o
>2.5 mins	2 mins	1.5 min	1 min	30 s

Nauwkeurigheid voorspelling

C11. De voorspelling gemaakt door de tool valt 95% van de tijd binnen een marge van …

o	o	o	o	o
1s	5s	10s	30 s	60s +

Interview | Feedback

Vragen

Opzet

1. Motivatie tooling
2. Suggesties, verbeterpunten

Vragen

1. Hoe zie jij nut/noodzaak van een tool voor holding support?
2. Wat voor voor- en nadelen zitten er aan hogere EAT adherence?
3. Hoe zie jij de link tussen het wel/niet hebben van een holding support systeem en de EAT adherence die minimaal gehaald kan worden?
4. En in het specifieke geval van deze tool?
5. Zou je de tool die je net hebt gezien willen gebruiken? Wat wel/niet?
6. Is je gevoel daarover veranderd of bijgesteld tijdens de loop van het experiment?

Suggesties

1. Wat voor suggesties of verbeterpunten zie jij voor een holding support tool?
2. Wat voor onderdelen vind je nog meer belangrijk?
3. Zou je een tool die al die features heeft willen gebruiken? En hoe veel?
4. Denk je dat er nog andere dingen zijn in een verbeter/innovatieproces die jouw motivatie om zo'n tool te gebruiken kunnen veranderen? Bv betrekken ACC, communicatie, implementatie….

B

Additional Results

B.1. Additional Quantitative Results

In Figure B.1 the complete EAT adherence for all participants are shown (not absolute) in a boxplot. It can be seen that for each scenario, the median is close to zero, both with and without tool. For both groups, the spread of EAT adherence is larger when the tool is not present. For both Group 1 and Group 2, it can be seen that the EAT adherence median with the tool is slightly larger than zero and the median with the tool is larger than the median without the tool. This indicates that with the presence of the tool, the EAT was overall crossed earlier than without the tool.

In Figures B.2 to B.6 histograms of the individual EAT adherence results per participant are shown. Some participants do show a large change in EAT adherence error with or without the tool, which can also be attributed to the width of the histogram bins. Especially participants who had a better overall score show less difference in these histogram plots, while participants who exhibited larger EAT adherence errors show more distinct results in these plots between their results with and without tool. These participants are: A1, A3, A4 and B4. None of these participant except for B4 show EAT adherence errors larger than $90s$. It can also be observed that several participants show large differences with and without the presence of the tool. These are: A2, B1, and B5. In all of these cases, the scores without the presence of the tool were much further away from the EAT (zero error) than the scores with the tool.

B.2. Survey and Interview Results on Acceptance of Technology

The participants were presented survey and semi-structured interview questions to explore their attitude toward innovation, using the framework of the Technology Acceptance Model (see Venkatesh and Davis [2000], Westin et al. [2015], Davis et al. [1989], Venkatesh and Bala [2008]). The results and their interpretation are discussed in this section. First, the external variables influencing technology acceptance are discussed based

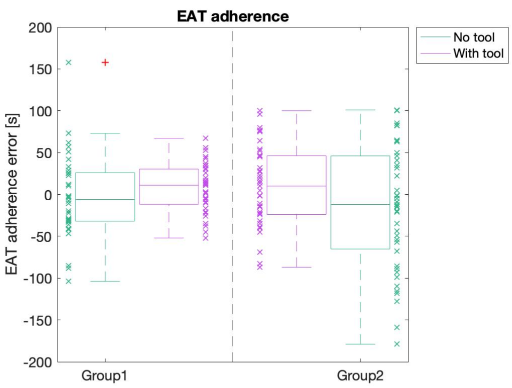
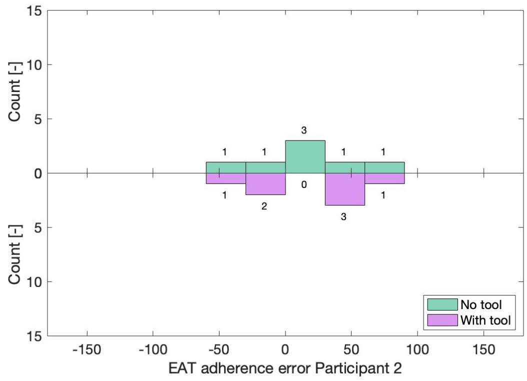
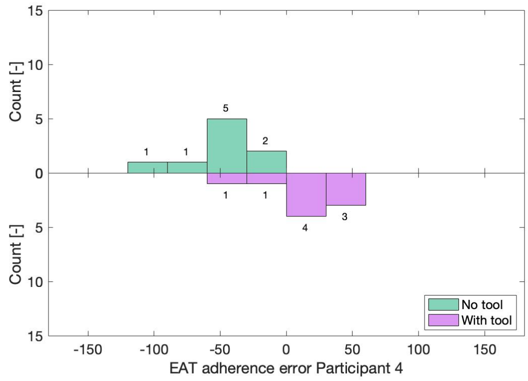


Figure B.1: Full EAT adherence results

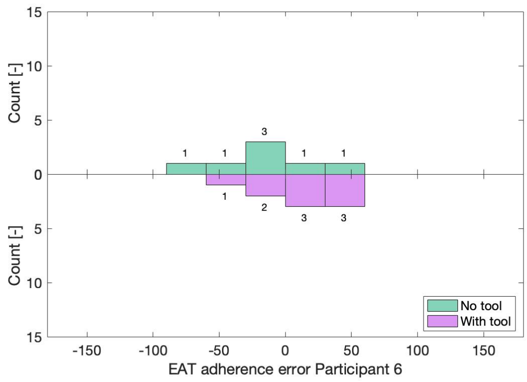


(a) Histogram of EAT adherence results for Participant A1

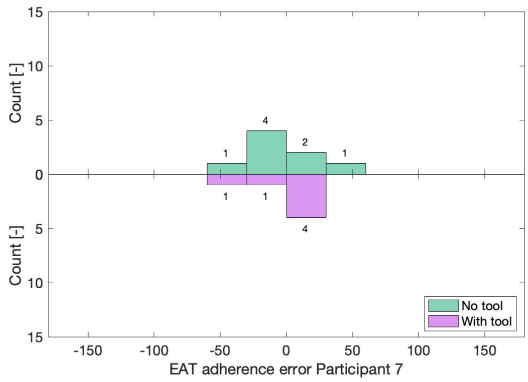


(b) Histogram of EAT adherence results for Participant A2

Figure B.2: Individual EAT adherence results for Participants A1 and A2

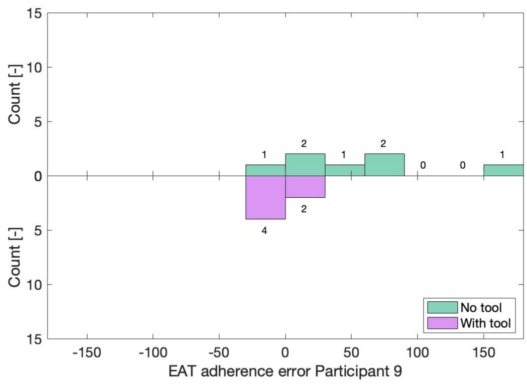


(a) Histogram of EAT adherence results for Participant A3

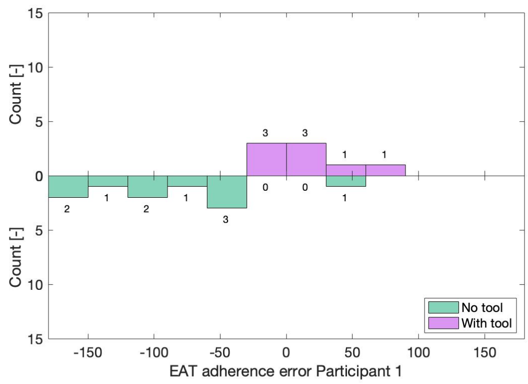


(b) Histogram of EAT adherence results for Participant A4

Figure B.3: Individual EAT adherence results for Participants A3 and A4

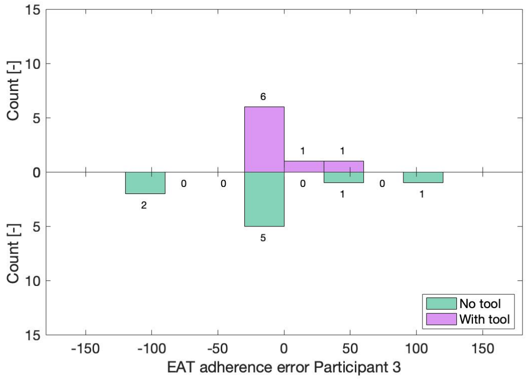


(a) Histogram of EAT adherence results for Participant A5

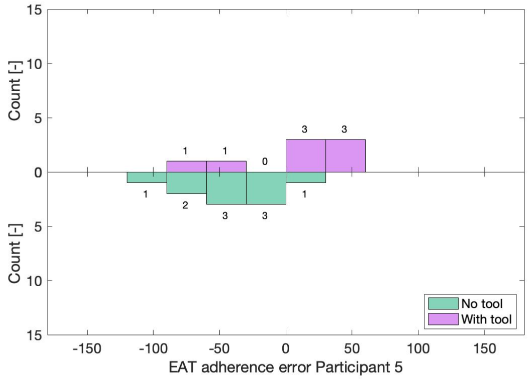


(b) Histogram of EAT adherence results for Participant B1

Figure B.4: Individual EAT adherence results for Participants A5 and B1

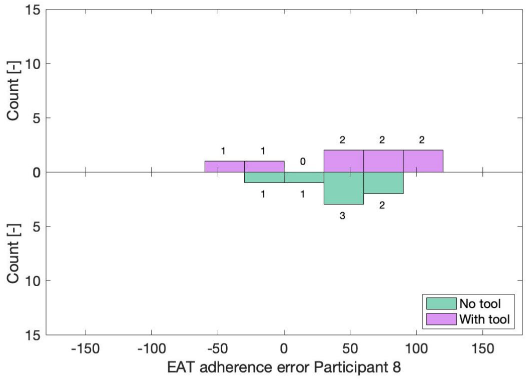


(a) Histogram of EAT adherence results for Participant B2

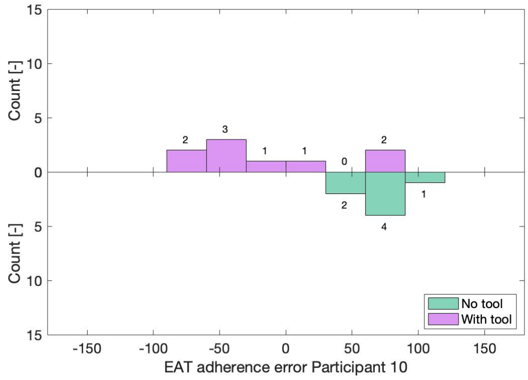


(b) Histogram of EAT adherence results for Participant B3

Figure B.5: Individual EAT adherence results for Participants B2 and B3



(a) Histogram of EAT adherence results for Participant B4



(b) Histogram of EAT adherence results for Participant B5

Figure B.6: Individual EAT adherence results for Participants B4 and B5 on the survey, interviews and EAT adherence scores. Then, the outcome of the survey questions on the perceived ease of use, perceived usefulness and satisfaction of the tool are presented. This is followed by quotes from the interview that can be related to participants' attitude toward use and intention to use. This section is followed by a discussion on the results.

B.2.1. Definition of Variables

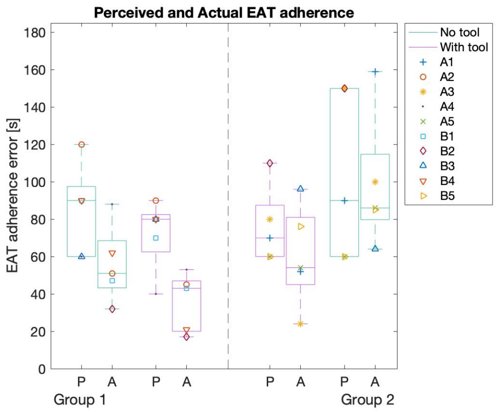
In Table B.1 the definition of the factors that were explored around technology acceptance in the context of this research are outlined.

Table B.1: Context-specific meaning of concepts and indicators for coding

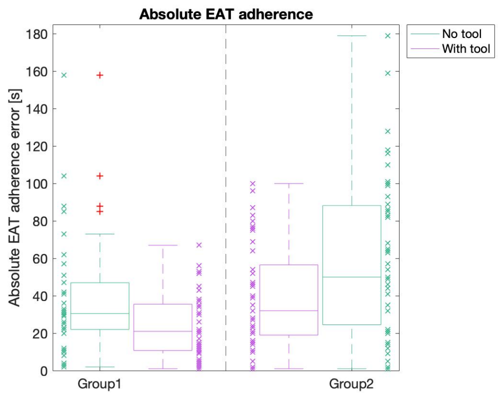
Theory	Concept/Code	Explanation
Mental model	Work domain	The mental picture an ATCo has of the tasks normally performed and support systems used for this; explaining the ATCos outlook on the work domain, EAT adherence
TAM: external variables	Result demonstrability	The level to which the impact on the EAT adherence is clear to the controller
	Output quality	The level of EAT adherence obtained with the tool
	Job relevance	The importance of improved EAT adherence from the perspective of the ATCo
	Image (social influence)	Whether or not an ATCo believes it is socially acceptable to use a support tool to perform their job in the entire organization of LVNL; in coding this also applies when participants think their status is influenced by the use of the tool
	Subjective norm	The idea that using a support tool is either approved or disapproved by the rest of the ACC group; in coding this applies when participants refer to what they think the rest of the group does
	Experience	The amount of experience the participant has as an ATCo in general, with holding stack management in specific, experience with using different strategies and tools for holding support, and general experience with system innovation processes within LVNL; it also refers to quotes about the familiarity of the simulation environment in general compared to actual LVNL systems
	Voluntariness	Whether there was any external factor (e.g., pressure from management, financial compensation) to participate in the study
	Age	Participant age
TAM: components	Perceived ease of use	Degree to which the ATCo believes working with the tool will be intuitive and free of (learning) effort
	Perceived usefulness	Degree to which an ATCo believes that using the particular system would enhance their job performance in terms of usefulness and satisfaction
	Attitude towards use	The emotion an ATCo feels toward the concept and using the proposed tool
	Intention to use	The estimate of the ATCo on their own subjective probability to use the tool when managing a holding stack
Interaction with systems	Autonomy	The degree to which the system allows the ATCo to take ownership in decision-making, so the degree to which the controller stays in charge of the active decision-taking
	Trust	The degree to which an ATCo believes the support system is reliant

B.2.2. External Variables

Result demonstrability It can be seen from the boxplot in Figure B.7a that the EAT adherence error that was obtained by the participants was actually much lower (=better) than the level of EAT adherence error the participants had perceived they managed to obtain. In other words, the participants assess their level of EAT adherence worse than it actually is. This is also reflected by the average perceived output prediction quality at



(a) Result demonstrability: level of actual VS perceived EAT adherence. The participants have been split per group here, where A1–A5 are P1–P5 and B1–B5 are P6–P10.



(b) Output quality: obtained levels of EAT adherence

Figure B.7: Results on actually obtained and perceived EAT adherence

28.5*s*, meaning that the participants expect that the prediction could be maximum 28.5*s* off from the actual optimal point. No direct feedback was given on performance during the experiment; however, it was possible to verify at what moment an aircraft actually crosses the IAF.

Out of all participants, only P6 did not indicate anything related to result demonstrability, with 18 quotes in total. The quotes can be summarized with the following three points: (1) the participants think it is very important that they can rely on the system to make the correct prediction, and they feel the need to be ensured of that [P1, P2, P3, P7, P10], (2) participants indicated that they did perform double checks when the tool was present [P2, P8, P9, P10] and (3) that it is important to be aware of the shortcomings of the system and that the system is able to show where these are [P2, P4, P5].

Participant P9 said, regarding their overall opinion and how the course of the experiment demonstrated the results of the tool: “No, I was sure that it is convenient and useful before. And so it turned out [during the experiment]. Emptying out the holding stack felt much more calm and structured with the tool than without.” On the topic of being able to rely on the system, many participants started their sentences with something similar to “If this works well, ...” [P3].

Output quality In Figure B.7b the EAT adherence that was actually obtained can be found, where the absolute EAT adherence scores are are plotted in boxplots. It can be seen that the scores improve when participants use the tool, and therefore the quality of the output when using the tool improves. The average quality of the solution without the tool was 44.55s error, and with the tool 29.91s error, which is an improvement of 49%. Regarding quality of output, no interview questions were asked but P5 indicated that “when you want to convince people to use the tool, it [the tool output] just has to be right”. Participants P2–P6 indicated similar things about output quality, which can be summarized as that it is important that the system needs to prove that the EAT actually improves; there were 9 quotes in total.

Job relevance All participants except for P7 indicated things about job relevance, with 21 quotes in total. These quotes (all participants except P7) can be summarized by participants outlining the relevance for their job and possibilities in terms of level of EAT adherence (aim at zero error) and the possibility to fly standard arrival routes (improving noise pollution and flight efficiency) as a result of improved adherence. Several participants [P3, P4, P5, P8, P10] indicated that more possibilities for validation using predictions is something they regard as positive. P10 did stress that holding is not something that is done often at LVNL, while P1 indicated that they expected to use the tool maximum a couple of times per year because of the frequency of holding stack presence at Schiphol. In general, however, the group did indicate that they thought both this tool as well as the goal (improve EAT adherence) would be relevant for their job. Quotes to sustain this are:

> “In my view, this is not a necessity but it is a sick [very good] support device” [P2]

All participants indicated that they believed the stack list did not support them in EAT adherence while managing a holding stack. The following quote gives an insight in the thoughts of this participant on the topic:

> "I did not use the stack list at all because it does not add value" [P1]

Image (social influence) Only five comments were made [P2, P4, P5, P8] that were in some way relevant for image, but each of these quotes can also be linked to another code. P2 indicated that they thought it was important to show you could comply exactly with what was agreed, such that in the case of a 2-minute error allowance the perfect score is 1 minute and 59 seconds error. P8 also said that the agreement that has been made with the entire group is important for the EAT adherence error score people aim for. P5 explained that within the group of ACC, there is "a shifting perspective on technological innovations" but that in this specific case, the problem is also that people aim to comply with the error bound instead of aiming for 0 seconds error.

Subjective norm In total, there were 22 quotes about subjective norm, by all participants but P1 and P9. These are summarized by the idea that many people within ACC (the colleagues of the participants) are very well able to aim at a certain error margin, e.g., 120s, and that because everyone within ACC does so, people keep aiming for this margin instead of at ± 0s error. P2 explained that "during the night, there is the night transition and at that moment it is not allowed to deviate from planning either", implying that if it is possible at that moment there should be no reason for higher adherence to planning during daytime to be impossible. Regarding the norm of exceeding the error margin, P2 said that if someone flies blue times [EAT prediction and time over IAF becomes blue the error margin is exceeded], the other ATCos will wonder or ask why that happened. In other words, there is a form of social control within the group.

Participants [P3, P4] both said that they thought some of their colleagues would show some resistance to new tooling, e.g., "people indicating they can manage without a tool" [P4] while they themselves did have a positive attitude toward improving EAT adherence and this innovation. P5 also indicated perceiving a general resistance within the group. These suggestions are not in accordance with what the participants of the experiment indicated regarding their stance on improving EAT adherence and system innovation.

Furthermore, several participants [P2, P5, P6, P8] said things about how the opinion of their colleagues mattered to them and to the way ACC deals with new tooling. P6 was quite literal in explaining this:

> "That is my opinion, but I do also wonder what others think about this" [tool]

As a final note on subjective norm within the ACC group, the learning process when becoming an ATCo is shaped by having a coach, who will instruct and coach new people on how they should manage the traffic and explain their own strategies. Coaching is part of the learning process. No research has been done on this topic but the interviews with [P2,P6,P7,P9] indicate that actually aiming for a +120s error instead of keeping this within the ± 120s error margin is maintained through the learning process.

Experience Since the experiment is a proof-of-concept experiment and none of the participants has worked with the tool before, there is no difference in experience between the participants. The years of experience in ATC is shown in Figure B.8; two of the participants are currently at the end of their education and therefore have not yet operated without supervision, which is the reason there are two participants with zero years of experience.

In total, there were 16 quotes on experience by all participants except for P3. Part of these regarded the fact that the system in which the test was performed was different from the systems they normally use. Getting used to working with the tool was linked to their performance by the participants, where some of them [P4, P8, P9] indicated that they expected that after getting used to the tool, they would have improved results on EAT adherence a next time. These people indicated that upon the first introduction of the tool, they spent some time processing what they actually saw on the screen and how to use that information. P5 indicated that the changed focus (on EAT adherence over separation) was also something that felt unnatural, and that getting used to it could have an impact on the way they would manage the traffic.

Voluntariness The participants were invited to participate in the experiment via email. Two emails were sent out; one of the participants sent a message in the ACC whatsapp group to notify people they were invited to participate in the experiment. Each participant voluntary participated in the experiment; eight of them

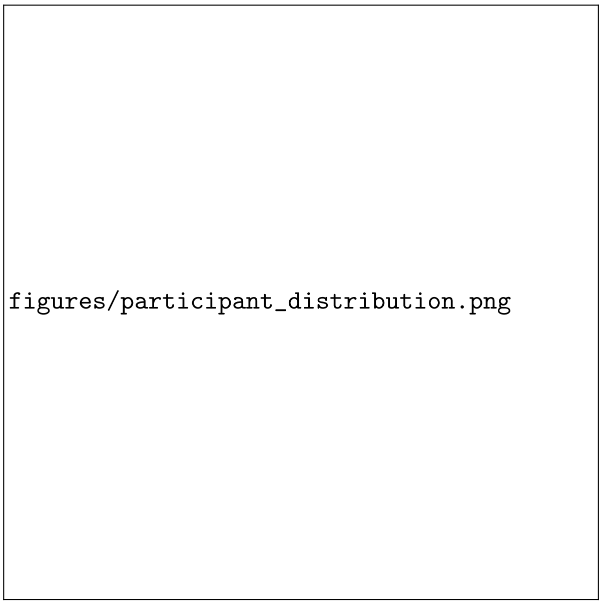


Figure B.8: Age and experience distribution of participants

replied to the invitation completely autonomously and two of the participants [P2, P7] were encouraged by their coach to partake in the experiment.

In total, there were 11 quotes regarding voluntariness by participants [P2, P5, P6, P8, P9]. Part of these regarded the voluntariness of participating in this experiment or innovation projects in general. P8 talked about the “involvement of operational staff at development projects” while P9 indicated the following:

> “What was remarkable is that yesterday we were talking about it [the experiment]. Half of the people said they did not get the email, or made comments like “I am never asked for anything anymore”, so there are still people who for some reason do not realize they can participate in these things.”

The other quotes regard the voluntariness of using (parts of) the tool, and participants indicated that they would like to be able to turn the features of the tool on and of when desired. For example, P2 indicated “when holding for another reason, like visibility conditions, it is important that those [the ECOL] dots can be turned off”.

Age The age distribution of the participants is shown in Figure B.8. It can be seen that 50% of the participants are under 30, but that there were also participants from the older age groups.

Only participant [P2] indicated something about age:

> “I think we have the advantage that we are relatively inexperienced. I think the older generation has more resistance [to technological/system innovations like this one]”

B.2.3. Perceived Ease of Use

In Figure B.9 the perceived ease of use as indicated by the participants can be found. Participants were asked to rate the ease of use of the interface after the first introduction during the training, after they had performed the scenario without the tool (as to measure the perceived ease of use of the simulation setup) and after using the tool. It can be seen that overall, participants did think using the interface was straightforward. During the first introduction, half of the participants indicated they thought it was straightforward, and another four that it was very straightforward. In the scenario without tool, these numbers were, respectively, five and three. When the participants used the full tool, three of them indicated that the use of the tool was somewhat difficult. Since the ease of use was also questioned for just the interface, it can be seen that some participants did find working with the tool somewhat difficult (and that this cannot be attributed for by the different interface than they normally work with).

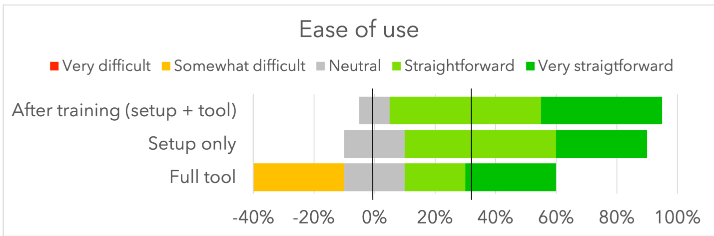


Figure B.9: Perceived ease of use, n=10

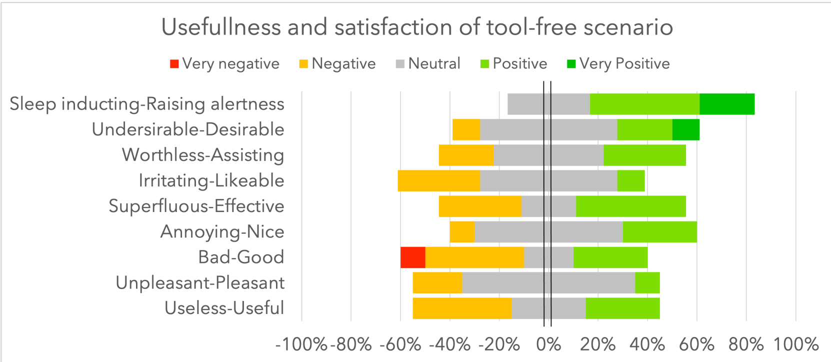


Figure B.10: Usefulness and satisfaction without tool

In total, 9 quotes were given on ease of use by participants [P1, P3, P5, P8, P10]. These indicated that the experiment setup was similar to the current work environment, which made it easy to use for them. For example, [P5] indicated "It aligns with what we are doing now, you do not need to work in any different manner, it just gives better insights in what you are already doing" where [P10] said "the simulation system looks good". Regarding implementation and ease of use, [P5] added about the delta-T that it could be "implemented tomorrow and everyone would use it as intended".

B.2.4. Perceived Usefulness

The results of the survey regarding usefulness and satisfaction are discussed in this section. Usefulness and satisfaction serve as a measure for perceived usefulness, substantiated with interview data. The scores for usefulness and satisfaction are determined as follows. A maximum positive score yields +2 points, neutral 0 points and a maximum negative score yields -2 points, while the in between scores yield +1 and -1 points. For usefulness, elements 1, 3, 5, 7 and 9 contribute to the score. The sum of these scores is then divided by 5 (elements). The satisfaction score is composed of the sum of elements 2, 4, 6 and 8, divided by 4. Both scores are normalized to fit in a 1-10 scale, meaning the score is multiplied by (9/40) and added to 5.5. The middle line representing the center of gravity of the overall survey results is determined by the average of the location of the middle of each bar.

In Figure B.10 the results of the survey on usefulness and satisfaction without tool can be found. It can be seen that many participants scored this scenario neutral, and extreme scores (maximum either positive or negative) have barely been given. The normalized score for usefulness is 5.77, and the normalized score for satisfaction is 5.61. The middle line of the overall survey results and zero-line are very close to each other, with the middle line only slightly to the right of the zero line. The overall score for raising alertness is seen to be much higher than the rest of the overall scores (position of the bar is skewed further to the right). This can be substantiated by the interview data, e.g., participant [P1] mentioned "without having a tool, I know that I must pay attention and therefore the absence of the tool raises alertness for me".

In Figure B.11 the results of the survey on usefulness and satisfaction after introduction of the tool and the training round can be found. It can be seen that zero very negative scores and only one negative score were given. The majority of the participants scored the overall tool positive at this moment. The normalized score for usefulness is 7.93, and the normalized score for satisfaction is 7.75. The middle line of the overall survey results is positioned relatively far to the right of the zero-line.

In Figure B.12 the results of the survey on usefulness and satisfaction of delta-T in the stack list after the

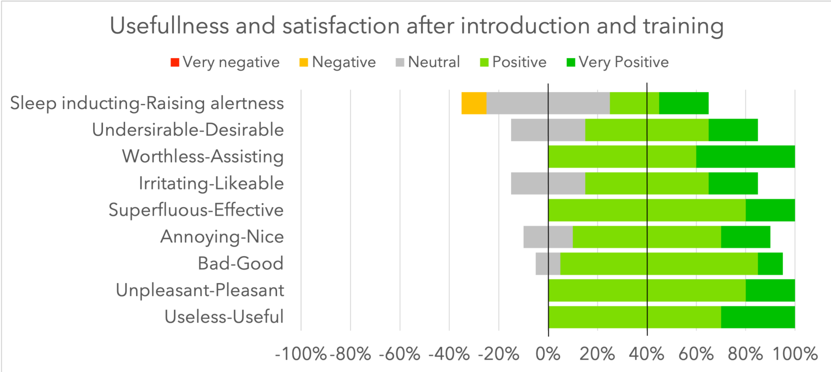


Figure B.11: Usefulness and satisfaction after introduction and training

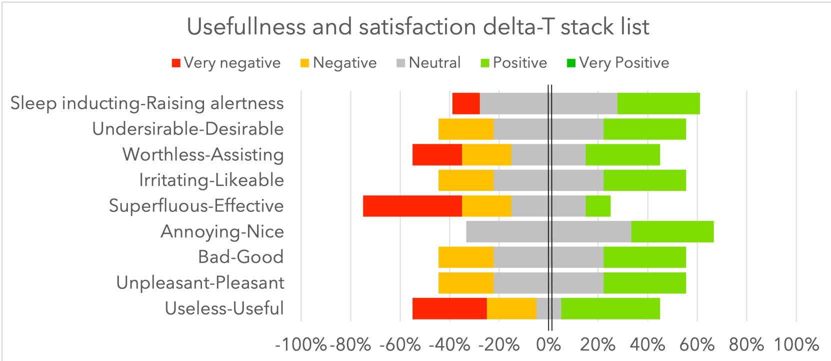


Figure B.12: Usefulness and satisfaction of delta-T in stack list

measurement run can be found. It can be seen that many very negative scores were given, especially in terms of usefulness; not a single very positive score was given. This is also reflected in the score for usefulness, being 4.87. The normalized score for satisfaction is 5.84. The middle line of the overall survey results and zero-line are very close to each other, with the middle line only slightly to the right of the zero line. The bar indicating superfluousness is furthest to the left (negative side), which is in accordance with interview results such as e.g.,

> "I did not use the stack list at all because it does not add value" [P7]

> "[after performing the scenario with tool] I haven't looked at the stack list and did not use those delta-T values" [P6]

This participant indicated both perceived ease of use in terms of how fast they thought they learned to work with the tool, as well as perceived usefulness.

> "I was pleasantly surprised, as I could lean on the tool to use it what it was meant for very quickly. I also noticed that I left the stack list since it did not provide any added value in this phase." [P10]

In Figure B.13 the results of the survey on usefulness and satisfaction of delta-T in the labels after the measurement run can be found. It can be seen that zero very negative scores were given, and the delta-T in the labels is scored slightly better in the satisfaction categories with respect to the usefulness categories. The normalized score for usefulness is 7.08, and the normalized score for satisfaction is 7.69. The middle line of the overall survey results and zero-line are located apart from each other, with the weight of the scoring positioned to the right of the zero-line.

In Figure B.14 the results of the survey on usefulness and satisfaction of the ecology dots after the measurement run can be found. It can be seen that many very positive scores were awarded to the ecology dots. Especially in the category of raising awareness, the ecology dots scored high. A quote from the interviews that can be linked to this phenomenon is "I've changed my strategy [...] and used the economy [ECOL] dots as a measure of when to turn in" [P10]. The normalized score for usefulness is 7.57, and the normalized score for

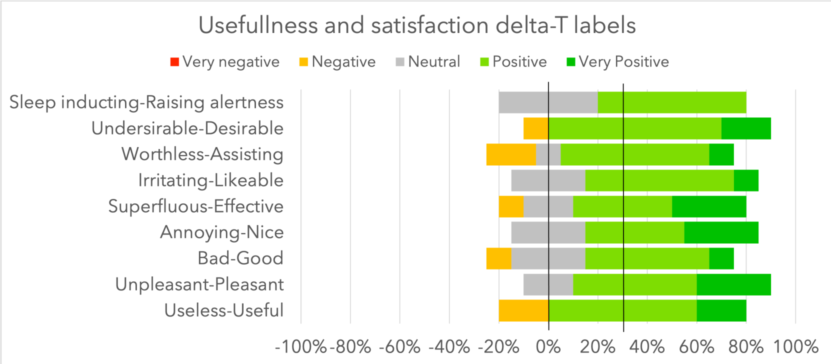


Figure B.13: Usefulness and satisfaction of delta-T in labels

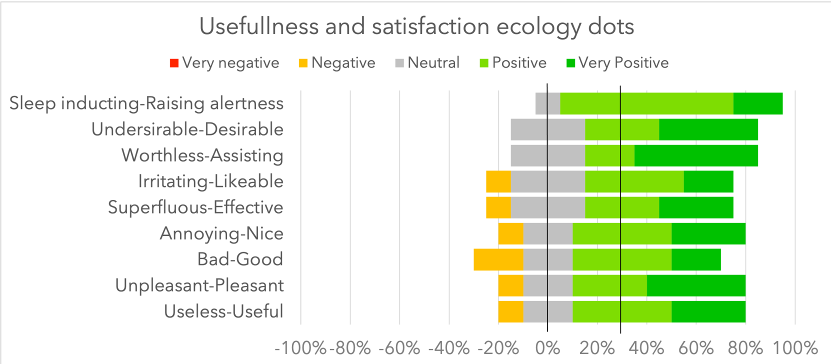


Figure B.14: Usefulness and satisfaction of ecology dots

satisfaction is 7.58. The middle line of the overall survey results is positioned the furthers to the right of the zero-line of all the results.

Overall, there were 35 quotes about perceived usefulness, by all participants. These can be summarized as that each participant thought the rationale behind the tool was good and they saw the use of it. All participants did mention something related to further development of the tool being needed before it could be used to full potential and implemented within their systems, but they did deem the concept useful. Participants linked the usefulness to improved EAT adherence [P1, P4, P6, P9], gaining time (increased workflow efficiency) [P2, P4, P5, P6, P9, P10] and increased predictability [P1, P4, P5]. Participant [P5] even indicated that they thought the concept would be useful beyond the context of a holding pattern, but also when they use vectoring for delaying aircraft.

B.2.5. Attitude Towards Use

In total, there were 37 quotes about attitude towards use by all participants. The participants all indicated a positive affect towards the use of the tool. Several participants indicated a preference toward the delta-T [P3, P6] or the ECOL dots [P1, P4, P5, P9].

Some participants linked their attitude toward use to other factors that are present within the TAM. The following participant [P1] links attitude towards use, perceived usefulness, ease of use and output quality.

> “If you want to convince people to use this, to have a positive feeling about using your tool, it [the prediction] just has to be right. It needs to be very user-friendly”

Another participant [P8] linked experience and result demonstrability to people’s attitude toward use.

> “For every change it is necessary to stress that it will cost time and effort, we will be asking something from you, or you’ll have to learn something new or unlearn something old, but in the end, this is what it will bring you. When people understand that, when you are really able to take them along and convince them, this program can be successful and in general will be successful.”

About the current setup of doing a proof-of-concept experiment and really taking on the opinions of ACC from very early on in the process, [P5] said:

> "Anyway, when you want people to accept your innovation, take people along [in the process]. Many people can make a valuable contribution and in this way, you use that to the full potential"

Finally, there were also some participants [P1, P2] who did indicate a positive attitude toward use but explained that during this proof-of-concept they also verified whether the system prediction was correct. They did use the tool in the first place, but did not fully rely on it.

B.2.6. Intention to Use

In total, there were 20 quotes about intention to use. All participants indicated a positive intention to use some aspects of the tool. Out of ten participants, nine indicated that they would use the decluttering feature, seven indicated they would use the full tool (both delta-T in label and ECOL dots) if available, two indicated they would just use the ECOL dots and one indicated an intention to use just the delta-T in the label.

> "If this would be implemented, I would use it." (P4)

The following quote shows both an intention to use as well as a perceived usefulness.

> "Definitely would want to use it, I think the feature where you can make the labels less visible is very useful: a top feature that I would really want to have in the hold. [...] The delta-T in the label is not very useful for me, which is because I prefer working with the vertical view, so personally it is okay if it is there but I don't need it." [P1]

This participant indicated an intention to use, combined with a perceived ease of use and perceived usefulness.

> "You could implement that tomorrow and I would use it. It matches what we do now and you don't have to learn anything new from it, so everyone would use it in the right way from the start; it just gives a better insight in what we are doing." [P5]

B.2.7. Interpretation of Results

This section discusses the results and explains the possible implications of the gathered data.

External variables The external variables that seem to play the largest role within the subject group are result demonstrability, output quality, job relevance and subjective norm.

The **result demonstrability** seems to be relevant in the sense that the participants indicate they have performed double checks during the current setup and are not ready to trust the system completely. In the current setup, people did indicate that they believed they did not yet receive perfect scores as the tool is still under development, but accepted that as it concerned a proof-of-concept experiment. Participant P1: "I believe that the results of your experiment do not do right to the potential of your tool". The results imply that the ATCos would want to experience a system first, have it prove the demonstrability of results obtained as well as its reliability, in order for the group to improve their stance toward using the innovation. **Output quality** was indicated by several participants as a crucial factor, also combined with being able to fully trust the prediction the system makes.

Image is a factor that seems to have a small impact as within the group, there seems to be little desire to stand out or have a different social status. On the other hand, the responses from the participants imply that **subjective norm** has a large impact. Participants saying things like "everyone does this or that" and "I'm wondering what others would think of it indicate that within ACC, the opinion of the group matters much for what is actually done. Several participants indicated they thought their colleagues would show some resistance to new tooling, while this is not in line with the results obtained within the participant group. Since the participant group made up only 10% of the total ACC group, and these are the people that voluntarily took place in the experiment, a possible explanation for this phenomenon could be that the subgroup who participated in the experiment has a more positive attitude toward technological innovation in general. However, the participant group did represent the ACC group well in terms of age and ACC experience. Therefore, another explanation could be that ATCos have an image of the stance of their colleagues toward innovation which is more negative than their stance is in reality. This also seems to be the case when interviewing people that have worked with ATCos: the resistance toward innovation is always mentioned, which is not in accordance with what the participants indicated. The participant group was to small to draw any conclusions from this that are representative for the entire group, but since subjective norm seems to partially determine the way ATCos execute their job, it is recommended to do a research on how people within the entire subject group view technological innovation.

Since there were no differences in **experience** between the participants, the effect of that on a positive outlook on the present innovation was not tested. From the interview results, it does seem as if gaining some experience with an innovation increases trust in the system and with that, the attitude towards using the system.

The **voluntariness** of participating in the experiment can potentially have caused that the people who participated had a more positive attitude toward innovation in general than the entire ATCo group. As explained above under subjective norm, this would require further research. In terms of voluntariness in using the (parts of) the tool, it seems as if being able to turn components of the tool on- and off when desired made the participants more at ease with the concept. This voluntariness seems to create a notion where the ATCos improve their attitude toward the system, as it lets them keep their autonomy and they have no need to fear the tool interfering or causing clutter in their screens when it is redundant.

As a final external factor, **age** was predicted to influence the stance on innovation in a negative way [Personal communications with Area Controllers, 2020, Personal communications with KDC, 2020–2021, Personal communications with NLR, 2021]. However, no link was found between age and the stance people had within this experiment on the innovation or their attitude and intention to use the tool. Again, this could be attributed to the size of the participant group and the possibility that people participating have a more positive stance toward innovation in general.

Influence of external variables on attitude toward use From the way participants linked external factors to the other components of the TAM, it seems as if not only perceived usefulness and perceived ease of use are influenced by the external variables as is suggested within the TAM, but also the attitude toward use. The affect a person has toward using a certain tool or system, and more general the introduction of all system innovations at LVNL seemed to be largely impacted by these factors from the interviews. This could be explained through the pride ATCos take in their job, which seems to be the main motivator for decreasing EAT adherence error, e.g., “I am capable of doing that” [P7].

Timing of involvement From the interview results it was found that multiple participants valued that they were taken along in the development process of the tool. Contrary to previous expectations, the fact that certain parts of the tool were still under development during the proof-of-concept test in the case study actually had a positive effect on how people indicated they appreciated being part of the test. One of the participants indicated that “the fact that it is not finished yet makes me feel much more comfortable with giving feedback on the concept” [P10]. An often-heard complaint was in line with the participant who indicated that “people are afraid to talk to OPS [operations; air traffic controllers]” [P5] which seemingly induced a vicious circle: “they always want to present us with something that is already finished, but then they get frustrated when we give them feedback on something in the core of their product” [P9]. In other words, the lack of feedback early on in a development process has on multiple occasions caused developers to produce unwanted products. In the interviews, ATCos suggested that giving this feedback to developers so late in the process causes the developers to fear negative feedback from operations, as it resulted in them having to re-develop their products or having to stop the development process after finding out it was not based on the right design requirements. Participants suggested that this situation has led to an even lower frequency of asking input from operations in development processes, making it impossible to change the direction of the project early on in the process, causing frustration within the organization. Since timing of taking people along in the development process of a tool was not part of this research, but findings are promising, it is recommended to do further research on this topic.

Recommendations are to perform a more thorough research on the common communication channels and to explore the communication possibilities within the organization, to research the link between timing of involvement (timing of involving operational staff in a system innovation project) and attitude toward use.

Perceived usefulness, ease of use, attitude and intention Based on the results, the participant group believes the delta-T, ECOL dots and decluttering feature are useful for holding stack management and EAT adherence. The group does not perceive the updated stack list as useful for these tasks. The participant group also showed an overall outlook on ease of use as being straightforward. In accordance with the TAM, where these positively influence attitude toward use and, both indirectly as well as indirectly (in the case of usefulness), intention to use the participants showed a positive attitude toward use and intention to use based

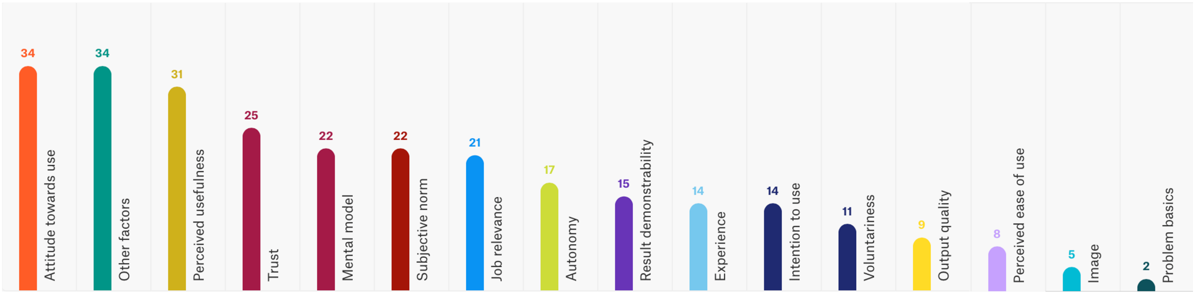


Figure B.15: Frequency of different codes in interviews

on the results. Therefore, it seems as if the concept of the tool will be positively accepted by ACC, especially when considering previous comments on subjective norm and the notion that already 10% of the group has a positive attitude toward and intention to use the tool.

B.3. Interview Code Frequency

The frequency by which participants mentioned several factors related to their attitude towards the innovation, perceived usefulness and intention to use can be found in Figure B.15.

Additional References Appendix B

F. Davis, R. Bagozzi, and P. Warshaw. User acceptance of computer technology: A comparison of two theoretical models. Management Science, 35:982–1003, 08 1989. doi: 10.1287/mnsc.35.8.982.

Personal communications with Area Controllers. Interviews with area controllers at LVNL, 2020.

Personal communications with KDC. interviews with f. dijkstra at kdc, lvnl, 2020–2021.

Personal communications with NLR. interviews with m. van apeldoorn, nlr, 2021.

V. Venkatesh and H. Bala. Technology acceptance model 3 and a research agenda on interventions. Decis. Sci., 39:273–315, 2008.

V. Venkatesh and F. Davis. A theoretical extension of the technology acceptance model: Four longitudinal field studies. Management Science, 46:186–204, 02 2000. doi: 10.1287/mnsc.46.2.186.11926.

C. Westin, C. Borst, and B. Hilburn. Strategic conformance: Overcoming acceptance issues of decision aiding automation? IEEE Transactions on Human-Machine Systems, 46:41–52, 11 2015. doi: 10.1109/THMS.2015.2482480.

C

Code Layout

C.1. General Code Structure

An overview of the elements added to the code is shown in Figure C.1.

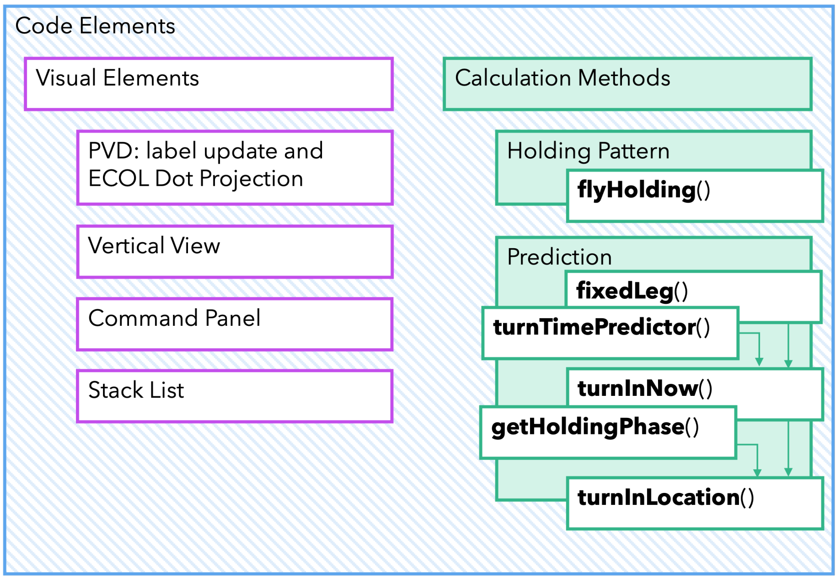


Figure C.1: Overview of elements added to code in SectorX

C.2. Aircraft Class and FMS Class Alterations

Additional variables the aircraft class can have are holdingPredictor (HoldingPredictor), turnInLocation (float[]) and turnInLocationTop (ArrayList<vector2D>). The flyHolding() method was added to the aircraft class. The appropriateBankForTurn() method was altered in the FMS class.

flyHolding() is used to fly a holding pattern. It algorithm functioning is presented in Algorithm 1.

appropriateBankForTurn() was changed in the FMS class such that it is capped at 25° since this is an ICAO regulation in holding.

C.3. HoldingPredictor Class

The HoldingPredictor Class is created to predict the delta-T and the ECOL dot locations.

Algorithm 1 flyHolding() method

Require: actual trajectory
Require: next waypoint is IAF and within 5NM, and the last waypoint on our route

```
A   —  B                              ▷ Generate racetrack pattern
|      |
IAF —  C
```

holdtime_s ← required leg time at FL
legLength_NM ← holdtime_s * aircraft ground speed
turnRadius_nm ← required turn radius for rate 1 turn capped at 25° bank
waypointA ← IAF + 2∗ turnRadius_nm
waypointB ← waypointA + legLength_NM
waypointC ← waypointB - 2∗ turnRadius_nm
FmsTrajectory ← waypointA, waypointB, waypointC, IAF
if holding entered (less than 5 NM from IAF) and not already directed to SPL **then**
    inHolding ← true
end if
if cleared to Schiphol **then**
    inHolding ← false
end if
if inHolding **then**
    fly the holding pattern
    holdingPredictor.turnInTime ← minimum time until IAF can be reached
    holdingPredictor.turnInLocation ← set of ECOL dot locations
end if

getHoldingPhase() Returns the holding phase of the aircraft, being inbound turn, inbound leg, outbound turn or outbound leg.

turnTimePredictor(GeoAngle inboundHeading) Computes the estimated turn time. It does this based on the current aircraft heading and the heading of the inbound leg, assuming a right-hand turn. It predicts the next heading after one time step (Δt based on the predicted angular velocity omega ω. Omega is found by estimated ground speed (follows from IAS, wind magnitude and relative wind heading at 0.5 time step from the current location, see turn time prediction algorithm plus Figure 13 in Part II, Section III.A). The method iteratively adds one time step to the predicted turn time until the difference (δ) between the next predicted heading and the inbound heading is smaller than $\omega\delta t$. Then it adds δ/ω to the predicted turn time. Since the algorithm starts using the current aircraft heading and computes the turn time of a right hand turn until reaching the inbound leg heading, it does not require knowledge of the holding phase.

fixedLeg(GeoAngle inboundHeading) Computes the leg time that needs to be traversed when an aircraft would fly directly toward the IAF. It either takes the current aircraft location, or, when on the inbound turn, it takes the location where the inbound turn started. From this, the minimum length of the inbound leg that needs to be flown to reach the IAF from the current position is calculated. Inbound ground speed is calculated based on IAS and wind. Then, the inbound leg length is divided by the inbound ground speed prediction to gain a prediction of (fixed) inbound leg time.

turnInNow(GeoAngle inboundHeading) Computes the estimated time the IAF is crossed if the aircraft would fly directly toward the IAF (flying straight tracks along the in- and outbound holding legs). It checks the holding phase, and based on this the minimum time before the IAF is reached is found by using fixed-Leg(inboundHeading) on the inbound leg, turnTime(inboundHeading) on the outbound turn, and a combination of both on the outbound leg and the inbound turn. The resulting time is added to the current scenario time.

turnInLocation(GeoAngle inboundHeading, float EAT) Computes the ideal turn-to-IAF location based on a given EAT and the inbound heading. It uses the earliest moment at which the IAF can be reached from the current position to find the delta-T, as computed by turnInNow(inboundHeading). If the delta-T is positive and the aircraft is either on the outbound turn or leg, it calculates ground speed for the in- and outbound leg. This is done to gain an estimate of the additional distance that needs to be covered before turning toward the IAF to reach it at the given EAT.

C.4. Visual Elements

The visual elements are divided into four categories: alterations to the PVD, the creation of the VV, the control panel, and the stack list.

C.4.1. Plan View Display Additions

The following alterations were made to the PVD: the addition of classes to visualize the ECOL dots and alterations in the label class.

GLAircraftLabel Has been changed such that the label layout and presented information were the same as in LVNL systems. An additional element was added to show the delta-T at the top line (line zero) of the label.

GLHolding Has been added, in which the ECOL dots are rendered based on the locations found at myEAT = [EAT-120s, EAT-110s, ..., EAT, EAT+10s, ... EAT+120s] by turnInLocation(inboundHeading, myEAT).

C.4.2. Stack (Vertical View)

The following classes were created under display/elements/stack to create the vertical view: **GLAircraft** (shows aircraft location), **GLAircraftLabel** (shows aircraft label and corresponding information. Delta-T in label is turned on- or off depending on whether the tool is turned on- or off), **GLAircraftLabelItem** (helper for the label), **GLAxisAltitude** (shows the IAF line and flight levels in the VV), **GLAxisNM** (shows the distance from SPL in NM at the bottom of the VV), **GLFlightTrails** (shows the history dots), **GLHolding** (shows the locations of the ECOL dots and the location where the aircraft will start the inbound turn if no control action is given), **GLListShapes** (creates the shape elements required for GLHolding), **GLStack** (adds all graphical layers to the VV and enables clicking it), **Stack** (creates the window for the VV) and **StackPanel** (makes the window resizable).

C.4.3. CDU

The **CMDFrame, CMDPage** and **CMDPane** classes were altered such that the command interface was always present during the simulation (experiment), it was not possible to accidentally switch off the command panel, and to change the command options such that the set only contained relevant command options for the present research. Two buttons were added, namely SPL and ATP to allow the ATCo to give a direct-to-(SPL or SPL via ARTIP) command. Clicking these buttons adds the said waypoints to the aircraft route and sets the inHolding boolean of the aircraft to false. This prevents the aircraft from continuing to fly the holding pattern and ensures it continues its route to the said waypoints.

C.4.4. Stack List

The stack list was created in display/elements/dialogs/LVNL. The class **LVNLStackList** contains the code for creating, rendering and updating the stack list.