The Effectiveness of Simulating Active Problem Solving in Pilot Training to Deal with Automation Surprises

Master of Science Thesis

J.K. van Leeuwen

21 October 2020

Faculty of Aerospace Engineering

TU Delft Delft University of Technology

Challenge the future

The Effectiveness of Simulating Active Problem Solving in Pilot Training to Deal with Automation Surprises

Master of Science Thesis

by

J.K. van Leeuwen

For obtaining the degree of Master of Science in Aerospace Engineering
at the Delft University of Technology,
to be defended publicly on Wednesday October 28, 2020 at 02:00 PM.

Student number:	4452305
Project duration:	January 27, 2020 – October 28, 2020
Graduation Committee:	Prof.dr.ir. M. Mulder
	Dr.ir. M.M. van Paassen
	Dr.ir. J.C.F. de Winter
	Dr.MSc. A. Landman
	Ir. O. Stroosma

TU Delft

Preface

Before you lies the thesis "The Effectiveness of Simulating Active Problem Solving in Pilot Training to Deal with Automation Surprises". It has been written to full fill the requirements for obtaining the degree of Master of Science in Aerospace Engineering, over the course of nine months.

When asked to describe me in a few words, most of my friends and family will mention my passion for aviation. This passion is impossible to explain or bring to words, as it is something that has developed over the years. It is this passion, however, that made some of the most important choices with respect to my (academic) career rather easy. It drove me to do my high school thesis in flight performance and aerodynamics with some help from the faculty Aerospace Engineering. It is because of this passion that I started my bachelor here at the Delft University of Technology in 2015, selected the master track Control & Simulation and ultimately chose this topic for my thesis. This thesis was the perfect coherence between the major topics of my master track: flight control, flight simulation and human performance. Whether it is a coincidence or not, the circle is complete again as one of my supervisors for this thesis was already helping me seven years ago with that high school thesis I mentioned before.

I would like to thank my supervisors for their excellent guidance and support during this process. Max, thanks for providing me with this challenging thesis topic and your valuable feedback throughout the last months. René, your lectures in human-machine systems combined with your insights helped me a lot during this thesis. Annemarie, I very much appreciate your help in setting up, running and analysing the experiment. Olaf, thanks for all the help around setting up the simulation in DUECA, testing the simulation and fixing daily problems we encountered during the experiment. Besides my supervisors, I would like to express my gratitude to Eric Groen, whose expertise in the field of aerospace human-machine experiments and wise words have greatly contributed to this project. I also wish to thank all of the 21 participants in the experiment, without whose cooperation I would not have been able to conduct this analysis. And if it would not be for Hans Mulder, I would have never found these motivated participants. My parents deserve a particular note of thanks: your wise counsel and kind words have, as always, served me well.

Jordy K. van Leeuwen
Delft, 21 October 2020

Contents

List of Figures

List of Tables

LIST OF SYMBOLS

Symbol	Meaning	Unit
C	PID controller output	-
e	Error between reference value and actual value	-
h	Altitude	[m] or [ft]
K_p	Proportional Gain	-
p	Roll rate	[rad/s]
q	Pitch rate	[rad/s]
T_d	Differential Gain	-
T_i	Integral Gain	-
V	Airspeed	[m/s] or [kts]
α	Angle of attack	[rad]
δ_a	Aileron deflection	[rad]
δ_e	Elevator deflection	[rad]
θ	Pitch angle	[rad]
ϕ	Roll angle	[rad]
ψ	Heading	[rad] or [deg]

Acronym	Meaning
AHRS	Attitude and Heading Reference System
AI	Attitude Indicator
ADC	Air Data Computer
ALT	Altitude Hold (autopilot mode)
ATC	Air Traffic Control
ATPL	Airline Transport Pilot License
CAA	Civil Aviation Autrority
CBT	Computer Based Training
CDI	Course Deviation Indicator
CDU	Control Display Unit
CRM	Crew Resource Management
EFIS	Electronic Flight Instrument System
FAA	Federal Aviation Authorities
FLC	Flight Level Change Mode (autopilot mode)
FMS	Flight Management System
GPS	Global Positioning System
GS	GlideSlope
HDG	Heading Hold (autopilot mode)
HSI	Horizontal Situational Indicator
IATA	International Air Transport Association
ILS	Instrument Landing System
LNAV	Lateral Navigation (autopilot mode)
LOC-I	Loss of Control In-Flight
LOFT	Line Oriented Flight Training
LOSA	Line Operation Safety Audit
MCP	Mode Control Panel
MFD	Multi-Function Display
NAV	Lateral Navigation (autopilot mode)
ND	Navigation Display
OOTL	Out-of-the-Loop
PBL	Problem-Based Learning
PBN	Performance Based Navigation
PFD	Primary Flight Display
PIT	Pitch Hold (autopilot mode)
RNAV	Area Navigation
ROL	Roll Hold (autopilot mode)
TC	Turn Coordinator
UPRT	Upset Prevention and Recovery Training
VNAV	Vertical Navigation (autopilot mode)
VNAV	Vertical Navigation (autopilot mode)
VHF	Very High Frequency
VOR	VHF Omnidirectional Range
VS	Vertical Speed Hold (autopilot mode)
VSI	Vertical Speed Indicator

Part I

Paper

The Effectiveness of Simulating Active Problem Solving in Pilot Training to Deal with Automation Surprises

Jordy Kelvin van Leeuwen
Faculty of Aerospace Engineering
Delft University of Technology
Kluyverweg 1, 2629 HS Delft
Email: jkvanleeuwen@student.tudelft.nl

***Abstract*—Participant feedback in previous research indicates a need for pilot training to handle non-routine situations with automation surprises. Therefore, we tested the effectiveness of using active problem solving during training on subsequent performance while dealing with automation surprises. We simulated a glass cockpit of a general aviation aircraft in a full motion flight simulator. An experimental group of private pilots (n = 10) was trained to actively diagnose and solve problems related to the autopilot, without foreknowledge of the training scenario. A control group (n = 10) received the same training scenarios with foreknowledge. The effectiveness of these pilots in dealing with both new and trained automation surprises was compared between groups. It was expected that the experimental group would be more effective in dealing with the new automation surprises, whereas the control group would be more effective in dealing with the repeated automation surprises. The experimental group indeed responded somewhat faster to the new automation surprises and were on average able to maintain the highest level of automation for a longer time than the control group. The opposite is true for the repeated automation surprises, where the control group was somewhat more effective in dealing with the automation surprises. Although these differences were of medium effect sizes (Cohens d = 0.5), they were not significant. Thus, the training approach taken in this study may need to be further enhanced and tested with more participants in order to have a significant effect.**

I. Introduction

Several reports by aviation authorities have indicated issues when flight crews have to deal with surprises directly related to flight deck automation (FAA, 1996; Fletcher & Bisset, 2017). Pilots sometimes make incorrect decisions when dealing with automation surprises as they do not fully understand the logic of the complex autoflight systems. According to the FAA, this is a contributing factor in over 40% of the investigated accidents (Flight Deck Automation Working Group, 2013). Pilot training focuses heavily on procedures and checklists, as these have proven to be extremely useful in time-critical and emergency situations (Degani & Wiener, 1998). As a result, pilots can reproduce the procedures they have been trained in very well, but often lack a higher-order understanding to make their own judgement about the current state of the systems.

Automation in aviation is intended to decrease both pilot workload and fatigue, allowing for increased productivity and a better execution of standard operations. Airliner policies state that “the level of automation used shall be the most appropriate for the task at hand with regard to safety, passenger comfort, regularity and economy” (p. 218) (Goteman, 2018). Practically, this means that pilots hardly ever use manual control, thus being automation managers and responsible for higher-order goals (Bourgeois-Bougrine, Gabon, Mollard, Coblentz, & Speyer, 2018). Many pilots report automation surprises (see De Boer and Hurts (2017)), a concept first introduced by Woods and Johannesen (1994) who defined it as “situations where crews are surprised by actions taken (or not taken) by the automatic system” (p. 56). Such an automation surprise occurs when there is a conflict between the pilot and the automation systems which the pilot detects but does not understand (Dehais, Peysakhovich, Scannella, Fongue, & Gateau, 2015). Indirect mode changes, inhibiting behaviour (where the automation cancels pilot actions) and degradation in automation are often mentioned as automation surprises (de Boer & Hurts, 2017; Dehais et al., 2015). These are indications that the flight crew misused or misunderstood the automation systems.

Automation surprises can jeopardise flight safety as the aircraft can get in an unrecoverable state if the surprise is not noticed and recovered in time (Sarter & Woods, 1997). With a survey among Dutch airline pilots, de Boer and Hurts (2017) found that on average a pilot experiences an automation surprise three times per year. Automation surprises are a result from vulnerabilities in pilot knowledge. The Flight Deck Automation Working Group (2013) mentioned that 40% of the analysed accidents show some knowledge deficit with the pilots. Another problem related to flight deck automation identified by Fletcher and Bisset (2017) in their report for the CAA is the lack of mode awareness and mode management among pilots. In addition, improper mode management, indirect mode changes and mode confusion are often mentioned by flightcrews (Fletcher & Bisset, 2017). In 27% of accidents reviewed by the Flight Deck Automation Working Group (2013), mode selection errors were mentioned.

These issues are directly related to pilot training, as they originate from insufficient knowledge. Pilot training focuses heavily on learning procedures, leaving limited time to teach problem solving and decision making. Due to this procedural approach, student pilots do not get the chance to make mistakes and learn from these mistakes. Since training organisations have a limited repertoire of training scenarios available, pilots who are employed for some time may find that refresher training becomes more and more predictable. As a result, the analysis by the Flight Deck Automation Working Group (2013) showed that pilots are not well equipped to handle non-routine situations. Pilots have indicated they require more system knowledge to understand how these systems operate and their underlying logic (Orlady, 2010). In a study for Boeing, Holder (2012) showed that pilots feel they do not receive sufficient information to operate the flight deck via training.

Dealing with an automation surprise in the flight deck can be seen as a diagnosis-solution problem. The key to learn troubleshooting and diagnosis-solution is by solving authentic troubleshooting problems as practise (Jonassen, 2010). By providing conceptual instruction, information is provided in a meaningful way. According to Jonassen and Strobel (2006), meaningful learning can only occur when students are actively working on authentic tasks. Students should interact with their environment and analyse the effects of their actions. In this way, the students construct their own interpretations and information becomes meaningful. Novak (2013) states that presented information and material must have meaning to the student. This requires the student to understand relevant concepts and prepositions of the environment.

Schaafstal, Schraagen, and van Berl (2000) recommend the use of a meta-strategy when training for troubleshooting, as well as functional models, and underlying domain knowledge. These functional models can be presented in the form of a simulation used during the training. The effectiveness of such simulations on learning problem-solving skills has been shown for both aviation-related problems (Brna, Ohlsson, & Pain, 1993), as well as for general engineering problems (White & Frederiksen, 1990). Rosa et al. (2020) found that experience with problem solving in dynamic environments can help pilots making decisions in unforeseen situations. Dynamic decision-making and problem-solving are crucial elements to the safety of aviation.

A teaching approach that incorporates the above is the student-centered Problem-Based Learning (PBL) method. In this approach, the tutor acts as a guide, since the student learns new information by self-directed learning (Dochy, Segers, den Bossche, & Gijbels, 2003). The student has to solve problems to gain the required knowledge and learn problem-solving skills. Students construct their own knowledge in an environment that is representative for a real world scenario (Savery & Duffy, 1995).

To summarise, literature indicates that current training is often insufficient in preparing pilots to effectively use the flight deck automation systems in non-routine scenarios. The current study was performed with the goal to test recommendations to extend pilot training with a section where pilots actively practice problem-solving skills in realistic scenarios. This should lead to pilots becoming more effective in dealing with automation surprises on the job. Additionally, the training should be comprehensive in such a way that it allows the trainee pilots to fully understand the on-board systems.

An experiment was designed in which private pilots with limited autopilot experience were trained for a modern general aviation flight deck automation system. This training was focused on problem-solving skills by letting the pilots respond to non-routine scenarios without foreknowledge. The effectiveness of these pilots in dealing with automation surprises after this training was compared to a control group who received the same training, only with foreknowledge about each scenario effectively removing the focus on problem solving from the training. For this purpose, several test scenarios were developed which included both new and trained automation surprises.

It was hypothesised that the experimental training, compared to the baseline training, would result in: (1) pilots being more effective in dealing with an automation surprise which they had not seen before during training, (2) pilots being less effective in dealing with an automation surprise which they had seen before during training, (3) no significant difference between the level of surprise experienced due to an automation surprise and (4) a higher perceived workload during the training. The main rationale for these hypotheses was that the experimental training would teach the participants problem-solving skills which they could apply when facing new automation surprises. However, applying a trained procedure for a trained situation would always be beneficial in terms of workload and performance.

II. Method

A. Participants

Twenty Dutch private pilots participated in the experiment. Pilots holding a type rating were not selected as they are likely extensively trained on automation systems. All pilots were in possession of an instrument rating (IR). Two balanced groups of each ten participants were formed in advance based on age, total flying hours and the characteristics listed in Table I. Pilots' trait anxiety and fluid intelligence were measured with the State-Trait Anxiety Inventory (STAI) test and Raven's Advanced Progressive Matrices at the start of the experiment, in order to re-balance to the group if required. However, no interventions to the initial group assignment were made. Table II gives an overview of the group comparison in hours, experience, trait anxiety and fluid intelligence and were analysed using an independent t-test. It should be noted that the control group scored 10% or about half a standard deviation higher on fluid intelligence, although this difference was not significant. Informed consent was obtained from each participant.

B. Experimental design and tasks

The participants were divided into two groups of each ten participants, from hereafter referred to as the experimental group and the control group. Both groups received a two hour

TABLE I
GROUP COMPARISON ON EXPERIENCE

	Experimental Group	Control Group
Commercial Pilot License	4/10	5/10
Glass cockpit experience	9/10	9/10
Generic autopilot experience	9/10	9/10
Specific avionics experience	5/10	5/10
Specific autopilot experience	3/10	4/10

TABLE II
GROUP COMPARISON FOR HOURS, AGE, ANXIETY AND INTELLIGENCE

	Experimental Group Mean (SD)	Control Group Mean (SD)	p
Age (years)	54.0 (8.31)	51.0 (9.81)	.470
Experience (hours)	694.5 (693.75)	666.1 (415.98)	.913
Trait anxiety (20-80)	26.9 (3.96)	26.9 (3.04)	.851
Intelligence (0-12)	9.0 (2.00)	9.9 (1.52)	.273

flight simulator training followed by a one hour flight simulator test which all took place in one day. The performance of both groups was compared in both new automation surprises and repeated automation surprises.

In each test scenario, the participants were tasked to complete a given flight plan while using as much of the automation systems as possible. In this way, the participants were incentivised to not simply disengage the autopilot as soon as they experience a partial failure. However, as some extreme cases may actually require to disengage the autopilot completely, participants were not instructed to keep the autopilot engaged at all costs.

C. Apparatus

The experiment was performed in the SIMONA Research Simulator at the Delft University of Technology (Stroosma, van Paassen, & Mulder, 2003). This is a full-motion simulator with a six-degrees-of-freedom hydraulic hexapod motion system[1]. Outside vision is rendered with FlightGear on a collimated display with a 180 degrees horizontal by 40 degrees vertical field of view. Sound effects of (autopilot) alarms, gear retraction, wind and engine noise were played on a 5.1 surround sound system installed in the simulator. The participants were not wearing headsets but could communicate via a two-way simplex intercom with the experiment coordinator who acted as the instructor during the training.

The aircraft model for this experiment was a Piper PA-34 Seneca III, a light multi-engine piston aircraft, as the non-linear, six-degrees-of-freedom software model developed by Muynck and Hesse (1990) for this aircraft was previously used in similar experiments in our research group. The system and sensor failures used for this study were added to the model, as well as a fully operational three-axis autopilot. Digital instruments were developed for this simulation and based on the Primary Flight Display and Multi Function Display of the Garmin G1000. The PFD with bezel is visualised in Figure 1. The PFD and the Multi Function Display (a moving map including a flight plan page) were both presented on a 1024 by 768 pixels touchscreen in the simulator. A digital standby instrument with airspeed, attitude and altitude information was provided on a third screen. The research simulator has a generic multi-crew flight deck including a control column and rudder pedals with control loading, electrical pitch trim, throttles, and a gear lever.

[1]Due to a malfunction with the motion system, one participant of the experimental group performed all the test scenarios without motion. One other participant from the experimental group performed test scenario 5 without motion.

Fig. 1. Primary Flight Display used in the simulator.

D. Procedure

The experiment started with a briefing in combination with 15 minutes of ground school in which the interface and basic working principles of the autopilot were discussed. This was followed by a familiarisation flight in the simulator in the form of a take-off and landing at what most participants considered to be their home airport. The purpose of this familiarisation flight was to get used to the dynamics of the aircraft, the layout of the generic cockpit and the avionics (as some pilots had no experience flying a glass cockpit before). Following the familiarisation flight was the training in which pilots received instruction on how to use the different modes of the autopilot and experienced five automation-related failures (Table III). After the training, the participants were asked to complete a questionnaire with questions regarding the experienced mental effort and the enjoyability of the training. To assure there was no confounding difference in factual knowledge between the two groups, a multiple-choice test consisting of ten questions about the usage and working principles of the automation was held after the training. The participants were informed that their performance would be evaluated in the following six test scenarios. After each test scenario, participants were asked to complete a questionnaire with subjective measures and a question about what happened in the scenario. No ATC communication or instructor feedback was included in the test scenarios. The procedure is visualised in Figure 2.

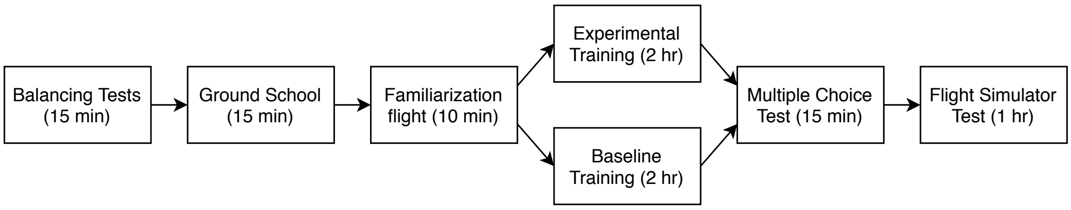


Fig. 2. Overview of the procedure for the experiment.

E. Experimental training

For the experiment, a training was developed focusing on Problem-Based Learning. The participants were exposed to several authentic automation failures, as recommended by Fletcher and Bisset (2017) and Nickolic and Sarter (2007). However, students are not likely to learn problem-solving skills completely on their own as discovery-based learning is not sufficiently effective (Kirschner, Sweller, & Clark, 2006; Mayer, 2004). Providing the student with a meta-strategy as scaffolding will likely result in better results of the training. Therefore, we developed a problem-solving strategy, to be applied when facing an automation failure, for the experimental training. Using such a strategy during training allows for students to learn effective problem-solving skills (White & Frederiksen, 1990).

Based on the literature of training for problem solving and expertise in the domain of flight deck automation, the following problem-solving strategy was created for this experiment:

1) Notice autopilot behaviour is off-nominal
2) Identify which sensor or system is faulty
3) Identify implications on autopilot performance
4) Switch to alternative information source if possible, or
5) Switch to lower level automation

This strategy was based on discrepancy detection as pilots were tasked in the first step to notice that the current autopilot behaviour is off nominal. After they had practised solving the failure, the participants from the experimental group received the information about the exact type of failure and the procedure that could have been used to solve it. Hereafter, the failure was repeated so that the participant got the chance to practise this preferred procedure at least once.

A control group was used to compare the results to a baseline. Similar to a highly predictable and procedural form of pilot training, this group of participants received specific procedures from the instructor for each failure in the training. Thus, these participants were trained using procedures and were not required to develop independent problem-solving and decision-making skills during the training. The participants in the control group received instructions about the procedures before the first time they experienced the failure and thus did not experience the process of diagnosing the situation and searching for the right solution.

Both groups experienced each training scenario twice. This allowed the experimental group to at least practise once with the correct procedure, and allowed the procedural group to have sufficient practise to become familiar with the procedures. Both the experimental training and the baseline training lasted two hours.

F. Training scenarios

The training was split into three sections; Lateral Flight Director training, Vertical Flight Director training and Navigation training. This division was created in order to provide the participants with an extensive training in the relevant autopilot modes, starting from the more basic intuitive modes progressing to the more complex modes. Each section covered a basic instruction and practice session in the simulator with the associated modes, which was referred to as the routine part. The routine part in each section was followed by non-routine scenarios. An overview of the types of failures for each non-routine scenario of the training is provided in Table III. Both the experimental group and control group were given the same failures, in the same order, during the training. All non-routine scenarios were stand-alone situations starting in stable cruise flight and consisted of a single simple task (i.e. climb and maintain 5,000 and fly heading 120). Weather conditions for all training scenarios were good visibility, low turbulence and no wind. Each non-routine scenario existed of a subsystem or sensor failure directly related to the modes that were covered in the routine part of the respective training section. The non-routine scenarios were designed in such a way that it would provide the necessary experience for the test scenarios.

In the first training section, the Lateral Flight Director training, participants were taught how to engage and disengage the autopilot, make use of the flight director in manual flight, how to use Control Wheel Steering and were introduced to Roll Hold mode and Heading mode. These modes were covered in one continuous simulation in cruise flight. The Lateral Flight Director training included one non-routine scenario, in which the complete primary flight display failed.

In the second section, the Vertical Flight Director training, participants practised with the Pitch Hold mode, Vertical Speed mode, Flight Level Change mode and Altitude Hold mode. The concept of selected altitude capture was also introduced. Again, these modes were covered in one continuous simulation in cruise flight. This section included two non-routine scenarios: an elevator servo failure and a blocked pitot tube.

The last section, the Navigation training, covered the more complex lateral and vertical navigation modes of the flight director. Participants were explained how to use the Navigation mode in combination with both VOR and GPS and how they could use the course deviation indicator. Finally, the use of Vertical Navigation mode was explained in a separate RNAV approach. The two non-routine scenarios for this section were a VOR receiver failure and an Air-Data-Computer failure.

TABLE III
TRAINING SCENARIOS

Scenario	Training	Failure	Phase of flight
Training 1	1 - Lateral FD	PFD failure	Cruise
Training 2	2 - Vertical FD	Elevator servo failure	Cruise
Training 3	2 - Vertical FD	Blocked pitot tube	Cruise climb
Training 4	3 - Navigation	VOR receiver failure	Cruise
Training 5	3 - Navigation	ADC failure	Cruise

G. Test scenarios

The test included a total of six scenarios, each containing one failure which affected the behaviour of the autopilot. Four failures were new to the participants, and two failure types were repeated from the non-routine training scenarios in a different setting. Each scenario took place at a different location, though in familiar airspace for the participants as airports in the Netherlands, Germany and Belgium were selected. To increase the realism and therefore the workload compared to the stand-alone training scenarios, the test scenarios covered a complete phase of flight with multiple way-points (excluding scenario 6 which was similar to the stand-alone training scenarios). For the scenarios that featured an approach, the appropriate approach plate was provided in paper form. Visibility was reduced in the test scenarios, turbulence was increased to moderate and wind speeds were varied up to 10 knots. Weather information was always provided at the start of the scenario. Table IV provides an overview of the test scenarios with the failure types.

TABLE IV
TEST SCENARIOS

Scenario	Failure	Phase of flight	Type
Test 1	GPS failure	Approach	New
Test 2	Air-Data-Computer failure	Approach	Repeated
Test 3	Roll servo failure	Departure	New
Test 4	Magnetometer (GMU) failure	Approach	New
Test 5	Blocked Static Port	Approach	New
Test 6	Elevator servo failure	Cruise	Repeated

The scenarios are visualised in Figure 3. Below, each scenario is briefly explained, including the preferred solution in terms of using the highest automation possible and the link to the training scenarios.

In test scenario 1, the participants were tasked to fly an RNAV approach. Right before reaching the first way-point of the approach, the GPS failed silently. This caused the autopilot to fly straight over this way-point, without initiating the turn to the next leg. The correct response in this scenario was to switch to VOR navigation as the RNAV approach was an overlay of the original VOR approach that included two VOR beacons which could be used to fly the approach. The vertical profile of the approach could preferably be flown with Vertical

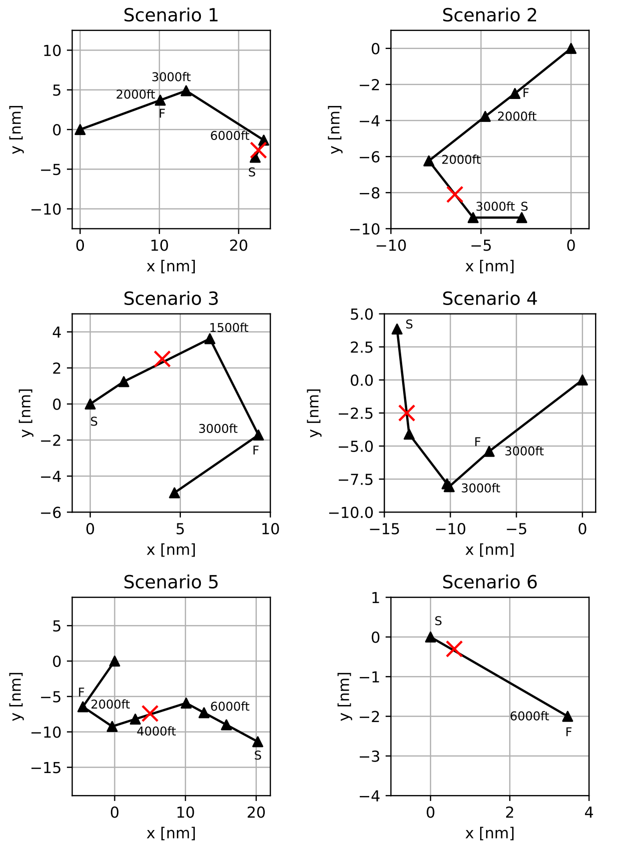


Fig. 3. Top down view of the six test scenarios with relevant altitude restrictions. The location of the failure is indicated with a red X and the start and finish are indicated with an S and F respectively.

Speed mode or Flight Level Change and Altitude Hold mode. Switching between navigation source was practised in training scenario 4 where instead of the GPS, the VOR radios had failed.

In test scenario 2, the Air Data Computer failed midway the second leg of a new RNAV approach. The preferred solution here was to continue with GPS navigation and Pitch Hold mode (as no other vertical mode was available due to the failure). The on-board automation detected this Air Data Computer Failure and automatically switched the autopilot to Pitch Hold mode. The pitch reference could be adjusted with Control Wheel Steering or the Nose UP or Nose DN buttons. This failure was practised in training scenario 5.

For test scenario 3, the roll servo failed and blocked the ailerons at a one degree deflection just after take-off. This failure manifested itself by the inability of the autopilot to follow the commands of the flight director. The preferred solution was to disengage the autopilot and continue to fly the departure manually. In training scenario 2, the participants practised with a blocked elevator servo, instead of a blocked aileron servo.

In test scenario 4, the magnetometer failed during an approach. This was again detected by the on-board automation which automatically switched the autopilot to Roll Hold mode. The preferred solution was to continue the approach in Roll Hold mode while using Control Wheel Steering for the turns and Vertical Speed or Flight Level Change in combination with Altitude Hold Mode to fly the vertical profile. Training scenario 1 prepared the participants to use the standby instrument.

In test scenario 5, the static port got blocked after the first several legs of the RNAV approach were successfully flown in GPS navigation and Vertical Path Tracking. As a result, the altitude on the PFD froze, the airspeed on the PFD became unreliable and VNAV mode would never level off. The standby instrument functioned as usual. The preferred solution was to continue with GPS Navigation and switch to Pitch Hold mode while changing the pitch reference with Control Wheel Steering or the Nose Up and Nose DN buttons. Training scenario 3 prepared the participants for a problem with the pitot-static-system.

In test scenario 6, the elevator servo blocked, causing the elevator to be stuck at one degree deflection. This was an exact copy of training scenario 2 with only one minor difference in the task. The task in the training to "maintain 5,000" was changed for this test into "climb and maintain 6,000".

H. Dependent measures

The state of the aircraft and pilot inputs were logged at 50 Hz for analysis. Participants were asked to fill in questionnaires regarding workload and the experienced failures following the training and following each test scenario. Participants were debriefed about the origin and consequence of the failures only after completing all the test scenarios. The following dependent measures were used:

- *Time after failure to first relevant interaction;* time in seconds between the moment when the failure has been activated and the first relevant interaction of the participant with the autopilot. Relevant interaction is defined for this purpose as any interaction with the autopilot that has an effect on the same axis as where the failure took place (either lateral or vertical).
- *Time after failure to select highest automation;* time in seconds between the moment when the failure has been activated and the first time the participant has selected the highest automation modes possible with respect to the failure. For the time to select the highest level of automation, scenarios 2 and 4 were not taken into account. The on-board automation automatically switches to the highest level of automation available in these two scenarios, thus making this measure irrelevant.
- *Time in highest automation;* fraction of time in seconds where the participant has selected the highest automation modes possible with respect to the failure over the total length of the scenario.
- *Deviation from flight plan;* the lateral and vertical deviation from the intended flight path and the final way-point.
- *Multiple choice test score;* number of the correct answers out of the ten multiple choice questions ranging the score between 0 to 10.
- *Subjective measures;* the participants rated their level of surprise as well as the difficulty of the scenario on a 0–10 Likert-type scale ranging from "not at all" to "extremely". For the measure of surprise, a scale based on the anxiety scale of Houtman and Bakker (1989) was used. The perceived tension was also measured using a horizontal version of this anxiety scale. Mental effort was measured with the Rating Scale Mental Effort (Dijkstra & van Doorn, 1985).

I. Statistical analysis

In order to analyse the performance measures over the different scenarios, z-scores of the time to relevant interaction, time to select highest automation and time in highest automation were calculated. As time to relevant interaction and time to select highest automation can be regarded as reaction time, it can be assumed that these variables are not normally distributed. Using a chi square test it was found that they were actually chi square distributed. Using a power transform these values were transformed to be normally distributed such that z-scores could be determined. These z-scores for the scenarios were averaged per participant for the four scenarios with new automation surprises and the two scenarios with repeated automation surprises (see Table IV). Effectiveness in dealing with the automation surprise was considered by having a high score on the measure 'time in the highest level of automation' and low scores on the measures 'time to first relevant interaction' and 'time to switch to the highest level of automation'. Besides evaluating effectiveness based on these three measures, the group responses were analysed for each scenario in more detail. The significance level α was set to 5% for this study. The criterion for a medium-sized effect size was .3 and for a large-sized effect was .5.

III. Results

A. Manipulation check of the training

The subjective measures of enjoyment and mental effort of the training are presented in Table V. Both groups rated the training as very enjoyable. Several participants commented positively on the high pace and intensity of the training, which are likely contributing factors to the high enjoyment scores. The experimental group experienced a significantly higher mental workload compared to the control group, as is in line with the hypothesis.

TABLE V
SUBJECTIVE MEASURES OF TRAINING

	Experimental Group Mean (SD)	Control Group Mean (SD)	p
Enjoyment (7-49)	44.8 (0.84)	44.1 (0.84)	.571
Mental Effort (0-150)	58.1 (5.99)	33.9 (2.60)	.002

Table VI shows the time each group was in control in each phase of the experiment. The experimental group required somewhat more time to solve the training scenarios since they were figuring out the problems by themselves.

TABLE VI
TIME SPENT IN CONTROL [IN SECONDS]

	Experimental Group Mean (SD)	Control Group Mean (SD)	p
Familiarisation flight	4.57 (1.09)	6.80 (2.20)	.172
Training	65.37 (7.12)	58.26 (4.06)	.015
Flight simulator Test	49.57 (6.04)	47.42 (4.15)	.348
Total	120.32 (10.05)	112.15 (8.16)	.060

As the participants in the experimental group were not guided in the non-routine training scenarios by the instructor, they had to solve the problems they were facing on their own. They succeeded in doing this on average for 42% of the training scenarios. Only one participant was not able to execute the preferred procedure in all the training scenarios. In all training scenarios, at least three participants executed the preferred procedure. After explaining the failure and procedure to the experimental group, they were able to comply and successfully finish the training scenarios the second time. The results show that the training scenarios were challenging, but not impossible.

On average, participants in the experimental group scored lower on the multiple choice test (M = 6.2, SD = 1.93), than the participants in the control group (M = 7.0, SD = 1.49). This difference, .8, BCa 95% CI [-2.42, .82], was not significant t(18) = -1.037, p = .314.

B. *Effectiveness in dealing with automation surprises*

For the new automation surprises, the time spent in the highest level of automation for the experimental group (Mean rank = 11.10) did not significantly differ from the control group (Mean rank = 9.90), U = 44.00, z = -.454, p = .684, r = -.102. For the repeated automation surprises, the time spent in the highest level of automation for the control group (Mean rank = 11.65) did also not significantly differ from the experimental group (Mean rank = 9.35), U = 61.50, z = .873, p = .393, r = .195. The distributions of the z-scores for the time spent in the highest level of automation are presented in Figure 4.

For the new automation surprises, the time to select the highest automation in the experimental group (Mean rank = 7.63) did not differ significantly from the control group (Mean rank = 11.0), U = 55.00, z = 1.333, p = .203, r = .314. However, it did represent a medium sized effect. For the repeated automation surprises, the time to select the highest automation in the control group (Mean rank = 17.17) did also not significantly differ from the experimental group (Mean rank = 5.83). The distributions of the z-scores for the time to first relevant interaction are presented in Figure 5. Two participants of the experimental group did not select the highest level of automation after the failure in any of the three test scenarios with new automation surprises used for this measure. Therefore, only eight data points were available.

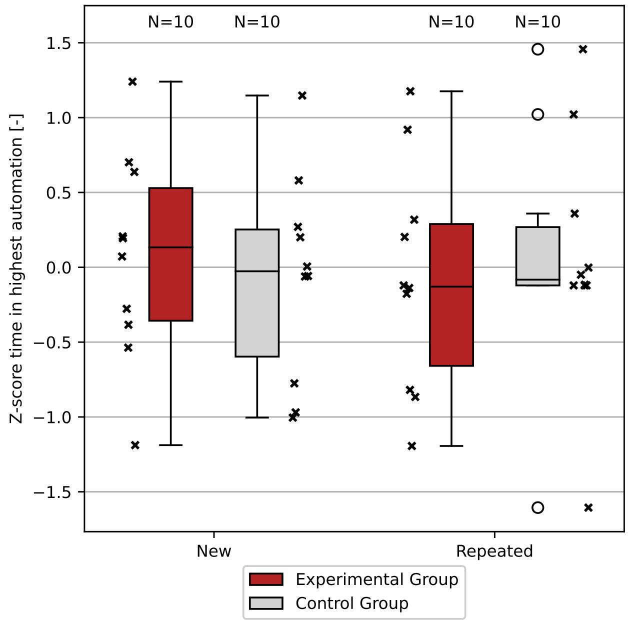


Fig. 4. Z-scores for the time the participants spent in the highest level of automation, for both the new and repeated automation surprises.

Four participants in both groups did not select the highest level of automation for the scenario with the repeated automation surprise used to analyse this measure, resulting in only six data points for each group.

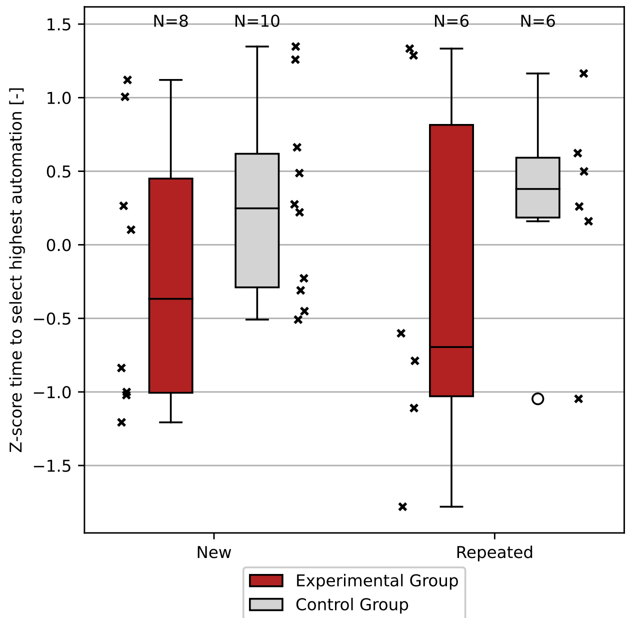


Fig. 5. Z-scores for the time it took to select the highest level of automation possible, for both the new and repeated automation surprises.

For the new automation surprises, the time to first relevant interaction in the experimental group (Mean rank = 8.80) did not significantly differ from the control group (mean rank = 12.20), U = 67.00, z = 1.285, p = .218, r = 0.287. However, this did represent a small to medium effect. For the repeated automation surprises, the time to first relevant interaction in the control group (Mean rank = 9.90) did not significantly differ from the experimental group (Mean rank = 11.10), U = 44.00, z = -.454, p = 0.684, r = -.102. The distributions of the z-scores for the time to select the highest automation modes are presented in Figure 6.

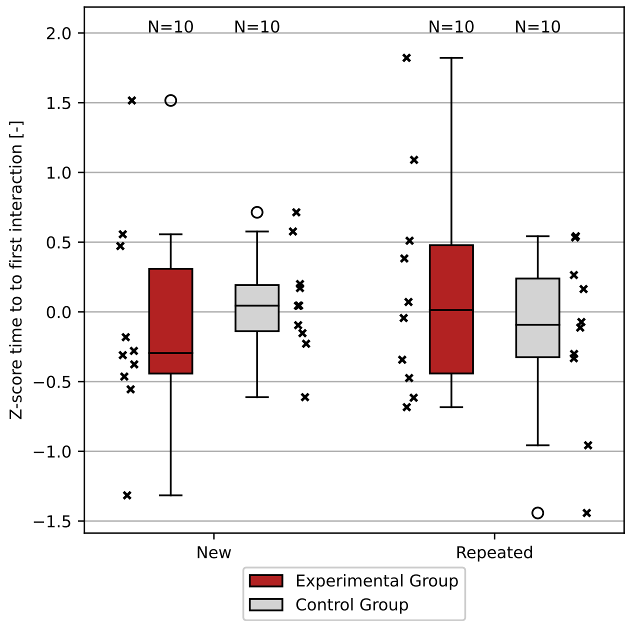


Fig. 6. Z-scores for the time after the failure to the first relevant autopilot interaction, for both the new and repeated automation surprises.

After each test scenario, the participants were asked to describe what has happened in a few words. Some responses mentioned the exact origin of the failure, whereas others only mentioned the consequence of the failure. Mentioning the consequence, though not as much as mentioning the failure, still indicates that the participant had some understanding of what was going on. The results of these questions are presented in Figure 7. For the new scenarios, the experimental group was able to mention the origin of the failure more frequently than the control group whereas both groups had a similar understanding of the scenarios. For the repeated scenarios, the control group shows a better understanding of what happened. The experimental group seems to have had a similar level of understanding of the repeated automation surprises compared to the new automation surprises.

These general observations show a slight difference between the two groups in the behaviour of the participants when faced with an automation surprise, in which the experimental group seems to be more effective with small to medium sized effects, yet not significant. The same is true for the scenarios with repeated automation surprises, as the control group seems more effective, but again not significant. The data shows a large spread in both groups, which is likely a contributing factor as why none of the effects are significant.

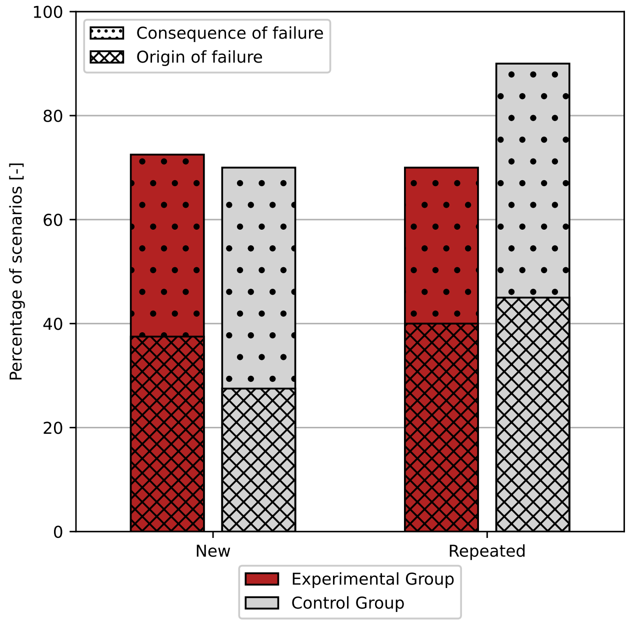


Fig. 7. Distribution of ability to identify the failure, for both the new and repeated automation surprises.

C. *Validation of test scenarios*

As a total of six scenarios were used for these measures of effectiveness in dealing with the automation surprises, a reliability analysis was performed. Based on the measure of time spent in highest level of automation, the six scenarios had high reliabilities, Cronbach's α = .72. For time to relevant interaction, Cronbach's α = .52 and would be increased to .78 if test scenario 3 was removed. For time to select highest level of automation, Cronbach's α = .16 and would be increased to 0.46 if test scenario 2 was removed.

D. *Subjective measures*

Subjective measures for surprise, scenario difficulty, experienced tension and mental effort are presented in Tables VII and VIII for the new automation surprises and the repeated automation surprises, respectively. The experimental group has given higher ratings for all these measures. Using a Mann-Whitney U test on all the subjective measures, it was found that only the mental effort in the repeated scenarios was significantly higher for the experimental group than the control group. It was expected that the experimental group would give lower scores on the measures difficulty and mental effort for the new automation surprises as they have practised solving unfamiliar surprises before, but the opposite was found to be true. These values for the subjective measures are comparable to a similar study by Landman et al. (2020) who simulated mechanical failures like a flap-asymmetry.

TABLE VII
SUBJECTIVE MEASURES FOR THE NEW AUTOMATION SURPRISES

	Experimental Group Median	Control Group Median	p
Surprise (0-10)	7.00	6.50	0.684
Difficulty (0-10)	6.00	5.00	0.143
Tension (0-10)	5.37	3.68	0.143
Mental Effort (0-150)	68.00	60.25	0.123

TABLE VIII
SUBJECTIVE MEASURES FOR THE REPEATED AUTOMATION SURPRISES

	Experimental Group Median	Control Group Median	p
Surprise (0-10)	4.75	4.25	0.529
Difficulty (0-10)	5.00	2.75	0.035
Tension (0-10)	2.95	2.53	0.315
Mental Effort (0-150)	57.50	41.00	0.095

E. Scenario-specific analysis

Figure 8 shows how many participants were able to identify the failure in each scenario. An overview of the three performance measures used to determine effectiveness in dealing with the automation surprises are presented per scenario in Figure 9. No significant differences between training groups were found when analysed per scenario either, apart for the time to first interaction in test scenario 3.

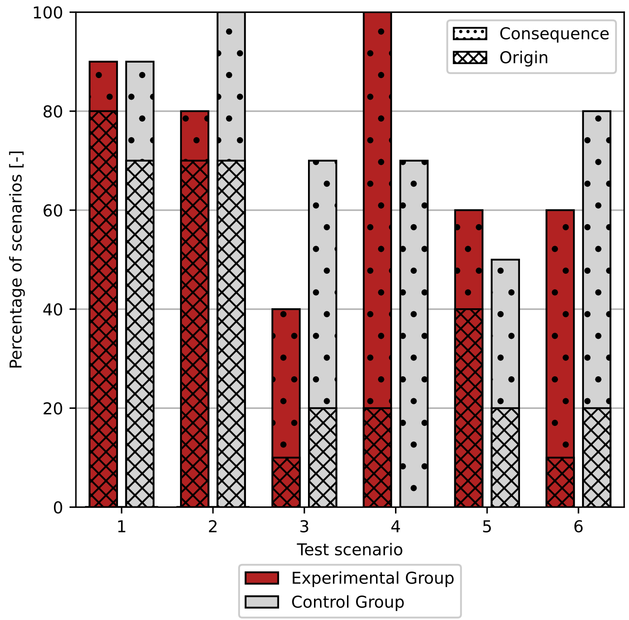


Fig. 8. Distribution of the ability to identify the failure, for all test scenarios.

For test scenario 1 (GPS failure), it was expected that the control group would be more likely to disengage the autopilot,

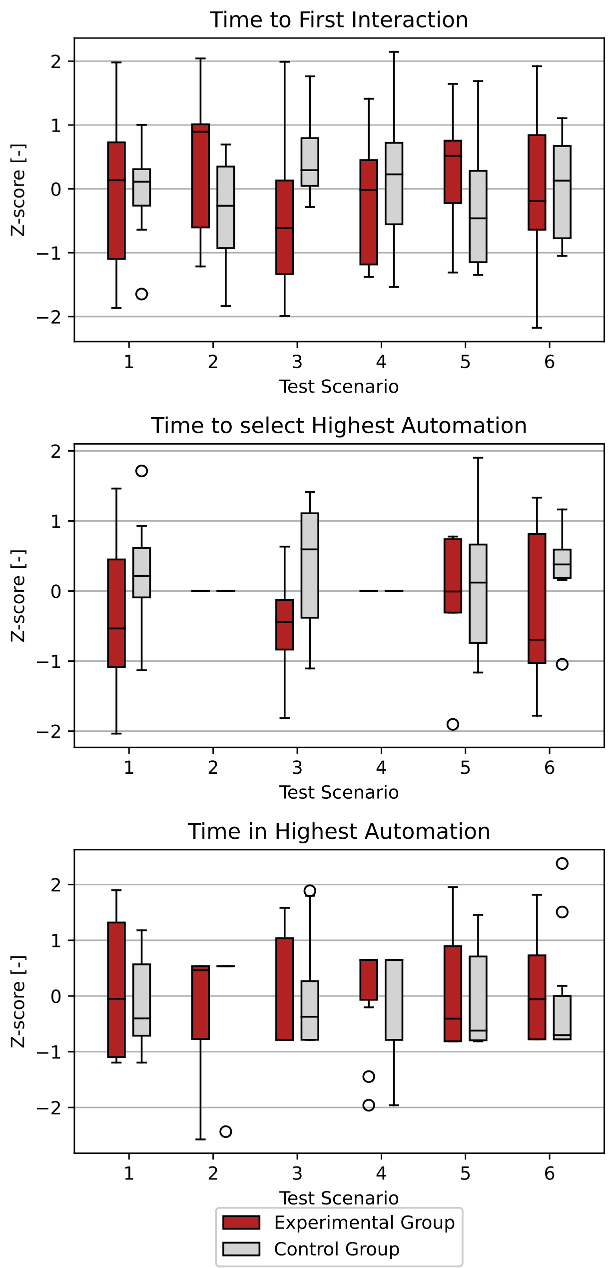


Fig. 9. Distribution of the performance measures per scenario. Note that test scenario 2 and 4 show no scores for time to select highest level of automation, as explained in the method.

end up in a wrong location as a result from not switching to VOR, and not being able to identify the problem. It was indeed found that only three participants of the experimental group compared to five participants in the control group ended up in the wrong location due to not switching to VOR navigation. In both groups only one participant disengaged the autopilot. Both groups had a good understanding of what was going on, as nine participants of the experimental group and eight participants of the control group were able to identify the origin or the consequence of the failure. The control group pressed more buttons (Mdn = 35.5) compared to the experimental group (Mdn = 26.5). One participant of the experimental group stalled the aircraft, while three participants of the control group found themselves overspeeding the aircraft.

For test scenario 2 (ADC failure), a repeated scenario, it was expected that the experimental group would need more time to diagnose the issue, as the control group already practised with this failure twice in the training. It was indeed observed that the median time to first interaction from the experimental group was 76.3 s compared to only 23.8 s for the control group. Besides, five participants of the experimental group disengaged the autopilot for at least some time, whereas only one participant from the control did so. Six participants from the experimental group manually changed away from the automatically selected mode after the failure, with two participants changing the mode within 10 seconds. Only one participant in the control group changed away from this automatically selected mode. The experimental group switched a combined total of 10 times between modes, of which five times lasted shorter than one second, against only two mode switches in the entire control group. All participants from the control group were able to mention the origin of consequence of the failure, whereas two participants of the experimental group failed to mention either. The participants from the control group were less surprised (Mdn = 3) than the control group (Mdn = 5). Two participants in the experimental group found themselves overspeeding the aircraft, compared to none in the control group.

For test scenario 3 (aileron servo failure), it was expected that the control group would respond later and would need more time before they disengaged the autopilot, while having more difficulties understanding the problem than the experimental group. The median of the time to first interaction was indeed much lower for the experimental group (11.9 s) compared to the control group (31.4 s). Surprisingly, six participants in the experimental group were not able to identify the failure or its consequence, compared to only three participants in the control group. As expected, the control group needed more time before deciding to disengage the autopilot as the median time to disengage to autopilot was 134.8 s for the control group and only 46.3 s for the experimental group. One participant in the experimental group requested to stop the simulation as he did not understand the scenario. In both groups, three participants continued to use the flight director for reference. While diagnosing the problem, five participants from the experimental group found themselves overspeeding the aircraft. The control group cycled the autopilot on and off a combined total of 13 times, whereas the experimental group only cycled the autopilot on and off a combined total of 5 times. This scenario had the highest ratings for scenario difficulty in the questionnaires with a 7.6 out of 10 from the experimental group and a 5.3 out of 10 by the control group.

For test scenario 4 (magnetometer failure), it was expected that the control group would take longer to respond to the failure and would be more likely to disengage the autopilot. Furthermore, it was expected that the control group would show a greater diversion from the intended flight plan. These expectations were not observed and the two groups reacted rather similarly with a median time to first interaction for the experimental group of 16 s and 18.8 s for the control group. One participant in the experimental group requested to stop the simulation as he did not understand the scenario. In the control group, four participants disengaged the autopilot compared to three participants in the control group. By analysing the cross track error, both groups show a similar diversion from the intended flight plan. All participants of the experimental group were able to identify the origin or at least the consequences of the failure, whereas none of the participants in the control were able to identify the origin. Three participants of the control group were also not able to identify the consequences of the failure.

For test scenario 5 (blocked static port), it was expected that the control group would be more likely to notice the problem later, spend more time diagnosing and trying out different modes and would disengage the autopilot more frequently than the experimental group. However, the control group was faster than the experimental group both in terms of time to first interaction and time to first select highest automation possible. The median for the time to select the highest level of automation for the experimental group was 30.6 s, and only 13.8 s for the control group. The control group did interact much more with the autopilot than the experimental group. Whereas the experimental group had a combined total of 21 mode changes (with a median of one), the control group had a total of 38 mode changes (with a median of four). As a result, the control group had a much larger average deviation of the flight plan compared to the experimental group. Since this is the only scenario in which the failure was introduced relatively late (after about 8 minutes), some interesting behaviour was observed during the time the automation was working as intended before the failure. Four participants in the experimental group never selected the highest vertical automation mode VNAV in this time. Only three participants in the experimental group and seven in the control group used this VNAV mode for the entirety of the time before the failure. A possible explanation for this could be that the participants were not comfortable using this rather complex mode. In line with that reasoning, it could be that the participants simply did not completely understand the usage of the autopilot. This is partially supported by the relatively low scores on the multiple choice test.

For test scenario 6 (elevator servo failure), a repeated scenario, it was expected that the experimental group would need more time to diagnose the issue, as the control group already practised with this failure with guidance in the training. Six participants from the experimental group and five participants from the control group disengaged the autopilot. Only three participants from each group continued to use the flight director for guidance. Despite this failure being part of one of the training scenarios, four participants in the experimental group and two participants in the control group where not able to identify the origin or consequences of the failure during the test. The experimental group cycled the autopilot on and off for a combined total of 10 times, whereas the control group only cycled the autopilot on and off for a combined total of five times. Two participants in the control group found themselves overspeeding the aircraft, compared to none in the experimental group.

F. Further observations

Two interesting correlations were found between the participant background and measures for effectiveness in dealing with the automation surprise. The time spent in the highest level of automation is significantly related to age, r = -.458, p = .042. This shows that older participants were less likely to select the highest form of automation. In addition, time to first interaction is significantly related to experience in flying hours, r = -.541, p = .014. This shows that participants with more flying experience reacted faster to the failures. These two correlations are visualised in Figure 10. This figure shows that the relation between participant experience in hours and the time to first relevant interaction may be exaggerated due to the two outliers in terms of experience.

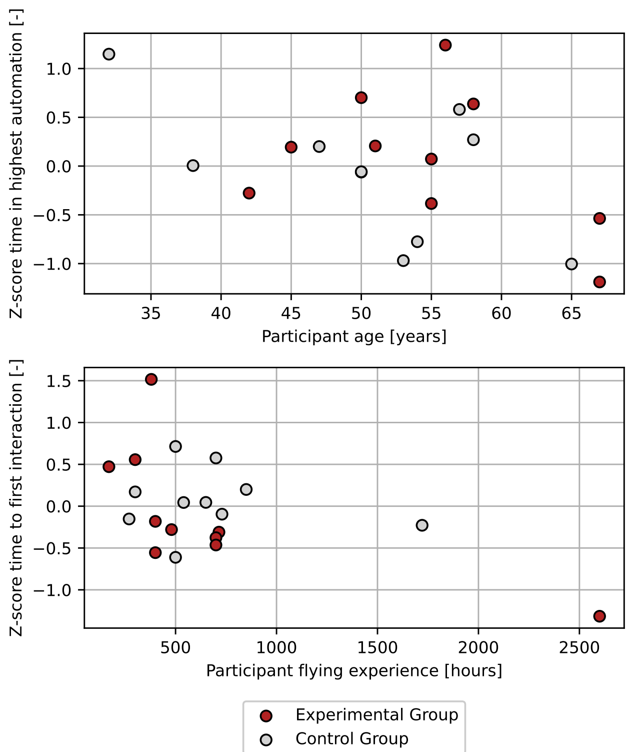


Fig. 10. Correlation of time in highest automation levels and age (top), and time to first relevant interaction and experience (bottom).

For the experimental group, no correlation was found between the number of training scenarios in which the participant executed the preferred procedure and the effectiveness in dealing with the automation surprises in the test. The one participant from the experimental group who did not manage to execute a single preferred procedure during the training scenarios did seem to perform below average in the test.

IV. Discussion

We have identified several possible reasons for the marginal non-significant differences. The problem-solving strategy that was provided to the experimental group may have been sub-optimal. Some participants indicated after the experiment that, partly due to the realism of the scenarios, they preferred to stick to their own habits when it comes to dealing with on-board problems. The strategy was not frequently applied in the training. This may have been a result of the type of problems selected for the training. They were developed to prepare pilots well for the upcoming failures in the test, but were not necessarily designed with the use of the problem-solving strategy in mind. This may have contributed to the low score of the experimental group in the training scenarios. The 42% of the training scenarios which were solved by the experimental group may have been too low for a proper transfer of training. In addition, some test scenarios were in retrospect better candidates to distinguish the two groups than others. The two scenarios in which warnings on the PFD occurred as a result of the failure (test scenario 2 and 4) were very similarly dealt with by both groups. Therefore, these scenarios may have been too easy.

When discussing the approach of the study and the hypotheses in the debriefing, most participants from both groups saw the potential benefit of simulating active problem solving in the training. However, the results are not conclusive enough to support this idea. The marginal differences that were observed can partly be attributed to some limitations of this experiment. Based on the results of the data analysis, a few recommendations can be made with respect to future experiments investigating this topic of pilot training for flight deck automation:

1) Increase training time; multiple participants were already struggling with the autopilot before the failure occurred in the test scenarios and several instances of mode confusion were observed.
2) Select a more homogeneous group; despite our efforts to balance the groups equally, individual differences in both age and experience between participants caused a large spread in most of the measures. This is a direct result from using private pilots, as they are in general a much more diverse group than airline pilots.
3) Decrease freedom in scenarios; although the scenarios created a realistic setting, they allowed for too much freedom, making it difficult to compare the dynamic pilot responses. Limiting the scope of the scenarios and the available options the participants have is likely to magnify the observed differences. One way to achieve this would be to limit the number of different modes to be used in the training and test.
4) Focus on non-obvious problems; the failures in test scenario 2 and test scenario 4 were rather obvious, as clear indications of these failures were displayed on the

PFD. In these two scenarios, the differences between the two groups were small to non-existent. Focusing on failures which do not trigger any sort of alarm in the cockpit, such as the GPS failure in test scenario 1 or the aileron servo failure in test scenario 3 is likely to result in more distinct differences.

V. Conclusion

It was expected that the experimental group would be more effective in dealing with automation surprises than the control group, as the experimental group had practised solving authentic problems in their training. The results show that the experimental group appeared to be slightly more effective in dealing with new automation surprises by selecting the highest automation level faster and remaining in it for a longer time than the control group. However, these results were not significant. In line with the expectation, the opposite was found for the repeated automation surprises. The control group appeared to be slightly more effective in dealing with the automation surprises which were repeated from the training scenarios, but again no significant differences were found. The main hypotheses regarding the effectiveness of active problem solving in pilot training cannot be accepted, as the results show only a small insignificant benefit of the experimental training for new automation surprises. A large spread of the results was observed in both groups, signifying individual differences between participants played a large role. Despite the marginal differences in the general observations, the expectations of the responses to the new automation surprises were mostly apparent as trends in the data.

Two significant correlations were found: a negative relation between participant age and the time spent in the highest automation and a negative relation between participant experience and time to first interaction. Older participants were less likely to select the highest form of automation, and participants with more flying experience reacted faster to the failures. It was found that all scenarios were reliable in terms of measuring time spent in the highest automation level possible.

In line with the hypotheses, the Problem-Based Learning type of training required significantly more mental effort than the baseline training, but both training types were similarly rated in terms of enjoyment. It can be argued that the two-hour training was not sufficient for the participants to completely understand the automation systems. Both groups were not able to pass the multiple choice test with distinction, and several participants struggled to select the highest level of automation even before the failure occurred. Furthermore, several instances of mode confusion were observed by the experiment coordinator.

This study shows that there is a slight benefit of simulating active problem solving in pilot training. However, as the differences between the two groups were not conclusive, the current results cannot be used to support a recommendation to aviation authorities to mandate active problem solving in pilot training at this time.

Acknowledgment

The author would like to thank prof. Dr. E. Groen from TNO for his advice on the design of the experiment and the data analysis. Further more, the author would like to thank NASA cognitive psychologists Dr. D. Billman and Dr. R. J. Mumaw for the crucial insights they provided on training.

References

Bourgeois-Bougrine, S., Gabon, P., Mollard, R., Coblentz, A., & Speyer, J.-J. (2018). Fatigue in aircrew from shorthaul flights in civil aviation: the effects of work schedules. In H. Muir & D. Harris (Eds.), *Human factors and aerospace safety* (pp. 177–187). Routledge. doi: 10.4324/9781315194035-5

Brna, P., Ohlsson, S., & Pain, H. (1993). The avionics job-family tutor: An approach to developing generic cognitive skills within a job-situated context. In *Artificial intelligence in education 1993 : proceedings of ai-ed 93, world conference on artificial intelligence in education* (pp. 513–520). Charlottesville, VA.

de Boer, R. J., & Hurts, K. (2017). Automation surprise: Results of a field survey of dutch pilots. *Aviation Psychology and Applied Human Factors*, *7*(1), 28-41. doi: 10.1027/2192-0923/a000113

Degani, A., & Wiener, E. L. (1998). Design and operational aspects of flight-deck procedures. In *the international air transport association (iata) annual meeting.* Montreal, Canada.

Dehais, F., Peysakhovich, V., Scannella, S., Fongue, J., & Gateau, T. (2015). "automation surprise" in aviation. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems - CHI 15.* New York, NY: ACM Press. doi: 10.1145/2702123.2702521

Dijkstra, F., & van Doorn, L. (1985). *The construction of a scale to measure subjective effort* (Unpublished doctoral dissertation). Delft University of Technology.

Dochy, F., Segers, M., den Bossche, P. V., & Gijbels, D. (2003). Effects of problem-based learning: a meta-analysis. *Learning and Instruction*, *13*(5), 533–568. doi: 10.1016/s0959-4752(02)00025-7

FAA. (1996). *The interfaces between flightcrews and modern flight deck systems* (Tech. Rep. No. 00784270). Federal Aviation Authority.

Fletcher, G., & Bisset, G. (2017). *Pilot training review – interim report: literature review.*

Flight Deck Automation Working Group. (2013). *Operational use of flight path management systems.*

Goteman, Ö. (2018). Automation policy or philosophy? management of automation in the operational reality. In S. Dekker & E. Hollnagel (Eds.), *Coping with computers in the cockpit* (p. 215 - 221). Routledge. doi: 10.4324/9780429460609-12

Holder, B. (2012). *Airline pilot perceptions of training effectiveness.* Seattle, WA.

Houtman, I., & Bakker, F. (1989, sep). The anxiety thermometer: A validation study. *Journal* *of Personality Assessment*, *53*(3), 575–582. doi: 10.1207/s15327752jpa5303_14

Jonassen, D. H. (2010). *Learning to solve problems*. Routledge. doi: 10.4324/9780203847527

Jonassen, D. H., & Strobel, J. (2006). Modeling for meaningful learning. In D. Hung & M. S. Khine (Eds.), *Engaged learning with emerging technologies* (pp. 1–27). Springer-Verlag. doi: 10.1007/1-4020-3669-8_1

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86. doi: 10.1207/s15326985ep4102_1

Landman, A., van Middelaar, S. H., Groen, E. L., van Paassen, M. M. R., Bronkhorst, A. W., & Mulder, M. (2020, jun). The effectiveness of a mnemonic-type startle and surprise management procedure for pilots. *The International Journal of Aerospace Psychology*, 1–15. doi: 10.1080/24721840.2020.1763798

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, *59*(1), 14–19. doi: 10.1037/0003-066x.59.1.14

Muynck, R. D., & Hesse, M. V. (1990). *The a priori simulator software package of the piper pa34 seneca iii.*

Nickolic, M. I., & Sarter, N. B. (2007). Flight deck disturbance management: A simulator study of diagnosis and recovery from breakdowns in pilot-automation coordination. *Human factors*, *49*(4), 553-563. doi: 10.1518/001872007X215647

Novak, J. D. (2013). Meaningful learning is the foundation for creativity. *Qurriculum. Revista de Teoría,Investigación y Práctica educativa*, 27–38. doi: 10.25145/j.qurricul

Orlady, L. M. (2010). Airline pilot training today and tomorrow. In B. G. Kanki, R. L. Helmreich, & J. Anca (Eds.), *Crew resource management* (pp. 469–491). Elsevier. doi: 10.1016/b978-0-12-374946-8.10020-2

Rosa, E., Dahlstrom, N., Knez, I., Ljung, R., Cameron, M., & Willander, J. (2020). Dynamic decision-making of airline pilots in low-fidelity simulation. *Theoretical Issues in Ergonomics Science*, 1–20. doi: 10.1080/1463922x.2020.1758830

Sarter, N. B., & Woods, D. D. (1997). Team play with a powerful and independent agent: Operational experiences and automation surprises on the airbus a-320. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(4), 553–569. doi: 10.1518/001872097778667997

Savery, J. R., & Duffy, T. M. (1995). Problem based learning: An instructional model and its constructivist framework. *Educational Technology*, *35*(5), 31–38. Retrieved from `www.jstor.org/stable/44428296`

Schaafstal, A., Schraagen, J. M., & van Berl, M. (2000). Cognitive task analysis and innovation of training: The case of structured troubleshooting. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *42*(1), 75–86. doi: 10.1518/001872000779656570

Stroosma, O., van Paassen, M. M., & Mulder, M. (2003, aug). Using the SIMONA research simulator for human-machine interaction research. In *AIAA modeling and simulation technologies conference and exhibit.* Austin,TX: American Institute of Aeronautics and Astronautics. doi: 10.2514/6.2003-5525

White, B. Y., & Frederiksen, J. R. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, *42*(1), 99–157. doi: 10.1016/0004-3702(90)90095-h

Woods, D. D., & Johannesen, L. J. (1994). *Behind human error: Cognitive systems, computers, and hindsight*. Crew Systems Ergonomics Information Analysis Center. doi: 10.1201/9781315568935

Part II

Preliminary Report

NOTE:
This part has already been graded under AE4020

1 INTRODUCTION

Commercial aviation is the safest form of travel, and its safety records have been significantly improved in the past decades. Looking at the most recent accidents in aviation, Loss of Control In-Flight (LOC-I) is found as the most prevalent cause. The International Air Transport Association (IATA) defines LOC-I as the flight crew being unable to maintain control of the aircraft in flight, resulting in an unrecoverable deviation from the intended flight path (IATA, 2019). Studies analysing LOC-I accidents have indicated that inappropriate crew responses are involved in 43% of the cases (IATA, 2019; Shappell et al., 2007).

Surprise and startle, due to unexpected events in the flight deck, are mentioned as crucial elements in recent flight safety events involving LOC-I (Landman, Groen, van Paassen, Bronkhorst, & Mulder, 2017a). The result of a surprise is that the flight crew needs solve the cognitive mismatch between their expectation and current situation, which is mentally challenging and can cause stress (Dismukes, Goldsmith, & Kochan, 2015). Landman, Groen, van Paassen, Bronkhorst, and Mulder (2017b) have shown that a surprise negatively influences pilot performance during stall recovery.

Even though the advances in the automated systems have greatly contributed to the safety of flying, it does has several drawbacks. In 1983, Bainbridge (1983) published her paper 'Ironies of Automation', in which she described issues with human-automation interaction. The task of the pilot has shifted from manually flying the aircraft into monitoring the automation, a task at which humans do not perform well (Bainbridge, 1983; Loss of Control Action Group, 2013). Manual flying skills are degraded by the use of automated systems since pilots do not practice their flying skills on a frequent basis. Finally, Bainbridge points out that although automation is meant to simplify the system, it achieves the opposite since the system and the automation are by definition more complex than just the system itself. Automation can hide system degradation, which together with complacency can create dangerous situations.

Several reports by aviation authorities have indicated issues when flight crews have to deal with surprises directly related to flight deck automation (FAA, 1996; Fletcher & Bisset, 2017). Pilots make incorrect decisions when dealing with automation surprises as they do not fully understand how the complex autoflight systems work. According to the FAA, this is a contributing factor in over 40% of the investigated accidents (Flight Deck Automation Working Group, 2013). Pilot training focuses heavily on procedures and checklists, as these have proven to be extremely useful in time critical and emergency situations (Degani & Wiener, 1998). As a result, pilots can only reproduce the procedures they have been trained in but lack a higher-order understanding to make their own judgement about the current state of the systems. Pilots are trained using a few highly standardised scenarios including little to no elements of surprise or novelty. The result is that pilots are not able to learn from their mistakes and consequently do not actively develop problem-solving skills during training. Fletcher and Bisset (2017) specifically mention in their review of pilot training for the CAA that there is a need to target decision making and problem solving for pilot automation training. As of 2019, Performance Based Navigation (PBN) and Upset Prevention and Recovery Training (UPRT) has been made mandatory parts of flightcrew training, as a result of recent accident reports, but no changes have been made in the flight deck automation training.

1.1. RESEARCH OBJECTIVE

The main research question for this thesis is:

By teaching a problem-solving strategy, will pilots become more effective in dealing with an automation surprise?

This main question is split into the following four sub-question:

1. What are the current training practices for flight deck automation systems and how effective are they in preparing pilots to operate the flight deck automation systems?
2. What is the effect of surprise on pilot behaviour in combination with the automation systems?
3. How can effective problem solving be trained?
4. How is the problem-solving capacity of a pilot affected by changing the training method for flight deck automation?

1.2. BACKGROUND

Previous research in the Human-Machine interface research group at the Delft University of Technology has focused on startle and surprise on the flight deck. Landman et al. (2017a) have introduced a model to understand the startle and surprise response among pilots. This model has been used to analyse surprise behaviour and reduce its effects during manual control. The model combines existing models on surprise, startle, perception and sense-making in an effort to understand the psychological effects flightcrews may experience during unexpected or threatening scenarios in flight. This model is visualised in Figure 1.1 (Landman et al., 2017a).

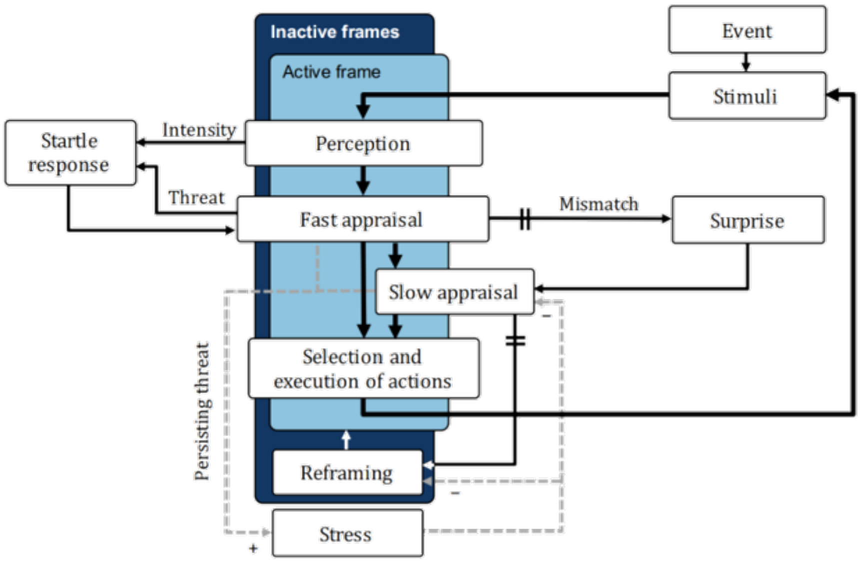


Figure 1.1: The conceptual model of startle and surprise

The core of the model consist of a perceptual cycle, which starts with an event causing stimuli, which in the case of automation surprises are generally either visual or aural. These stimuli are perceived by the pilot who uses this information for sense-making. This is modelled to be either fast and almost automatic, which is the case for automaticity, or slow and requiring effort from the pilot. Based on appraisal, the pilot selects and executes the appropriate actions which causes new stimuli.

Left of the perceptual cycle is the startle response. Startle can be seen as a reflex to threatening stimuli. In contrast, a surprise is an emotional reaction triggered by a mismatch between the current situation and the mental model of the pilot. Startle can occur without surprise, when an event triggers the startle response but is in line with the mental model of the pilot. Similarly, a surprise can occur without the need for startle.

However, if the startling event is unexplainable, a surprise will be experienced by the pilot.

Perception and appraisal followed by execution of selected actions all rely on the active frame, which is the current mental model the pilot has of the situation. Frames are based on previous experience and are important for proper situation awareness. Frames filter and select incoming data to provide meaningful information and consequently skill based behaviour. A mismatch between the active frame and the incoming information will cause a surprise. A surprise will require (slow) appraisal to understand the exact mismatch. Reframing is required when the current frame is inconsistent with the incoming information. If the appropriate frame can be activated, the surprise will be over and the pilot understands the situation once again. However, when no appropriate frame is available or the appropriate frame can not be activated, the pilot loses situation awareness. Difficulty in reframing can be explained by increased levels of stress as a result from the surprise event. Besides, stress negatively influences slow appraisal as the filtering effect by the active frame is reduced and preference is given to simple and familiar risk avoiding solutions.

1.3. REPORT PURPOSE AND STRUCTURE

The purpose of this report is to present the literature study and the experiment plan for the research on using a problem-solving strategy in training to help pilots effectively deal with automation surprises. The results of this research will be presented in a paper. The structure of this report is as follows. Chapter 2 contains the literature study which presents the state-of-the-art in flight deck automation training. The topic of automation on the flight deck is introduced and operational problems are discussed. Besides, different views on training from the field of educational psychology are covered. Chapter 3 introduces the experiment to answer to main research question, states the hypotheses and presents the experiment plan. The design of the scenarios used in the experiment is discussed in Chapter 4. Based on a set of requirements, a series of ideas for potential scenarios was generated. In Chapter 5, the implementation of the flight simulation is explained. This includes the aircraft model choice, the general implementation and a detailed description of the autopilot and avionics which were created for this project. Finally, Chapter 6 presents the conclusion of this preliminary thesis report.

Chapter 4 and 5 represent the current state of the development of the scenarios and the flight simulation, as this is a work in progress. Updates to these sections will be provided in appendices of the final report of this thesis.

2
LITERATURE

This chapter presents the Literature Study that was performed to determine the state-of-the-art in flight deck automation training and to answer some of the sub-questions. The first section is an introduction to flight deck automation and the problems that arise from automation usage. This is followed by a section covering pilot training and its associated problems. Finally, theory from the field of educational psychology is provided on the topic of training and learning.

2.1. FLIGHT DECK AUTOMATION

The term automation originates from the Greek 'auto', meaning self, and 'matos', meaning moving or acting. In their paper on human-machine interaction, Sheridan and Parasuraman (2005) define automation as the full or partial replacement of functions previously performed by humans. Reasons for automation include increasing safety and performance, and to control dangerous or difficult processes (Wiener & Curry, 1980). Automation in aviation is supposed to decrease both pilot workload and fatigue, allowing for increased productivity and a better execution of standard operations. Physical workload is reduced as pilots do not need to control the aircraft manually for long periods of time. Mental workload is decreased as routine computations associated with navigation are now performed by the automated systems and displayed on screens. Besides, automation allows for more accurate control which results in more comfort and less fuel use compared to manual control by a pilot (Bourgeois-Bougrine, Gabon, Mollard, Coblentz, & Speyer, 2018). The use of automation can increase mental workload as well. Harris and Harris (2016) mention in their book Human Performance on the Flight Deck that automation simply changes the nature of the workload instead of reducing it. The introduction of automation systems in the flight deck has greatly contributed to the safety records by reducing accident rates resulting in a decline in fatalities of 50% over the last decade (Dobie, 2014).

Different systems to quantify the level of automation used, or taxonomies, have been developed since 1970. These taxonomies are based on the division and responsibility of both autonomy and authority between the human and the machine. Assuming different levels of automation exist allows to train pilots in the different levels of involvement required in controlling the aircraft (Goteman, 2018).

2.1.1. USE OF AUTOMATION

Flight Deck automation can be divided into autoflight systems, flight instruments, warning systems and navigation systems. The autoflight systems include the autopilot and flight director, auto-throttle and flight management system (FMS). The flight instruments in modern cockpit consist primarily of the Primary Flight Display (PFD), Navigation Display (ND) and the Electronic Flight Instrument System (EFIS). When the EFIS information is displayed on the same screen as the ND, it is considered to be a Multi-Function Display (MFD). Information displayed on the PFD includes, but is not limited to, attitude, indicated airspeed, altitude, vertical speed and heading information. The navigation systems include Global Positioning System (GPS), Instrument Landing System (ILS), localizer, VOR, Area Navigation (RNAV) systems and the vertical and lateral navigation modes (VNAV and LNAV respectively) (Flight Deck Automation Working Group, 2013).

Bourgeois-Bougrine et al. (2018) state that the pilot controls the aircraft using the interface to the flight deck automation systems as well as the primary flight controls. They identified three control modes:

1. Manual Control; the pilot is flying the aircraft manually using the primary flight controls. The autopilot is turned off and has no influence on the control of the aircraft.

2. Tactical Control; the pilot enters long-term commands into the FMS using the Mode Control Panel (MCP) mid-flight. The autopilot controls the aircraft to maintain the commanded heading and altitude. In this way, the pilot is able to respond to real-time events such as vectoring by Air Traffic Control (ATC).

3. Strategic Control; the flight-plan is entered into the FMS before flight and is executed by autopilot throughout the flight. The FMS commands the autopilot for both lateral and vertical navigation.

Airliner policies state that "the level of automation used shall be the most appropriate for the task at hand with regard to safety, passenger comfort, regularity and economy" (Goteman, 2018). Practically, this means that pilots hardly ever use manual control, thus being automation managers and responsible for higher-order goals (Bourgeois-Bougrine et al., 2018). The pilot is pushed into a supervisory roll, where the task of the pilot has shifted from manually flying the aircraft into monitoring the automation, a task at which humans do not perform well (Bainbridge, 1983; Loss of Control Action Group, 2013). After take-off, commercial aircraft are almost always flown using autoflight systems which are controlled by the FMS. In airliners, pilots interact with the FMS through the Control Display Unit (CDU) and Mode Control Panel (MCP). The CDU is used for strategic control and the MCP is used for tactical control.

In general, aircraft control has a hierarchical nature, since the flight path is controlled by changing lower order states. The control surfaces of an aircraft are used to influence the rotational acceleration around an axis; ailerons control the roll rate, elevators control the pitch rate and the rudder controls the yaw rate. In order to use roll rates to get to a change of flight path, cascaded control is required. For a change in altitude, the elevators are used to change the pitch rate. A change in pitch rate results in a change in pitch angle, which on its own results in a change in vertical speed required to obtain the change in altitude. The internal structure of the autopilot is based upon this cascaded control. As a result, different autopilot modes exist for both the pitch and the roll axis. The basic modes for the pitch axis are pitch angle hold and altitude hold. The basic modes for the roll axis are roll angle hold and heading hold. In strategic control, the Lateral Navigation (LNAV) mode and Vertical Navigation (VNAV) mode use a combination of the basic modes to track the intended flight-path automatically. LNAV determines the desired track course based on the flight plan and selects the appropriate navigation source to track the flight plan by selecting roll control modes. With VNAV active, altitude and speed targets (provided that the aircraft is installed with an auto-throttle) are automatically selected based on the flight plan entered in the FMS. Pitch and thrust control modes are automatically selected to obtain and maintain these targets. Finally, VNAV is able to provide the optimum descent path as reference for the altitude and speed targets (Sherry, Polson, Mumaw, & Palmer, 2001). These general autopilot modes are summarised in Table 2.1.

Table 2.1: General autopilot modes

Level of automation	Vertical modes	Lateral modes
Low	Pitch angle hold	Roll angle hold
Intermediate	Vertical Speed hold	-
Intermediate	Altitude hold	Heading hold
High	VNAV	LNAV

2.1.2. Identified Problems with Automation Usage

In 1996, the Federal Aviation Authorities (FAA) charted a human factors team to investigate the human-machine-interface between flightcrews and advanced flight decks. They concluded that vulnerabilities existed among flightcrews in managing the automation systems and maintaining proper situation awareness. With respect to automation management, they showed their concerns regarding pilot understanding of the automation systems and incorrect pilot decisions regarding the appropriate automation level to use during the different phases of flight. Concerns regarding situation awareness included mode awareness and flight path awareness (FAA, 1996). In the years after, many improvements have been made with respect to automation design, training and operational use. However, the FAA noticed that flight incident and accident reports still suggested flightcrews having difficulties with the autoflight systems, partly because of the introduction of new concepts such as Area Navigation (RNAV) and Performance-Based-Navigation (PBN). Therefore, the FAA founded a team of experts to investigate and address these new concerns. The Flight Deck Automation Working Group (2013) analysed data from recent accident and incident reports. Besides, they performed interviews with aircraft manufacturers, airlines and pilot training organisations. They have reviewed airline policies, flight operations, and pilot qualification and training. Based on this analysis, 28 findings were published regarding flight deck automation, pilot training, flight deck design, flight-crew management, procedures and regulations and data availability. Below, the findings with regard to flight deck automation are summarised:

- Managing Malfunctions; system failures which occur in normal flight operations are successfully identified as threats and managed accordingly. However, when no procedures or checklists exist for certain failures, a pilot's capacity to take appropriate actions is reduced. This effect is increased due to insufficient system knowledge and lack of understanding of the current state of the aircraft.

- Automated Systems; although automated systems have greatly contributed to aviation safety and operational efficiency, several vulnerabilities were identified. These include over-reliance in automated systems, autoflight mode confusion, wrong implementations of information automation and FMS programming and usage errors.

- Operator Policies for Flight Path Management; the automation policies used by airliners should be improved to focus on tasks related to flight path management. This would allow pilots to make more effective use of the automated systems.

- Pilot Knowledge and Skills for Flight Path Management; pilots do not possess the knowledge and skills required for efficient and effective use of the flight path management systems and related tasks.

As of 2019, PBN and Upset Prevention and Recovery Training (UPRT) have been made mandatory parts of flightcrew training, partly as a result of the report published by the joint working group, but no changes have been made in the flight deck automation training.

The Civil Aviation Authority commissioned a project as a "review of recent training studies and activities and potential improvements in order to inform policy on taking the matter forward internationally."(Fletcher & Bisset, 2017) They mainly focused on reviewing pilot training in the UK. The report mentioned that pilots over-delegate authority to the autoflight systems which reduces their situation awareness and ultimately makes the flightcrew over-reliant on the automation. Situation awareness on the flight deck is described as the perception, understanding and prediction of environmental and situational cues. In order to preserve proper situation awareness, expertise in monitoring, anticipation, recognition and situational assessment are mentioned as crucial elements. A reduced situation awareness lowers a pilot's ability to predict and anticipate aircraft behaviour. This increases the difficulty of controlling the aircraft in unexpected events. As mentioned before, the main task of pilots during normal flight is monitoring the automation. This observing role in combination with the over-reliance causes pilots to become "out-of-the-loop" (OOTL). Gouraud, Delorme, and Berberian (2017) defines this OOTL problem as "a deterioration of the operator's attention when interacting with highly automated systems". This means a reduced or total loss of situation awareness. Another result from OOTL situations is vigilance failure, where concentration is not maintained over a longer time period (Sarter & Woods, 1995). A pilot experiencing OOTL performance problems could take longer to identify and intervene automation failures. According to the Flight Deck Automation Working Group (2013), high reliability of automation systems and operational policies by airliners contribute to the over-reliance in the flight deck automation. Their report shows that pilots were too confident in about a quarter of the accidents and a reduced situation awareness was found in over 55% of analysed accidents.

Another problem related to flight deck automation identified by Fletcher and Bisset (2017) in their report for the CAA is the lack of mode awareness and mode management amongst pilots. Pilots should always select the appropriate autopilot mode depending on the circumstances. However, pilot responses to unexpected events are often incorrect or inappropriate (Martin, 2013). Whereas upset recovery procedures state to reduce the level of automation used, a study by Nickolic and Sarter (2007) showed that pilots did the exact opposite, and several of their participants reported to have inaccurate knowledge about the different modes of the flight path management system. Besides, improper mode management, indirect mode changes and mode confusion are often mentioned by flightcrews (Fletcher & Bisset, 2017). The FMS can automatically change between different modes, which causes confusion among pilots. This is known as lack of mode awareness. Sherry et al. (2001) analysed the VNAV function in modern aircraft and concluded that this function is overloaded. The VNAV button on the MCP can command one of six possible behaviours and these will change autonomously.

In their analysis, pilots reported unexpected performance of the VNAV function during descent. In 27% of accidents reviewed by the Flight Deck Automation Working Group (2013), mode selection errors were mentioned.

The Flight Deck Automation Working Group (2013) has mentioned startle while the pilot is monitoring the automated systems to be a factor for LOC-I accidents. Another issue experienced by many pilots is automation surprise (see de Boer and Hurts (2017)), a concept firstly introduced by Woods and Johannesen (1994) who defined it as "situations where crews are surprised by actions taken (or not taken) by the automatic system". Such an automation surprise occurs when there is a conflict between the pilot and the automation systems which the pilot detects but does not understand (Dehais, Peysakhovich, Scannella, Fongue, & Gateau, 2015). Indirect mode changes, inhibited behaviour (where the automation cancels pilot actions) and degradation in automation are often mentioned as automation surprises (de Boer & Hurts, 2017; Dehais et al., 2015). These are indications that the flight crew misused or misunderstood the automation systems. Automation surprises can jeopardise flight safety as the aircraft can get in an unrecoverable state if the surprise is not noticed and recovered in time (Sarter & Woods, 1997). With a survey among Dutch airline pilots, de Boer and Hurts (2017) found that on average a pilot experiences an automation surprise three times per year. However, most of these automation surprises remain inconsequential. They also found that the likelihood of an automation surprise is significantly higher during flight phases with high workload. Besides, most of the automation surprises from this survey occurred when using a high level of automation, which is the case during Strategic Control. Automation surprises are a result from vulnerabilities in pilot knowledge. The Flight Deck Automation Working Group (2013) mentions that 40% of the analysed accidents show some knowledge deficit with the pilots. They found the following areas in which pilot knowledge generally is insufficient:

- Understanding of the complex relations of the autoflight systems
- Crew Resource Management (CRS)
- Upset recovery
- Operating limits and energy management

Besides, understanding of how the FMS computes the optimal flight path is lacking. Line Operation Safety Audits (LOSA), a programme airlines use to identify threats to operation safety due to human errors, have even indicated that this insufficient knowledge about flight path computations seem to be the standard among pilots, and that "crews that were able to demonstrate such anticipation were rated highly by their peers" (Flight Deck Automation Working Group, 2013).

2.2. STATE-OF-THE-ART PILOT TRAINING

This section covers the organisation and focus of current pilot training. Even though pilot training has been revised and updated several times since the introduction of the glass cockpit, human factors remain to be a prevalent cause in many accidents (Flight Deck Automation Working Group, 2013). As a result, several studies and reports have indicated pitfalls and problems with current pilot training. Partially based on these reports, studies have been performed to improve training by adding new features.

2.2.1. TRAINING PHASES

Pilot training is split into different phases. Trainee pilots traditionally follow ab initio training followed by an aircraft specific type rating (Rigner & Dekker, 2000). During the ab initio training, trainee pilots follow clearly defined phases utilising ground school and flight training in both simulators and real aircraft (Fletcher & Bisset, 2017). During ground school, the trainee pilots receive standardised training in courses like general aircraft knowledge, performance and operational procedures. Flight training starts in a single engine piston aircraft and is continued in a twin engine piston aircraft. When the ab initio training is successfully finished, the trainee pilot will receive the (frozen) Airline Transport Pilot License (ATPL). With this license, pilots are able to work and fly commercially. When a pilot has accumulated 1500 flying hours, the ATPL is unfrozen and the pilot can work as a captain. However in order to operate an airliner, ATPL pilots first have to get their type rating training in which they learn the skills to operate one specific type of aircraft. The type rating training is the phase were pilots learn and practise with the actual aircraft they will be flying (may it be in a simulator). Thus, in this phase of the training, pilots are actually trained for the flight deck automation systems. The type rating training is divided into two parts: systems training and procedural training. Systems training is provided over the course of three weeks of ground school. During ground school, system knowledge and system integration knowledge of all the on-board systems are covered (Hawkins & Orlady, 1993). After completing ground school, pilots continue to the procedural training part of the type rating. This takes place in flight simulators, where pilots repeatedly practise standard scenarios. They are trained in normal flight operations, abnormal flight operations and emergency flight operations. Pilots develop and enhance expert skills which helps to act in a instinctive and natural way during non-routine operations (Kiss, 2019). The type rating training also includes Crew Resource Management (CRM), a non-technical course to achieve safe and efficient operation in communication, teamwork and leadership (Rigner & Dekker, 2000).

In order to keep the flight crew adequately trained and proficient for the aircraft they are operating, annual recurrent training is required. Recurrent training consist of a review to assess knowledge, instruction in subjects required for initial ground training, a competence check and CRM training (according to 14 CFR 121.427[1]). The delivery of the recurrent training is type dependent and defined by the operators. This normally includes LOFT in a simulator to refresh standard operating procedures and emergency procedures (Fletcher & Bisset, 2017).

2.2.2. Identified Problems with Training

Airline pilot training is a huge expense for airlines (Qi, Bard, & Yu, 2004). Commercial pilots are required to perform recurrent training lasting one to three days annually, with costs for a Full Flight Simulator ranging between €350 and €550 per hour(FAA, 2004). Airlines design their training programs based on operational needs and fleet requirements whilst complying with regulatory requirements and FAA approval. This results in strict time and budget constraints on pilot training (Orlady, 2010). Due to these time constraints, both flight instructors and trainee pilots have indicated several problems with the training. Little time is spent on the workings of the complex flight deck automation systems. Instead, pilot training is focused on passing check-rides by rehearsing procedures extensively. Moebus (2009) showed in a study commissioned by EASA that the theoretical questions for pilot exams can be rote-learned. They identified the risk of rote-learning by trainee pilots for traditional pilot training since EASA publishes the database of questions used for the theoretical pilot exams.

The Flight Deck Automation Working Group (2013) reported the following two findings from their analysis of pilot training:

- Current Training Time, Methods and Content; the current training programs are not sufficient to teach pilots the required knowledge, skills and judgement for effective use of the flight path management systems.
- Flight Instructor Training and Qualification; flight instructors lack the experience and line-operation familiarity to effectively train pilots for effective use of the flight path management systems.

Pilot training focuses heavily on learning procedures, leaving very little time to teach decision making and judgement. Due to this procedural approach, trainee pilots do not get the chance to make mistakes and learn from these mistakes. Since no unpredictability is used and only standard scenarios are simulated, trainee pilots are not trained for diagnosing problems. As a result, the analysis showed that pilots are not well equipped to handle non-routine situations. Nikolic and Sarter (2007) at Boeing performed a simulator experiment with 12 airline pilots flying three scenarios relying heavily on automation knowledge. The scenarios were designed in such a way that they would introduce automation disturbances which the pilots had to deal with by recovering to a safe state. The objective of this study was to determine how pilots diagnose and respond to automation problems. Even though all pilots were able to successfully finish the scenarios, difficulties with disturbance management were found as little diagnosis was shown due to knowledge gaps and time constraints. Instead, less effective and generic recovery strategies were used as pilots did not often switch to a lower automation level. The authors highlight the need for improvements in automation training in the area of error and disturbance management.

In their Automation Training Practitioners' Guide, written for the FAA, Lyall, Boehm-Davis, and Jentsch (2008) recommend proper systems knowledge of the automation systems in order to operate effectively in the flight deck. However, this knowledge seems to be lacking. Pilots have indicated they need more system knowledge to understand how these systems operate and their underlying logic (Orlady, 2010). In a study for Boeing,

[1] Code of Federal Regulations, title 14 Aeronautics and Space by the FAA, reviewed on March 30 2020

Holder (2012) showed that pilots feel they do not receive the information to operate the flight deck via training. Three quarters of the 800 pilots interviewed for this research indicated they experienced difficulties in understanding the FMS in their first six months after training. Training lacks in teaching pilots to operate the FMS, as an estimated mere 38% of FMS learning comes directly from training. Learning to operate the FMS occurs mostly on the line. This study showed that pilots want training to improve upon automation surprises and transitions between modes, two key areas identified as problems with flight deck automation. The Flight Deck Automation Working Group (2013) emphasises the need to include overall flight path management philosophy, workings of individual components and the integration of different systems into airline pilot training. Flight instructor Kiss (2019) has found that "candidates were not procedurally prepared because they had lost time trying to find system information during the systems study. Therefore, there was less time available to study procedures."

2.2.3. PREVIOUS RESEARCH ON PILOT TRAINING

Roper, Baxley, Swieringa, and Hubbs (2019) at NASA used the introduction of a new type of avionics as an opportunity to study pilot training. Interval Management, a concept used by ATC and flight crews to efficiently manage aircraft spacing, was added into the flight deck. As NASA required test pilots to validate this new avionics system in flight, an enhanced training program was required for the test pilots. NASA used a multi-tier approach with increasing levels of complexity in the training. This multi-tiered training consisted of three phases: computer based, classroom and simulator training. The Computer Based Training (CBT) consisted of reading material, videos and a touch-screen interface which pilot could use from home. This concept of CBT is not part of traditional pilot training. The results of this additional CBT were successful as they found that the classroom instructions were more constructive. Other factors mentioned to contribute to the success of this newly developed training were repetitive physical programming to learn the new system and progressive training with more realistic scenarios in the simulator. Although the addition of CBT was novel in this enhanced training, the traditional concept of emphasising the separate tasks on how to interface with the system were still applied, rather than focusing on the bigger picture. The concept of Computer Based Training was also studied specifically for the longitudinal modes of the autopilot by Mitchell (2000) and for the autopilot in general by Javaux and Sherry (1999). Both studies showed positive results, where immediate feedback, experiential learning and the use of part-task scenarios were mentioned as contributing effects to its success.

A study by Plat and Amalberti (2000) focused on training to deal with automation surprises. Two Line Oriented Flight Training (LOFT) sessions were conducted where the crews experienced several software bugs, without being instructed explicitly how to deal with these problems. A strategy commonly used by the participants when facing an automation failure was to cycle the automation to reset it. Other strategies included resetting the circuit breakers, searching for a procedure in the provided documentation and reverting to manual control by disengaging the autopilot. The authors mention that the pilots' situation awareness was acceptable, but that most of the pilots indicated a higher workload compared to traditional training. Besides, the pilots were mostly positive about this new training and mentioned that this training will be useful when facing automation problems on the line. In their conclusions, they recommend to train future crews on the risks of potential problems in the automation. Although all the aforementioned studies have proposed and tested alternatives for traditional pilot training, no single study included a control group to see if the new training program was actually an improvement over current pilot training.

The effectiveness of adding variability in pilot training on manual flying skills was shown by Huet et al. (2011). Pilots who received training in which flight and environment conditions varied were better able to follow a glide slope compared to pilots who received a training were conditions were fixed. A study by Landman et al. (2018) concluded that adding variability and unpredictability in pilot training has a positive effect on pilot performance in surprise situations. A simulator experiment was performed in which two groups of pilots were trained to deal with failures and events. One group received unpredictable and variable scenarios during the training, whereas the other group received highly predictable scenarios. Pilots who received the unpredictable and variable training showed better performance compared to the control group when faced with new surprise situations, although performance during training was worse compared to the control group.

Both studies focused on manual flying skills, and did not include any form of flight deck automation.

2.3. A THEORETICAL VIEW ON TRAINING AND PROBLEM SOLVING

If an automation surprise is a result of a failure in the automation systems, a pilot needs to take appropriate actions to solve this problem. While for some failures, such as an engine failure during take-off, specific procedures exist, this is generally not the case for failures within the automation systems. The current approach in training is to rote-learn specific procedures for a specific set of common or detrimental failures. In the past decades, research in the field of educational psychology has provided completely different approaches to teach skills and knowledge for all sorts of applications.

2.3.1. TEACHING APPROACHES FROM EDUCATIONAL PSYCHOLOGY

One teaching approach is the student-centered Problem-Based Learning (PBL) method. In this approach, the tutor is merely a guide since the student learns new information by self directed learning (Dochy, Segers, den Bossche, & Gijbels, 2003). The student has to solve problems to gain the required knowledge and learn problem-solving skills. The focus is not necessarily on finding a predefined solution to the problem, but rather allows for development of skills and knowledge throughout the process of solving the problem. Students construct their own knowledge in an environment that is representative for a real world scenario (Savery & Duffy, 1995). Although PBL could theoretically be used for many different kind of training purposes, studies on this topic tend to be limited to (elementary) class room examples. Discovery-Based Instruction is a pedagogy where not even a problem is provided but only a realistic environment to interact with. It is completely up to the student to find the information and or skills required to effectively manipulate the learning environment (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011). Discovery-Based Instruction is applied in different fields ranging from simulations to following manuals independently. However, the effect of providing minimal to no guidance is debated in literature (Kirschner, Sweller, & Clark, 2006; Mayer, 2004). Alfieri et al. (2011) showed, using a meta-analysis, that explicit instruction is indeed favourable over Discovery Based Instruction. Jong and Joolingen (1998) recommends providing additional support for the student with Discovery Based Instruction which can include direct access the domain knowledge and support for regulative learning process by asking questions.

Based on these concepts, the 5E instructional model was developed three decades ago (Bybee et al., 2006). This model was specifically developed to be used in teaching STEM subjects and has become popular in the science education community (Marshall, Horton, & Smart, 2009). The 5E model, consisting of the steps engagement, exploration, explanation, elaboration and evaluation, uses inquiry-based learning to teach students knowledge and skills in fundamental scientific concepts. Bybee et al. (2006) showed that already in 2006, over 200.000 lesson plans and more than 70.000 curriculums have been successfully created and executed based on this instructional model. In the first phase engagement, the student becomes committed to learning by combining prior knowledge and experiences to the new topic. In the exploration phase, the students explore the topic by predicting, designing, testing and asking questions. After the exploration phase, the teacher can use the experiences of the student to focus on specific concepts, processes or skills to provide the student with a profound understanding of the material. A deeper understanding is developed in the elaboration phase, where students are asked to apply their conceptual knowledge in new scenarios. Finally, in the evaluation phase, students should judge their own understanding and performance.

2.3.2. PROBLEM-SOLVING TRAINING

In order to tackle the above mentioned problems with pilot knowledge and specifically pilot training, it is crucial to understand the cognitive processes pilots use to comprehend the situation in the flight deck. This understanding can be modelled using frames. A frame is a explanatory structure such as a mental model which can vary from really superficial to very in-depth knowledge (Landman et al., 2017a). In order to improve the knowledge pilots possess of the automation systems, pilot training should focus on improving the quality of the pilots' frame or frames of the automation systems.

Problem solving can be seen as a cognitive process. According to Jonassen (2010), problems encountered in the real world have five characteristics: structuredness, context, complexity, dynamicity and domain specificity. Based on these characteristics, he identified eleven different types of problems. One type, commonly associated with problem solving in general, is troubleshooting. Troubleshooting requires both domain and system knowledge. Dealing with an automation surprise in the flight deck can be seen as a diagnosis-solution problem. Similar to troubleshooting, diagnosis-solution starts out by identifying the fault state. After the fault has been identified, a solution must be chosen from a solution space consisting of multiple solutions and solution options. The key to learn troubleshooting and diagnosis-solution is by solving authentic troubleshooting problems as practise, as Jonassen (2010) mentions in his handbook for designing problem-solving learning environments. Besides, he also emphasises just-in-time instruction. Teachers should not provide all knowledge beforehand, but should rather provide conceptual instruction. Conceptual instruction, as opposed to procedural instruction, focuses on domain knowledge (Hiebert & Lefevre, 2013). The effectiveness of providing conceptual instruction to teach problem solving was shown in a study performed by Fyfe, DeCaro, and Rittle-Johnson (2014).

By providing conceptual instruction, information is provided in a meaningful way. The educational theory that supports this is meaningful learning, where the information is completely understood by the learner in such a way that connections with other existing knowledge and information can be made (Ausubel, 2000). A training focusing on meaningful learning should be student-based. According to Jonassen and Strobel (2006), meaningful learning can occur when students are actively working on authentic tasks. Students should interact with their environment and analyse the effects of their actions. In this way, the students construct their own interpretations and information becomes meaningful. This is enhanced when the environment and tasks are authentic and realistic. Novak (2013) states that presented information and material must have meaning to the student. That requires the student to understand relevant concepts and prepositions of the environment. Besides, the student must choose to combine and link the new knowledge with existing ideas and step away from rote learning.

In order to bring structure to the problem-solving process, either domain-specific or general strategies can be trained. Five common troubleshooting strategies, according to Jonassen and Strobel (2006), are:

1. Trial and error: randomly check each part of the system
2. Exhaustive: check for all possible failures
3. Topographic: identify the failure by performing a series of checks
4. Split Half: split the system in half and check which part is still operating
5. Discrepancy Detection: detect mismatches between expected behaviour and the actual behaviour of the system

Schaafstal, Schraagen, and van Berl (2000) recommend the use of a meta-strategy when training for troubleshooting, as well as functional models, and underlying domain knowledge. These functional models can be presented in the form of a simulation used during the training. White and Frederiksen (1990) described such an environment in their research on intelligent learning. They simulated electric circuit behaviour based on qualitative models. With the premise that expertise is captured in mental models or frames, the simulation allowed students to learn problem-solving strategies. The trials showed that novices were able to learn to troubleshoot using this method. In a similar way, Brna, Ohlsson, and Pain (1993) showed that airmen trained with an avionics simulator showed model-based reasoning which they were able to apply in troubleshooting. Their participants were able to develop expertise in minimal training time due to the focus of generic skills in the job centered environment. Rosa et al. (2020) found that experience with problem solving in dynamic environments can help pilots making decisions in unforeseen situations. They emphasise the importance of dynamic decision making in pilot training, as creative problem solving is a crucial element to the safety of aviation.

The six cognitive processes involved in transfer of knowledge and skills are understanding, applying, analysing, evaluating and creating (Mayer, 2002). These processes can be used to construct a meaningful training, as opposed to rote learning, which only supports the retention of knowledge and skills. Lintern and Boot (2019) performed a meta study to analyse the transfer of Computer Based Training (CBT). Their analyses showed no transfer of cognitive skills from CBT and video games in particular. However, training specifically designed around operating vehicles like cars and aircraft show the opposite result. The focus on part-task training with a specific focus on skill elements in an environment that shares critical features to the operational system is what allows for this general transfer of cognitive skills seen in flight training.

3
Experiment Plan

In this chapter, the experiment that will be performed to answer the main research question is explained. After the experiment is introduced, based on the literature study, the hypotheses for the experiment are presented. These are followed with the Experiment Methodology. The experiment procedure, which traditionally is part of the methodology, is provided a separate section, to explain each part of this procedure in more detail. In order to verify some of the initial thoughts and to receive feedback on the experiment procedure, a preliminary experiment will be performed. The details of this preliminary experiment are explained in the last section of this chapter.

3.1. Experiment Introduction

The experiment for this research is developed based on the literature study. It was found that pilots have difficulty understanding the flight-path management systems (Flight Deck Automation Working Group, 2013), which can be traced back to a lack of declarative knowledge and little to no room to learn decision making and problem-solving skills during pilot training (Fletcher & Bisset, 2017). For the experiment, a new automation training will be developed focusing on problem-based learning. The participants will be exposed to several automation failures, as recommended by Fletcher and Bisset (2017) and Nickolic and Sarter (2007). Rather than simulating a complete flight, each scenario will focus on a specific phase of flight, following the success of part-task training in studies performed by Mitchell (2000) and Javaux and Sherry (1999). Practising with non-routine situations should result in more effective problem-solving skills and improved automation knowledge, management and handling skills (Fletcher & Bisset, 2017).

In the initial idea, participants would not receive any help from the instructor during the training, as they should independently solve the problems during the training. However, NASA cognitive psychologists dr. D. Billman and dr. R. J. Mumaw provided the crucial insight that these participants will require some scaffolding in order to make this training meaningful. Participants are not likely to learn problem-solving skills completely on their own as discovery based learning is not effective enough (Kirschner et al., 2006; Mayer, 2004). Providing the participants with a meta strategy as scaffolding will likely result in a better results of the training. Therefore, participants will receive a general problem-solving strategy which can be applied when facing an automation failure. During the training, the participants will be able to learn effective problem-solving skills by practising with this problem-solving strategy (White & Frederiksen, 1990).

Similar to experiments performed by Landman et al. (2018) and Huet et al. (2011), a control group will be included to compare the results to a baseline. This control group will consist of a group of pilots who will be trained using the same automation failures as mentioned above. Similar to current day pilot training, this group of participants will receive a specific procedure from the instructor for each failure. Thus, these participants will be trained using procedures and will not be required to develop independent problem-solving and decision-making skills during the training.

The participants who will receive the new training (from here on mentioned as the strategical group) and the participants from the control group (from here on mentioned as the procedural group) will fly several test scenarios after the training to compare their problem-solving skills. Neither group will receive any type of guidance during these test scenarios. The difference in training between the two groups is regarded as parallel manipulation.

3.2. HYPOTHESES

The following hypotheses have been set up for this experiment:

- Pilots who received the strategic training will be more effective in dealing with an automation surprise which they have not seen before during training.
- Pilots who received the procedural training will be more effective in dealing with an automation surprise they have practised during the training.
- The level of surprise which the pilots experience will not be significantly different between the two training programs.
- The strategic training will result in a higher perceived workload compared to the procedural training.

3.3. EXPERIMENT METHODOLOGY

3.3.1. PARTICIPANTS

Thirty participants will be invited for this experiment. The participants will be private pilots in possession of a Private Pilot License (PPL) and an Instrument Rating but have preferable no experience flying aircraft which are equipped with a glass cockpit. By making use of participants who are already in the possession of a pilots license and the Instrument Rating, concepts such as general flying skills and instrument flying do not need to be included in the training. Thus, the training can be solely focused on the new type of autopilot. If too few pilots who match these conditions are available to participate with this experiment, it can be decided to invite pilots with a PPL and IR who have glass cockpit experience. Though it is not required, pilots in possession of an Instrument Rating are likely to have experience with a glass cockpit. In that case, all pilots participating should have some glass cockpit experience.

3.3.2. TASKS AND INSTRUCTION

The task of the participants in the flight simulator test will be to follow a flight plan to a specific way-point in each scenario. Pilots will be briefed that the experiment will investigate glass cockpit conversion training and will be instructed to always use as much of the automation systems as possible. In this way, the participants will be incentivised to not simply disengage the autopilot as soon as they experience a partial failure. However, as some extreme cases may actually require to disengage the autopilot completely, no explicit instruction about keeping the autopilot engaged will be provided.

3.3.3. INDEPENDENT VARIABLES

The independent variable is the type of training which the two groups will receive. One group will receive the strategical training and the other group, which acts as a control group, will receive the procedural training. Both training programmes will focus on non-routine situations such as sensor or subsystem failures. The distinction between the two training programmes is the way the participants are trained to deal with an automation failure. In the procedural training, the student will be informed about the exact failure he or she will experience and will be guided throughout the situation by the instructor. As each scenario during the training is repeated, the pilots from the procedural group will receive the same procedure for the scenarios twice. In the strategical training, the student is initially only told he or she will experience some sort of failure, without specifying the type of failure. The student should actively interact with the system and try to figure out the problem on their own by making use of the problem-solving strategy.

Based on the literature of training problem solving and expertise in the domain of flight deck automation, the following problem-solving strategy is created:

1. Notice autopilot behaviour is off nominal
2. Identify which sensor or system is faulty
3. Identify implications on autopilot
4. Switch to alternative information source if possible, or
5. Switch to lower level automation

This strategy is based on discrepancy detection as pilots are tasked in the first step to notice that the current autopilot behaviour is off nominal. The first three steps of the strategy are related to problem diagnosis, whereas the remaining two steps are solution-finding tasks. It is only after the scenario that the students of the strategical group will receive the information about the exact type of failure and the procedure that could have been used to solve it. Hereafter, the student will get the chance to practise this procedure during the second time the scenario will be trained. The students in the procedural group will directly receive the solutions and thus will not experience the process of diagnosing the situation him/herself and searching for the right solution. However, both groups will experience each training scenario twice. This allows the strategical group to at least practise once with the correct procedure, and allows the procedural group to have sufficient practise to become familiar with the procedures.

3.3.4. CONTROL VARIABLES

The control variables in the experiment will be:

- *Aircraft type;* the exact same aircraft, configuration and avionics layout will be used for both groups.
- *Instruction;* both groups will receive the same instruction in the form of ground school to teach the basics of the autopilot operations.
- *Normal operations training;* the normal operations training for the autopilot will be the same for both groups. To minimise deviations, this training will be partly scripted.
- *Training time in the simulator;* both groups will receive a similar time in the flight simulator for training. Although this will vary slightly for each participant, effort will be made to assure that every participant will receive similar training time. The total training time spend in the simulator will be recorded.
- *Scenarios in the training;* the scenarios in the training will be the same for both groups. This includes geographic location, weather settings, failure type and failure timing. The order of the training scenarios will also be the same for both groups. Both groups will have access to the same quick reference card in the simulator which covers the most important topics of the instruction.
- *Scenarios in the test;* the scenarios in the test will be the same for both groups. This includes geographic location, weather settings, failure type and failure timing. The order of the test scenarios will also be the same for both groups. Both groups will have access to the same quick reference card in the simulator which covers the most important topics of the instruction.
- *Briefing and debriefing;* both groups will receive the same briefing and debriefing for the test.

3.3.5. DEPENDENT MEASURES

The following dependent measures will be used:

- *Time to notice failure;* this is the time in seconds between the moment when the failure has been activated by the experimenter and the first response of the participant to the failure. This first response will be identified either by the first interaction with the autopilot or by the first verbal notice of the participant that something is wrong, provided that the participants are instructed to think out loud. As the latter can disrupt the workflow of the participants, it is still to be determined whether the pilots are instructed to think out loud or not.
- *Time to take appropriate action;* this is the time in seconds that the participant is diagnosing the failure and finding the solution, and will be measured from the first interaction with the autopilot until the participant has selected the mode in which he or she will continue to fly the scenario.
- *Selection of correct mode;* each scenario has a preferred sequence of actions to take to successfully deal with the failure. The autopilot interactions of the participants will be compared to the preferred actions to see if they were able to deal with the failure in an efficient way.
- *Deviation of flight plan;* as the participants are tasked with flying a specific flight plan, a specific way-point in the flight plan will be chosen as a performance measure. The lateral and vertical deviation at this way-point will be an indication of how efficiently the participant was able to deal with the failure.

- *Workload;* the workload will be measured using a workload scale in an interview after the test.
- *Written exam score;* as the pilots take a multiple choice exam, a score (between 0% and 100%) will be rewarded to each pilot's theoretical knowledge of the autopilot and the subsystems.

3.4. PROCEDURE

The procedure for the experiment is visualised in Figure 3.1. Each section of the procedure is explained in more detail below. The experiment will only start when the Informed Consent procedure has been completed. After the experiment, the participants will be informed about the purpose of the experiment. In due time, permission will be sought with the Human Research Ethics Committee of the TU Delft for this experiment with pilot participants. The entire experiment is expected to take 4 hours per participant.

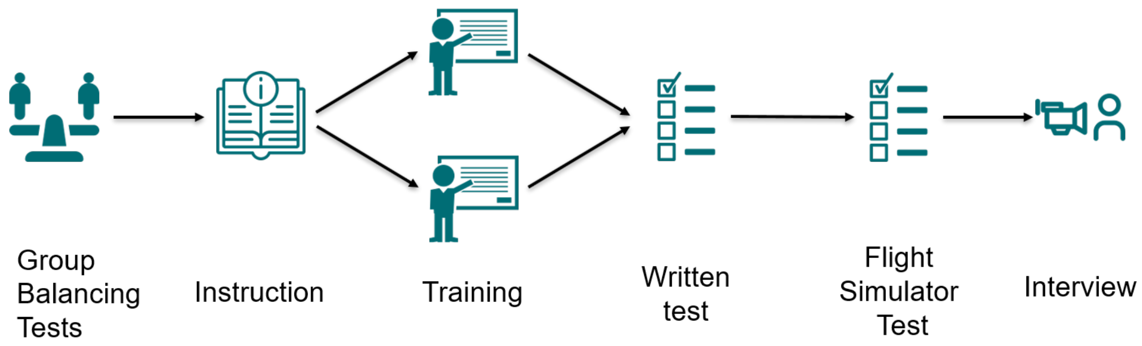


Figure 3.1: Experiment procedure

3.4.1. GROUP BALANCING

A small batch of tests before the experiment starts will determine in which group the participant will be placed. These tests will at least include a fluid intelligence test and a stress resistance test. For the fluid intelligence test, a standardised laboratory test will be used such as the commonly used Raven's Progressive Matrices (Raven & Raven, 2003). The stress resistance will be tested by the use of a small scenario flown in the simulator which includes an unrelated event such as an aerodynamic stall, comparable to the scenarios used by Landman et al. (2018). Based on the results of these tests and the flying experience of the participant, the participant will be placed in either the strategical group or the procedural group. In this way, individual differences between the participants will be balanced out between the two groups such that the effect that these factors may have on the results will be minimised.

3.4.2. INSTRUCTION

Both groups will receive an instruction for the new type of autopilot before the flight simulator training starts, which will be comparable to ground school in current day pilot training. This instruction will take place outside of the flight simulator using either a deck of slides or an instruction manual. The instruction will include the basic operation of the autopilot, the different modi and the integration of the G1000 avionics. This instruction will be based on the Quick Reference Guide by Garmin[1]. During this instruction, the problem-solving strategy will be explained only to the strategical group. The final part of the instruction will be a short familiarisation flight in the flight simulator. Participants shall take-off and fly a traffic pattern whilst they have to opportunity to get used to the aircraft and the new type of avionics.

3.4.3. FULL FLIGHT SIMULATOR TRAINING

After the instruction, each group will receive their respective flight simulator training program; either the strategical training or the procedural training. The focus of both training programmes will be on non-routine situation by introducing several partial automation failures. The procedural group will receive the correct procedure to handle each failure and will be guided by the instructor. The strategical group will receive a

[1] `https://static.garmincdn.com/pumac/G1000:CessnaNavIII_CockpitReferenceGuide_SystemSoftwareVersion0563.00orlater_.pdf` (Visited on 17 May 2020)

general problem-solving strategy which they should apply in each training scenario. They will not be guided by the instructor.

3.4.4. WRITTEN EXAM

Before the participants will take part in the full flight simulator test, they will be asked to take a written exam in the form of multiple choice questions. These questions will focus on the theoretical knowledge of the new automation system. These multiple choice questions will be inspired by the questions from Garmin's Instructor's Reference [2].

3.4.5. FULL FLIGHT SIMULATOR TEST

Both groups will fly the same five to ten scenarios which will each include a failure that triggers an automation surprise, but they receive no information about the failure they will experience. By using failures, all pilots will experience an automation surprise. Pilots are instructed to complete the given mission using as much of the available automation as possible. The first scenario includes a GPS failure during a GPS approach. The correct response is to switch to a different navigation source such as VOR. The second scenario consists of a servo failure of the Automatic Flight Control System which restricts a change in aileron deflection with the autopilot engaged. The correct action would be to turn off the autopilot in this case, whilst leaving the Flight Director on. Pilots who are trained with a problem-solving strategy are expected to be able to act appropriately in these scenarios since they have practised diagnosing and solving similar problems before in the training. Since the procedural group has only learned procedures for some specific situations, they are expected to be unable to handle novel scenarios and will likely fail to take appropriate action of switching to a different information source or to a lower level of automation. Other scenarios are still being designed at the time of writing.

3.4.6. INTERVIEW

After the test, the pilots will be receive a questionnaire. They will be asked to rate the mental workload of the training and test using the Rating Scale Mental Effort (Dijkstra & van Doorn, 1985). Other subjective measures such as enjoyment of the training and level of surprise during the test scenarios will also be included in the questionnaire. The rating scales and exact questions that will be included are still to be determined. Finally, the pilots will be asked about their thinking and reasoning during the automation surprises in order to understand and compare how pilots from the different groups diagnosed the automation problems they have faced. As certain subjective measures such as level of surprise may pass off with time, it is still to be determined if these questions will be asked immediately after each test scenario or all together after the test.

3.5. PRELIMINARY EXPERIMENT

Two weeks before the experiment is planned, a preliminary experiment will be performed. This preliminary experiment will be based on the actual experiment but reduced in overall complexity by scaling down the actual experiment. A small number of participants (two or three) will be invited to this experiment. This experiment will only include the instruction, the training and the flight simulator test. During the training and test, only a couple of scenarios will be included. Therefore, the preliminary experiment is expected to take 1 to 2 hours per participant. The preliminary experiment has the following objectives:

- Receive feedback from the participants on the instruction and the briefing;
- Receive tips from the pilots and experts which will join this preliminary experiment on how to improve the actual experiment;
- A practise run to fine tune the procedure and remove any inconsistencies, and
- A quick verification that the pilots are able to learn the autopilot operations in the little time available for training.

[2] http://static.garmin.com/pumac/G1000:Non-AirframeSpecific_PilotsTrainingGuide_InstructorsReference-06_.pdf (Visited on17 May 2020)

4
EXPERIMENT SCENARIOS DESIGN AND IMPLEMENTATION

In this chapter, the approach taken to design the scenarios for the test and the training is explained. The first section covers the requirements that have been set up to aid the scenario design. This is followed by a section in which the options for scenarios are explored. Finally, the implementation of these scenarios in DUECA is explained.

4.1. REQUIREMENTS FOR SCENARIOS

It is important that the scenarios used in the training and the test are well designed. The training should allow the pilots of both groups to become familiar with the new type of autopilot while having a clear distinction in the way they approach an automation related problem. Besides, the relation between the scenarios used in the test and the training should be in such a way that they allow for transfer of the training task. That means that the strategical group should be able to apply the problem-solving strategy learned during the training, and that the procedural group should be able to apply the procedures they have learned during the training. Thus, both novel scenarios and scenarios already practised during the training should be present in the test.

In order to design training and test scenarios properly, such that the experiment will have clear results and little to no ambiguity, the following requirements have been set up to aid the scenario design process:

SCE-1 Each scenario shall trigger an automation surprise with all participants.
SCE-2 The problem in each shall not be too straightforward to solve.
SCE-3 The scenario shall allow for problem diagnosing.
SCE-4 Each scenario shall be distinctive and unique from one another.
SCE-5 Each scenario shall require some interaction with the autopilot.
SCE-6 Each scenario shall be realistic.
SCE-7 The training scenarios shall prepare pilots for the scenarios in the test.
SCE-8 The test scenarios shall induce a different reaction between the two groups.

4.2. SCENARIO DESIGN

Table 4.1 presents an overview of the dependence and interactions between the different subsystems, sensors and flight director modi of the G1000 Avionics and the GFC700.

4.2.1. OPTIONS FOR AUTOMATION FAILURES

In order to trigger an automation surprise with all participants, a partial automation failure can be introduced. Most automation surprises experienced by pilots on the line are the result of a working automation system in combination with insufficient pilot knowledge. However, this effect cannot be guaranteed to take place with all participants. Therefore, the following three types of failures can be used in the training and the test:

- Sensor failures; an overview of possible sensor failures that affect the behaviour of the autopilot is provided in Table 4.2.
- Subsystem failures; an overview of the potential subsystems of the G1000 that can fail with the corresponding result is provided in Table 4.3.

- Software bugs; these can include incorrect automatic switching between modes, a mismatch between pilot inputs such as selected altitude and the actual reference value for the PID controllers or wrong information being displayed on the PFD or the MFD.

Table 4.1: Relation between flight director modi and aircraft systems

Mode	Gauges			Sensors/Receivers		Subsystems		
VNV	altimeter	VSI	-	Static port	-	ADC	AHRS	-
ALT	altimeter	VSI	-	Static port	-	ADC	AHRS	-
VS	VSI	altimeter	AI	Static port	-	ADC	AHRS	-
FLC	VSI	altimeter	AI	Pitot tube	Static port	ADC	AHRS	-
PIT	AI	altimeter	VSI	-	-	AHRS	-	-
NAV (GPS)	CDI	map	HSI	GPS receiver		GIA	GMU	AHRS
NAV (VOR)	CDI	map	HSI	VOR radio	VOR receiver	GIA	GMU	AHRS
HDG	HSI	-	-	-	-	GMU	AHRS	-
ROL	AI	TC	HSI	-	-	AHRS	-	-

The subsystems of the G1000 avionics systems are summarised in Appendix B. When one of these subsystems fails, a clear marking on the PFD is presented, as can be seen in Figure 4.1 (Retrieved from the Integrated Flight Deck Guide[1] by Garmin). The following subsystem names are used in table 4.3: GDC74 ADC (Air Data Computer), GMU44 (tri-axial magnetometer), GIA (Integrated Avionics Unit) and GEA71 (Engine Indication System). Due to these clear indications on the PFD, subsystem failures will be relatively easy to diagnose. However, since they have a detrimental effect on the autopilot behaviour and therefore require dynamic problem-solving skills to cope with, these type of failures are still likely to trigger an automation surprise and can thus be included as scenarios for the training and the test. Individual sensor failures will generally not result in a clear warning message and are therefore harder to diagnose. The effect of a sensor failure is likely to be not as severe as a complete subsystem failure, but can still have a major effect on the autopilot behaviour.

Table 4.2: Sensor failures with corresponding effect on autopilot behaviour

Failure	Result	Affected a/p mode	Appropriate Action
Static port	Incorrect climb/dive	VS	Switch to FLC/PIT
Pitot tube	Incorrect climb/dive	FLC	Switch to VS
HSI	Not following heading bug	HDG	Switch to ROLL
Gyro	Fast wing rock	HDG	AP off
Control system	Uncommended climb/dive	ALT	AP off
Static port	Not holding altitude	ALT	Switch to PIT
GPS failure	Not following GPS, incorrect VPATH	VNAV	Switch to ALT and HDG
Radio	Not following CRS pointer	NAV	Switch to HDG
GS antenna	Not capturing glideslope	APR	Switch to GPS approach
ADC	No GPS approach possible	APR	Switch to ILS approach

4.2.2. TRAINING AND TEST SCENARIO DESIGN

As the participants of the experiment will preferable have no experience in using the G1000 in combination with the GFC700 autopilot, they need to learn the basic operation during the training. If only pilots with experience in the autopilot and avionics can participate (due to difficulties in finding participants who match these criteria), a basic instruction will still be given to make sure all participants have at least the same basis. Therefore, the training will be split up into three phases: basic lateral autopilot training (covering HDG mode), basic vertical autopilot training (covering VS, FLC and ALT mode) and advanced autopilot training (covering NAV and VNAV mode). Each phase will start with a steady flight condition in which the participants are able to interact with the respective autopilot modes with instructor guidance. The second part of each phase is the non-routine training. In this part, the participants will experience some automation failures related to the modes which have just been introduced. This again will start from a steady flight in which, after

[1]https://static.garmincdn.com/pumac/190-00498-07_0A_Web.pdf (Visited on 17 May 2020)

some time, an automation failure is introduced by the experimenter. Depending on their training program, the pilots will be asked to either apply the general problem-solving strategy or the procedure specific to that type of failure during this part of the training. Each non-routine scenario is supposed to take several minutes as the only objective is to deal with that specific failure. The training will include around five partial automation failures.

For the test, more elaborate scenarios will be created. Pilots will receive a flight plan for each scenario. The scenarios will start with the appropriate autopilot mode engaged. As the scenarios in the test will each mimic a phase of a normal flight, these scenarios are expected to take between five and ten minutes. At a specific time or point along the flight plan, the failure will be introduced. The participant should deal with these failures in such a way that one is still able to finish the mission of following the flight path. The test shall include five to ten scenarios.

Table 4.3: G1000 subsystem failures and corresponding effect on autopilot behaviour

Failure	Effect	Affected a/p mode	Appropriate Action
PFD	GIA 1 and GFC700 a/p lost	All	MFD in reversionary mode
MFD	GIA 2 lost	None	PFD in reversionary mode
ADC	No altitude, vertical speed and airspeed, vertical a/p lost	All vertical modes	Use standby gauges
AHRS	GFC700 a/p and HSI lost	All	Use standby gauges
GMU44	Bearing pointer lost, HSI becomes CDI, lateral a/p lost	All lateral modes	Use GPS and Wet Compass
GIA (partial)	No effect, replaced by the other unit	None	-
GIA (both)	GFC700 a/p lost	All	-
Audio-panel	No use op a/p allowed	All	-
GEA71	All engine/airframe data lost	None	-
Transponder	Transponder lost	None	-

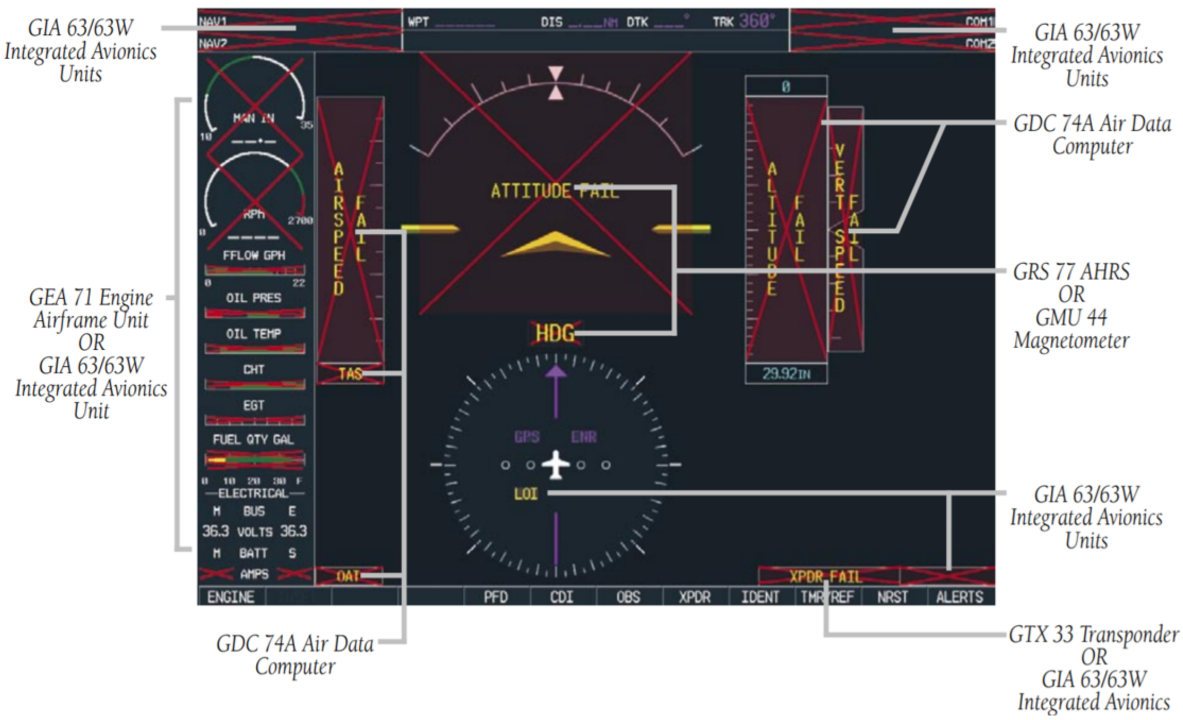


Figure 4.1: System Failure Annunciations

4.2.3. Rationale and Expected Pilot Behaviour

As the pilots in the strategical group are trained using a general problem-solving strategy for automation failures, it is expected that they will be able to more effectively deal with automation surprises in the test. For any new scenario they will face, they will be able to apply the trained strategy and find a solution. Pilots which have learned specific procedures for the failures they have experienced during training will likely not be able to deal with new automation failures. As they have only applied procedures during the training, they will quickly have more problems with diagnosing novel issues and taking the appropriate action when facing novel automation failures. These pilots may be trying to apply one of the procedures from the training if they experience an automation failure with comparable effects on the autopilot. In this way, they will apply the wrong procedure and will likely not be able to complete the given mission in the test.

For the scenarios in the test that are repeated training scenarios, it is likely that the procedural group performs better. As they have been trained thoroughly to deal with that specific failure, they will be able to quickly execute the trained procedure. Since the strategical group has not been trained so thoroughly in dealing with that specific scenario, they will not be able to immediately apply the correct procedure but rather start applying the problem-solving strategy. In this way, they will probably still be able to take the appropriate action, but will not be as effective in dealing with the automation surprise as the procedural group. This expected behaviour of the pilots is the foundation of the hypotheses which were explained in Chapter 3.

4.3. Scenario Implementation in DUECA

The Experiment Control Interface (ECI), which is part of the simulation, will be extended to include options for the specific automation failures of the scenarios. For each specific failure, the resulting effect of the autopilot will be determined and modelled. These failures can be triggered based on timing, pilot action or once a specific point along the flight plan is reached. The ECI module will send an event update to the Autopilot module. Depending on the type of failure, the autopilot behaviour will be modified by overwriting the state vector received from the PA34 dynamics module or by interfering with the output of the autopilot in some other way.

These failure events will also be read by the PFD and the MFD modules to show the required information on the displays that result from the failure. The dynamics module will remain unaffected by these automation failures.

5
Implementation of the Piper Seneca Model with Autopilot and Avionics

In this chapter, the implementation of the flight simulation is explained. The first section covers the chosen aircraft model and the general structure of the flight simulation. This is followed by a section focusing on the autopilot and a section focusing on the avionics. Finally, the next steps in the development of the simulation will be discussed.

5.1. The Piper Seneca Model

For this project, an existing flight simulation of the Piper Seneca II will be used. Figure 5.1 shows a picture of the aircraft [1]. The selection of this aircraft type and flight simulation were based upon the following arguments:

1. A rather complete non-linear simulation including several failure modes of this aircraft model is available to use from previous projects;
2. Several relateable experiments within the Startle & Surprise research group have used the same aircraft type and simulation;
3. This aircraft can be equipped with modern-day avionics and automated systems similar to those found in airliners, yet it does not require an ATPL or type rating to operate it.

Figure 5.1: The Piper Seneca V

[1] Retrieved from `https://www.piper.com/model/seneca/` (Visited on 17 May 2020)

5.1.1. Piper Seneca Aircraft

The Piper Seneca is a twin prop aircraft developed in the early 1970s by Piper with over 4700 aircraft build as of 2019. The latest variant has been the Piper Seneca V which has been in production since 1997 and has a unit price of $1.030.000. It is a single pilot aircraft which seats five to six passengers or a useful load of 604 kg. This aircraft is mainly popular with air charter companies, small regional carriers and with private individuals. The maximum range is 828 nm and the maximum cruise speed is 200 kts.

This aircraft used to be equipped with the iconic Piper Legacy Autopilot. This is an electric two-axis rate-based autopilot consisting of a command console, a lateral guidance systems and servos. Modern versions of this aircraft are by default equipped with the Garmin G1000 glass cockpit interface which includes the Garmin GFC700 autopilot. More on the workings of these avionics is explained in the sections below.

5.1.2. Simulation

The simulation of this aircraft is an extension to the project *SenecaTraining* by van Oorschot (2017) which was based on the original *Asym1* project by Koolstra (2017). For this Asym1 project, a dynamic model of the Piper Seneca was created using lookup tables in Simulink. This Simulink model was converted to C++ code using the Real-Time Workshop. This model includes options for engine failures, mass shifts and control surface limitations. Van Oorschot extended this simulation by adding the Experiment Control Interface, which is a Graphical User Interface used to select predefined experiment scenarios. The project used for this experiment is saved on the server as *SenecaAutomationTraining*

The simulation is created in the DUECA Framework, which is middleware for communication and timing developed by the TU Delft (van Paassen, Stroosma, & Delatour, 2000). DUECA is a C++ library for Linux/GNU and follows a data driven and modular approach. The modules used in the simulation are presented with a description in Table 5.1. The layout of the project with the relation between the modules and the channels is presented in two tables in Appendix A.

Table 5.1: DUECA modules in the main project directory

Module	**Type**	**Description**
Autopilot	Calculation	Autopilot and Flight Director computing control surface deflections based on autopilot settings
CitationLogger	Output	Data logger
CitationIncoSelector	Input	GUI to select the initial conditions of the simulation
CvCalculation	Calculation	Module to compute Cv
ECI	Input	GUI to select one of the preset scenarios
Enginesound	Output	Plays the sounds associated to the engines and propellers
FCSAdapter	Calculation	Couples the flight controls input to the flight dynamics module PA34
G1000	Output	Primary Flight Display visualisation
GFC700GUI	Input	GUI to interact with the autopilot
Malfunctions	Input	GUI to select malfunctions
MFD	Output	Multi Function Display visualisation
PA34	Calculation	Dynamics module computing the new state vector based on the previous state and the (auto-) pilot input
WAVPlayer	Output	Required module for playing engine sounds
WeatherProxy	Calculation	Computes weather data

5.2. GFC700 Autopilot Simulation

The GFC700 is an automatic flight control system, integrated in the G1000, which provides a flight director, digital two-axis autopilot and a yaw damper. The GFC700 and G1000 are based on a build-in Attitude Reference and Heading Reference (AHRS) system, an Air Data Computer (ADC), a magnetometer and two Integrated Avionics Units (GIA). The pitch and roll modes of the Flight Director are presented respectively in Table 5.2 and 5.3. The italicised modes are currently not supported in the simulation.

Table 5.2: Pitch Modes Flight Director

Pitch Mode	**Control**	**Annunciation**
Pitch Hold	default	PIT
Selected Altitude Capture	-	ALTS
Altitude Hold	ALT Key	ALT
Vertical Speed	VS Key	VS
Flight Level Change	FLC Key	FLC
Vertical Path Tracking	VNV Key	VPTH
VNAV Target Altitude Capture	-	ALTV
Glidepath	*APR Key*	*GP*
Glideslope	*APR Key*	*GS*
Go Around	*GA Switch*	*GA*

Table 5.3: Roll Modes Flight Director

Roll Mode	**Control**	**Annunciation**
Roll Hold	default	ROL
Heading Select	HDG key	HDG
Navigation (GPS)	NAV Key	GPS
Navigation (VOR)	NAV Key	VOR
Navigation (LOC)	*NAV Key*	*LOC*
Backcourse	*BC Key*	*BC*
Approach (GPS)	*APR Key*	*GPSa*
Approach (VOR)	*APR Key*	*VAPP*
Approach (ILS)	*APR Key*	*LOC*
Go Around	*GA Switch*	*GA*

To assure that the autopilot will not unintentionally perform actions belonging to a wrong mode, a Finite State Machine (FSM) is incorporated for both the pitch and the roll axis. These two FSMs are set up like a Flight Director (FD), as they compute the respective pitch angle and roll angle based on the input settings and the active autopilot mode. These reference pitch and roll angles are then used to determine the corresponding aileron and elevator commands if the autopilot is engaged. The aileron and elevator commands will be sent to the PA34 dynamics module where they overwrite the pilot input.

Switching between states is performed using a State Transition Matrix (Randhawa, Jeppu, Nayak, & Murthy, 2010) for both the lateral and vertical flight director, presented in Table 5.4 and 5.5. The STMs contain the simulated flight director states as rows and their respective transitions based on events. Events are the buttons which the pilot can press on the autopilot module, and are presented as columns in the STMs. For example, consider the current vertical flight director mode to be altitude hold (ALT HOLD). Now, the pilot decides to press the ALT key. As the current vertical flight director state equals 3, the corresponding transition to the event of the ALT key would be 1, thus the flight director switches to state 1, which is pitch hold (PIT HOLD).

Table 5.4: State transition matrix for the vertical flight director

STATE	**AP key**	**VS key**	**ALT key**	**VNAV key**	**FLC key**
OFF (0)	1	2	3	4	5
PIT HOLD (1)	0	2	3	4	5
VS HOLD (2)	0	1	3	4	5
ALT HOLD (3)	0	2	1	4	5
VPATH (4)	0	2	3	1	5
FLC (5)	0	2	3	4	1

Table 5.5: State transition matrix for the lateral flight director

STATE	AP	HDG	NAV	VNAV key
OFF (0)	1	2	3	4
ROL HOLD (1)	0	2	3	4
HDG HOLD (2)	0	1	3	4
NAV (3)	0	2	1	4
VPATH (4)	0	2	3	1

The Flight Director and Autopilot consist of cascaded PID controllers in the form of:

$$C = K_p\left(e - T_d\frac{dV}{dt} + \frac{1}{T_i}\int e(t)dt\right) \tag{5.1}$$

Here, e is the error between the reference value and the actual value, K_p is the proportional gain, T_d is the differential gain and T_i is the integral gain. These gains were tuned in the Simulink model and manually converted to the C++ code. The control gains used in the current version of the flight director and autopilot (May 2020) are presented in Appendix C. Servo dynamics are simulated as a first order lag with a time constant of 0.1 (for both the aileron and elevator servo). Euler integration is used for the integral control loops and for the servo lag terms. The step response for a step in pitch angle and roll angle are presented in Figure 5.2. These plots were generated from the data directly taken from the DUECA project simulation. Step responses for the other control loops are presented in Appendix C. The automatic transition between flight director states was matched with the G1000 Cockpit Reference Guide by Garmin[2]. Besides, the behaviour will be verified by one of the pilots from the section Control & Simulation.

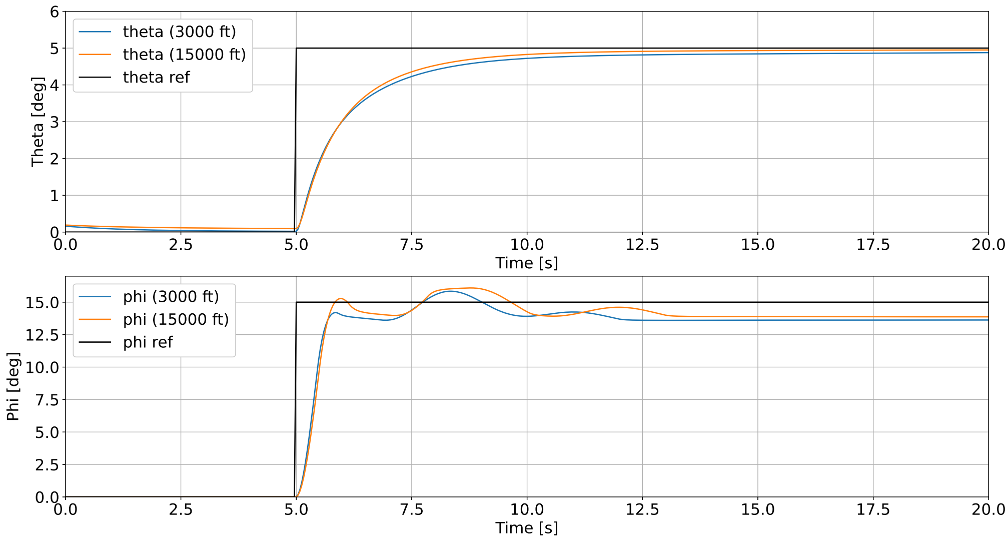


Figure 5.2: Step response for θ_{ref} and ϕ_{ref}

For the pilot interface, a Graphical User Interface (GUI) has been added to the simulation. This GUI includes all the push buttons associated to the GFC700 and spin buttons for the reference altitude, reference heading and course. These spin buttons will likely be replaced by two simple buttons to increase and decrease the reference value. Another option is to use the rotary encoders on the Mode Control Panel (MCP) which is installed in the Simona Research Simulator. The participants of the experiment will be able to interact with the GUI using a touchscreen interface. The GFC700GUI module in simulation is responsible for recording each button press. This module sends this button press as an event to the Autopilot module which determines the resulting action of that button press. The current version of the GUI is presented in Figure 5.3.

[2]`https://static.garmincdn.com/pumac/190-00384-12_0A_Web.pdf` (Visited on 12 June 2020)

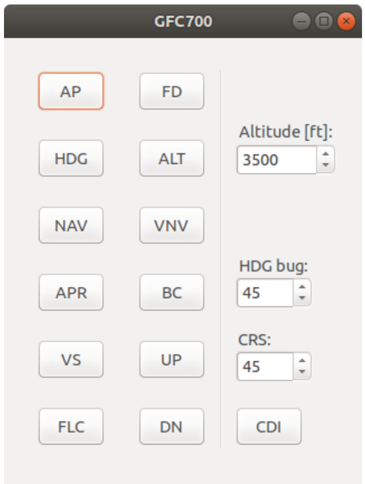


Figure 5.3: GUI for the GFC700 autopilot

5.3. GARMIN G1000 AVIONICS SIMULATION

The Garmin G1000 is an avionics system comprising of a Primary Flight Display (PFD) and a Multi Function Display (MFD). The PFD is similar to those found in airliners as it includes an altitude tape, an indicated airspeed tape, a vertical speed indicator, an attitude indicator, a horizontal situational indicator, flight director and a flight mode annunciator. Besides, is shows the navigation and communication radios.

For this project, a G1000 module developed by Comans (2017) will be used and extended. This module is extended to include a flight director, correct autopilot annunciations, a fully operational horizontal situational indicator and some several small additions to make it resemble the actual G1000 even more such as airspeed and altitude trend vectors and correct reference bugs. Besides, the entire MFD is added as the original module did not have a MFD. The engine instruments were changed to resemble those found in a twin prop and moved from the PFD to the MFD. The simulated PFD is visualised in Figure 5.4.

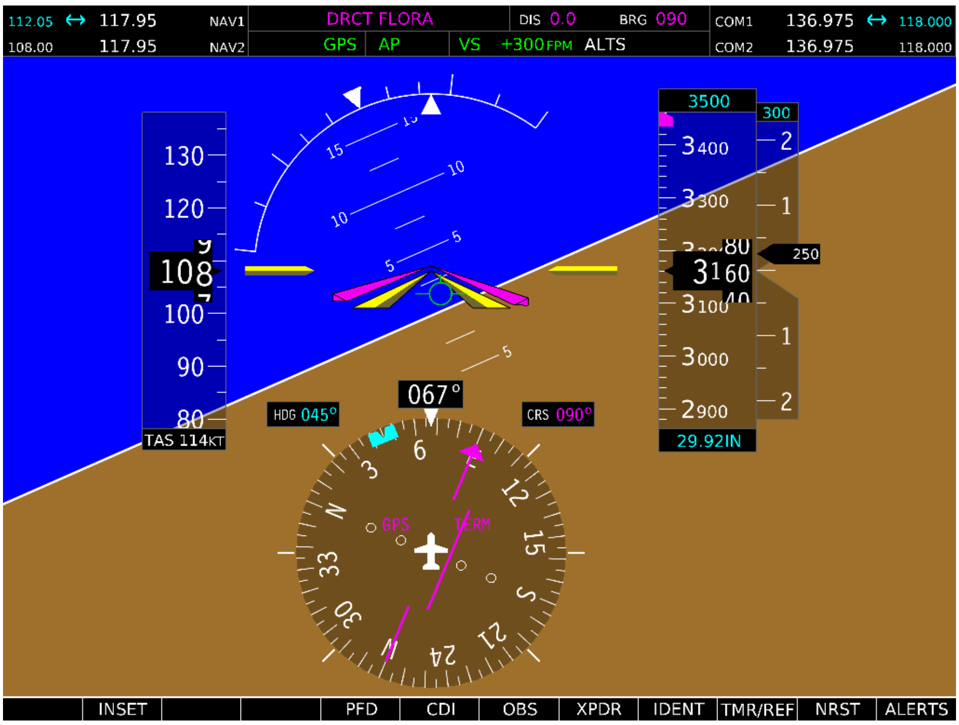


Figure 5.4: The simulated Primary Flight Display of the G1000

5.4. NEXT STEPS

The following features will be added to the flight simulation in the weeks after the due date for this report:

1. Flight Plan functionality, such that the NAV mode can follow a flight plan consisting of multiple legs.
2. VNAV mode capability, based on the Flight Plan, target vertical speeds will be computed and send to the flight director.
3. Autopilot/sensor failure modes, which will be controlled using the existing structure for the other types of failures.

The results of these next steps will be documented in appendices of the final report of this thesis.

6
Conclusion

Human factors continue to play an important role in the flight deck. Unexpected events which are difficult to explain may trigger an emotional and cognitive reaction in pilots, known as a surprise response. This preliminary report proposes a research to find out if pilot training can be improved by teaching problem-solving skills in the form of a strategy that can be applied when facing with an automation surprise. The main research question of this research is defined as:

By teaching a problem-solving strategy, will pilots become more effective in dealing with an automation surprise?

This research question was split into four sub-questions, of which the first three have been answered with the literature study. This literature study was performed to determine the state-of-the art in flight deck automation training and to identify problems in automation usage. The sub-questions are elaborated upon below.

1. *What are the current training practices for flight deck automation systems and how effective are they in preparing pilots to operate the flight deck automation systems?*

To answer this first sub-question, current day pilot training and its effectiveness was investigated. Trainee pilots traditionally follow ab initio training followed by an aircraft specific type rating. The ab initio training phase consists of ground school and flight training up to twin propeller engine aircraft. Students learn the main principles of flight and are trained in manual flying skills and instrument flying for general aviation aircraft. In the type rating, the student will be trained for the one specific airliner he or she will be flying commercially during their pilot career. This type rating consists of three weeks of ground school followed by full flight simulator training. The focus of the ground school is on descriptive system knowledge while the full flight simulator training focuses on procedural training. Standard operating procedures as well as emergency procedures are practised until the student passes the check ride. Reports by the FAA and the CAA have indicated that this training is not sufficient to teach pilots the knowledge and skills required to effectively operate the flight path management systems. There is too much focus on procedures and little to no training time is spent on training decision making and problem solving. Both pilots and flight instructors have substantiated these claims as they indicate that pilots have knowledge deficits with respect to the flight deck automation systems.

2. *What is the effect of surprise on pilot behaviour in combination with the automation systems?*

For this second sub-question, the literature regarding Startle & Surprise in the research group at the Faculty of Aerospace Engineering at the Delft University of Technology has been extended to include the use of automation. Pilots experience an automation surprise on average three times a year, though it should be noted that most of these remain without consequence. An automation surprise can be triggered if the system fails to take an expected action, if the system carries out an unexpected action or by a (partial) automation failure. Most automation surprises occur with high levels of automation active such as Vertical Navigation and Lateral Navigation. Confusion of the active auto-flight mode is often mentioned by pilots and is a direct result of the knowledge deficits mentioned above. It is known that pilots show difficulty dealing with non-routine situations, especially if no procedure is available. Pilots have indicated that including automation surprises in pilot training would be a major improvement. This is also the recommendation of scientific research focusing on automation surprises and pilot training.

3. *How can effective problem solving be trained?*

As problem solving can be seen as a cognitive process, literature on training from the field of educational psychology was used to answer this third sub-question. Dealing with an automation surprise in the flight deck can be seen as a diagnosis-solution problem. Similar to troubleshooting, diagnosis-solution starts out by identifying the fault state. After the fault has been identified, a solution must be chosen from a solution space consisting of multiple solutions and solution options. The key to learn troubleshooting and diagnosis-solution is by solving authentic troubleshooting problems as practise. This can be trained in a learning environment such as a simulation. Based on the problem-solving strategy of discrepancy detection, the following problem-solving strategy to deal with automation failures has been developed:

1. Notice autopilot behaviour is off nominal
2. Identify which sensor or system is faulty
3. Identify implications on autopilot
4. Switch to alternative information source if possible, or
5. Switch to lower level automation

This is a student-centered, Problem-Based Learning approach in which the strategy acts as additional support to achieve the required knowledge and learn problem-solving skills.

4. *How is the problem-solving capacity of a pilot affected by changing the training method for flight deck automation?*

In order to answer this fourth research question, an experiment investigating the effect of teaching this problem-solving strategy during pilot training will be performed. A total of 30 pilot participants who are in possession of a private pilot license and an instrument rating will be invited to take part in this experiment. These pilots will be trained for a type of autopilot preferable unknown to them, namely the Garmin GFC700 as part of the Garmin G1000 avionics. The participants will be equally divided over two groups. One group will receive a training which is focused on learning procedures like current day pilot training. This group of 15 pilots will act as the control group for the experiment. The other group of 15 pilots will receive a training in which the problem-solving strategy will be explained. The training will be split up into three phases: basic lateral autopilot training, basic vertical autopilot training and advanced autopilot training. Each phase will start with a steady flight condition in which the participants are able to interact with the respective autopilot modes with instructor guidance. The second part of each phase is the non-routine training. In this part, the participants will experience some automation failures related to the modes which have just been introduced. Depending on their training program, the pilots will be asked to either apply the general problem-solving strategy or the procedure specific to that type of failure during this part of the training.

After the training, the pilots will fly another five to ten scenarios which include failures that influence the autopilot behaviour as part of the full flight simulator test. The pilots will be asked to complete the scenarios using as much of the automation systems as possible. It is hypothesised that the pilots who received the strategical training are more effective in dealing with these automation surprises in the test as they have practised dealing with unknown problems before during the training. The control group has only been trained with some procedures for very specific cases and is likely to show difficulty in dealing with novel scenarios.

The experiment will be performed in the Simona Research Simulator located at the Delft University of Technology. The existing flight simulation model of the Piper Seneca II aircraft is extended to include the simulated autopilot and avionics. The reason for this choice was that this rather complete non-linear model already existed and has been used by several relateable experiments within this research group.

References

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1–18. doi: 10.1037/a0021017

Ausubel, D. P. (2000). *The acquisition and retention of knowledge: A cognitive view*. Springer Netherlands. doi: 10.1007/978-94-015-9454-7

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6). doi: 10.1016/0005-1098(83)90046-8

Bourgeois-Bougrine, S., Gabon, P., Mollard, R., Coblentz, A., & Speyer, J.-J. (2018). Fatigue in aircrew from shorthaul flights in civil aviation: the effects of work schedules. In H. Muir & D. Harris (Eds.), *Human factors and aerospace safety* (pp. 177–187). Routledge. doi: 10.4324/9781315194035-5

Brna, P., Ohlsson, S., & Pain, H. (1993). The avionics job-family tutor: An approach to developing generic cognitive skills within a job-situated context. *Artificial intelligence in education 1993 : proceedings of AI-ED 93, World Conference on Artificial Intelligence in Education*, 513–520.

Bybee, R., Taylor, J., Gardner, A., Scotter, P., Powell, J., Westbrook, A., & Landes, N. (2006). The bscs 5e instructional model: Origins and effectiveness. *Office of Science Education National Institutes of Health*, 1-80.

Comans, J. (2017). *Visualizing rules, regulations, and procedures in ecological information systems* (Doctoral dissertation, Delft University of Technology). doi: 10.4233/uuid:9b3f9bb6-ef1b-41ed-803a-7e7976784b85

de Boer, R. J., & Hurts, K. (2017). Automation surprise: Results of a field survey of dutch pilots. *Aviation Psychology and Applied Human Factors*, *7*(1), 28-41. doi: 10.1027/2192-0923/a000113

Degani, A., & Wiener, E. L. (1998). *Design and operational aspects of flight-deck procedures* (Tech. Rep.). NASA. Retrieved from `https://ti.arc.nasa.gov/m/profile/adegani/Design%20and%20Operational%20Aspects.pdf`

Dehais, F., Peysakhovich, V., Scannella, S., Fongue, J., & Gateau, T. (2015). "automation surprise" in aviation. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems - CHI '15*. ACM Press. doi: 10.1145/2702123.2702521

Dijkstra, F., & van Doorn, L. (1985). *The construction of a scale to measure subjective effort* (Doctoral dissertation, Delft University of Technology). Retrieved from `https://www.researchgate.net/publication/266392097_The_Construction_of_a_Scale_to_Measure_Perceived_Effort`

Dismukes, R. K., Goldsmith, T. E., & Kochan, J. A. (2015). Effects of acute stress on aircrew performance: Literature review and analysis of operational aspects..

Dobie, G. (2014). *Global aviation safety study: A review of 60 years of improvement in aviation safety* (Tech. Rep.). AGCS Allianz. Retrieved from `https://www.agcs.allianz.com/content/dam/onemarketing/agcs/agcs/reports/AGCS-Global-Aviation-Safety-2014-report.pdf`

Dochy, F., Segers, M., den Bossche, P. V., & Gijbels, D. (2003). Effects of problem-based learning: a meta-analysis. *Learning and Instruction*, *13*(5), 533–568. doi: 10.1016/s0959-4752(02)00025-7

FAA. (1996). *The interfaces between flightcrews and modern flight deck systems* (Tech. Rep.). Federal Aviation Authority. Retrieved from `http://www.tc.faa.gov/its/worldpac/techrpt/hffaces.pdf`

FAA. (2004). *Economic values for faa investment and regulatory decisions, a guide* (Tech. Rep.). Federal Aviation Authority. Retrieved from `https://www.faa.gov/regulations_policies/policy_guidance/benefit_cost/media/050404%20Critical%20Values%20Dec%2031%20Report%2007Jan05.pdf`

Fletcher, G., & Bisset, G. (2017). *Pilot training review – interim report: literature review* (Tech. Rep.). Civil Aviation Authority. Retrieved from `http://publicapps.caa.co.uk/docs/33/CAP1581bLiteratureReview.pdf`

Flight Deck Automation Working Group. (2013). *Operational use of flight path management systems* (Tech. Rep.). Federal Aviation Authority. Retrieved from `https://www.skybrary.aero/bookshelf/books/2501.pdf`

Fyfe, E. R., DeCaro, M. S., & Rittle-Johnson, B. (2014). An alternative time for telling: When conceptual instruction prior to problem solving improves mathematical knowledge. *British Journal of Educational Psychology*, *84*(3), 502–519. doi: 10.1111/bjep.12035

Goteman, Ö. (2018). Automation policy or philosophy? management of automation in the operational reality. In S. Dekker & E. Hollnagel (Eds.), *Coping with computers in the cockpit* (p. 215 - 221). Routledge. doi: 10.4324/9780429460609-12

Gouraud, J., Delorme, A., & Berberian, B. (2017). Autopilot, mind wandering, and the out of the loop performance problem. *Frontiers in Neuroscience*, *11*(541). doi: 10.3389/fnins.2017.00541

Harris, D., & Harris, D. (2016). *Human performance on the flight deck.* CRC Press LLC. doi: 10.1201/9781315252988

Hawkins, F. H., & Orlady, H. W. (1993). *Human factors in flight.* Taylor& Francis Ltd.

Hiebert, J., & Lefevre, P. (2013). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge* (p. 1-27). Routledge. doi: 10.4324/9780203063538

Holder, B. (2012). *Airline pilot perceptions of training effectiveness* (Tech. Rep.). P. O. Box 3707. Seattle, Washington 98124, USA: Boeing Commercial Airplanes.

Huet, M., Jacobs, D. M., Camachon, C., Missenard, O., Gray, R., & Montagne, G. (2011). The education of attention as explanation of variability of practice effects: Learning the final approach phase in a flight simulator. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(6), 1841–1854. doi: 10.1037/a0024386

IATA. (2019). *Loss of control in-flight accident analysis report* (Tech. Rep.). International Air Transport Association. Retrieved from https://www.iata.org/contentassets/b6eb2adc248c484192101edd1ed36015/loc-i_2019.pdf

Javaux, D., & Sherry, L. (1999). How a cognitive tutor can improve pilot knowledge of mode transitions. In *Gateway to the new millennium. 18th digital avionics systems conference. proceedings (cat. no.99ch37033)* (p. 4.B.4-4.B.4.). IEEE. doi: 10.1109/DASC.1999.863727

Jonassen, D. H. (2010). *Learning to solve problems.* Routledge. doi: 10.4324/9780203847527

Jonassen, D. H., & Strobel, J. (2006). Modeling for meaningful learning. In D. Hung & M. S. Khine (Eds.), *Engaged learning with emerging technologies* (pp. 1–27). Springer-Verlag. doi: 10.1007/1-4020-3669-8_1

Jong, T. D., & Joolingen, W. R. V. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, *68*(2), 179–201. doi: 10.3102/00346543068002179

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86. doi: 10.1207/s15326985ep4102_1

Kiss, D. M. (2019). Enhanced pilot learning interface. In S. Nazir, A. M. Teperi, & A. Polak-Sopińska (Eds.), *Advances in human factors in training.* Springer, Cham. doi: 10.1007/978-3-319-93882-0_3

Koolstra, H. (2017). *Preventing aircraft loss of control* (Doctoral dissertation, Delft University of Technology). Retrieved from 10.4233/uuid:96053b94-0f1a-45cd-9e45-1f39e78c55dey

Landman, A., Groen, E. L., van Paassen, M. M. R., Bronkhorst, A. W., & Mulder, M. (2017b). The influence of surprise on upset recovery performance in airline pilots. *The International Journal of Aerospace Psychology*, *27*(1-2), 2–14. doi: 10.1080/10508414.2017.1365610

Landman, A., Groen, E. L., van Paassen, R., Bronkhorst, A. W., & Mulder, M. (2017a). Dealing with unexpected events on the flight deck: A conceptual model of startle and surprise. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *59*(8), 1161-1172. doi: 10.1177/0018720817723428

Landman, A., van Oorschot, P., van Paassen, M. M. R., Groen, E. L., Bronkhorst, A. W., & Mulder, M. (2018). Training pilots for unexpected events: A simulator study on the advantage of unpredictable and variable scenarios. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *60*(6), 793–805. doi: 10.1177/0018720818779928

Lintern, G., & Boot, W. R. (2019). Cognitive training: Transfer beyond the laboratory? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 001872081987981. doi: 10.1177/0018720819879814

Loss of Control Action Group. (2013). *Monitoring matters: Guidance on the development of pilot monitoring skills* (Tech. Rep.). Civil Aviation Authorities. Retrieved from https://publicapps.caa.co.uk/docs/33/9323-CAA-Monitoring%20Matters%202nd%20Edition%20April%202013.pdf

Lyall, E. A., Boehm-Davis, D. A., & Jentsch, F. (2008). *Automation training practitioners' guide.* Retrieved from https://pdfs.semanticscholar.org/9149/bee48c1fb779db690b32187b1437b12337cd.pdf?_ga=2.173525795.104298600.1585751851-1616464184.1580316847 (unpublished)

Marshall, J. C., Horton, B., & Smart, J. (2009). 4e × 2 instructional model: Uniting three learning constructs to improve praxis in science and mathematics classrooms. *Journal of Science Teacher Education*, *20*(6), 501–516. doi: 10.1007/s10972-008-9114-7

Martin, W. L. (2013). *Pathological behaviours in pilots during unexpected critical events: The effects of startle, freeze and denial on situation outcome* (Doctoral dissertation, Griffith University Queensland). Retrieved from https://research-repository.griffith.edu.au/bitstream/handle/10072/366319/Martin_2014_01Thesis.pdf?sequence=1&isAllowed=y

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory Into Practice*, *41*(4), 226–232. doi: 10.1207/s15430421tip4104_4

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, *59*(1), 14–19. doi: 10.1037/0003-066x.59.1.14

Mitchell, C. M. (2000). Horizons in pilot training: Desktop tutoring systems. In R. A. Nadine B. Sarter (Ed.), *Cognitive engineering in the aviation domain.* Lawrence Erlbaum Associates, Publishers. doi: 10.1201/b12462

Moebus, K. (2009). *Impact assessment of the publication of questions of theoretical examinations for part 66 and part fcl* (Tech. Rep.). European Union Aviation Safety Agency. Retrieved from https://www.easa.europa.eu/sites/default/files/dfu/Final%20Report%20on%20publication%20assessment.pdf

Nickolic, M. I., & Sarter, N. B. (2007). Flight deck disturbance management: A simulator study of diagnosis and recovery from breakdowns in pilot-automation coordination. *Human factors*, *49*(4), 553-563. doi: 10.1518/001872007X215647

Nikolic, M. I., & Sarter, N. B. (2007, aug). Flight deck disturbance management: A simulator study of diagnosis and recovery from breakdowns in pilot-automation coordination. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *49*, 553–563. doi: 10.1518/001872007x215647

Novak, J. D. (2013). Meaningful learning is the foundation for creativity. *Qurriculum. Revista de Teoría,Investigación y Práctica educativa*, 27–38. doi: 10.25145/j.qurricul

Orlady, L. M. (2010). Airline pilot training today and tomorrow. In B. G. Kanki, R. L. Helmreich, & J. Anca (Eds.), *Crew resource management* (pp. 469–491). Elsevier. doi: 10.1016/b978-0-12-374946-8.10020-2

Plat, M. V., & Amalberti, R. (2000). Experimental crew training to deal with automation surprises. In N. B. Sarter & R. Amalberti (Eds.), *Cognitive engineering in the aviation domain.* Lawrence Erlbaum Associates, Publishers. doi: 10.1201/b12462

Qi, X., Bard, J. F., & Yu, G. (2004). Class scheduling for pilot training. *Operations Research*, *52*(1), 148–162. doi: 10.1287/opre.1030.0076

Randhawa, P., Jeppu, Y., Nayak, C., & Murthy, N. (2010). Mode logic design of a vertical autopilot for aircrafts. *International Journal of Advances in Engineering and Emerging Technology (IJAEET)*, *1*(1), 7–15. Retrieved from http://erlibrary.org/papers/ijaeet/v1/i1/ERL-101242.pdf

Raven, J., & Raven, J. (2003). Raven progressive matrices. In *Handbook of nonverbal assessment* (pp. 223–237). Springer US. doi: 10.1007/978-1-4615-0153-4_11

Rigner, J., & Dekker, S. (2000). Sharing the burden of flight deck automation training. *The International Journal of Aviation Psychology*, *10*(4), 317–326. doi: 10.1207/s15327108ijap1004_1

Roper, R., Baxley, B., Swieringa, K., & Hubbs, C. (2019). Participant training for a flight test evaluation of interval management. In S. Nazir, A. M. Teperi, & A. Polak-Sopińska (Eds.), *Advances in human factors in training.* Springer, Cham. doi: 10.1007/978-3-319-93882-0_2

Rosa, E., Dahlstrom, N., Knez, I., Ljung, R., Cameron, M., & Willander, J. (2020). Dynamic decision-making of airline pilots in low-fidelity simulation. *Theoretical Issues in Ergonomics Science*, 1–20. doi: 10.1080/1463922x.2020.1758830

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? mode error and awareness in supervisory control. *Human factors*, *37*(1), 5-19. doi: 10.10.1518/001872095779049516

Sarter, N. B., & Woods, D. D. (1997). Team play with a powerful and independent agent: Operational experiences and automation surprises on the airbus a-320. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(4), 553–569. doi: 10.1518/001872097778667997

Savery, J. R., & Duffy, T. M. (1995). Problem based learning: An instructional model and its constructivist framework. *Educational Technology*, *35*(5), 31–38. Retrieved from www.jstor.org/stable/44428296

Schaafstal, A., Schraagen, J. M., & van Berl, M. (2000). Cognitive task analysis and innovation of training: The case of structured troubleshooting. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *42*(1), 75–86. doi: 10.1518/001872000779656570

Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., & Wiegmann, D. (2007). Human error and commercial aviation accidents: An analysis using the human factors analysis and classification system. *Human factors*, *49*(2), 227-242. doi: 10.1518/001872007X312469

Sheridan, T. B., & Parasuraman, R. (2005). Human-automation interaction. *Reviews of Human Factors and Ergonomics*, *1*(1), 89 - 129. doi: 10.1518/155723405783703082

Sherry, L., Polson, P. G., Mumaw, R. J., & Palmer, E. (2001). *A cognitive engineering analysis of the vertical navigation (vnav) function* (Tech. Rep.). NASA Ames. Retrieved from `https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20010046486.pdf`

van Oorschot, P. (2017). *Using unpredictability and variety in pilot training to improve performance in surprise situations* (Master's thesis, Delft University of Technology). Retrieved from `http://resolver.tudelft.nl/uuid:537c5257-f785-4121-827c-8d5bbfeaeb92`

van Paassen, M., Stroosma, O., & Delatour, J. (2000). DUECA - data-driven activation in distributed real-time computation. In *Modeling and simulation technologies conference.* American Institute of Aeronautics and Astronautics. doi: 10.2514/6.2000-4503

White, B. Y., & Frederiksen, J. R. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, *42*(1), 99–157. doi: 10.1016/0004-3702(90)90095-h

Wiener, E. L., & Curry, R. E. (1980). Flight deck automation: Promises and problems. *Ergonomics*, *23*(10), 995 - 1011. doi: 10.1080/00140138008924809

Woods, D. D., & Johannesen, L. J. (1994). *Behind human error: Cognitive systems, computers, and hindsight.* Crew Systems Ergonomics Information Analysis Center. doi: 10.1201/9781315568935

Part III

Appendices

A

LAYOUT OF THE DUECA PROJECT

Table A.1: Modules and Channels in the simulation project (part 1)

Channel	G1000	CitationLogger	WeatherProxy	Autopilot	CitationIncoSelector	GFC700GUI	CvCalculation
CitationOutput	read	read		read	read		read
NavData							
DisplayEvent		read					
Vc_Channel							write
SecondaryControls	read						
AP2G1000Channel	read			write			
AutopilotChannel	read			write			
AutopilotTargets	read						
Trim_inco					write		
INCOSelect		read			read		
MassEvent		read					
EngineEvent		read					
ControlEvent		read					
WindEvent		read	read		w&r		
FGWeatherInput			write				
GFC700Event				read		write	
Snapshot					write		
STARSelect					write		
SnapshotRequest					write		
Reposition					write		
CitationPilotInput					read		read
PrimarySwitches							read
VehicleCabPosition							
VehicleCabMotion							
PropSoundData							
PrimaryControls							
SecondarySwitches							
SideStickOutput							
StickSwitches							
QFeel							
SideStickInput							
PilotControlConfig							

Table A.2: Modules and Channels in the simulation project (part 2)

Channel	PA34	ECI	FCSAdapter	WAVPlayer	Malfunctions	CitationNavigator	MFD	Enginesound
CitationOutput	write	read	read			read	read	
NavData						write		
DisplayEvent		write			w&r			
Vc_Channel								
SecondaryControls			read				read	
AP2G1000Channel							read	
AutopilotChannel	read		read				read	
AutopilotTargets							read	
Trim_inco	read							
INCOSelect		write						
MassEvent	w&r	write			w&r			
EngineEvent	w&r	write			w&r			
ControlEvent	read	write			w&r			
WindEvent	read	write						
FGWeatherInput								
GFC700Event								
Snapshot	write							
STARSelect						read		
SnapshotRequest	read							
Reposition	read							
CitationPilotInput	w&r		w&r					
PrimarySwitches			read					
VehicleCabPosition	write							
VehicleCabMotion	write							
PropSoundData	write			read				read
PrimaryControls			read					
SecondarySwitches			read					
SideStickOutput			read					
StickSwitches			read					
QFeel			write					
SideStickInput			write					
PilotControlConfig			w&r			w&r		

B

AVIONICS SUBSYSTEM OVERVIEW

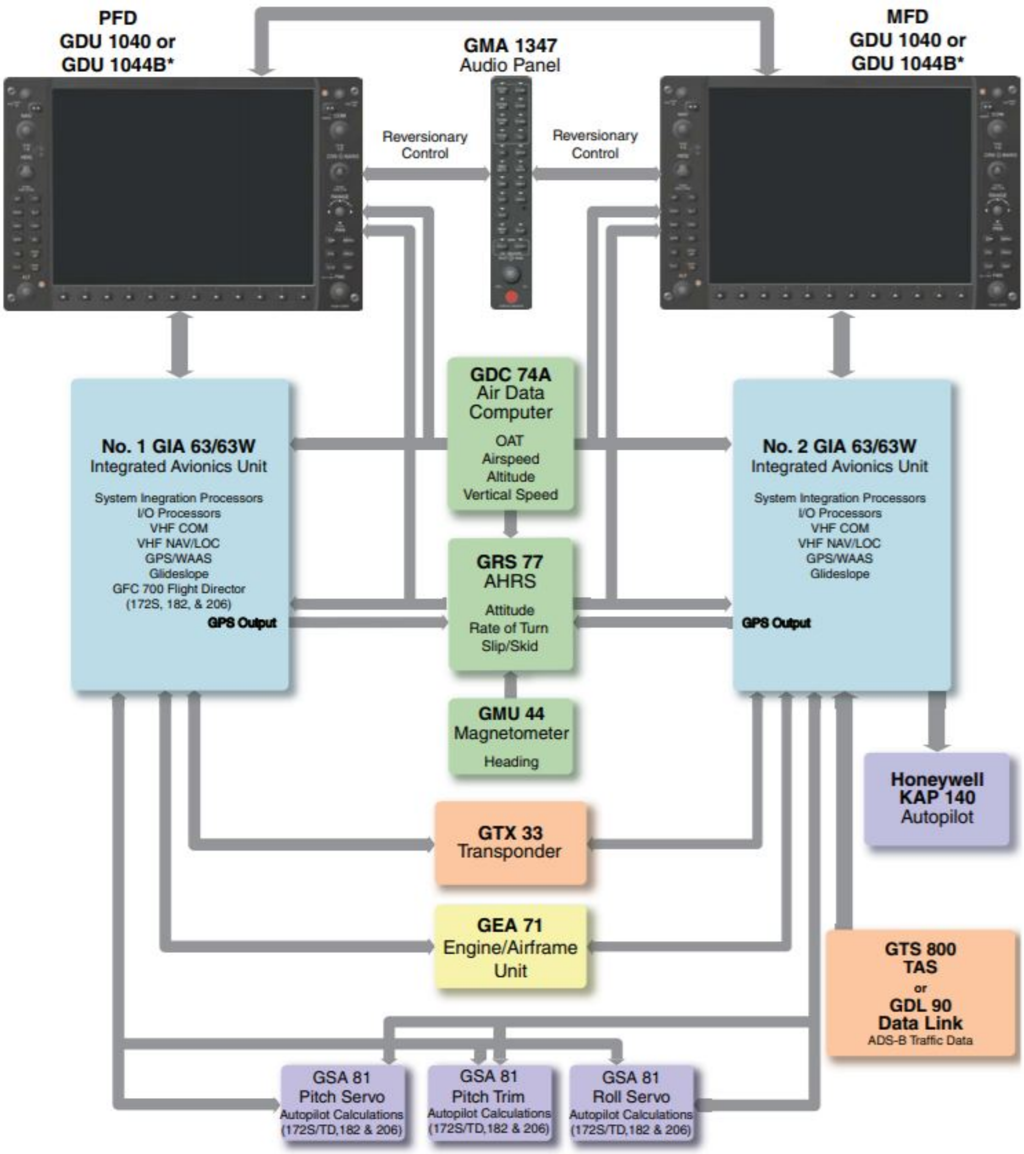


Figure B.1: G1000 Subsystems Overview

Figure B.1 was retrieved from the the G1000 Quick Reference Guide by Garmin[1]

[1] https://static.garmincdn.com/pumac/G1000:CessnaNavIII_CockpitReferenceGuide_SystemSoftwareVersion0563.00orlater_.pdf (Visited on 17 May 2020)

C
AUTOPILOT CONTROL LOOPS AND STEP RESPONSES

The control loops used for the flight director and the autopilot are summarised in Table C.1. If these gains will be tweaked or tuned after this report was handed in, updated gains will be published in the final report of this thesis.

Table C.1: Control loop gains used for the autopilot (version May 2020)

Control	Type	Error e	Kp	Td	Ti	Control e	Note
PIT	-	-	-	-	-	θ_{ref} [rad]	θ_{ref} equals θ at activation
VS	-	-	-	-	-	θ_{ref} [rad]	$\theta_{ref} = asin(V_{ref}/V_{tas}) + \alpha$
ALT	PD	$h_{ref} - h$ in [m]	0.017	1.0	-	θ_{ref} [rad]	h_{ref} equals selected altitude
VPATH	PD	$h_{ref} - h$ in [m]	0.017	1.0	-	θ_{ref} [rad]	h_{ref} equals h at next waypoint
FLC		$V_{ref} - V_{ias}$ in [m/s]	-0.070	2	-	θ_{ref} [rad]	V_{tas} is converted to V_{ias}
ALTS	PD	$h_{ref} - h$ in [m]	0.017	1.0	-	θ_{ref} [rad]	θ_{ref} is clipped at θ at activation
ROL	-	-	-	-	-	ϕ_{ref} [rad]	phi_{ref} equals θ at activation
HDG	PID	$\psi_{ref} - \psi$ in [deg]	0.0264	0.8	1000.0	ϕ_{ref} [rad]	ψ_{ref} equals selected heading
NAV	PID	$\psi_{ref} - \psi$ in [deg]	0.0264	0.8	1000.0	ϕ_{ref} [rad]	ψ_{ref} determined from flightplan
q	PD	$\theta_{ref} - \theta$ in [rad]	0.80	-0.20	-	q_{ref} [rad/s]	Outer loop vertical autopilot
δ_e	PID	$q_{ref} - q$ in [rad/s]	-4.0	0.050	0.5	δ_e [rad]	Inner loop vertical autopilot
p	P	$\phi_{ref} - \phi$ in [rad]	5.73	-	-	p_{ref} [rad/s]	Outer loop lateral autopilot
δ_a	PD	$p_{ref} - p$ in [rad/s]	-40	0.050	-	δ_a [rad]	Inner loop lateral autopilot

The following page contains the step response plots for a step in vertical speed and selected airspeed, and for a step in altitude in heading in Figures C.1 and C.2 respectively. No response plots are presented for the flight director modes VNAV, ALTS and NAV as the control loops used are equivalent to those of the other modes for which step responses are presented. These modes only compute the reference value which they forward to one of the other modes. For example, VNAV computes the required vertical speed based on the horizontal and vertical distance to the next way point, which is read by the control loop of VS to determine θ_{ref} and eventually δ_e.

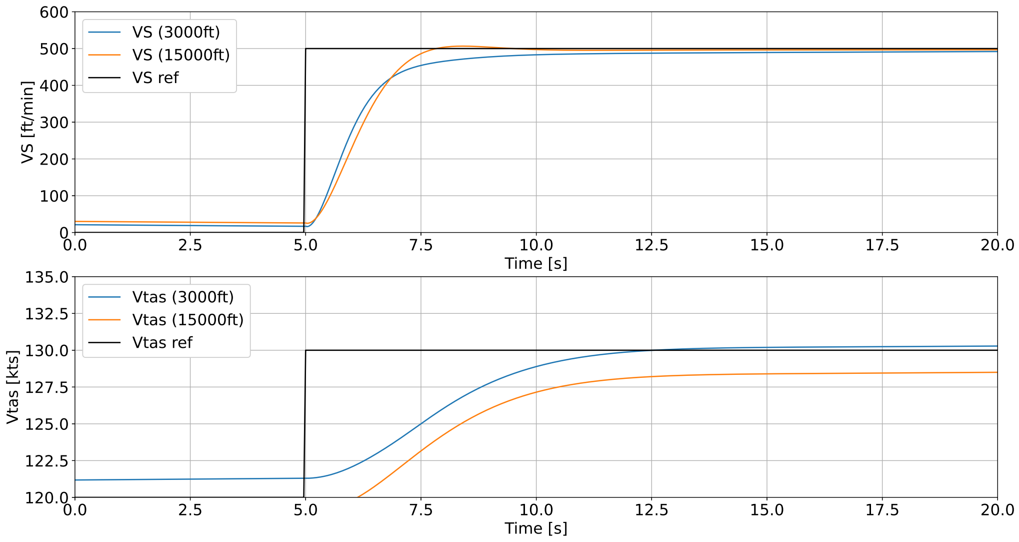


Figure C.1: Step response for Vertical Speed Hold and Flight Level Change

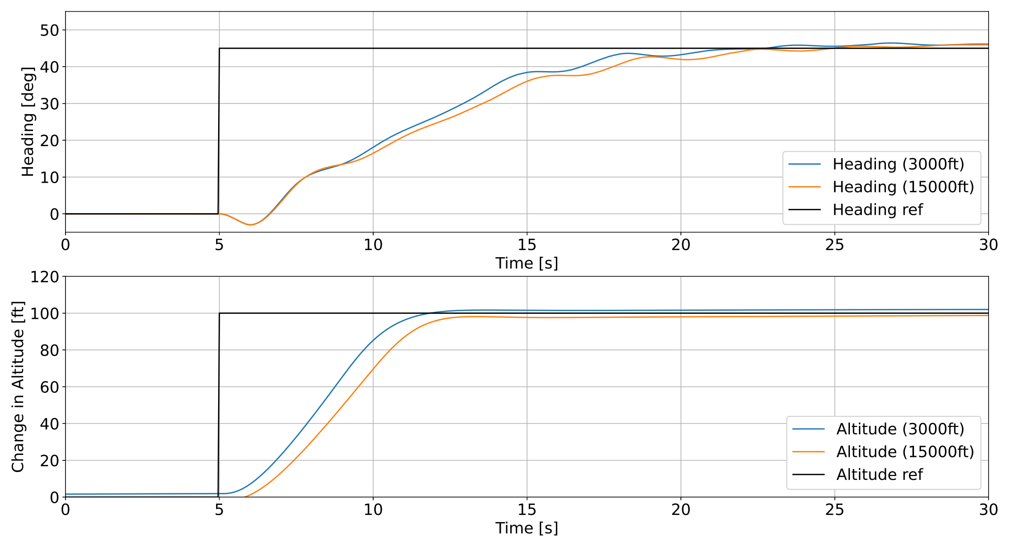


Figure C.2: Step response for altitude and heading

D
Informed Consent

The informed consent form can be found on the following page.

Toestemmingsverklaringformulier (informed consent)

J. (Jordy) van Leeuwen

Titel onderzoek: Flight Deck Automation Training
Verantwoordelijke onderzoeker: prof. dr. Ir. M. Mulder

In te vullen door de deelnemer

Ik verklaar op een voor mij duidelijke wijze te zijn ingelicht over de aard, methode, doel en de risico's en belasting van het onderzoek. Ik weet dat de gegevens en resultaten van het onderzoek alleen anoniem en vertrouwelijk aan derden bekend gemaakt zullen worden. Mijn vragen zijn naar tevredenheid beantwoord.

Ik verklaar dat de onderzoeker mij gedetailleerde veiligheidsinstructies heeft gegeven om te garanderen dat het experiment uitgevoerd kan worden in lijn met de huidige RIVM COVID-19 richtlijnen en dat ik deze instructies begrijp.

Ik begrijp dat ik voor mijn vervoer en het experiment ten alle tijden de huidige RIVM COVID-19 richtlijnen zal volgen. Ik verklaar dat ik naar de Faculteit Luchtvaart- en Ruimtevaarttechniek aan de TU Delft ofwel met mijn eigen auto, per fiets of lopend ben gekomen.

Ik stem geheel vrijwillig in met deelname aan dit onderzoek. Ik behoud me daarbij het recht voor om op elk moment zonder opgaaf van redenen mijn deelname aan dit onderzoek te beëindigen.

Ik begrijp dat geluidsopnames uitsluitend voor wetenschappelijke analyse zullen worden gebruikt. In aanvulling daarop ga ik:

- wel / niet* akkoord met het gebruik van geluidsopnames voor presentaties door de onderzoekers. Dergelijke opnames zullen eerst met mij besproken worden, alvorens de onderzoekers ze op deze manier willen gebruiken. Hierbij heb ik de mogelijkheid om mijn toestemming alsnog in te trekken.*doorhalen wat niet van toepassing is

Naam deelnemer: ____

Datum: __________________ Handtekening deelnemer: _____________________

In te vullen door de uitvoerende onderzoeker

Ik heb een mondelinge en schriftelijke toelichting gegeven op het onderzoek. Ik zal resterende vragen over het onderzoek naar vermogen beantwoorden. De deelnemer zal van een eventuele voortijdige beëindiging van deelname aan dit onderzoek geen nadelige gevolgen ondervinden.

Naam onderzoeker: ________________________________

Datum: __________________ Handtekening onderzoeker: ___________________

E

Multiple Choice Test

These were the multiple choice test questions. The correct answer is shown in bold here.

1
U leest op de Autopilot Status Box:
HDG | AP | VS +300FPM ALTS
Beweringen:
1) De huidige laterale modus is heading mode
2) De autopilot staat aan
3) ALTV is armed
Welke beweringen zijn juist?
a) Alleen bewering 1 is juist
b) Bewering 1 en 2 zijn juist
c) Bewering 1 en 3 zijn juist
d) Alle beweringen zijn juist

2
U leest op de Autopilot Status Box:
HDG | AP | VS +300FPM ALTS

Wat gebeurt er als de AP key wordt ingedrukt?
a) de autopilot schakelt uit
b) de autopilot schakelt aan
c) de flight director schakelt aan
d) de flight director schakelt uit

3
Welke instrumenten zijn afhankelijk van de Air Data Computer?
a) airspeed indicator, altimeter, vertical speed indicator
b) airspeed indicator, altimeter, vertical speed indicator, attitude indicator
c) airspeed indicator, altimeter, attitude indicator
d) attitude indicator, altimeter, vertical speed indicator

4
Op welke manieren kan de autopilot ontkoppeld worden?
a) door de AP key in te drukken
b) door de AP key of de FD key in te drukken
c) door de AP key of de FD key in te drukken, of door een systeem failure
d) door de AP key in te drukken of door een systeem failure

5
Beweringen:
1) NAV modus switcht naar ROL modus wanneer de CDI key wordt ingedrukt
2) Armed modi worden in het groen weergeven in de autopilot status box
Welke beweringen zijn juist?

a) alleen bewering 1 is juist
b) alleen bewering 2 is juist
c) beide beweringen zijn juist
d) geen bewering is juist

6
Welke informatie geeft de richting naar het active waypoint aan vanaf de huidige positie?
a) Desired Track (DTK) op de MFD
b) Track (TRK) op de MFD
c) Bearing (BRG) op de PFD
d) deze informatie is op de flight plan pagina op de MFD te vinden

7
Hoe kan er van navigation source geswitcht worden?
a) Door de NAV key in te drukken
b) Door de CDI softkey in te drukken
c) Door de HDG key in te drukken
d) Door de CWS button in te drukken

8
Wat is de functie van Control Wheel Steering (CWS)?
a) De roll en pitch referentie hoeken naar nul brengen voor straight and level flight
b) De flight director command bars synchroniseren met de huidige pitch en roll hoeken
c) De laterale modus roll hold (ROL) en de verticale modus pitch hold (PIT) activeren
d) De autopilot ontkoppelen

9
De autopilot staat in roll modus (ROL), huidige heading is 150. De volgende informatie is beschikbaar: heading bug = 170, selected course = 20, desired track (flight plan) = 210. De huidige navigation source voor de Horizontal Situational Indicator (HSI) is VOR 1. Vervolgens wordt de NAV key ingedrukt. Welke heading kunnen we na enige tijd aflezen?
a) 170 graden
b) 20 graden
c) 210 graden
d) 150 graden

10
Hoe lang blijft de modus Flight Level Change (FLC) aan?
a) Totdat de piloot een andere laterale modus selecteert
b) Totdat de piloot de autopilot uitschakelt
c) Totdat de selected altitude gecaptured wordt
d) Totdat de selected airspeed bereikt is

F

Quick Reference Card

The following two pages contain the Quick Reference Card which both groups were allowed to bring into the SIMONA during both the training and the test.

==
KNEEPAD
==

Take-off
Flaps……....UP
V_r…………..80 kts
V_{mc}…...........80 kts

After take-off
Pitch……….~13 deg
V_2…………..92 kts
Gear……….UP

Downwind
V_{dw}….............115 kts

Base leg
V_{app}………...90 kts
Flaps……....25
Gear……….DOWN

Final
Flaps………LAND

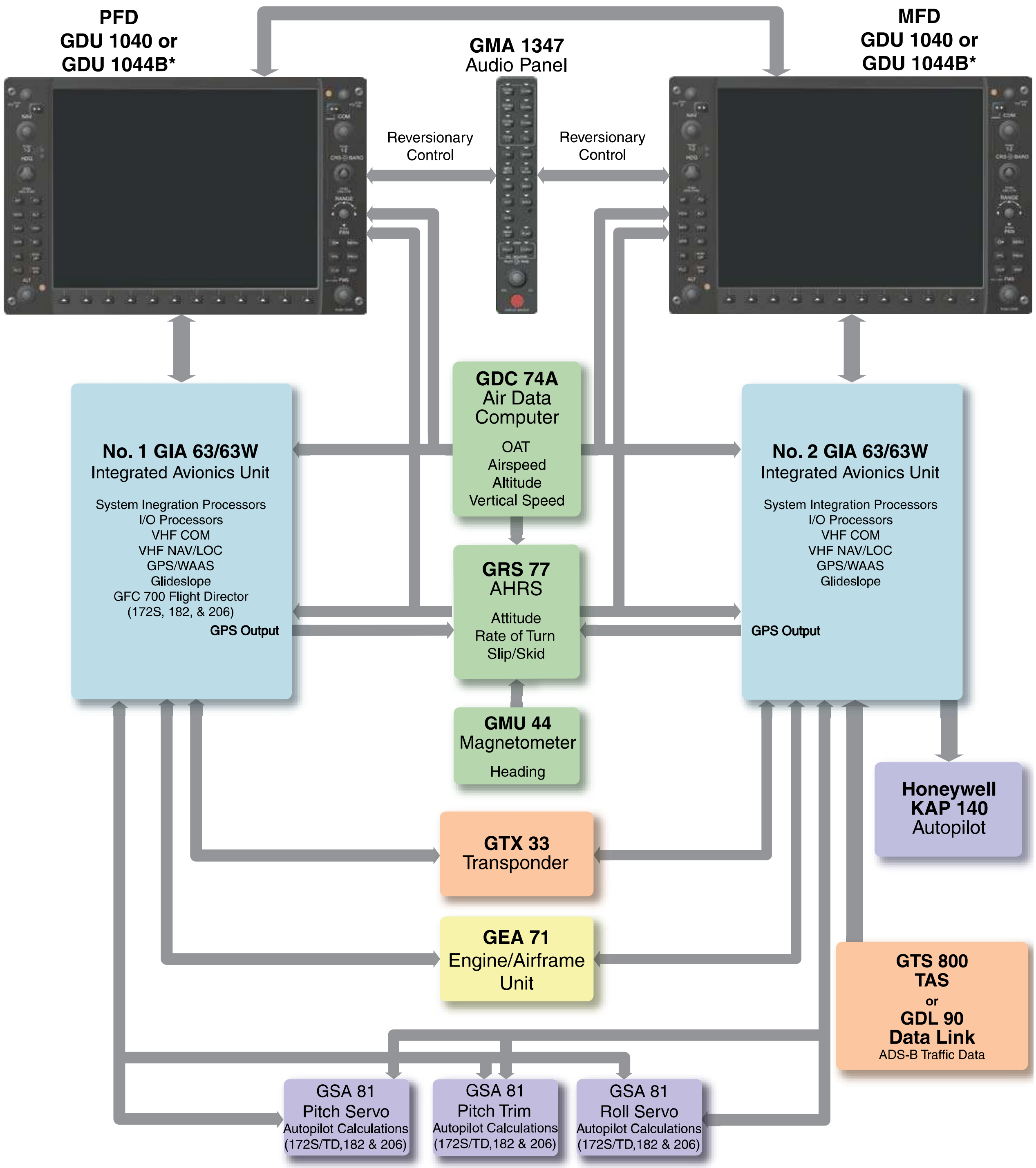
==
AUTOPILOT SUMMARY
==

Pitch Modes:

- **Pitch Hold** (default mode)— Holds the current aircraft pitch attitude; may be used to climb/descend to the Selected Altitude
- **Selected Altitude Capture** — Captures the Selected Altitude
- **Altitude Hold** — Holds the current Altitude Reference
- **Vertical Speed** — Maintains the current aircraft vertical speed; may be used to climb/descend to the Selected Altitude
- **Flight Level Change** — Maintains the current aircraft airspeed while the aircraft is climbing/ descending to the Selected Altitude
- **Vertical Path Tracking** — Follows an active vertical profile for enroute and terminal phases of flight
- **VNAV Target Altitude Capture** — Captures the VNAV Target Altitude

Roll Modes:

- **Roll Hold** (default mode) — Holds the current aircraft roll attitude or rolls the wings level, depending on the commanded bank angle
- **Heading Select** — Captures and tracks the Selected Heading
- **Navigation** (GPS, VOR) — Captures and tracks the selected navigation source



* The GDU 1040 is available in systems not using the GFC 700 Automatic Flight Control System.
The GDU 1044B is available in systems using the Garmin GFC 700 Automatic Flight Control System.

Figure 1-1 Basic G1000 System

G

Ground School Slides

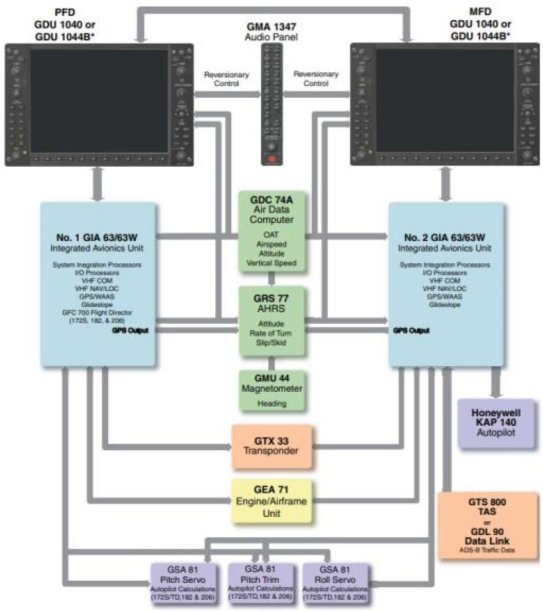
The following pages contain the slides used for the ground school of the participants. The last two slides were only used for the experimental group, as they explain the problem-solving strategy.

Theorie

Vergelijkbaar met:
G1000

TUDelft

G1000 System



TUDelft

2

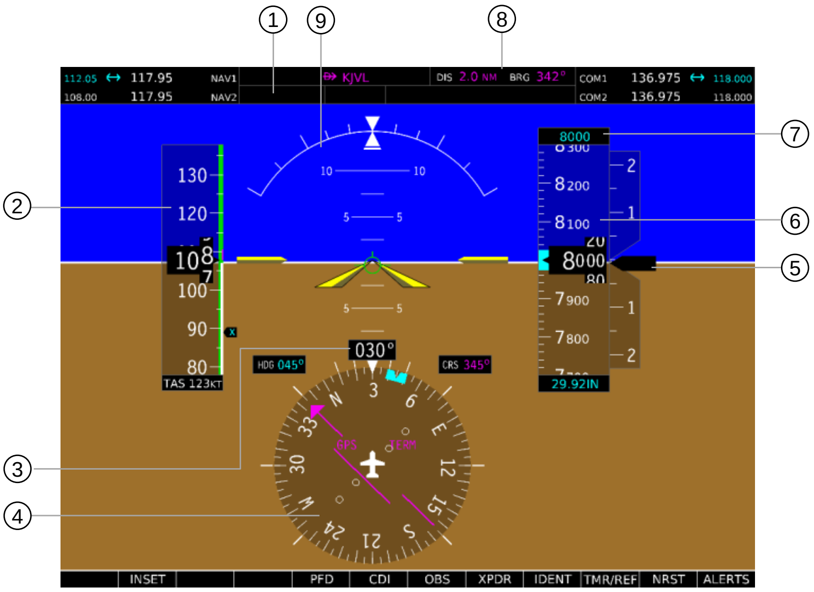
G1000 System

3

G1000 System

4

Flight Instruments



1. Autopilot Status Box
2. Airspeed Indicator
3. Heading Box
4. Horizontal Situational Indicator
5. Vertical Speed Indicator
6. Altimeter
7. Selected Altitude Box
8. Navigation Status Box
9. Attitude Indicator

TU Delft

5

Automatic Flight Control

TU Delft

6

Autopilot Operation

Engaging the Autopilot

When the **AP** Key is pressed, the autopilot and flight director (if not already engaged) are activated. Engagement is indicated by a green 'AP' annunciation in the center of the AFCS Status Box. The flight director engages in Pitch and Roll Hold modes when initially activated

Autopilot Engaged

Disengaging the Autopilot

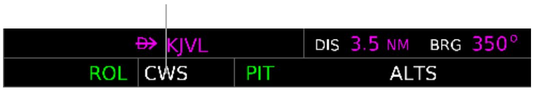
The autopilot is manually disengaged by pushing the **AP** Key on the MFD. Manual disengagement is indicated by a five-second flashing yellow 'AP' annunciation and a three-second autopilot disconnect aural alert.

Autopilot Manually Disengaged

Control Wheel Steering

During autopilot operation, the aircraft may be hand-flown without disengaging the autopilot. Pressing and holding the **CWS** Button disengages the pitch and roll servos from the flight control surfaces and allows the aircraft to be hand-flown. At the same time, the flight director is synchronized to the aircraft attitude during the maneuver. The 'AP' annunciation is temporarily replaced by 'CWS' in white for the duration of CWS maneuvers.
In most scenarios, releasing the **CWS** Button reengages the autopilot with a new reference. Refer to the flight director modes section for CWS behavior in each mode.

Control Wheel Steering



TU Delft

7

Flight Director Operation

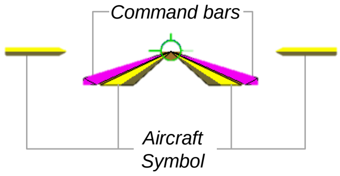
With the flight director activated, the aircraft can be hand-flown to follow the path shown by the Command Bars. The flight director also provides commands to the autopilot.

Activating the Flight Director

Pressing the **FD** or **AP** Key (when the flight director is not active) activates the flight director in default pitch/roll modes. The flight director may be turned off by pressing the **FD** Key.

Command Bars

Upon activation of the flight director, Command Bars are displayed on the PFD as a single cue. If the attitude information sent to the flight director becomes invalid or unavailable, the Command Bars are removed from the display. The Command Bars do not override the aircraft symbol.



TU Delft

8

Flight Director Modes

Pitch Modes

- **Pitch Hold** (default mode)— Holds the current aircraft pitch attitude; may be used to climb/descend to the Selected Altitude
- **Altitude Hold** — Holds the current Altitude Reference
- **Vertical Speed** — Maintains the current aircraft vertical speed; may be used to climb/descend to the Selected Altitude
- **Flight Level Change** — Maintains the current aircraft airspeed while the aircraft is climbing/ descending to the Selected Altitude
- **Vertical Path Tracking** — Follows an active vertical profile for enroute and terminal phases of flight
- ***VNAV Target Altitude Capture*** *— Captures the VNAV Target Altitude*
- ***Selected Altitude Capture*** *— Captures the Selected Altitude*

Roll Modes

- **Roll Hold** (default mode) — Holds the current aircraft roll attitude or rolls the wings level, depending on the commanded bank angle
- **Heading Select** — Captures and tracks the Selected Heading
- **Navigation** (GPS, VOR) — Captures and tracks the selected navigation source

TU Delft

Flight Director Pitch Modes

Pitch Modes

The mode reference (shown with default measurement units) is displayed next to the active mode annunciation for Altitude Hold, Vertical Speed, and Flight Level Change modes. The **NOSE UP**/ **NOSE DN** Keys can be used to change the pitch mode reference while operating under Pitch Hold, Vertical Speed, or Flight Level Change Mode.

Pitch Mode	Control	Annunciation
Pitch Hold	(default)	PIT
Altitude Hold	**ALT** key	ALT nnnn ft
Vertical Speed	**VS** key	VS nnn fpm
Selected Altitude Capture	*	ALTS
Flight Level Change	**FLC** key	FLC nnn kts
Vertical Path Tracking	**VNV** key	VPTH
VNV Target Altitude Capture	**	ALTV

* ALTS is armed automatically when PIT, VS, FLC, or GA is active, and under VPTH when the Selected Altitude is to be captured instead of the VNAV Target Altitude.

** ALTV is armed automatically under VPTH when the VNAV Target Altitude is to be captured instead of the Selected Altitude.

TU Delft

Flight Director Roll Modes

Roll Modes

The following table relates each roll mode to its respective control and annunciation.

Roll Mode	Control	Annunciation
Roll Hold	(default)	ROL
Heading Select	**HDG** key	HDG
Navigation, GPS/VOR	**NAV** key	GPS/VOR

TU Delft

11

Non-routine

Problem-solving strategie

TU Delft

Problem-solving Strategie

1. Notice autopilot behaviour is off nominal
2. Identify which sensor or system is faulty
3. Identify implications on autopilot
4. Switch to alternative information source if possible, or
5. Switch to lower level automation

TU Delft

13

H
Experiment Script

The following pages contain the script that the experimenter used for the training and the test. During the training, the experimenter was the instructor meaning that interaction between the experimenter and the participant was allowed (and encouraged). During the test, the experimenter only explained the things written down in the script, and no more.

Einde introductievlucht
Start autopilot training

Onderdeel 1: horizontale modi

START RECORDING

- Eerste onderdeel, horizontale modi: roll hold, heading hold
- Oefenen met deze modi in een doorlopende simulatie
- Start: 5000 voet, getrimd met a/p uit. We beginnen manueel.
- Klaar om te beginnen?

You have control [start 201]

- Opdracht: autopilot aanzetten met AP knop

drukt AP knop in

- De autopilot start standaard in ROLL en PITCH hold mode
- Indicaties Autopilot Status Box: ROL, AP, PIT
- Leg uit: magenta command bars van FD
- FD automatisch ingeschakeld omdat a/p dit nodig heeft als referentie
- Leg uit: CWS om nieuwe roll en pitch referentie in te stellen
- Opdracht: gebruik CWS om een nieuwe roll en pitch hoek als referentie in te stellen (5 graden pitch, +20 graden roll)

gebruikt CWS

- Indicatie Autopilot Status Box: CWS

- Heading niet constant in roll modus
- Leg uit: heading bug en heading hold
- a/p zal bocht inzetten met max bank 20 graden, wings level bij heading bug
- Opdracht: HDG selecteren

drukt HDG knop in

- Heading vastgehouden ondanks wind en turbulentie.
- Indicaties Autopilot Status Box: HDG
- Modus cancelen door nogmaals op knop te drukken
- Laterale modi gaan terug naar ROL
- Opdracht: huidige roll modus cancelen

drukt HDG knop in

- Leg uit: werking heading select draaiknop
- Opdracht: selecteer heading 180, selecteer HDG modus

verandert heading select en drukt HDG knop in

- Selected heading kan ook in Heading hold veranderd worden
- Opdracht: selecteer 90 graden heading

selecteert 90 graden heading

- Met SYNC wordt hdg bug huidige heading
- Vragen?
- Opdracht: probeer nu zelf eens een aantal selected headings met en zonder HDG hold actief, kijk naar het effect

- FD kan zonder a/p, blijft standaard aan
- Command Bars blijven referentie aangeven
- Als a/p uitschakelt: audio signaal, AP geel en dan weg
- Opdracht: autopilot uitschakelen met AP knop

Drukt AP knop in

- Indicaties Autopilot Status Box: HDG | | PIT
- Wijs op manual control,
- Opdracht: stukje dalen en aflevelen.

Duwt stuurkolom naar voren

- Flight director als referentie gebruiken om manual naar heading 150 te vliegen
- Opdracht: Gebruik hiervoor eerst de heading select draaiknop.

selecteert 150 graden heading

- Wijs op het niet reageren van vliegtuig want AP off.
- Opdracht: volg de command bars (follow the needle) naar hdg 150

Heading 150 bereikt

- Ook met AP off kunt u de HDG modus cancelen. De flight director gaat dan weer naar ROL modus.

- De flight director kan uitgeschakeld worden met de FD knop.
- Opdracht: Schakel de flight director uit met de FD knop

drukt de FD knop in

- Recap: de flight director bepaalt de referentie roll en pitch hoek, de autopilot probeert deze roll en pitch hoek aan te houden wanneer ingeschakeld.
- AP en FD ingeschakeld door AP knop
- De autopilot wordt uitgeschakeld door nogmaals op de AP knop te drukken. De flight director blijft dan ingeschakeld.
- De flight director wordt uitgeschakeld door op de FD knop te drukken.

- Opdracht: Gebruik de autopilot om een heading van 270 te vliegen en schakel dan de autopilot en flight director uit.

Na uitvoeren oefening: stop simulatie

2x PFD failure scenario

- Non-routine fase van de laterale modi
- Onbekende failure in het volgende scenario
- Het is uw taak om hier zo effectief mogelijk mee om te gaan, door gebruik te maken van de problem-solving strategy
  - 1.Notice autopilot behaviour is off nominal
  - 2.Identify which sensor or system is faulty
  - 3.Identify implications on autopilot
  - 4.Switch to alternative information source if possible, or
  - 5.Switch to lower level automation

- Start: 5000ft, met AP en FD uit.
- Opdracht: Vlieg heading 270 op 5000 ft vliegen

Your controls [start 201]

Na uitvoeren oefening: stop simulatie

- Uitleg: PFD failure: volledige PFD valt weg, a/p schakelt uit
- Te herkennen aan: PFD scherm gaat op zwart
- Procedure: pak stuurkolom direct vast en vlieg manueel (trim werkt nog), maak gebruik van het standby instrument en de MFD (track)
- We gaan dit scenario nog een keer herhalen
- Opdracht: vlieg heading 270

Your controls [start 201]

Na uitvoeren oefening: stop simulatie

<u>Onderdeel 2: Verticale modi</u>

- Tweede onderdeel verticale modi: Pitch hold, Vertical Speed Mode, Altitude Hold Mode en Flight Level Change Mode.
- Start: 5000 ft, met AP en FD uit.
- <u>Opdracht</u>: na starten mag u de autopilot en flight director aanzetten.

Your controls [start 301]

Start simulatie, schakelt AP en FD in

- Indicaties: ROL en PIT.
- Pitch Hold Mode actief
- Met PIT actief volgen de AP en/of de FD een geselecteerde pitch hoek.
- 1x op UP zorgt voor 0.5 deg pitch up, 1x op DOWN voor 0.5 deg pitch down.
- <u>Opdracht</u>: gebruikt PIT om de neus 2 graden omhoog te pitchen.

drukt UP knop een paar keer in

- <u>Opdracht</u>: gebruik PIT om weer 2 graden omlaag te pitchen.

drukt DOWN knop een paar keer in

- Als één van die andere modi actief is, hebben de UP en DOWN knoppen een iets andere betekenis. Echter: UP zorgt altijd voor een 'nose up' commando, en DOWN zorgt altijd voor een 'nose down' commando.

- Het is vaak handiger om altitude constant te houden ipv de pitch hoek.
- Dat kan met de Altitude Hold modus
- Opdracht: activeer de altitude hold modus met de ALT knop.

  drukt de ALT knop in

- Indicatie: ALT XXXX ft
- Omdat ALT altijd de huidige altitude als referentie neemt bij het inschakelen,
- U kunt de referentie altitude niet aanpassen.
- Als u wilt stijgen of dalen, kunnen we gebruik maken van de Vertical Speed Modus. Deze volgt een opgegeven vertical speed.
- Deze is bij activeren altijd 0 ft/min
- Opdracht: Activeer Vertical speed mode door op de VS knop te drukken

  drukt VS knop in

- Indicaties: VS +0fpm, selected VS boven VSI en VS bug
- Met UP verhoogt u de referentie VS met 100 ft/min en vise versa voor DOWN.
- Opdracht: verhoog de vertical speed naar 400 ft/min klimmen.

  verhoogt VS naar +400fpm

- Opdracht: breng vertical speed terug naar 0 ft/min.

  verlaagt VS naar 0fpm

- Selected altitude capture in wit -> gearmd
- Selected altitude boven altimeter
- De autopilot en flight director zullen deze geselecteerde hoogte automatisch 'capturen' en vervolgens automatisch overschakelen in Altitude Hold modus.
- Selected aan te passen met draaiknop.
- Opdracht: Selecteer altitude capture op 7000 voet.

  selecteert 6000 voet

- Als u straks dmv VS gaat klimmen, zal u nu automatisch aflevelen op 6000 voet.
- Volgorde: VS (ALTS) - ALTS (ALT) - ALT
- Opdracht: selecteer 700 ft/min vertical speed klimmen en let op de status box als u bij 6000 voet aankomt.

  selecteert +500fpm

- Opdracht: probeer nu naar 5500 voet te dalen dmv ALTS en VS maar let erop dat uw snelheid hierdoor kan oplopen. Corrigeer eventueel thrust.

  selecteert 5500 voet

- Opdracht: probeer zelf naar een paar verschillende hoogtes te vliegen met PIT, ALTS en VS.

- Let erop dat u zelf verantwoordelijk bent voor powermanagement.
- Laatste verticale modus: Flight Level Change modus.
- Gebruikt als referentie een opgegeven airspeed en controleert pitch om deze te halen. Pitch wordt dus automatisch aangepast als u vermogen toevoegt of wegneemt.
- Ook hier zelf verantwoordelijk voor powermanagement.
- Referentie in te stellen met UP en DOWN knoppen (stappen van 5 knopen)
- Opdracht: Activeer de Flight Level Change modus activeren door FLC knop

  drukt FLC knop in

- De huidige airspeed van XXX knopen is nu ingesteld als de referentie airspeed.
- Indicaties: FLC XXX kts
- Opdracht: verlaag de referentie airspeed naar 120 knopen en let op de neusstand en snelheid.

  drukt twee keer op UP

- Opdracht: Stel referentie airspeed in op 125 en zet gas volledig open om te zien wat dit met de pitch hoek doet.

  drukt twee keer op UP en duwt gashendel naar voren

- Opdracht: Zet gas weer terug op 60-70%
- Selected Altitude Capture kan ook in deze modus gearmd en gecaptured worden.
- Opdracht: Probeer met behulp van FLC en ALTS een aantal verschillende altitudes te bereiken.

  Na uitvoeren oefening: stop simulatie

2x Elevator servo failure scenario

- Non-routine fase van de verticale modi
- Onbekende failure in het volgende scenario
- Het is uw taak om hier zo effectief mogelijk mee om te gaan, door gebruik te maken van de problem-solving strategie
- Start: 5000 ft, ap fd staat uit
- Opdracht: Vlieg heading 120 en maintain 5000 ft, door zo veel mogelijk gebruik te maken van de automatisering

Your controls [start 302]

Na uitvoeren oefening: stop simulatie

- Uitleg: Geblokkeerde elevator servo: zolang autopilot aanstaat, geen verandering in elevator deflectie
- Te herkennen aan: autopilot reageert in geen enkele verticale modus op FD, mogelijk pitch up of pitch down moment gecreëerd -> AP volgt FD niet!
- Procedure: schakel de a/p meteen uit en vlieg manueel. FD kan aanblijven als referentie, selecteer de benodigde laterale en verticale modus.
- We gaan dit scenario nog een keer herhalen
- Opdracht: Vlieg heading 120 en maintain naar 5000 ft, door zo veel mogelijk gebruik te maken van de automatisering

Your controls [start 302]

Na uitvoeren oefening: stop simulatie

2x Blocked pitot tube during climb scenario

- Nieuwe onbekende failure in het volgende scenario
- Het is uw taak om hier zo effectief mogelijk mee om te gaan, door gebruik te maken van de problem-solving strategie
- Start: 5000 ft, ap staat uit
- Opdracht: klim naar 6500 ft door gebruik te maken van Flight Level Change

Your controls [start 303]

Na uitvoeren oefening: stop simulatie

- Uitleg: dit is een geblokkeerde pitot tube, indicated airspeed is hierdoor onbetrouwbaar (neemt toe met hoogte)
- Te herkennen aan: vreemd gedrag van ASI, airspeed neemt toe met hoogte
- Procedure: switch naar Vertical Speed modus en zit het klimmen door. Gebruik het standby instrument als ASI
- We gaan dit scenario nog een keer herhalen
- Opdracht: klim naar 6000 ft door gebruik te maken van Flight Level Change

Your controls [start 303]

Na uitvoeren oefening: stop simulatie

Onderdeel 3: navigatie

- Laatste onderdeel van de training: laterale en verticale navigatie. Navigatie modi (horizontaal en verticaal) in combinatie met VOR en GPS.

- We zullen beginnen met VOR navigatie.
- Selectie van VOR baken normaliter mogelijk door frequentie in NAV 1 of NAV 2 te selecteren. In deze simulatie is dit niet mogelijk. Ik zorg ervoor dat in alle scenario's de juiste frequenties ingesteld zijn.
- We gebruiken altijd bestaande VOR bakens, met de juiste frequentie, locatie en naam.
- HSI staat ingesteld op GPS, met softkey nr 6 (onder CDI) kan tussen nav source geswitcht worden: GPS - NAV1 - NAV2.
- Radiaal te selecteren met de course draaiknop (follow-the-needle)
- Start Noord-Holland op 5000 voet, noordwestelijk van het VOR baken Spijkerboor.
- VOR baken al ingesteld op de G1000.
- Het VOR baken wordt ook weergegeven op de MFD. Gebruik range -/+ om in een uit te zoomen
- Opdracht: selecteer NAV1, vlieg 360 inbound Spykerboor (radiaal 180) en intercept handmatig

Your controls. [Start scenario 401]

vliegt radiaal

Na uitvoeren oefening: stop simulatie

- Dezelfde opdracht herhalen maar nu uitvoeren dmv AP.
- AP en FD zullen radiaal automatisch onderscheppen.
- <u>Opdracht</u>: Na starten, AP inschakelen, selecteer radiaal 180, NAV modus

Your controls. [Start scenario 401]

als NAV knop ingedrukt

- Indicatie VOR
- DME door softkey nr 7 (schakelt tussen nav1, nav2 en uit)
- Course deviation indicator naald erg gevoelig als u dicht bij het baken komt, net als in het echt.

vliegt scenario

- NAV 2 ingesteld op Pampus VOR baken.
- Switchen van NAV 1 naar 2 door 1 keer drukken op CDI knop. Nog een keer switcht naar GPS.
- Bij switchen valt de NAV modus terug naar ROL, dus opnieuw activeren met NAV knop.
- <u>Opdracht</u>: bekijk op welke radiaal we t.o.v. Pampus vliegen. Vlieg vervolgens weer verder op radiaal 180 naar Spykerboor.

- Nog vragen over VOR NAV?
- Laatste onderdeel: GPS flight plan navigatie.
- In de G1000 kan een flightplan worden geprogrammeerd.
- In de simulatie is dit niet aan te passen. Voor elk scenario juiste vliegplan ingeladen. Deze kunt u natuurlijk nalopen.

- Start: noorden van het vliegveld Midden-Zeeland, ap aan in roll en pit
- flightplan met waypoints: ODVIL en OMASA en EHRD (allebei rond de kustlijn).
- <u>Opdracht</u>: druk op FPL knop om flight plan op MFD te tonen.

drukt FPL knop

- Actieve leg is vanaf uw huidige positie naar het waypoint ODVIL. Deze is magenta, aankomende legs zijn wit.
- De waypoints staan ingesteld als fly-by.
- Op de flight plan pagina ziet u de waypoints in het flightplan, met de afstand en de richting van elke leg.
- Bovenin de MFD kunnen we onze ground speed, desired track, GPS track en ETE aflezen.
- Bovenin de PFD zien we de distance en bearing naar het actieve waypoint
- Flight Director en AP intercepten active leg (wordt ook geupdate)

- Opdracht: straks GPS navigatie modus activeren. Eerst verifiëren dat de HSI op GPS staat (GPS magenta). Zoniet, met CDI knop GPS selecteren. Dan op NAV drukken.

Your controls [Start scenario 402]

drukt NAV knop in

- Indicatie: GPS
- De naald op de HSI werkt hetzelfde als met VOR navigate. De autopilot en flight director zullen nu de actieve leg intercepten.
- Kijk wat er gebeurt en monitor de autopilot.

Na uitvoeren oefening: stop simulatie

- Heeft u nog vragen over GPS navigatie?

- Laatste modus: VNV
- Bij de waypoints in het flightplan kunnen ook target hoogtes worden opgegeven. De flight director berekent wanneer de daling moet worden ingezet om de volgende vertical waypoint te behalen (top of descent) en wanneer daling moet eindigen (bottom of descent).
- De onderste helft van de flight plan pagina geeft informatie voor de Vertical Navigation modus: actieve vertical waypoint, tijd tot top of descent.
- Hoogte hoeft niet opgegeven te worden bij waypoints. Dan automatisch naar hoogte volgende waypoint.
- VNV alleen mogelijk als GPS NAV actief is.
- Als u op dan op de VNV knop drukt wordt vertical path tracking gearmd. De hoogtes van de waypoints zijn altijd door mij van tevoren opgegeven.
- Zodra dan de top of descent bereikt is, wordt vertical path tracking automatisch geactiveerd, en wordt VNV target altitude capture gearmd.
- Deze VNV target altitude capture werkt hetzelfde als de selected altitude capture in ALT, maar nu met de VNV target altitude (rechts boven de selected altitude op de PFD in het paars).
- Bij inschakelen van VNV: de vertical separation indicator, links van de altimeter. Vergelijkbaar met glideslope indicator: verticale afwijking van het berekende vertical path laat zien.
- Simulatie start in approach runway 22 Schiphol.
- Eerste waypoint in het flightplan: AGOGO, 3000 voet. Tweede waypoint, EH671: geen hoogte. De initial approach fix BLUSY en de final approach fix EH661: beide 2000 voet.
- Approach plate: EHAM RNP Rwy 22
- Opdracht: check dit straks op de flight plan pagina van MFD.
- Opdracht: Activeer straks GPS NAV en VNV om de approach te vliegen.

Your controls [start scenario 403]

voor bereiken AGAPO

- Na passeren AGAPO, 3000 voet wordt aangehouden tot top-of-descent voor het volgende verticale waypoint (BLUSY 2000 ft). Deze zal al plaatsvinden tussen AGOGO en EH617 omdat EH617 geen hoogte heeft.

Na uitvoeren oefening: stop simulatie

2x VOR Receiver failure scenario

- Non-routine fase van de verticale modi
- Volgende scenario: VOR training. U mag radiaal 235 naar Costa vliegen, en vervolgens radiaal 110 naar Nikky. Als back-up staat deze route ook als GPS flightplan ingesteld. Start 3000 voet, net opgestegen van Midden-Zeeland met de a/p aan in rol en pit modus. NAV1 staat ingesteld op VOR COA en NAV2 staat ingesteld op VOR NIK. U kunt dit verifiëren op uw displays.
- Onbekende failure in het volgende scenario
- Het is uw taak om hier zo effectief mogelijk mee om te gaan, door gebruik te maken van de problem-solving strategie
- <u>Opdracht</u>: Vlieg naar 55 inbound COSTA (radiaal 235) en vervolgens naar 290 inbound NIKKY (radiaal 110) door gebruik te maken van VOR navigation modus

Your controls [start 404]

Na uitvoeren oefening: stop simulatie

- Uitleg: dit is een navigation radio receiver failure: geen VOR meer
- Te herkennen: rode kruizen door de NAV radio's op PFD en MFD
- Procedure: VOR is niet meer beschikbaar. Schakel over naar GPS navigation modus door de HSI te veranderen naar GPS met de CDI knop
- We gaan dit scenario nog een keer herhalen
- Opdracht: Vlieg naar COA op radiaal 235 en vervolgens naar NIK op radiaal 110 door gebruik te maken van VOR navigation modus

Your controls [start 404]

Na uitvoeren oefening: stop simulatie

2x ADC failure scenario

- Nieuwe onbekende failure in het volgende scenario
- Het is uw taak om hier zo effectief mogelijk mee om te gaan, door gebruik te maken van de problem-solving strategie
- We starten de simulatie op 5000ft boven zeeland, met de a/p aan in rol en pit modus
- <u>Opdracht</u>: vlieg naar het waypoint OMASA 3000 ft

Your controls [start 405]

Na uitvoeren oefening: stop simulatie

- Uitleg: ADC failure: geen ASI, ALT en VS. Verticale a/p valt terug in Pitch Hold modus
- Te herkennen door: rode kruizen door ASI, ALT en VS.
- Procedure: laat a/p aan (laterale modus naar keuze, verticale modus in pitch), maak gebruik van de UP/DOWN keys of CWS om de pitch angle te veranderen. Gebruik het standby instrument voor air data.
- We gaan dit scenario nog een keer herhalen

- Opdracht: vlieg naar het waypoint OMASA, 3000ft. Maak zoveel mogelijk gebruik van de automatisering.

  Your controls [start 405]

  Na uitvoeren oefening: stop simulatie

  U mag nu formulier 1 invullen

  PAUZEERT RECORDING

Test

START RECORDING

- Dit was het einde van de training van non-routine situaties. Nu begint de experimentele test waarin uw prestatie gemeten zal worden. Doe dus goed uw best.
- Experimentele groep: probeer de problem-solving strategie te gebruiken als mogelijk.
- Ik ben geen instructeur meer, u bent op uzelf aangewezen.
-
- Zoveel mogelijk automatisering gebruiken waar mogelijk
- We gaan evalueren hoe goed u met het systeem om kan gaan
- 6 scenario's, elk met een briefing vooraf en een paar korte vragen achteraf

Scenario 1

Briefing:

- Scenario: Approach EBBR Rwy 25L
- Chart: EBBR RNP Rwy 25L
- Begin: zuid-oostelijk van VOR FLO op 6500ft, ap aan in rol en pitch modus. Over 1 minuut komen we aan bij FLO.
- G1000: De route FLO - DIKBO - GIKNU - EBBR staat als flightplan in de G1000. NAV 1 staat ingesteld op VOR FLO, NAV 2 staat ingesteld op VOR BUB, wat vlak voor de runway ligt
- Weer: matige turbulentie, matig zicht, wind 4m/s 260
- Opdracht: vlieg deze approach tot aan GIKNU 2000ft met zoveel mogelijk automatisering als mogelijk. Zet vervolgens de a/p uit en maak een handmatige landing op runway 25L

Your Controls [start 501]

Na uitvoeren oefening: stop simulatie

- U mag nu formulier A invullen

Scenario 2

Briefing:

- Scenario: Approach EHGG Rwy 05
- Chart: EHGG RNP Rwy 05
- Begin: Oostelijk van waypoint VEXAR
- G1000: route VEXAR - TUVOX - EH509 - BANDU - EHGG staat als flightplan in de G1000. Geen NAV radio's ingesteld
- Weer: matige turbulentie, matig zicht, wind 5m/s 030
- Opdracht: vlieg deze approach tot aan BANDU 2000ft met zoveel mogelijk automatisering als mogelijk. Zet vervolgens de a/p uit en maak een handmatige landing op runway 05

  Your Controls [start 502]

  Na uitvoeren oefening: stop simulatie
- U mag nu formulier B invullen

Scenario 3

Briefing:

- Scenario: take-off EHRD Rwy 06
- Chart: FlightPlan EHRD
- Begin: Rwy 06 EHRD
- G1000: geen flightplan, waypoints als referentie in MFD. NAV1 op VOR RTM
- Weer: matige turbulentie, matig zicht, wind 5 m/s 070
- Opdracht: Vlieg het opgegeven flightplan van de chart met zo veel mogelijk automatisering als mogelijk tot aan EH161. Voor referentie is op de achterkant de route weergeven. Bij een hoogte van 500ft kan u de a/p inschakelen

  Your Controls [start 503]

  Na uitvoeren oefening: stop simulatie
- U mag nu formulier C invullen?

Scenario 4

Briefing:

- Scenario: Approach EDDL Rwy 05
- Chart: FLIGHPLAN EDDL
- Begin: Halverwege de leg EKDUS - DL525 op 4500ft, ap aan in roll en pitch hold modus
- G1000: geen flightplan, waypoints als referentie in MFD. Geen NAV radio’ s
- Weer: matige turbulentie, matig zicht, wind 5m/s 040
- Opdracht: Vlieg het opgegeven flightplan van de chart tot aan DIKMI 3000ft met zo veel mogelijk automatisering als mogelijk. Voor referentie is op de achterkant de route weergeven. Ze de autopilot uit bij DIKMI en maak een manuele landing op Rwy 05

  Your Controls [start 504]

  Na uitvoeren oefening: stop simulatie

- U mag nu formulier D invullen

Scenario 5

Briefing:

- Scenario: Approach EHEH Rwy 03 (SOPVI1D)
- Chart: EHEH RNP Rwy 04
- Begin: waypoint SOPVI op 7200 ft, ap aan in roll en pitch modus
- G1000: route SOPVI - EH586 - EH587 - SOMEM - EH588 - MITSA - ERSUL - EHEH als flightplan in G1000
- Weer: matige turbulentie, matig zicht, wind 5m/s 040
- Opdracht: vlieg de approach tot ERSUL met zo veel mogelijk automatisering als mogelijk. Zet de autopilot uit en maak een manuele landing op Rwy 04

Your Controls [start 505]

Na uitvoeren oefening: stop simulatie

- U mag nu formulier E invullen

Scenario 6

Briefing:

- Scenario: Cruise
- Chart: geen
- Begin: cruise op 5000ft, a/p en FD staan uit
- G1000: geen flightplan en geen NAV radios
- Weer: matige turbulentie, matig zicht, wind 5m/s 040
- Opdracht: Vlieg heading 120 en klim naar 6000 ft, door zo veel mogelijk gebruik te maken van de automatisering

Your Controls [start 506]

Na uitvoeren oefening: stop simulatie

- U mag formulier F invullen

PAUZEERT RECORDING

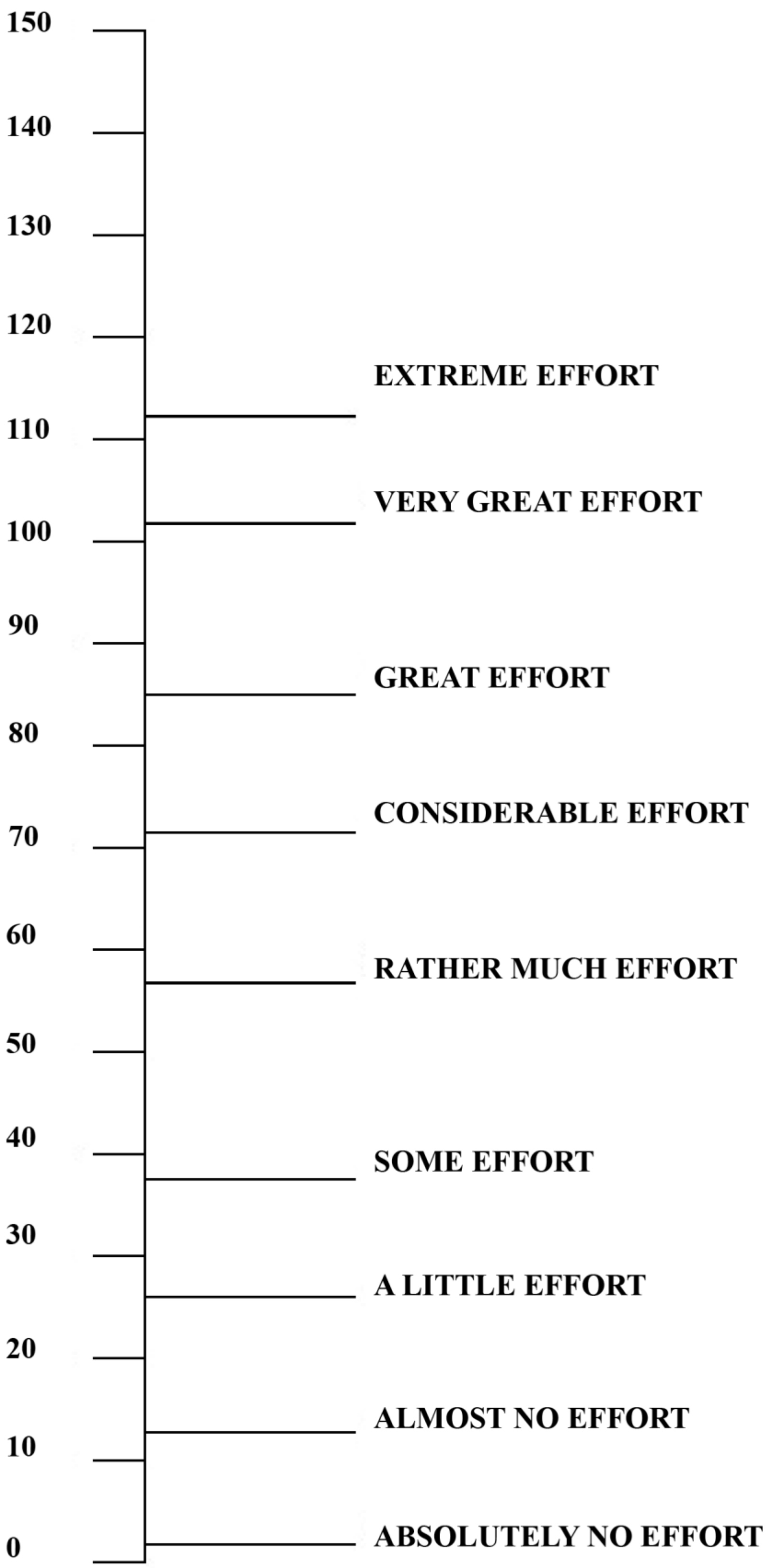
PAK FLES WIJN

I
QUESTIONNAIRES

The following pages contain the questionnaires that were used after the training (Formulier 1), and after each test scenario (Formulier A).

Formulier 1

How much mental effort did the training require? (Please place a cross on the line)



z.o.z.

	Not at all true			Somewhat true			Very true
This training was fun to do.	1	2	3	4	5	6	7
I would describe this training as very interesting.	1	2	3	4	5	6	7
This training did not hold my attention at all.	1	2	3	4	5	6	7
I thought this training was quite enjoyable.	1	2	3	4	5	6	7
While I was doing this training, I was thinking about how much I enjoyed it.	1	2	3	4	5	6	7
I thought this was a boring training.	1	2	3	4	5	6	7
I enjoyed doing this training very much	1	2	3	4	5	6	7

Formulier A

How surprised were you by the event? (With surprise we mean the extent to which the events mismatched with your expectations.)

Not at all Extremely

0	1	2	3	4	5	6	7	8	9	10

How difficult was it to understand what had happened?

Not at all Extremely

0	1	2	3	4	5	6	7	8	9	10

How much tension or anxiety did you feel during the scenario? (please place a cross on the line).

Extremely little Maximum

◆————————————————◆

To what extend did your previous experience or training help you to deal with the issue?

Not at all	Somewhat	Moderately	Much	Very much

Please describe in a few words what happened in this scenario:

____

____

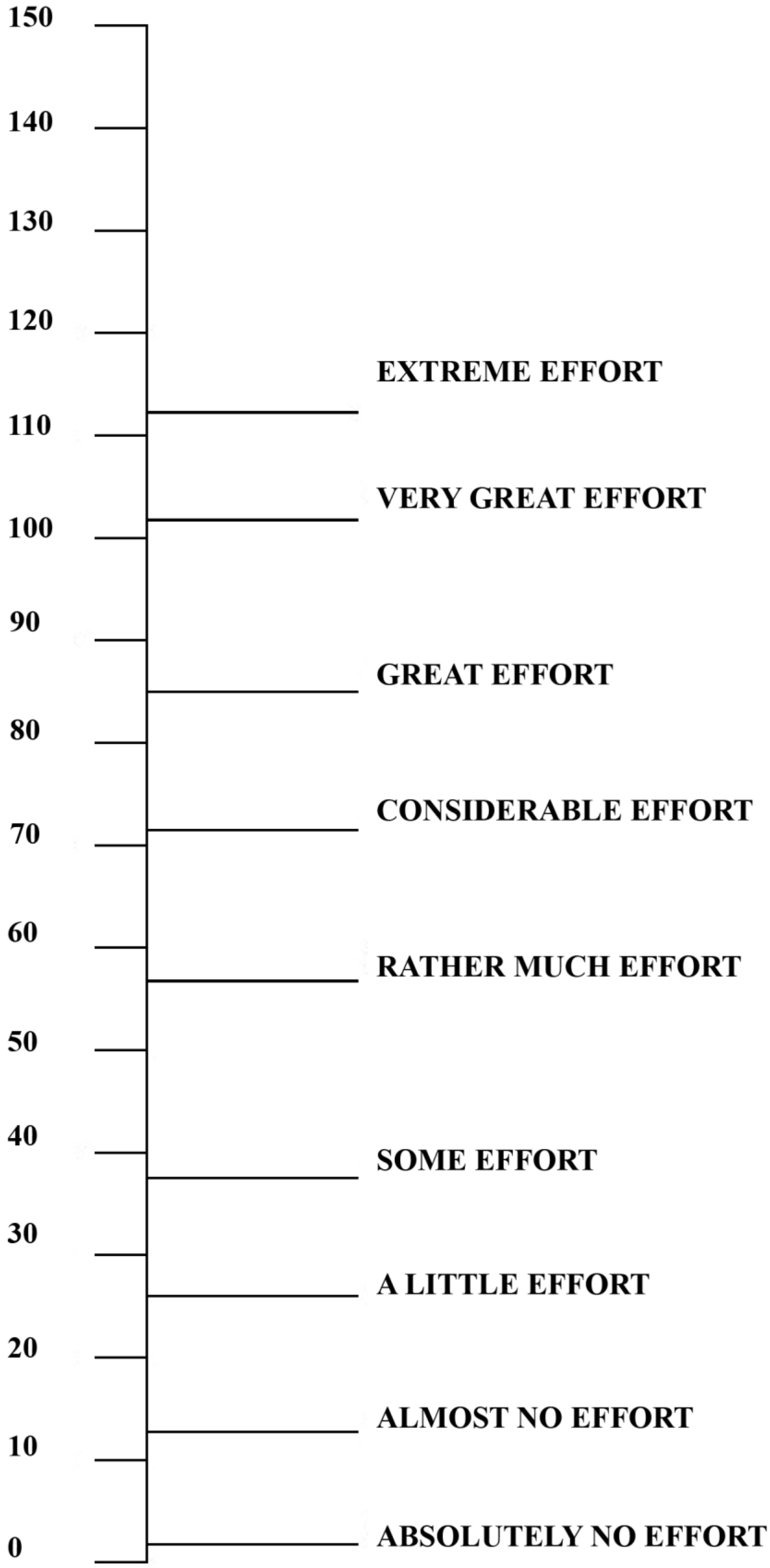
How did you respond to this?

____

____

z.o.z.

How much mental effort did the scenario require? (Please place a cross on the line.)



J
EXTENDED RESULTS

The following pages contain the extended results from the data analysis of this experiment. In Figure J.3, test scenario 4 and 2 are not used as the autopilot immediately switched to the highest level possible, making this measure irrelevant.

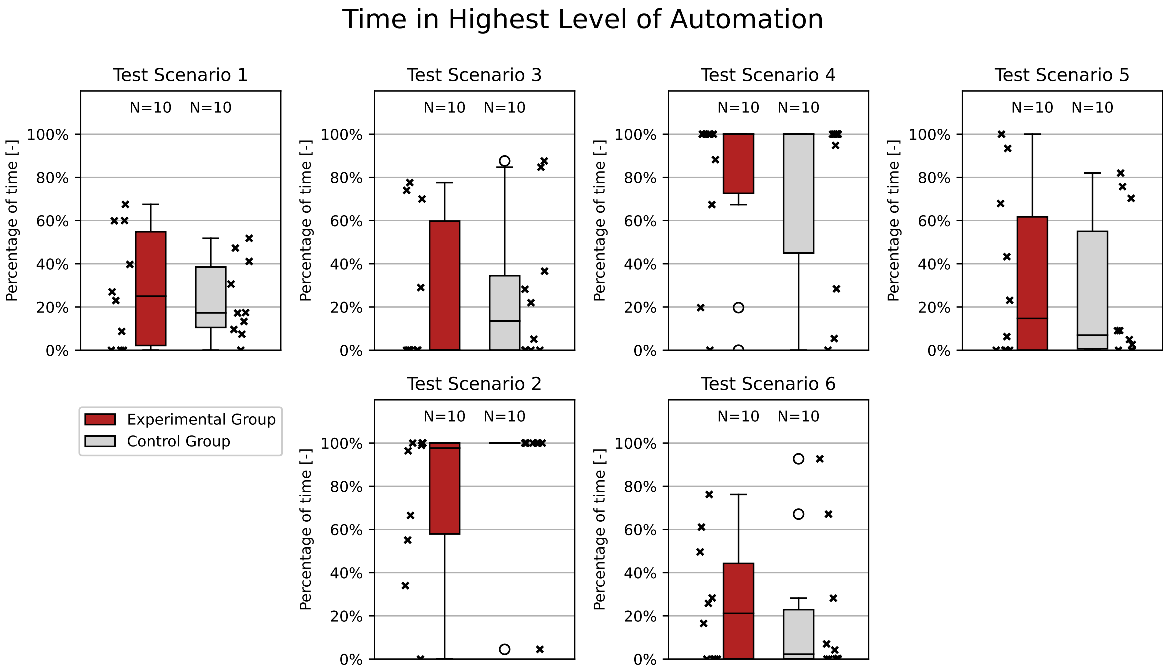


Figure J.1: Time spend in highest level of automation

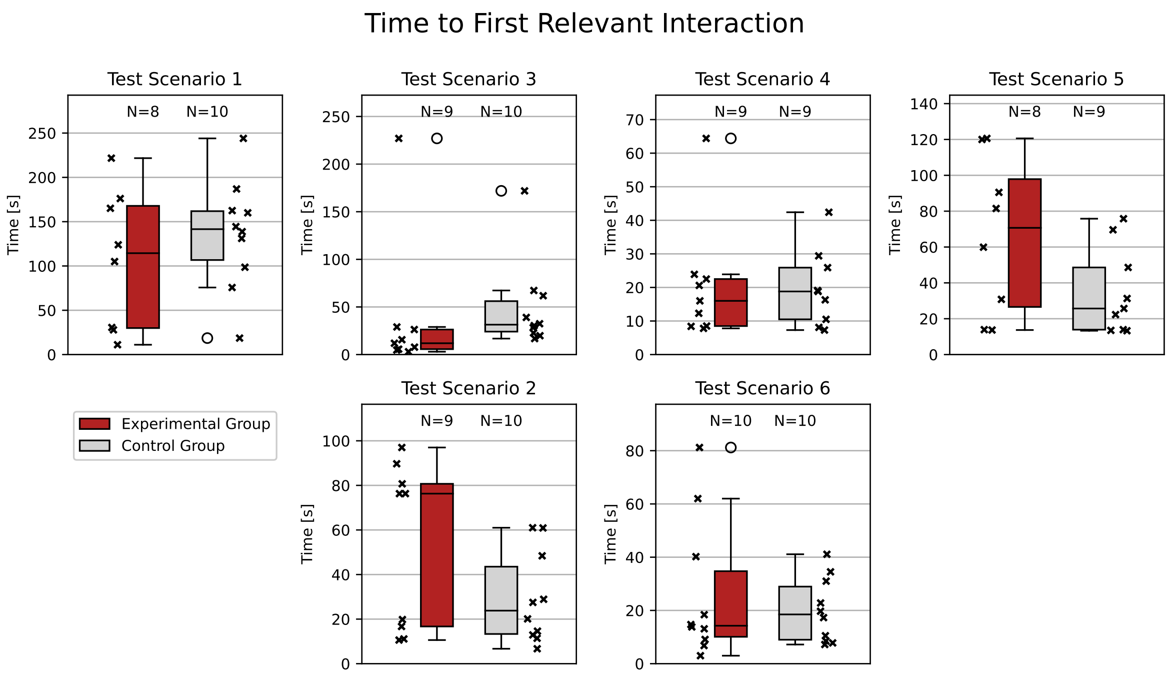


Figure J.2: Time to first relevant interaction following the failure

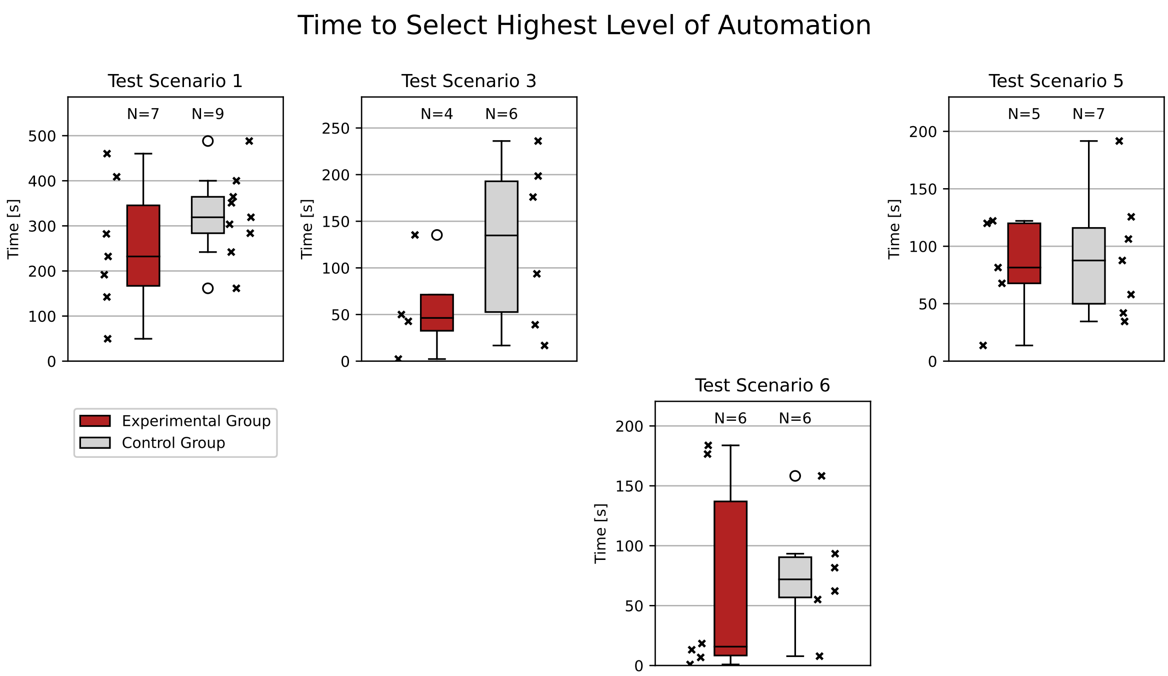


Figure J.3: Time to select highest level of automation.

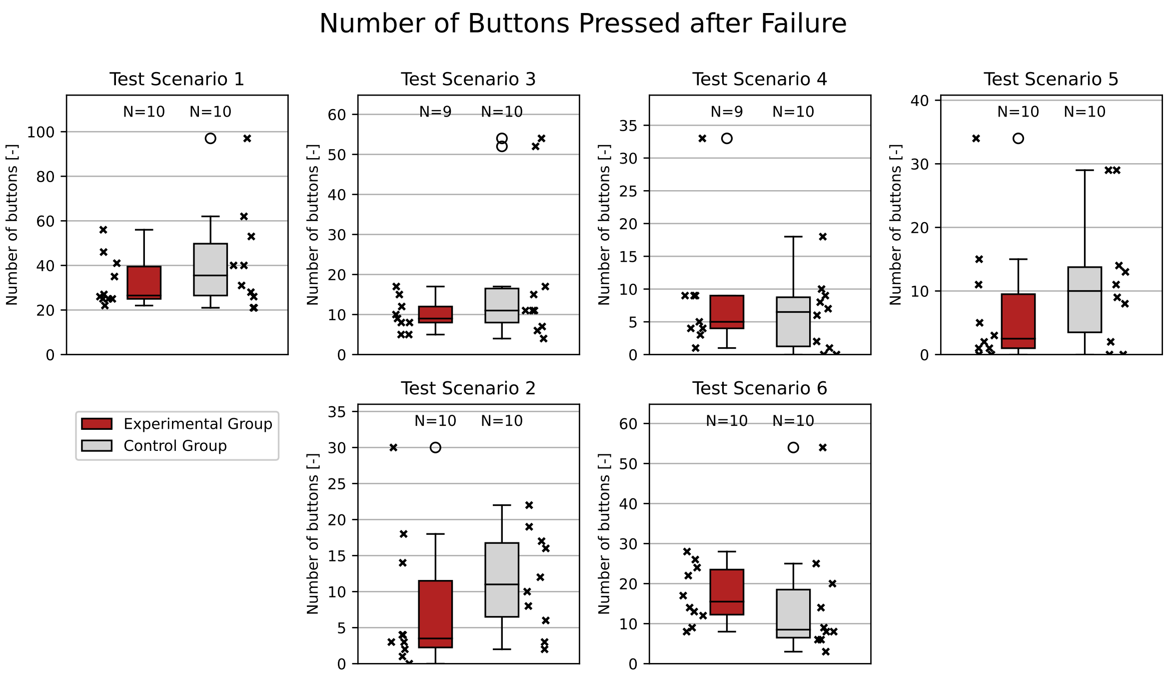


Figure J.4: Number of buttons pressed after the failure

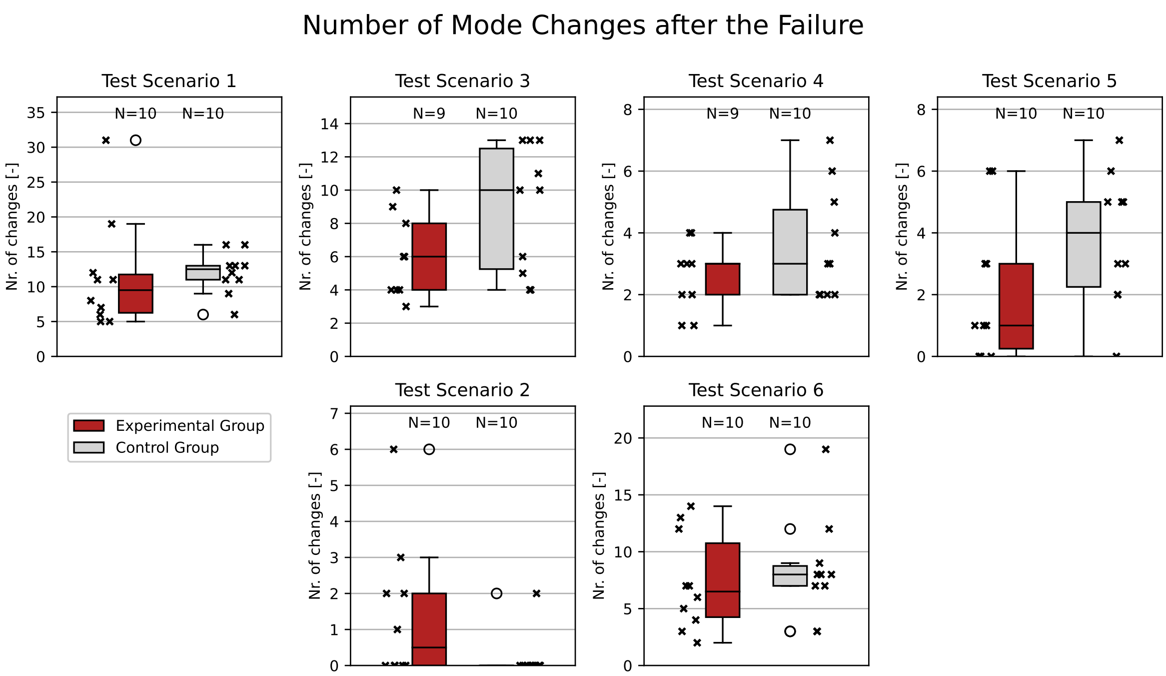


Figure J.5: Total number of Mode Changes following the failure

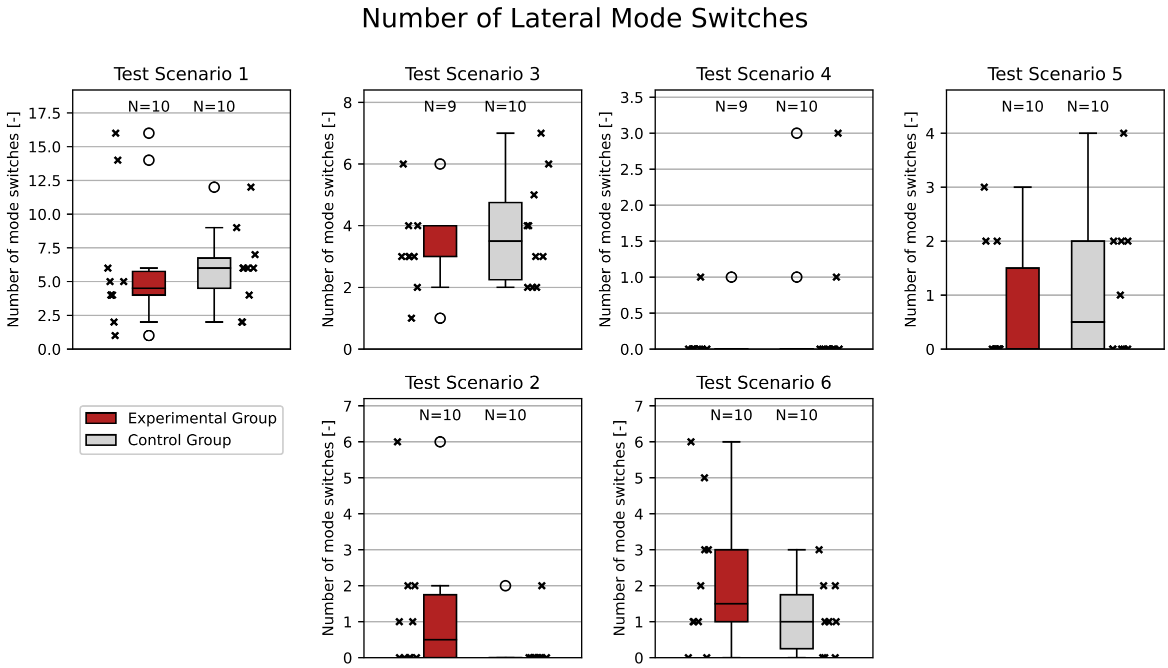


Figure J.6: Number of Lateral Mode Switches following the failure

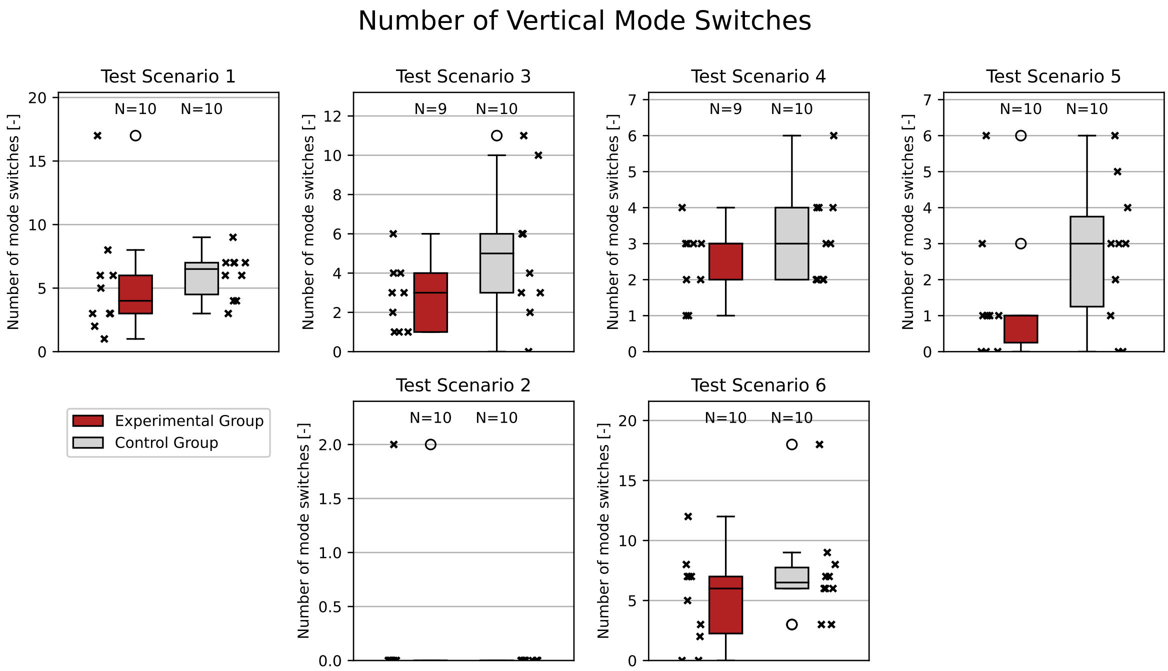


Figure J.7: Number of Vertical Mode Switches following the failure

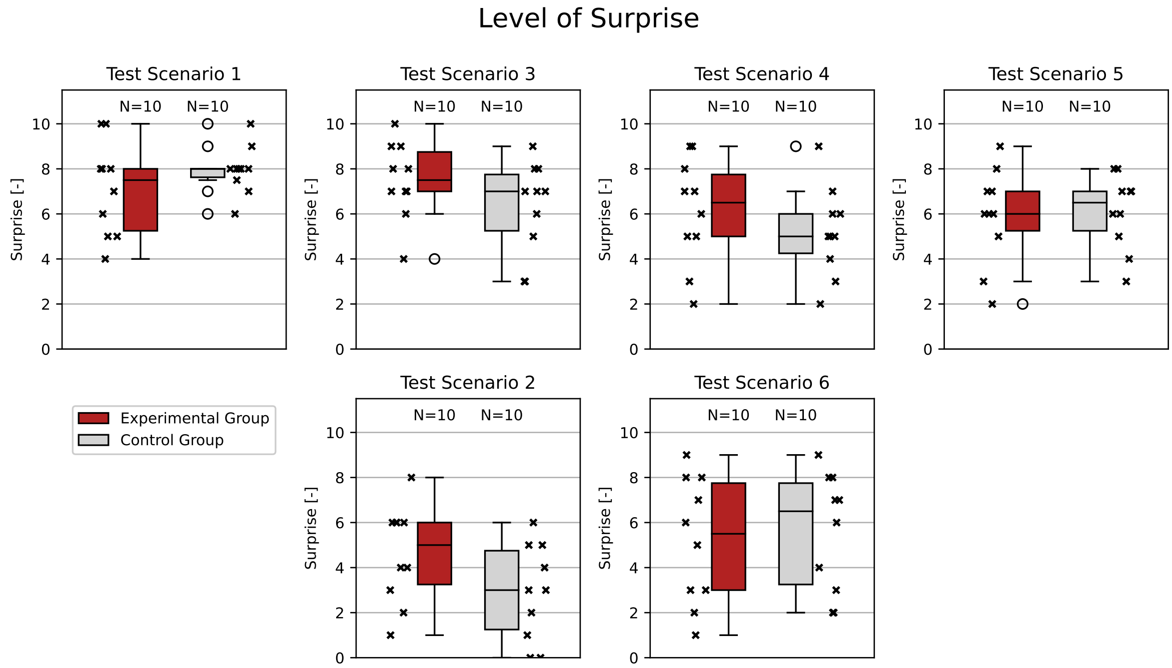


Figure J.8: Subjective measure: Level of Surprise

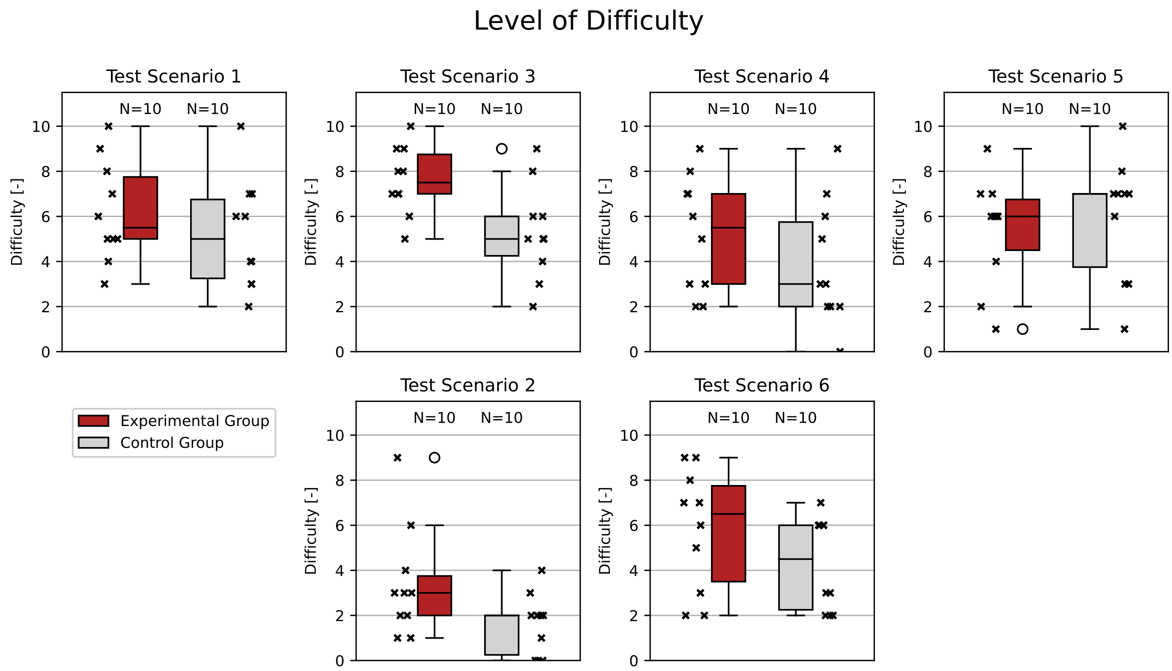


Figure J.9: Subjective measure: Difficulty of the Scenario

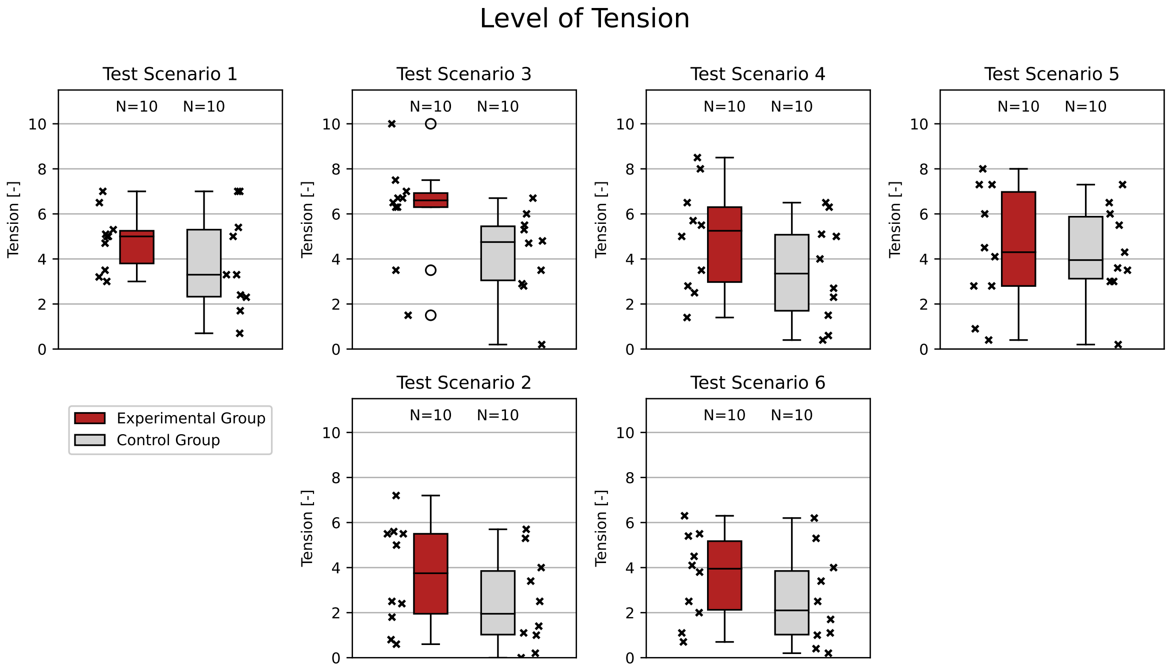


Figure J.10: Subjective measure: Experienced Tension

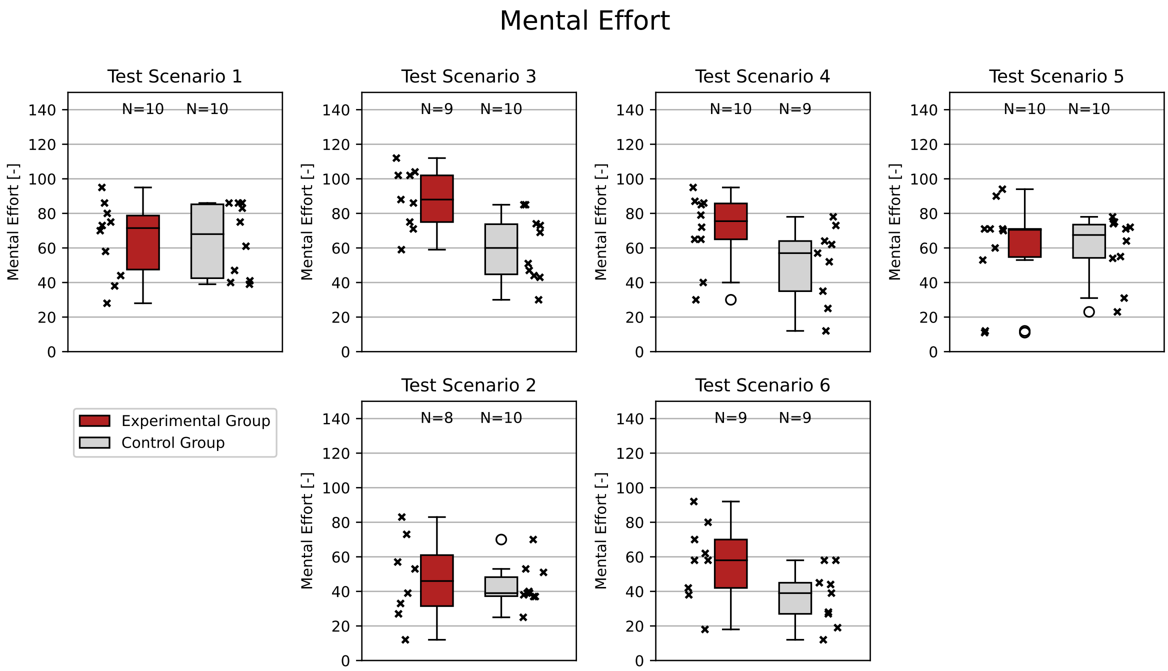


Figure J.11: Subjective measure: Mental Effort

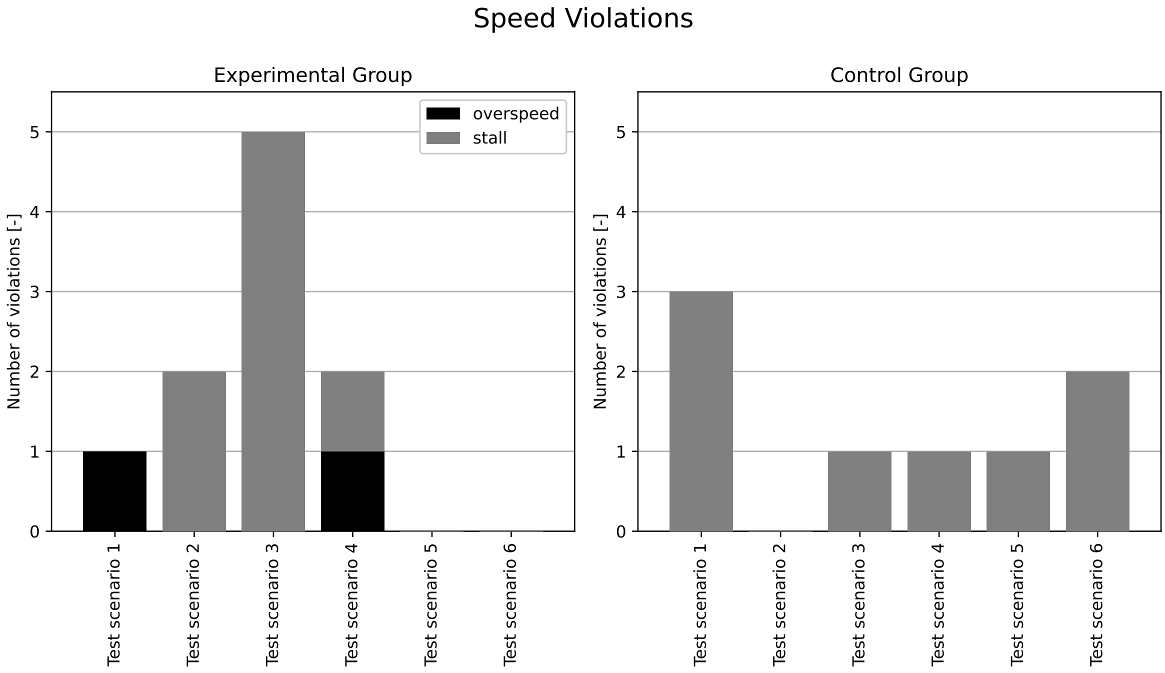


Figure J.12: Speed Violations

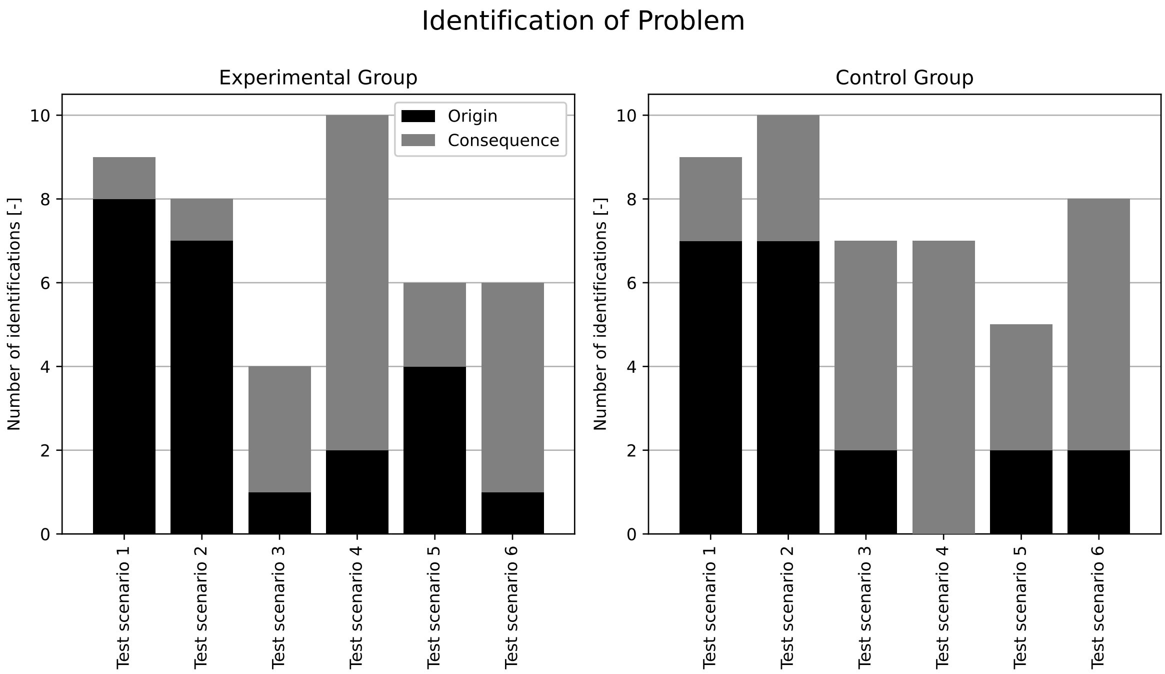


Figure J.13: Identification of the Problem

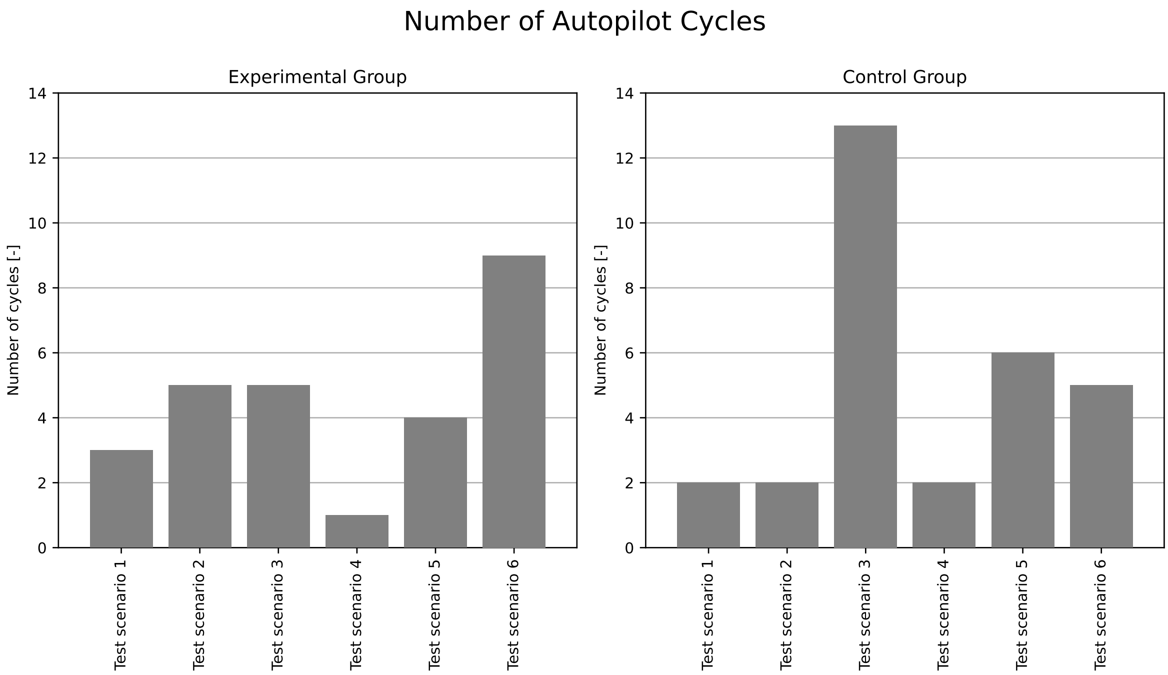


Figure J.14: Number of Autopilot Cycles

K

DUECA Project Explained

The project on the DUECA repository for this project is called *SenecaAutomationTraining*. It is based on the project *SenecaTraining* by Peter van Oorschot. It is recommended to consult his thesis *Using Unpredictability and Variety in Pilot Training to Improve Performance in Surprise Situations* (August 2017) for detailed information about his work. Here, the autopilot and the avionics (based on the Garmin G1000) are further explained. Four modules have been added to the original *SenecaTraining* project:

- **Autopilot**
- **G1000**
- **MFD**
- **GFC700GIU**

Remarks on New/Changed Modules

Autopilot (new)

This module translates all button presses, updates the autopilot settings and calculates the control surface deflections to be used in the dynamics module (PA34). The autopilot consists of three finite state machines: one for the lateral axis, one for the vertical axis and for the autopilot state itself (off, on, only flight director on, Control Wheel Steering). Based on the settings and the selected mode, the lateral finite state machine and the vertical finite state machine compute the required roll and pitch angle using PID controllers. The third final state machine computes the required aileron, rudder and elevator deflection based on these required roll and pitch angles. A first order lag term is added to smoothen this output. This module also performs all the navigation calculations for the VOR or GPS.

G1000 (new)

This module registers all the button presses and sends it to the Autopilot module, and displays the Primary Flight Display with the associated autopilot settings.

MFD (new)

This module displays the Multi Function Display, also known as the Navigation Display. It can only display a route and waypoints that have been loaded using a navigation file (more on this later).

GFC700GUI (new)

Module that spawns a graphical user interface (GUI) which can be used to interact with the autopilot. Note that this is purely for debugging, as during the experiment the buttons on the bezel of the PFD were used to interact with the autopilot.

PA34 (changed)

The dynamics module reads the autopilot state. If the autopilot is turned on, the pilot input is overwritten with the control surface deflections that are coming directly from the Autopilot module.

Malfunctions (changed)

The autopilot failures are added in this module such that they can be operated using the GUI that is created by this module.

ECI (changed)
The ECI is updated in order to deal with all the automation failures that were created for this study, and to read the updated format of the scenario files (more on this later).

B747PFD (changed)
This module has been converted in order to be used as a standby instrument, and currently does not represent a proper Boeing 747 PFD anymore. The actual PFD used in this project is created by the G1000 module.

AFCSAdapter (changed)
This module is updated to properly handle the autopilot commands. Note that the control column in the SRS will not move when the autopilot is engaged. When we tried to make the column move according to the control surface deflections calculated by the autopilot, the entire system became unstable. This is because the entire dynamics involved around the autopilot changes if the control column is introduced.

EngineSound (changed)
The sound for autopilot disconnect is added to this module.

Required Text Files
This project requires three types of text files: scenario files (.sce), navigation files (.nav) and intitial conditions files (.inco). The scenario file lay-out has been updated from the project *SenecaTraining*, and the navigation files have been added for this project. Scenario files are stored in the directory /run/run-data/scenarios. The navigation files are stored in the directory /run. Furthermore, initial conditions files are used to initialise the state of the aircraft. The way these files work is not changed (autopilot initialisation happens in the scenario files) and therefore not explained in this section.

Scenario files contain some initialisation and are used to 'program' a scenario using different events (like weather or failures). Example of a scenario file:

```
id 501
inco Clean_125kts_ebbr.inco
APstate 1
verticalFDstate 1
lateralFDstate 1
CDI 0
fp_name fp_ebbr.nav
lat_zero 0.8883539330470371
lon_zero 0.07894646163111112
alt_zero 50.0
psi_zero 0.0

windEvent
eventtime 0
enable_turb 1
turb_int 0.3
wind_vel 4
wind_dir 260
enable_windshear 0
# visibility and clouds
fg_visibility 20000
fg_cloud0_alt 10000

failureEvent
eventtime 30
GPS_failure 1
```

Please note the following: line 1 (starting with id) until 11 (starting with psi_zero) are mandatory in this exact order! They are explained below:

1. id: ID number for the scenario that will be saved by the logger
2. inco: initial conditions file that is to be loaded
3. APstate: initial state of the final state machine for the autopilot in module Autopilot
4. verticalFDstate: initial state (initial mode) of the final state machine for the lateral axis in the module Autopilot
5. lateralFDstate: initial state (initial mode) of the final state machine for the lateral axis in the module Autopilot
6. CDI: initial state for the Course Deviation Indicator (0=GPS, 1=VOR1, 2=VOR2)
7. fp_name: name of the navigation file that is to be loaded
8. lat_zero: initial latitude in radians to be used by FlightGear
9. lon_zero: initial longitude in radians to be used by FlightGear
10. alt_zero: initial altitude in meters to be used by FlightGear
11. psi_zero: initial psi in radians to be used by FlightGear

Following these mandatory lines, one can program their own scenario using events like windEvent, or engineEvent (not included in the example above) as explained by the README file in the directory /run/run-data/scenarios by Peter van Oorschot. The option to program a failureEvent was added for this project. After calling a failureEvent, one can select the trigger and the type of automation failure. Important: these scenario files should always end with a white line (otherwise things break)!

Navigation files are used to hard-code the GPS flighplan, the radios, the VORs and the waypoints. Example of a navigation file:

```
nav1active 117.950
nav1stby 108.000
nav2active 117.950
nav2stby 108.000
use_vors 0 0
vor1 0 0
vor2 0 0
nr_wps 6
D 3877.4 24347.5 -1
AGOGO 6877.4 20347.5 3000
EH671 11646.8 16837.9 -1
BLUSY 11866.1 13504.5 2000
EH661 7525.8 8581.7 2000
EHAM 0 0 -1
```

Please note the following: line 1 (starting with nav1stby) until 7 (starting with nr_wps) are mandatory in this exact order! They are explained below:

1. nav1active: frequency to be displayed on the active navigation radio 1 on the PFD and MFD (purely aesthetic)
2. nav1stby: frequency to be displayed on the standby navigation radio 1 on the PFD and MFD (purely aestetic)
3. nav2active: frequency to be displayed on the active navigation radio 2 on the PFD and MFD (purely aesthetic)
4. nav2stby: frequency to be displayed on the standby navigation radio 2 on the PFD and MFD (purely aesthetic)
5. use_vors: this is followed by two binary numbers to indicate if VOR1 and VOR2 are used respectively
6. vor1: location (x and y) in meters from the start of the scenario (as defined by the inco file) where the VOR 1 beacon is located
7. vor2: location (x and y) in meters from the start of the scenario (as defined by the inco file) where the VOR 2 beacon is located
8. nr_wps: the number of waypoints that will follow. These are the actual waypoints for the flightplan that will be drawn on the MFD and that will be used for the navigation calculations in the Autopilot module.

This is followed by the waypoints that are used for the flightplan. Each waypoint needs for entries: the name, the x location in meters, the y location in meters and the altitude in feet. This altitude can be used by the VNV mode to fly RNAV approaches. If one does not want an altitude with a waypoint, the value of `-1` should be given. It is also possible to add waypoints without adding them to the flighplan. This can be done in a similar way, by adding the line `nr_map_items` followed by the waypoints.

POSSIBLE FUTURE IMPROVEMENTS

It was found during the experiment that it is not possible to change any autopilot settings when the simulation is paused (holdCurrent). The autopilot module can be improved by correctly handling all the button presses during the holdCurrent.

Currently, the radios cannot be changed and serve no purpose. One could expand the avionics such that it is possible to select radio beacons with their frequency. This would require a lot of work as currently the beacons are fixed by the navigation files (and therefore thus hard-coded for each scenario).