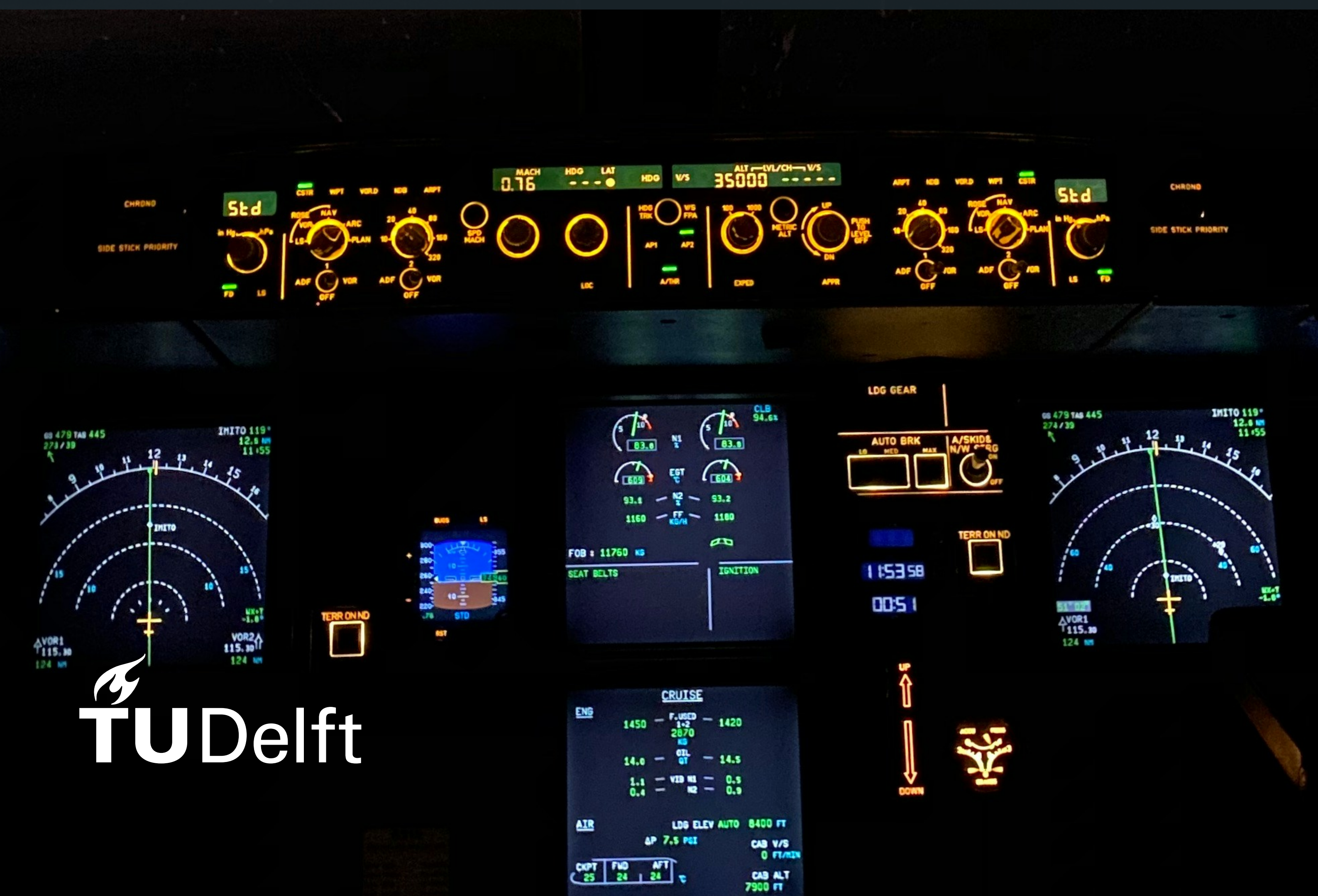
Determining the Effects of Performance Pressure on Learning During Recurrent Pilot Training

MSc Thesis

A. Sepulcri



TU Delft

Determining the Effects of Performance Pressure on Learning During Recurrent Pilot Training

MSc Thesis

Thesis report

by

A. Sepulcri

to obtain the degree of Master of Science
at the Delft University of Technology
to be defended publicly on June 13, 2024 at 13:00

Thesis committee:	
Supervisors:	Dr. Ir. M.M. van Paassen
	Prof. Dr. Ir. M. Mulder
	Dr. H.M. Landman
	Ir. O. Stroosma
External examiner:	Dr. F. Yin
Place:	Faculty of Aerospace Engineering, Delft
Project Duration:	June, 2023 - June, 2024
Student number:	4803493

An electronic version of this thesis is available at https://repository.tudelft.nl.

Faculty of Aerospace Engineering · Delft University of Technology

Preface

After six years of studying my time has finally come to write one of these prefaces. If you had told me when I first moved here from Hong Kong that I would become so passionate about a topic as unglamorous as pilot training, I would have been extremely disappointed in my future self. Yet this is one of those things that studying Control and Simulation and working on a thesis for a year will do to you. I would therefore like to invite you, dear reader to enjoy my master's thesis "Determining the Effects of Performance Pressure on Learning During Recurrent Pilot Training".

This thesis would never have been possible without many people whom I'll list off now. Firstly I would like to thank the seventeen pilots who participated in my experiment without whom it would have been impossible to complete the experiment. I would also like to sincerely apologise to those who were sorted into the high-pressure group for my enthusiasm in raising their state anxieties.

I would also like to thank my supervisors: René, Max, Olaf and Annemarie for their guidance and expertise during this thesis. It has been an immense privilege to work with the all-star team of the Aerospace Engineering faculty. I appreciate your patience and keeping your faith in me however many times it may have felt misplaced. Next I would like to thank all my friends I met here that helped me turn the Netherlands into somewhere that I'm proud to call home now. Finally, I would like to thank my parents and my brother for unconditionally supporting me all these years. Every time I felt like giving up you were what kept me going.

Andrea Sepulcri
Delft, 23 May 2024

Contents

Nomenclature

List of Abbreviations

ATPL	Airline Transport Pilot Licence
CA	Competent Authority
CRM	Crew Resource Management
DUECA	Delft University Environment for Communication and Activation
EASA	European Union Aviation Safety Agency
EBT	Evidence-Based Training
EU	European Union
PFD	Primary Flight Display
RSME	Rating Scale Mental Effort
SET	Simulation for Experiential Training
SOP	Standard Operating Procedure
SRS	SIMONA Research Simulator
STAI	State-Trait Anxiety Inventory

List of Figures

List of Tables

Part I

Scientific Paper

Determining the Effects of Performance Pressure on Learning During Recurrent Pilot Training

Andrea Sepulcri
Faculty of Aerospace Engineering
Delft University of Technology
Delft, The Netherlands

***Abstract*—Recurrent training is crucial to helping pilots maintain or upgrade their flying skills. However, constantly monitoring and scrutinising pilots during training can lead to performance pressure. If pilots who are under pressure feel that their behaviour is affected it may ultimately reduce the effectiveness of recurrent training. To study the effects of performance pressure on pilot behaviour and learning, pilots (n=17) holding an airline transport pilot licence and type rating for large multi engine aircraft were invited to participate in a simulator experiment where they would learn a new autopilot system. The low-pressure group flew a series of training scenarios under no performance pressure whereas the high-pressure group flew the same training scenarios with prompts designed to induce performance pressure and raise their anxiety. It was hypothesised that the high-pressure group would be less incentivised to explore different functions of the autopilot due to the performance pressure, therefore leading to a lower quality of training which would be reflected in poorer knowledge of the autopilot system. Ultimately, whilst the high-pressure group did in fact undergo training with a significantly higher state anxiety. This was found to have no significant effect on their behaviour during training or their performance in a test scenario. The findings also suggest that the experiment would benefit from being repeated with more guidance given during training and more constrained scenarios to reduce the variation in the data.**

I. Introduction

Previous research has indicated that the most benefit for learning autopilot logic and behaviour comes when pilots are allowed to learn in an exploratory fashion rather than relying on procedures and rote instruction (van Leeuwen et al., 2024). However, traditional recurrent training has utilised a task-based approach (Burratto & Graef, 2021) emphasising the use of standard operating procedures (SOPs) which while effective for situations with objectively correct answers, ultimately limits pilots' problem-solving abilities in unknown situations.

To prevent accidents, it is crucial for pilots to regularly undergo recurrent training to maintain or upgrade their skills (EASA, 2016). It is also critical to monitor pilots during training and to report underperforming pilots for remedial training as allowing pilots who consistently underperform in training to continue flying can lead to accidents such as in the case of CommutAir Flight 4933 (NTSB, 2022), where the captain was cited by the National Transportation Safety Board (NTSB) as a principle source of the accident due to having: "*an inadequate foundation for being a Captain, which CommutAir did not effectively address*" (NTSB, 2022, p. 3).

Modern aviation regulations encourage emphasis on startle and surprise, and crew resource management training during recurrent training (EASA, 2015). However, the focus on task-based training rather than evidence-based training combined with competent authorities' low level of familiarity with modern aviation challenges turns recurrent training into a box ticking exercise (Ranieri et al., 2019), putting pressure on pilots to simply memorise SOPs rather than demonstrate a genuine understanding of technical systems aboard aircraft. The inadequacy of current training regulations is reflected by the fact that airlines report that a significant amount of air transport pilot license (ATPL) holders do not meet their basic entry requirements (European Commission, 2019).

Monitoring pilots during training and imposing consequences on them based on their performance induces performance pressure on pilots which may encourage them to adhere to SOPs rather than utilise genuine problem-solving and diagnosis skills reducing the effectiveness of exploratory training. This indicates the need to investigate the effects that performance pressure has on a pilot's ability to learn during automation training.

II. Background

The literature on the effects of performance pressure on performance and learning are mixed. As found by Eisenberger and Aselage (2009), expecting a higher reward for high performance promotes a compulsion to perform well which subsequently leads to increased perceived self-determination and increased intrinsic job interest. In a repeated manual landing task, performance pressure was found to increase anxiety and reduce attention but was ultimately inconsequential to performance (Allsop & Gray, 2014). However, in a stimulus-response task any change in performance pressure was found to have a detrimental effect to performance (Cassell et al., 2018). In an experiment studying learning scientific content under pressure by Hinze and Rapp (2014), performing retrieval practice whilst under performance pressure was detrimental to long-term retention and understanding of scientific content. However, participants were able to overcome performance pressure in a test (Hinze & Rapp, 2014).

The following hypotheses can therefore be proposed to be investigated in an experiment:

1) Pilots who underwent training with increased performance pressure will be less inclined to explore and

utilise different functions of the autopilot in the training scenarios

2) Pilots who underwent training with increased performance pressure will receive a lower quality of training which will be reflected by a poorer performance in the testing scenarios

The first hypothesis tests the effects that performance pressure has on the behaviour of pilots during exploratory training. Pilots who undergo training with performance pressure may be less inclined to attempt using different autopilot modes, as they may feel incentivised to adhere to utilising modes that they have previously learnt to use and know will behave as expected. The second hypothesis tests the effects that the high pressure training would have on the pilots when undergoing the test scenarios. Pilots who explore less in training would have a worse understanding of the autopilot functions and therefore when encountering instrument failures would be less adept at resolving them due to their poorer understanding.

III. Method

A. Experiment Design

The experiment conducted was a between-subjects experiment with participants being divided into two groups. Both groups flew the same scenarios in the same order. The low-pressure group underwent the briefing and training phase under a low pressure manipulation, whilst the high-pressure group received prompts intended to elevate their anxiety into a high pressure state. All participants underwent the test with the same prompts and instructions.

B. Participants

For this experiment, active airline pilots (n=17) were invited to participate. All pilots involved were active airline pilots with the exception of two pilots. All pilots possessed an ATPL and were type rated for commercial aircraft, however no participating pilots had any experience operating the Garmin G1000 with the GFC700 autopilot which the simulation was based on. The participants were divided into two groups balanced over three factors: age, total flight hours and trait anxiety (STAI-T) measured using the State-Trait Anxiety Inventory from Spielberger et al. (1983). All measures were reported by the pilots ahead of their experiment sessions. A comparison of the two groups can be found in Table I.

TABLE I: Comparison of the low-pressure and high-pressure group for age, flight hours and trait anxiety

	Low-pressure (n=8) Mean (SD)	High-pressure (n=9) Mean (SD)	p
Age (years)	50.6 (8.44)	52.8 (6.27)	0.581
Flight experience (Hours)	12400 (4760)	13700 (3190)	0.526
STAI-T (20-80)	27.1 (5.93)	29.8 (11.8)	0.597

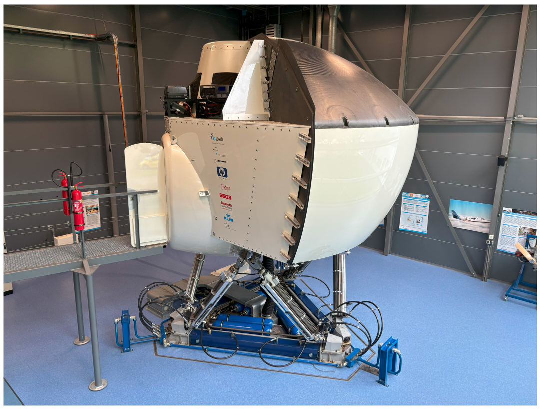


Fig. 1: The SIMONA Research Simulator

Fig. 2: The interior of the SRS cabin used during the experiment

C. Equipment

The experiment was performed in the SIMONA Research Simulator (SRS) at the aerospace engineering faculty at the Delft University of Technology (Stroosma et al., 2003) shown in Figure 1. The SRS is a full motion flight simulator featuring a six degrees-of-freedom hydraulics motion system. For the purposes of this experiment the interior was configured with a control column, rudder pedals, throttle and electrical pitch trim shown in Figure 2. The outside visuals were rendered using FlightGear on a collimated display with a 180 $^\circ \times$ 40 $^\circ$ field of view. Participants were able to communicate with the experiment organiser in the control room at all times via a headset.

The aircraft modelled in the simulator was a Piper PA-34 Seneca III which was previously used in other experiments testing startle and surprise, and automation training. The avionics used were based on the Garmin G1000 Primary Flight Display (PFD) and Multi Function Display (MFD) simulation utilised in previous automation training experiments (van Leeuwen et al., 2024). The avionics were extended with features from the Garmin G1000 pilot's guide as needed for the experiment.

D. General Experiment Procedure

The full procedure of each experiment session is outlined in Figure 3. Each experiment session began with a short briefing to ensure that participants were aware of the goal of the study. To avoid confounding effects, the title of the study that was told to the participants was: “Investigating the Effectiveness of Exploratory Learning in Automation Training”. This was found to be similar enough to the real goal of the study that they would perform as desired. Subsequently, participants underwent a fifteen minute ground school presentation where they were taught the basic operation of the autopilot system. Participants were told which modes were available and how to activate them, further interactions within the autopilot were left up to the participants to discover.

This was followed by the training scenarios where the participants in the low-pressure and high-pressure group both flew training scenarios in the simulator. The low-pressure group received instructions intended to keep their state anxiety low and the high-pressure group received instructions designed to raise their state anxiety as detailed in Section III-H. Before each scenario, the participants received a flight plan with waypoints, courses, altitudes and distances that they were instructed to fly to the best of their abilities. The training phase was further split up into two phases. In the first phase, the aircraft flew without any instrument failures to help the participants learn and familiarise themselves with the operation of the autopilot. In the second phase, the participants encountered instrument failures where they were asked to continue flying using the next highest mode of automation available. To prevent confounding effects the pilots were not told what the instrument failures were or when they would happen rather they were told in the briefing that there was a possibility for an instrument failure to occur during any scenario. After each training scenario, pilots were then asked to rate the mental effort required for the scenario using the rating scale mental effort (RSME) from Zijlstra and van Doorn (1985). At the end of the training scenarios, both groups then rated their state anxiety throughout the training phase using the state-trait anxiety inventory state form (STAI-S) from Spielberger et al. (1983) to ensure the pressure manipulation was successful.

Subsequently, both groups moved on to the testing phase where they all flew the same testing scenarios. After each testing scenario, the pilots were asked again to fill out an RSME form. Pilots were also asked several questions regarding their behaviour during the scenario and their motivations for the actions they took. At the end of the testing scenarios, pilots were again asked to fill in an STAI-S form to evaluate their anxiety during testing. To end the session, all participants completed a ten question multiple-choice test designed to test their knowledge of the automation based on the scenarios that occurred during training and testing as well as other non-scenario specific general questions. It was chosen to have this paper test at the end of the experiment to avoid confounding effects from the participants possibly learning new information during the paper test.

TABLE II: Overview of the available vertical modes ranked from lowest to highest complexity

Autopilot mode	Description
Pitch hold	Hold the set pitch angle
Altitude hold	Hold the current altitude
Vertical speed	Climb or descend at the set vertical speed
Flight level change	Climb or descend at the set airspeed
Vertical navigation	Follow the programmed vertical profile

TABLE III: Overview of the available lateral modes ranked from lowest to highest complexity

Autopilot mode	Description
Roll hold	Hold the set roll angle
Heading select	Fly at the selected heading
Navigation (VOR)	Intercept the VOR beacon at the set radial
Navigation (GPS)	Follow the programmed flight plan

E. Available Autopilot Modes

In the ground school, the pilots were told the available autopilot modes in the experiment ranked in terms of complexity as shown in Table II and Table III. The level of complexity was determined based on how much workload and pilot input was needed to operate the mode.

In the vertical modes, pitch hold was ranked the lowest complexity as it simply held a set pitch angle but did not account for any further parameters. One level higher was altitude hold which albeit useful to hold a constant altitude also did not provide any more benefits. The next highest mode was vertical speed mode which climbed and descended at a set vertical speed. This was deemed to be a higher level of automation since pilots could be more precise in their climb and descent rate compared to pitch hold. One level higher than vertical speed mode was flight level change which flew the aircraft at the set airspeed. This was considered to be a higher level of automation since it would account for airspeed avoiding stalling or over-speeding and therefore provided safety benefits when compared to vertical speed mode. Finally, the highest vertical mode of automation was vertical navigation which followed the vertical profile in the flight plan. This was the most complex mode since the pilot did not have to take any action when flying the aircraft in this mode.

In the lateral modes, roll hold was the least complex mode since it flew the aircraft at a set roll angle. One level higher was heading select mode which flew at the set heading which whilst useful did not provide any navigational benefits such as compensating for crosswinds. VOR navigation was the next highest mode as it intercepted the VOR beacon at the desired radial and thus would provide more precise navigation. Finally, GPS navigation was the most complex lateral mode as it would simply fly the waypoints according to the flight plan therefore requiring the least input from the pilot. All participating pilots agreed with the rankings of the modes in terms of complexity.

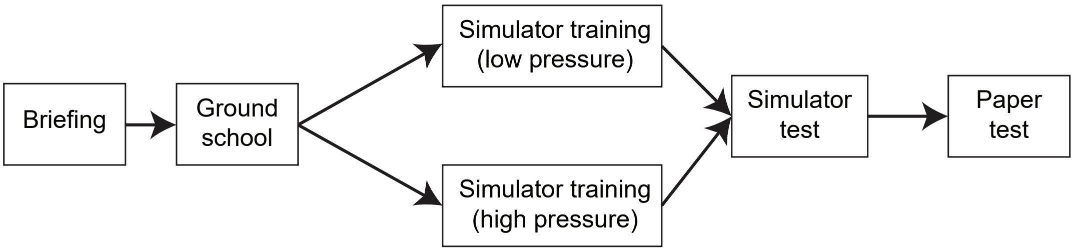


Fig. 3: Flowchart showing the procedure of every experiment session

F. Scenario Design

The design of training scenarios is crucial in order to enable effective evidence-based training (Dahlstrom et al., 2022). Moreover, in this experiment the training scenarios would be the only opportunities for pilots to learn how to operate the autopilot ahead of the testing scenarios. An effective scenario design was therefore dependent on several requirements that were laid out before the experiment. Firstly, no scenarios would have an optimal solution of manual flight. Since pilots are trained to initially disconnect autopilot and fly manually when encountering automation failures, it was decided that scenarios with manual flight as a solution would not provide any meaningful results or differences between the two groups.

Furthermore, all training scenarios had to be linked to one of the eventual testing scenarios. This would allow for knowledge transfer from the training to the testing scenario therefore rewarding pilots who learnt more in the training scenarios and consequently had a better overall knowledge of the operation of the autopilot. Moreover, all scenarios would contain periods of low workload where the primary focus was on monitoring the autopilot followed by a period of high workload where pilots would be required to diagnose failures and resolve them. Finally, where possible the scenarios should have multiple possible solutions to the failures where pilots could use different autopilot modes to fly the rest of the scenario again rewarding pilots who had more effective training who would select the more optimal autopilot mode. Using these guidelines, six autopilot training scenarios were designed. Three scenarios focused on using the vertical modes of the autopilot and three scenarios focused on using the lateral modes. Moreover, the highest available mode of automation at any point in the scenarios was not necessarily the highest autopilot mode overall therefore encouraging pilots to explore and utilise the different modes. The flight plans for these scenarios can be seen in Figure 4. A more detailed explanation of the scenarios is given in the following section.

G. Training and Testing Scenarios

A summary of all the scenarios and associated failures is shown in Table IV. A summary of the highest automation mode available before failure and the highest automation mode available after failure is given in Table V.

1) Introductory Manual Flight: In the first training scenario, pilots flew a short manual flight scenario with instructions from the control room to allow them an opportunity to familiarise themselves with the aircraft model used in the simulator had they decided to switch to manual flight for safety at any point during the experiment.

2) Vertical Mode Familiarisation: Subsequently, pilots flew another familiarisation scenario which involved a descent from 6,000 feet to 2,000 feet. Lateral navigation could be handled using GPS navigation. Vertical navigation was available for the first three waypoints after which pilots could complete the descent using any of the other available vertical modes.

3) Lateral Mode Familiarisation: In the final familiarisation scenario, pilots navigated a short route at a constant altitude. The altitude could be maintained using altitude hold. GPS waypoints were available for the first three waypoints after which pilots would have to select the correct VOR frequency and navigate the final two waypoints using VOR navigation.

4) Lateral Mode Training: In the lateral mode training scenario, pilots navigated several waypoints using VOR beacons. After the second VOR waypoint the VOR receiver would fail rendering it unusable. Pilots who navigated the waypoints well could then see that a GPS route was also available and navigate the rest of the scenario using GPS. However, pilots who had missed certain waypoints by too great a distance were unable to use GPS navigation for the rest of the scenario and therefore had to fly using heading select which would be less accurate.

5) Vertical Mode Training: The final training scenario was the vertical mode training scenario. Pilots flew a climbing scenario from 2,000 feet to 6,000 feet where at 4,192 feet of altitude the pitot tube would clog causing the airspeed to become unreliable and rapidly increase as the altitude increased. If the aircraft was in flight level change, the flight director would begin indicating an increasingly higher pitch angle threatening to stall the aircraft if appropriate action was not taken. Since the altitude tape was still functioning as expected, pilots could then switch modes and complete the rest of the scenario climbing in vertical speed mode.

6) Lateral Mode Testing: In the first testing scenario, the pilots flew a route navigating using GPS. Just before the third waypoint, the GPS would silently fail causing the map on the MFD, and the ETA and distance to the next waypoint to freeze. If no action was taken the aircraft would then hold the heading it was flying at and overshoot the waypoint. Pilots would therefore have to recognise that the GPS had failed and take action. This could be seen from the fact that the MFD would be frozen along with the estimated time

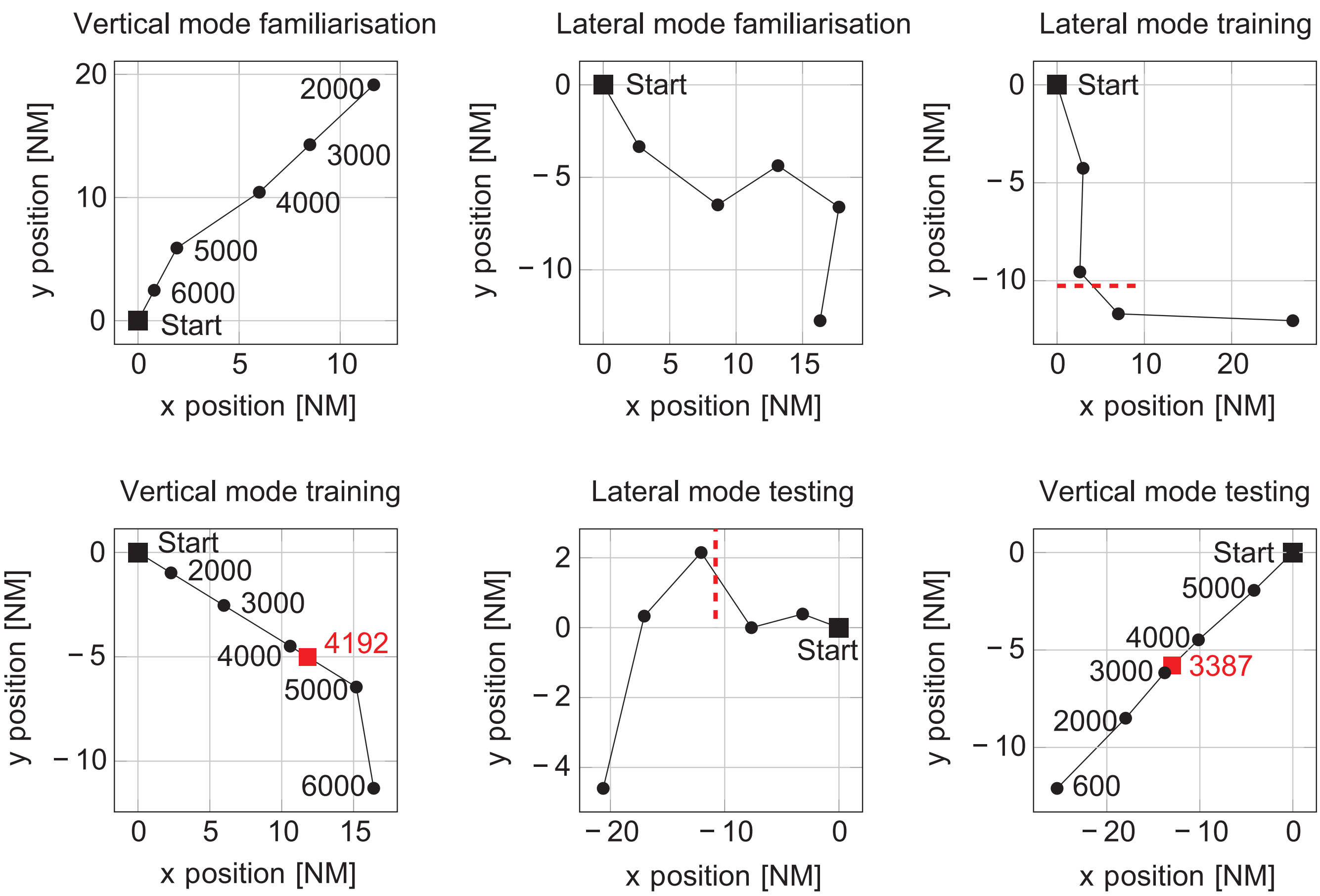


Fig. 4: All the training and testing scenarios used in the experiment as viewed from above. Altitude for each node is indicated if the waypoint altitudes were included in the flight plan. The square nodes indicate the beginning of each scenario and the red nodes/lines indicate locations where the instrument failures occur

of arrival and distance from the waypoint never updating contradicting the information from the distance measuring equipment (DME). Learning to read information about the GPS routes and using the DME was possible in the lateral familiarisation and training scenarios. The two standby VORs were tuned to the final two waypoints of the scenario and therefore the optimal action was to change the navigation source to VOR and use the DME information to ascertain their relative position.

7) Vertical Mode Testing: In the second testing scenario, pilots were tasked with a descent from 5,000 feet to 600 feet. At 3,387 feet the static port would clog causing the altitude tape and vertical speed indicator to freeze and the airspeed indicator to increase as the altitude decreased. In vertical navigation mode, this would cause the aircraft to continue descending overshooting the altitude of the next waypoint. The optimal solution was therefore to disregard the instruments on the PFD and instead rely on the altitude and airspeed indication given in the standby instruments and complete the scenario using pitch hold mode and manually changing the pitch reference to complete the descent. In the vertical training scenario, pilots would have learnt that the flight level change mode would not work with an unreliable airspeed indication and that the vertical speed mode relied purely on the altitude indication making both of those modes unviable.

TABLE IV: Training and testing scenarios and failures

Scenario	Failure
Introductory manual flight	Nominal flight
Vertical mode familiarisation	Nominal flight
Lateral mode familiarisation	Nominal flight
Lateral mode training	VOR receiver failure
Vertical mode training	Blocked pitot tube
Lateral mode testing	GPS failure
Vertical mode testing	Blocked static port

H. Pressure Manipulation

The only difference between the low-pressure and high-pressure group was the performance pressure manipulation that was applied to them during the briefing, ground school and training scenarios. The pressure manipulation was based on interviews with active pilots and from experiences in previous

TABLE V: The highest possible mode per training and testing scenario and the corresponding optimal solution after instrument failure

Scenario	Highest Mode Available	Highest Mode After Failure
Lateral mode familiarisation	Navigation (VOR)	Navigation (GPS)
Vertical mode familiarisation	Flight level change	Vertical speed
Lateral mode testing	Navigation (GPS)	Navigation (VOR)
Vertical mode testing	Vertical navigation	Pitch hold

experiments. This pressure manipulation was applied in two phases of the experiment, firstly during the briefing and the ground school, and then again in the training phase.

During the briefing, participants in the high-pressure group were told that they would be monitored and their data would be recorded for analysis throughout the whole experiment, and that the number of training scenarios they would undergo would depend on their performance during the training scenarios. During the training phase, participants in the high-pressure group were once again reminded before each scenario that their data were being recorded for analysis and that the number of training scenarios they underwent was dependent on their performance. Moreover, in the low-pressure group the participants were given all the training scenario flight plans at once whereas in the high-pressure group participants first received the two familiarisation flight plans in a folder with "Training" written on it. After completing those, they were told that due to their performance they would be required to undergo some more training scenarios and received a new folder with "Extra" written on it containing four more flight plans.

By telling the high-pressure group that their performance was being monitored and recorded for analysis, it was anticipated that they would feel pressured about being judged based on their performance and raise their concerns about making mistakes or not performing well. Moreover, it was also anticipated that they would feel pressured to perform well as they might be compared with their peers and thus have to meet certain expectations. The second pressure manipulation of telling them that the number of training scenarios they would undergo being dependent on performance was also anticipated to work similarly to the first. This provided the pressure of having to meet certain standards before being able to progress to the next phase of the experiment.

At the end of the experiment, all participants were told the real title of the study and the high-pressure group were told that they were sorted into the high-pressure group and the prompts heard in the experiment were unrelated to how they may have actually performed.

I. Dependent Measures

The aircraft states and pilot inputs were logged for the duration of the experiment for analysis. The participants were also asked to fill in subjective questionnaires during the experiment and were interviewed after the testing scenarios for the motivations of their actions. These dependent measures are divided into five general categories, those being: exploration, problem-solving and diagnosis, performance, system knowledge and subjective measures. The measures used in each category are explained below.

- **Exploration:**
  - Number of mode switches in the familiarisation scenarios
  - Number of mode switches before failure in the training scenarios
- **Problem-solving and diagnosis:**
  - Time after failure to switch modes
  - Time after failure to select final automation mode
  - Number of mode switches after failure
- **Performance:**
  - Final automation mode selected
- **System knowledge:**
  - Paper test score
- **Subjective measures:**
  - State anxiety in training
  - Mental effort

To test the first hypothesis that pilots undergoing training under increased performance pressure will be less inclined to explore, the exploration category of data was used. If the pilots in the high-pressure group were in fact less willing to explore different autopilot modes, this would be reflected by the fact that they would switch modes fewer times in the familiarisation scenarios and before failure in the training scenarios.

To test the second hypothesis that pilots who underwent training under increased performance pressure receive a lower quality of training the problem-solving and diagnosis and system knowledge categories were used. Pilots who received a lower quality of training would theoretically be less familiar with the intricacies of the autopilot. Therefore, if a failure occurred they would firstly take longer to diagnose it and switch modes and subsequently take longer to resolve the issue and possibly erroneously utilise several different modes resulting in a longer time to select the final mode of automation and switching modes more often after failure. Moreover, a lower quality training may be reflected by a worse system knowledge leading to those participants scoring lower in the paper test.

Finally, to see if a lower quality of training results in poorer performance, the final mode selected by pilots after failure can be analysed. Pilots who perform worse would be less likely to

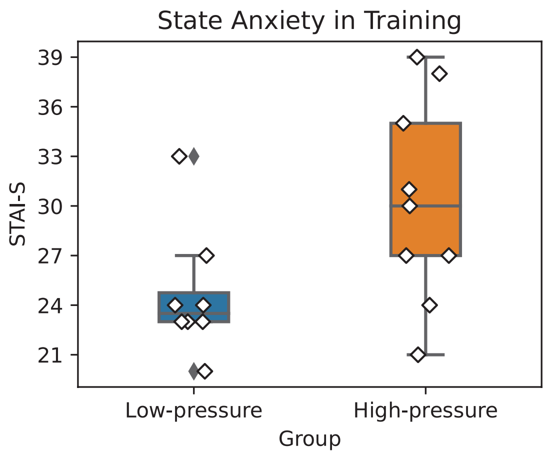


Fig. 5: State anxiety in training

select the correct next highest mode of automation available after an instrument failure.

IV. Results

A. State anxiety

In order to test whether the pressure manipulation was successful, at the end of training all participants were asked to rate their state anxiety using the state-trait anxiety inventory (Spielberger et al., 1983). A Mann-Whitney U Test indicated that the high-pressure group (mean=24.6, SD=3.64) did have a significantly higher state anxiety during training compared to the low-pressure group (mean=30.2, SD=5.83) (U = 15.0, $p < 0.0469$) the data are shown in Figure 5.

B. Exploration

To determine the difference between the two groups in terms of exploratory behaviour, the number of mode changes performed by each group in the familiarisation scenarios and the number of mode changes before failure in the training scenarios were logged. In the scenarios focusing on the lateral modes, only the lateral mode changes were logged and vice versa with the vertical modes. All the data are summarised in Table VI. In the vertical and lateral familiarisation scenarios, the low-pressure group switched modes more frequently compared to the high-pressure group. Whereas in the lateral and vertical training scenarios the high-pressure group switched modes more frequently before failures. A Mann-Whitney U test showed that the difference was not significant in the vertical mode familiarisation scenario (U = 42.5, $p < 0.562$) or the lateral mode familiarisation scenario (U = 40.5, $p < 0.699$). The difference was not significant either in the lateral mode training scenario (U = 31.0, $p < 0.660$) or the vertical mode training scenario (U = 20.0, $p < 0.132$).

C. Problem-solving and diagnosis

To determine the pilots' diagnosis skills, the time to first change autopilot modes after instrument failure was logged. In both the lateral and vertical mode testing scenarios, the high-pressure group switched modes faster than the low-pressure group. A Mann-Whitney U test revealed no differences between the groups in the lateral mode testing scenario (U = 39.0, $p < 0.815$) or the vertical mode testing scenario (U = 51.0, $p < 0.167$). The summary for these measures is given in Table VII.

To quantify the participants' problem-solving skills, two measures were used. First, how long it took the pilot to select their final mode of automation after the instrument failure and the number of mode changes after the instrument failure. The high-pressure group selected the final mode of automation earlier in the lateral mode testing scenario whereas in the vertical mode testing scenario the low-pressure group selected the final automation mode earlier. A Mann-Whitney U test revealed that the difference between the two groups was not significant in the lateral mode testing scenario (U = 46.0, $p < 0.370$) or the vertical mode testing scenario (U = 32.0, $p < 0.743$).

When looking at the number of mode changes after failure; the low-pressure group switched between fewer modes in the lateral and vertical mode testing scenarios. A Mann-Whitney U test showed no significant differences however in the lateral mode testing scenario (U = 34.0, $p < 0.884$) or the vertical mode testing scenario (U = 30.0, $p < 0.593$). The data for both of these measures are summarised in Table VIII and Table IX.

D. Performance

In the horizontal mode testing scenario, the low-pressure and high-pressure group selected the optimal mode with the exact same frequency. In the vertical mode testing scenario, the high-pressure group selected the optimal mode more frequently. When analysed using Pearson's Chi-Square test. In both the lateral mode testing scenario ($\chi^2(1) = 0$, $p < 1$) and the vertical mode testing scenario ($\chi^2(1) = 0$, $p < 1$) the selection of the final modes was found to be completely independent of the type of training that the participants received. The data are shown in Table X.

E. System knowledge

In the paper test that the participants completed after the experiment, the low-pressure group (mean=5.13, SD=2.03) scored lower than the high-pressure group (mean=5.44, SD=2.00) on average. However, a Mann-Whitney U test showed that the difference was not significant (U = 33.0, $p < 0.808$).

F. Workload

The participant workload was measured after each scenario in the experiment using the RSME (Zijlstra & van Doorn, 1985). The high-pressure group consistently reported a lower workload on average in each scenario except for the lateral mode training scenario. However none of the differences between the groups in the scenario was found to be different. Furthermore, a Mann-Whitney U test showed no significant difference in average workload between the groups (U = 40.5, $p < 0.700$). The full data are shown in Table XI

TABLE VI: Number of mode switches in training scenarios

	Low-pressure Mean (SD)	High-pressure Mean (SD)	p
Vertical mode familiarisation	19.0 (6.10)	18.3 (6.34)	0.562
Lateral mode familiarisation	12.3 (3.49)	11.7 (4.85)	0.699
Lateral mode training (pre-failure)	8.38 (3.77)	8.89 (3.38)	0.660
Vertical mode training (pre-failure)	11.6 (2.69)	15.2 (5.01)	0.132

TABLE VII: Time taken to first switch modes after failure in the testing scenarios

	Low-pressure Mean (SD)	High-pressure Mean (SD)	p
Lateral mode testing	86.8 (71.2)	78.9 (60.9)	0.815
Vertical mode testing	16.2 (5.37)	13.5 (9.60)	0.167

TABLE VIII: Time taken to select the final mode after the instrument failure in the testing scenarios

	Low-pressure Mean (SD)	High-pressure Mean (SD)	p
Lateral mode testing	260 (126)	205 (71.9)	0.370
Vertical mode testing	93.8 (70.7)	106 (65.4)	0.743

TABLE IX: Number of mode changes after failure in the testing scenarios

	Low-pressure Mean (SD)	High-pressure Mean (SD)	p
Lateral mode testing	3.00 (1.80)	3.22 (1.40)	0.884
Vertical mode testing	3.00 (2.24)	4.00 (3.13)	0.593

TABLE X: Frequency of the final selected modes in the testing scenarios

Scenario	Group	Optimal mode	Other modes
Lateral mode testing	Low-pressure	5	3
	High-pressure	5	4
Vertical mode testing	Low-pressure	4	4
	High-pressure	5	4

TABLE XI: RSME for each group per scenario

	Low-pressure Mean (SD)	High-pressure Mean (SD)	p
Vertical mode familiarisation	27.1 (18.8)	26.7 (18.2)	0.923
Lateral mode familiarisation	41.4 (24.9)	33.3 (25.4)	0.123
Lateral mode training	35.5 (21.1)	40.2 (23.7)	0.809
Vertical mode training	47.1 (21.8)	42.1 (23.4)	0.772
Lateral mode testing	62.0 (13.4)	46.6 (23.8)	0.208
Vertical mode testing	72.9 (19.4)	60.6 (25.2)	0.630
Mean RSME	45.3 (17.2)	41.7 (19.4)	0.700

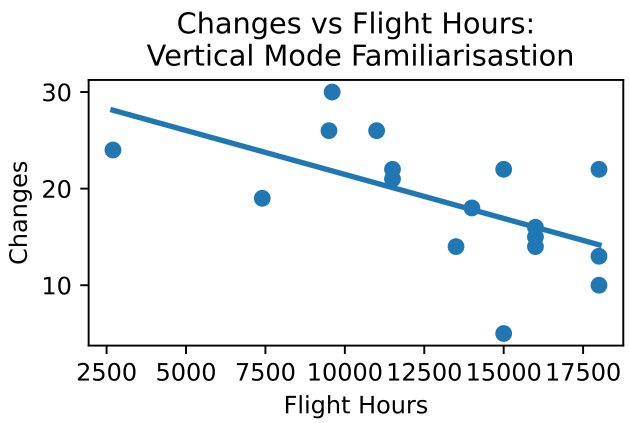


Fig. 6: Number of mode changes vs flight hours in the vertical mode familiarisation scenario

G. Additional Measures

One measure that did yield significant results was when the number of mode changes in the first familiarisation scenario was plotted against the total number of flight hours per participant. The number of mode changes made in the vertical familiarisation scenario was found to have a strong negative correlation with the number of flight hours which was confirmed when looking at the Pearson's correlation coefficient ($r = -0.593$, $p = 0.012$).

H. Scenario validation

Participants were questioned after the testing scenarios on the motivations for their actions during the testing scenarios. Their answers could be used to assess the validity of the experiment in providing a similar experience to recurrent pilot training. Many pilots successfully diagnosed the instrument failures in the testing scenarios and stated that similar instrument failures are utilised in recurrent pilot training. Moreover, when asked to motivate their actions in testing some pilots cited the training scenarios as having provided insight into how certain autopilot modes functioned influencing their choices in the testing scenarios. This indicated that for some participants, knowledge transfer from the training scenarios to the testing scenarios did occur from the exploratory training.

Some pilots commented on the lack of procedures or contact with air traffic control as being unrealistic and affecting their behaviour. This occurred for example in the blocked static port scenario where one pilot remained level until the end of the scenario and stated that they would have contacted air traffic control to obtain their altitude. However, other participants who also commented on those issues ultimately did not have their behaviour affected by them.

V. Discussion

The findings show that in the experiment that was conducted, the performance pressure that was applied to the high-pressure group did cause their state anxiety to be significantly higher than the low-pressure group. However, training under increased performance pressure ultimately did not have significant effects on the pilots' behaviour nor the quality of training. There was no significant difference in the number of mode switches between groups in the familiarisation and training scenarios pre-failure. This indicated that pilots who underwent the training with the high-pressure manipulation did not feel sufficiently disincentivised from exploring the different autopilot modes compared to the group who were not manipulated.

There was also no significant difference between the groups when it came to performance in the testing scenarios with both groups taking equally as long diagnosing and resolving automation problems, switching automation modes the same amount of times after failure and utilising the next most optimal mode equally as frequently. Furthermore, undergoing training under a high-pressure manipulation did not result in any significant differences in workload or technical knowledge.

However, this does not necessarily suggest that performance pressure has no effect on pilot behaviour or learning. Rather it is likelier that the complete sample from all the pilots contained too much variation in the data to provide meaningful results. This was due to several observations made during the experiment. Firstly, the lack of guidance during training meant that pilots formed their knowledge of the autopilot in the first two familiarisation scenarios and afterwards rarely deviated from the modes they knew how to use. This caused many pilots to hold incorrect preconceptions about the operation of the autopilot and be hesitant to explore new modes even if they were in the low-pressure group.

Moreover, the performance pressure manipulation may have not been sufficient to override personal preferences and individual differences in the pilots. For example, exploration in the first scenario which was the scenario with the most freedom, was strongly correlated to pilot experience. Furthermore, older pilots disconnected the autopilot more often than younger pilots as pilot training recommends; despite receiving the same instructions to utilise the highest mode of automation possible. When questioned about this behaviour, pilots who switched to manual flight stated that they preferred a poorer performance (regarding use of automation) rather than flying unsafely. This further indicates that whilst the performance pressure manipulation was effective in raising the high-pressure group's state anxiety, it may have been unsuccessful in providing sufficient performance incentive for the pilots to fly using the highest mode of automation possible.

The following recommendations can therefore be made for future studies into the effects of performance pressure on learning:

1) Train and test participants on a more focused task with fewer degrees of freedom to make comparisons between pilots easier
2) Conduct the experiment with a more homogeneous group of participants to minimise the effects of individual differences

3) Train and test participants on a task with which they are less familiar with to avoid personal preferences overriding the performance pressure manipulation

VI. Conclusion

The two hypotheses proposed for this study were that pilots who underwent training under a high-pressure manipulation would be disincentivised from exhibiting exploratory behaviour during training scenarios and that pilots who underwent training under a high-pressure manipulation would receive a lower quality of training followed by worse performance in test scenarios. Ultimately, the results of the experiment demonstrate that in the context of the experiment; training under a high-pressure manipulation ultimately did not have any significant effect on the pilots' behaviour in training nor their performance in a test or system knowledge. However, it cannot be concluded that performance pressure does not have any effect on learning rather that other factors such as individual differences between pilots and the structure of the training scenarios may have had a large contribution to the variation in the data. The recommendation is therefore to continue the research into the effects of performance pressure on learning.

Acknowledgment

The author would like thank all of the experiment participants especially those sorted into the high-pressure group as well as the department of Control and Simulation at the TU Delft faculty of Aerospace Engineering.

References

Allsop, J., & Gray, R. (2014). Flying under pressure: Effects of anxiety on attention and gaze behavior in aviation. *Journal of Applied Research in Memory and Cognition*, *3*, 63–71. https://doi.org/10.1016/j.jarmac.2014.04.010

Burratto, F., & Graef, R. (2021). Training pilots for resilience. *The Airbus Safety Magazine*. https://safetyfirst.airbus.com/training-pilots-for-resilience/

Cassell, V. E., Beattie, S. J., & Lawrence, G. P. (2018). Changing performance pressure between training and competition influences action planning because of a reduction in the efficiency of action execution. *Anxiety, Stress and Coping*, *31*, 107–120. https://doi.org/10.1080/10615806.2017.1373389

Dahlstrom, N., Cameron, M., & Kennedy, R. J. (2022). Simulation for experiential training (set) as an enabler for evidence based training (ebt). https://www.researchgate.net/publication/363376160_Simulation_for_Experiential_Training_SET_as_an_enabler_for_Evidence_Based_Training_EBT_A_white_paper_from_CAE_and_Emirates

EASA. (2015, September). *Annex ii to ed decision 2015/022/r*. https://www.easa.europa.eu/en/document-library/agency-decisions/ed-decision-2015022r

EASA. (2016). *Part-fcl*. https://www.easa.europa.eu/sites/default/files/dfu/Part-FCL.pdf

Eisenberger, R., & Aselage, J. (2009). Incremental effects of reward on experienced performance pressure: Positive outcomes for intrinsic interest and creativity. *Journal of Organizational Behavior*, *30*, 95–117. https://doi.org/10.1002/job.543

European Commission. (2019). *Report from the commission to the european parliament, the council, the european economic and social committee and the committee of the regions - aviation strategy for europe: Maintaining and promoting high social standards*. https://ec.europa.eu/transport/modes/air/25years-eu-aviation_en

Hinze, S. R., & Rapp, D. N. (2014). Retrieval (sometimes) enhances learning: Performance pressure reduces the benefits of retrieval practice. *Applied Cognitive Psychology*, *28*, 597–606. https://doi.org/10.1002/acp.3032

NTSB. (2022, July). *Aviation investigation final report: Commutair flight 4933*. https://data.ntsb.gov/carol-repgen/api/Aviation/ReportMain/GenerateNewestReport/99050/pdf

Ranieri, A., Gregorini, S., Moreno, F., & Cano, M. (2019). *Evaluation report: Evaluation of the applicable rules for initial and recurrent pilot training, testing and checking*. https://www.easa.europa.eu/sites/default/files/dfu/evaluation_report_on_the_applicable_rules_for_initial_and_recurrent_pilot_training_testing_and_checking.pdf

Spielberger, C. D., Gorssuch, R. L., Lushene, P. R., Vagg, P. R., & Jacobs, G. A. (1983). Manual for the state-trait anxiety inventory. *Consulting Psychologists Press*.

Stroosma, O., van Paassen, M. M., & Mulder, M. (2003). Using the simona research simulator for human-machine interaction research. *AIAA Modeling and Simulation Technologies Conference and Exhibit*. https://doi.org/10.2514/6.2003-5525

van Leeuwen, J. K., Landman, A., Groen, E. L., Mumaw, R. J., Stroosma, O., van Paassen, M. M., & Mulder, M. (2024). Using problem-based exploratory training to improve pilot understanding of autopilot functions. *Cognition, Technology & Work*. https://doi.org/10.1007/s10111-024-00751-z

Zijlstra, F. R. H., & van Doorn, L. (1985). *The construction of a scale to measure perceived effort*.

Part II

Preliminary Thesis Report

This part has already been graded for the course AE4020 Literature Study

1
Introduction

In the field of aviation, maintaining and updating pilot skills through recurrent training is paramount to safety. Furthermore, it is critical that pilots are monitored throughout training and have underperforming pilots undergo remedial training as letting them continue to fly can in some cases lead to accidents (NTSB, 2022). As aircraft become more complex and pilots become more reliant on autopilot. The role of the modern pilot has shifted from manually flying the aircraft to monitoring and diagnosing automation issues. There is therefore a need to critically examine the adequacy of current practices in pilot training and their ability to match the modern challenges that a pilot faces.

Cognitive skills such as crew resource management (CRM), and autopilot monitoring and management experience a higher decay rate compared to other competencies (Childs & Spears, 1986). Although those skills have been emphasised in newer regulation for recurrent pilot training (EASA, 2015a). Many members of the industry still believe that current regulations do not adequately prepare pilots for modern challenges (Ranieri et al., 2019). One reason for this is that the best manner in which to learn troubleshooting and diagnosis-solution skills is to solve authentic problems (Jonassen, 2010) which is not currently the case in recurrent pilot training. This has been partially addressed by shifting from task-based training to evidence-based training (EBT) which has proved effective (Burratto & Graef, 2021).

Another issue in recurrent pilot training is the performance pressure that pilots experience. It is important to monitor pilots during training and to report cases of under-performance for remedial training to avoid preventable accidents such as in the case of CommutAir Flight 4933 (NTSB, 2022) where the captain was cited as having quote (NTSB, 2022, p. 3): "*an inadequate foundation for being a Captain*" due to her repeated problems in training. However, this monitoring during training could possibly lead to hesitation, less communication about uncertainties and stricter adherence to standard operating procedures (SOPs) which are not applicable to all scenarios (Casner et al., 2013) hence leading to pilots performing worse when confronted with real issues in flight. In previous research it has been shown that unexpected and varied training that was more exploratory resulted in a better understanding for pilots rather than procedural training (Landman et al., 2018; van Leeuwen, 2020). The proposal for this thesis is therefore to study to what extent performance pressure affects pilots' abilities to learn during exploratory training.

1.1. Research Objective

This leads to the research objective for this thesis:

What are the effects of performance pressure during recurrent pilot training?

From the main research question, some sub-questions can also be devised as:

1. What are current practices in recurrent pilot training and what deficiencies are there?
2. Through what sources do pilots experience performance pressure from during training?
3. To what extent does pressure affect a pilot's exploratory behaviour during unexpected and varied training?

1.2. Report Structure

To achieve the research objective, an experiment will be conducted where the effects of performance pressure on pilots will be studied and quantified if possible. This report serves to provide a background on the research as well as how the experiment will be conducted. In Chapter 2 a review of relevant literature on recurrent pilot training, current deficiencies in pilot training and the effects of pressure on performance and learning is given. This is followed by Chapter 3 where the scenarios that will be used in the experiment are explained and motivated and finally, the full experiment plan is shown in Chapter 4.

2
Literature Review

In this chapter, a summary of current standards and practices in recurrent pilot training is given in Section 2.1 followed by literature on current issues in pilot training in Section 2.2. Finally, performance pressure in pilot training and the effects of performance pressure on learning are covered in Section 2.3.

2.1. Recurrent Pilot Training

As the landscape of aviation continues to evolve, the role of recurrent pilot training has become increasingly vital to help pilots maintain and enhance their skills. In this section, the role of recurrent pilot training and current issues with training are explored, as well as how pilots are evaluated in recurrent training.

2.1.1. The Role of Recurrent Training and Current Issues

On the 9th of March 2019, CommutAir Flight 4933 took off from Newark Liberty International Airport. On the approach to its final destination, Presque Isle International Airport; poor judgement and decision making by the captain led to the aircraft missing the runway and touching down to the right of the runway. According to the accident report, the captain was subject to multiple checks before the eventual accident having been monitored for 9 months in 2016 and failing multiple trainings and proficiency checks in September of 2017. This was followed by remedial training in March of 2018 (NTSB, 2022, p. 13). Despite all this, the captain was still identified as a principal source of the accident with her repeated training problems as reflected by the investigation report which stated that the captain had: "*an inadequate foundation for being a Captain, which CommutAir did not effectively address*" (NTSB, p. 3).

This incident highlights the need for effective recurrent training for the safety of passengers and pilots and the necessity for airline pilots to regularly undergo training and checks in order to maintain or upgrade their flight skills. Recurrent pilot training is also necessary in order to comply with aviation authorities and for pilots to maintain their type ratings (EASA, 2016). Currently, pilots undergo recurrent training every six to twelve months which will consist of a training and assessment program with instruction typically happening in a classroom environment followed by flight simulator training and assessment sessions. The simulator instructor will then assess the pilots' performances (Mavin, 2016, p. 169). As seen in the case of CommutAir Flight 4933's pilot, sub-par performance in training will lead to increased scrutiny and require further checks, with repeatedly poor performances leading in some cases to termination.

To best make use of recurrent training, regulations therefore have to be adjusted to most accurately reflect the real challenges a pilot faces on the job. Since use of automation in aircraft has seen a large increase in recent decades due to its advantages increasing safety and decreasing workload allowing for increased productivity. The role of the modern airline pilot has thus shifted to one more focused on monitoring automation and resolving issues a task which humans struggle with. This was first identified in the seminal paper by Bainbridge which documented the paradoxical way in which increasing automation to theoretically simplify a job actually increases difficulty for the operator who has to supervise it (Bainbridge, 1983). Modern recurrent pilot training regulations emphasise learning automation and the philosophy of the use of automation as well as monitoring and troubleshooting automation problems (EASA, 2015a). However despite the advances in technology and training, this paradoxical issue remains today (Strauch, 2018).

Moreover, recurrent training courses are often scheduled after pilots' shifts when they are already exhausted. This leads to recurrent training being less effective and useful as it should be (Ranieri et al., 2019, p. 41). Consequently, a significant portion of airlines operating in the European Union (EU) have reported that a large proportion of pilots holding an airline transport pilot licence (ATPL) do not meet the airlines' basic entry requirements (European Commission, 2019). Current members of the aviation industry believe that pilots receive less training today and with lower quality highlighting the need for studies on improving pilot training to be conducted (Ranieri et al., 2019, p. 41).

2.1.2. Evaluation of Pilot Performance in Recurrent Training

Traditionally, recurrent training utilised a task-based approach which was based on evaluating a pilot's performance based on how they performed on a particular set of tasks (Burratto & Graef, 2021). While task-based training can be effective with simple systems where there are objectively correct answers, it is extremely limited with complex systems as not all aspects can be trained and tested. The introduction of EBT in recurrent training was intended to replace task-based training (ICAO, 2013). In EBT, pilots are not evaluated on their performance in specific tasks but rather on how they demonstrate the use of key competencies in operations and training (EASA, 2015b). The European Union Aviation Safety Agency (EASA) has since defined nine pilot competencies which are used by flight examiners as a guide to evaluate the pilot's overall performance (EASA, 2021). These competencies are listed below:

- Application of procedures and compliance with regulations **(PRO)**
- Communication **(COM)**
- Aircraft Flight Path Management, automation **(FPA)**
- Aircraft Flight Path Management, manual control **(FPM)**
- Leadership and teamwork **(LTW)**
- Problem solving and decision making **(PSD)**
- Situation awareness and management of information **(SAW)**
- Workload management **(WLM)**
- Application of knowledge **(KNO)**

Since EBT is not designed to test pilots' abilities on specific pre-determined scenarios, it allows training to increase pilots' "resilience" which is defined by Burratto and Graef as: "*the ability of a flight crew member to recognize, absorb and adapt to disruptions*" (Burratto & Graef, 2021, p. 2) and gives trainees the opportunity to receive more relevant feedback and subsequently receive more attention in training in specific areas of improvement (Cameron et al., 2022).

2.2. Issues in Pilot Training

Due to the rapid increase at which aviation technology has evolved, reports and studies have found that authorities remain on the back foot when it comes to determining appropriate regulations for pilot training. In this section, issues in training such as lack of unpredictability and variability, and insufficient training of CRM and autopilot management are explored.

2.2.1. Unpredictability and Variability

One particular issue that occurs in recurrent pilot training is the lack of training of authentic problem solving skills when confronted with abnormal events. When pilots then encounter these events in flight, despite having rehearsed for such emergencies they deviate from procedures causing accidents (Casner et al., 2013). Such was the case in Colgan Air flight 3407 where despite having received stall recovery training, the captain still did not apply proper procedure leading to loss of control in flight (NTSB, 2010). This was also the case in CommutAir flight 4933, the first officer was cited as having noticed an incongruency in the instruments guiding the landing. However, despite this according to the investigation: "*neither flight crew member considered that a navigational air error could be occurring*" (NTSB, 2022, p. 3) and decided to proceed with landing.

According to Casner et al., this stems from deficiencies in regulations since compentent authorities (CAs) choose events for training that are most likely to occur and carry severe consequences if resolved incorrectly which at its surface appears to be a logical manner to do so. However, this is ultimately detrimental to pilot learning as pilots then know what to expect during recurrent training and since their career depends on their performance during training they then adhere to SOPs rather than attempt to use diagnosis and problem-solving skills.

Whilst these SOPs are effective when the pilot is expecting an abnormality in flight such as during recurrent training. In the experiment conducted by Casner et al. (2013), when presented with such situations in an unexpected fashion; pilots' behaviours would vary drastically. This is due to the fact that SOPs often rely on the experience of individual operators hence the dramatic variation in behaviour between pilots despite the procedures being the same (Huijbrechts & van Paassen, 2023). The recommendation from the experiment by Casner et al. was therefore to implement for unexpected and varied scenarios in pilot training.

The evidence on the effectiveness of exploratory learning compared to more structured methods is mixed. A meta-analysis by Carolan et al. was performed on the effectiveness of learner control and exploratory learning. Learner control provides learners with active control over training variables such as the pace of training or what feedback to receive whereas exploratory learning requires learners to discover relationships and interactions by exploring the learning content. From this analysis it was found that both methods provided differing levels of success depending on the conditions (Carolan et al., 2014). Exploratory learning may provide more benefit in a task-based environment where the focus is more on procedural skills or for domain experienced learners adapting learned skills to new tasks or environments. On the other hand, learner control may be more effective for tasks that involve developing cognitive skills.

Both strategies were found to be more effective for cognitive skill learning rather than for knowledge learning tasks. However, exploratory learning benefited far transfer for skill-based tests. This suggests that with exploratory learning, the emphasis in training on finding procedures and strategies for generating rules and solutions is ultimately beneficial (Carolan et al., 2014).

The benefits of unexpected and varied scenarios during pilot training was explored in an experiment conducted by Landman et al. (2018). Two groups of pilots underwent flight scenarios in a simulator with the control group receiving highly predictable training scenarios whilst the experimental group received unpredictable and varied training scenarios. Both groups then underwent a test with an unrelated scenario and a test with a related scenario. It was found that in the related test scenario, the group that received the unpredictable and varied training performed better than the control group; however no significant differences were found between the groups in the unrelated test scenario. This study concentrated on manual flight for all scenarios therefore no conclusions on unexpected and varied training on autopilot management can be made.

A followup study by van Leeuwen (2020) investigated the effects of unexpected and varied scenarios on autopilot training. Once again, two groups of pilots underwent two different types of training with the experimental group being trained to actively diagnose and solve automation surprises without any knowledge of the training scenarios and the control group receiving knowledge of the training scenarios. It was found that the experimental group in fact responded faster to new automation surprises whereas the opposite case was true for repeated automation surprises. However, ultimately the differences between the groups were not significant.

2.2.2. Crew Resource Management and Autopilot Management

CRM, and monitoring and management are elements of piloting which have long been identified as a critical aspect of recurrent pilot training (Noon & Murphy, 1984) as appropriate CRM will often prevent potential air disasters. Moreover, studies have proven that cognitive skills such as monitoring autopilot in flight undergo greater and more rapid decay over time than control-oriented skills (Childs & Spears, 1986) hence the need for more frequent and in-depth training.

Despite the changes in regulations to place emphasis on autopilot management in recurrent training (EASA, 2015a), according to a report by Ranieri et al. (2019) only 19% of pilots have a positive opinion on the adequacy of the current rules to meet modern aviation challenges. Therefore, there is a need for research to be performed into current issues in pilot training and how it can be improved going forward. One identified issue in recurrent training is the lack of coverage of CRM and monitoring skills. In a study conducted for Boeing it was found that despite over 99% of pilots identified autopilot monitoring and cross checking as an important skill, only 47% had it explicitly covered during training (Fletcher & Bisset, 2017; Holder, 2013).

This lack of diagnosis skills was shown in an experiment conducted by Nikolic and Sarter (2007) where airline pilots were invited to fly a simulator scenario with three challenging automation tasks likely to cause potential errors and misjudgements. The results indicated that pilots rarely adhered to the reference path when addressing scenario events and there was often a delay in disturbance detection. Furthermore, pilots seldom diagnosed issues, primarily due to knowledge deficiencies and time pressures and they frequently resorted to generic, sub-optimal recovery methods and leaned heavily on advanced automation to mitigate error consequences, highlighting the need for further developments in diagnosis and error management training.

2.3. Performance Pressure in Training and Effects on Learning

Current literature on the effects of performance pressure on pilots undergoing training is limited. Nonetheless, there are several studies that aimed to quantify the effects of performance pressure on motor skills and overall pilot performance. In this section, the sources from which pilots experience performance pressure during training are explored as well as the effects of pressure on performance and learning.

2.3.1. Performance Pressure in Recurrent Pilot Training

During recurrent pilot training, it is important to monitor pilots' performances and test them as letting pilots who repeatedly fail checkrides and recurrent trainings continue flying can in some cases lead to disastrous consequences (NTSB, 2010, p. 10). However, through personal interviews with active pilots, it was found that anecdotally, monitoring of performance often leads to a feeling of pressure for the pilots which is amplified when they fail and have to undergo checkrides again. This presents a further problem as pilots who fail checkrides should be further tested to ensure safety. Moreover, in a study by Demerouti et al. (2019) it was found that 40% of pilots surveyed experienced high levels of burnout due to various factors such as job demands. This leads to an indirect negative effect on simulator training performance as pilots then lack the resources to adjust their work characteristics in a beneficial manner.

This is exacerbated by the low level of proficiency of CAs and their staff along with their lack of understanding of the real challenges of modern aviation. This leads to the role of CA staff in recurrent training simply being one of checking off checklists rather than a supervisory one (Ranieri et al., 2019, pp. 65–66), putting pressure on pilots to simply memorise SOPs rather than demonstrate a genuine understanding of technical systems on board an aircraft which would be beneficial in flight.

2.3.2. Effect of Pressure on Performance and Learning

An experiment performed by Allsop and Gray (2014) tested the effects of inducing an anxiety situation on pilots' gazes measured via a head mounted eye tracking system and their subsequent performance. In this case there were two groups, a control and an experimental group. The pilots were required to achieve and maintain a certain flight path and complete a landing. This consisted of a practice phase, a pre-test phase and an anxiety phase. For the control group, nothing changed in the anxiety phase whereas the experimental group were put in a high-pressure environment by being told that their flight results would be shared to everyone participating in the experiment with a monetary prize at stake. Moreover, the participants were videoed during the experiment and flew in a mock online environment with other aircraft. The control group on the other hand did not have this anxiety manipulation applied and simply performed the test under the same conditions as the training.

Afterwards, it was found that inducing performance pressure can effectively manipulate anxiety and while anxiety did lead to a reduction in attention, it ultimately did not lead to a significant decrease in performance (Allsop & Gray, 2014). However, it is important to note that the primary focus of the study was the pilots' gazes and performances over a repetitive task which may not be as good a representation for recurrent training which consists of various types of tasks that require a higher cognitive understanding of aircraft systems and performance does not necessarily correlate with learning.

In terms of varying pressure during a task, a study conducted by Cassell et al. (2018) had 4 groups complete a two-dimensional stimulus-response task. These groups were split into 2 control groups (control-control and anxiety-anxiety) and 2 experimental groups (control-anxiety and anxiety-control) with the experimental groups experiencing a change in pressure conditions between training and testing going from either low to high pressure or vice versa. Ultimately it was found that the experimental groups experienced a detrimental effect to performance regardless of whether the pressure was increased or decreased, whereas the control groups had a similar performance throughout the experiment.

This experiment used reaction time and movement time as the dependent measures therefore focusing more on motor control rather than the effects of pressure on the performance of a cognitive task as will be done in this project. As a result, while the findings are intriguing, they would only apply to the physical aspects of the pilots' responses in training rather than the cognitive aspects of learning.

Increased performance due to performance pressure may be linked to intrinsic job interest (Eisenberger & Aselage, 2009). As found by Eisenberger and Aselage, the higher expected reward for high performance promotes a compulsion to perform well which subsequently leads to increased perceived self-determination and increased intrinsic job interest.

The effects of performance pressure on learning and more specifically, retrieval practice was studied by Hinze and Rapp (2014). The experiment involved participants studying a text that they had no prior knowledge of before having to recall content from the text using prompts from a quiz with the purpose of encouraging long-term retention. The low-stakes group performed this recall activity with no performance pressure whilst the high-stakes group did so with a monetary incentive at stake. The groups were then tested on their knowledge of the text and graded. Ultimately it was found that participants could overcome pressure effects during a test but performing retrieval practice whilst under high-stakes performance pressure was ultimately detrimental to long-term retention and understanding of scientific content.

From this literature review, it can be seen that there is clearly a need to study the extent to which performance pressure affects pilot behaviour and learning during recurrent training. Hence, the proposal in this thesis for the design of an experiment to test the effects of performance pressure on learning. For such an experiment, the scenario design is crucial as that will determine factors such as whether the experiment is a between-subjects or within-subjects experiment and the necessary types of participants. The scenario design is explained and motivated in the following section.

3
Scenario Design

In this chapter the scenarios that were designed for the experiment are explained. First a framework for the scenario design is explained in Section 3.1 followed by the requirements for the scenarios in Section 3.2. The training and test scenarios are then described in Section 3.3 with a brief overview of the scenario execution in Section 3.4 and finally a look at the secondary tasks in Section 3.5.

3.1. Simulation for Experiential Training

One of the most crucial aspects of recurrent pilot training with regards to EBT is the training scenario design (Dahlstrom & Kennedy, 2022). An effective training scenario will require the appropriate level of attention from the trainee allowing proper learning and skill transfer to a wider variety of situations

An effective design philosophy to enable EBT in pilot training was explored in a white paper by CAE and Emirates on simulation for experiential training (SET) (Dahlstrom et al., 2022). In this paper, Dahlstrom et al. identified several key principles for the design of simulation training scenarios, those being:

- Use of information sources
- Variation in information density
- Periods of high and low workload
- Concurrent and parallel tasks and priorities
- Competing options at decision points
- Follow up of decisions and “effect control”

It is important to note that the scope of this project and those of CAE and Emirates are different and as such not all concepts from Dahlstrom et al. (2022) can be used as some may not be as applicable. In the scenario design for this experiment, variation in workload, concurrent tasks and competing options were used as guiding principles.

3.2. Requirements

In order to ensure that the training and test scenarios would encourage the desired behaviour from the pilots and adhere to the previously identified principles, a set of requirements must be defined. These are shown in Table 3.1 where the identifier prefix **GEN** indicates a general requirement for all scenarios and the identifier prefix **DIAG** indicates a requirement for scenarios that require failure diagnosis such as training and test scenarios.

3.3. Scenarios

Utilising these requirements and based off the scenarios created for the experiment by van Leeuwen (2020), the scenarios for the proposed experiment can be created. These are shown in Table 3.2. As can be seen, the scenarios fall into three categories those being: familiarisation flight scenarios, training scenarios and test scenarios.

The familiarisation flight scenarios will allow the participants of the experiment to fly the aircraft in a nominal state firstly to get used to the aircraft dynamics should they decide that switching to manual flight

Table 3.1: Table showing the requirements for the scenario designs

Identifier	Requirement
GEN-1	Every scenario shall be unique
GEN-2	Every scenario shall be realistic
DIAG-1	The failure diagnosis scenarios shall utilise a noticeable and realistic aircraft instrument failure
DIAG-2	The failure diagnosis scenarios shall have an optimal solution which requires use of autopilot
DIAG-3	The failure diagnosis scenarios shall have multiple acceptable solutions
DIAG-4	The failure diagnosis scenarios shall contain variations in workload
DIAG-5	Each failure diagnosis scenario shall have a unique optimal solution
DIAG-6	The test scenarios shall have optimal solutions which utilise higher automation modes than the training scenarios
DIAG-8	The failure diagnosis scenarios shall not have manual flight as an optimal solution

is the best course of action. Moreover, they will also then be able to learn the vertical and lateral modes in a low-stakes scenario where they can fully focus on learning. However, no guidance will be given on which modes to use rather the participants will select and switch between modes of their own volition.

In the training scenarios, two basic failures scenarios will be presented to the participants. Once again, no guidance will be given on how to resolve the failures that will occur. Participants will instead have to rely on the knowledge gained from the familiarisation flight scenarios and the troubleshooting training that was given in the ground school. Finally, the participants will undergo two testing scenarios. Participants will be expected to use the knowledge gained in the training scenarios as well as their own diagnosis and problem solving skills to resolve the test scenarios.

Shown in Table 3.3 and Table 3.4 are the automation modes ranked based on their levels of complexity and automation. It can be seen that there are multiple valid modes in lower levels of automation that the participants can use to resolve the failures in the training and test scenarios. Whilst there is no "correct" solution, there are certain features within the autopilot such as automatic capture of the selected altitude that make using the optimal alternative mode require a lower workload compared to lower modes of automation.

Table 3.2: Table showing the selected scenarios for the experiment

Phase	Scenario	Active mode before failure	Optimal alternative mode
Nominal flight	Manual flight	None	None
	Vertical mode familiarisation	None	None
	Lateral mode familiarisation	None	None
Training	VOR receiver failure	NAV (VOR)	HDG
	Blocked static port	VS	PIT
Testing	Navigation with GPS failure	VNV/NAV (GPS)	ALT/VS/FLC + NAV (VOR)
	Blocked pitot tube	FLC	VS

Table 3.3: Table showing the vertical autopilot modes ranked by the level of automation

Vertical Mode	Level of automation
Manual flight	Lowest
Pitch hold	
Altitude hold	
Vertical speed, flight level change	
Navigation modes	Maximum

Table 3.4: Table showing the lateral autopilot modes ranked by the level of automation

Lateral Mode	Level of automation
Manual flight	Lowest
Roll hold	
Heading select	
Navigation modes	Maximum

3.4. Scenario Execution

Before each scenario, the participants will receive a flight plan of the scenario so that they know ahead of time what will be expected of them to complete the scenario. They will also be told when in the scenario they will be expected to complete a high workload secondary task. A diagram illustrating the scenario execution is shown in Figure 3.1.

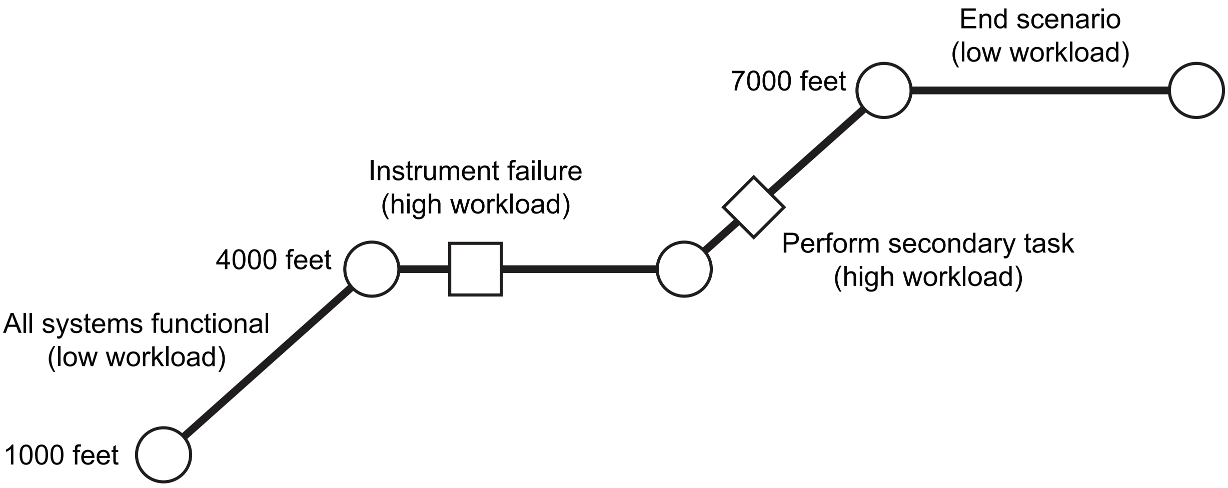


Figure 3.1: Diagram showing the execution of a hypothetical climbing scenario

Figure 3.1 illustrates a scenario that could be applied to the blocked static port training scenario. The participants will begin in a nominal flight state where the workload is relatively low only requiring them to ensure they're climbing to the correct altitude and monitoring the relevant subsystems. After a sufficient amount of time in the low workload state, the relevant instrument failure will then be triggered requiring the participant to diagnose and resolve it in an acceptable manner. Since participants will be informed of when the secondary task will have to be performed, they will be motivated to resolve the failure in as effective a way as possible in order to shift their focus as much as possible for the secondary task. After the failure has been resolved and the secondary task completed, the participant will then be able to return to a low workload state and complete the scenario.

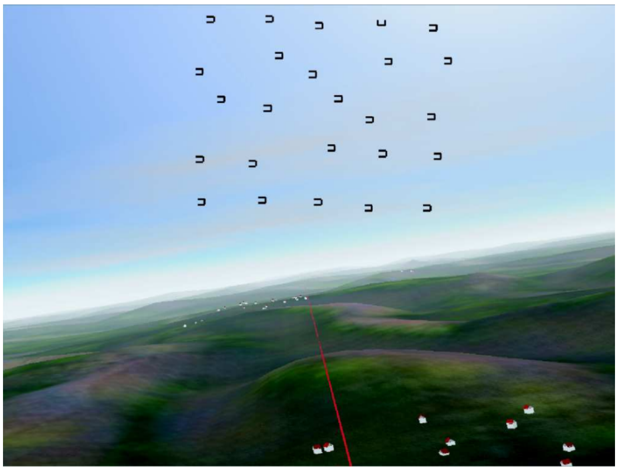


Figure 3.2: Secondary task overlaid on the simulator's outside visuals (Brouwer et al., 2004)

3.5. Secondary Tasks

In an experiment by Brouwer et al. (2004), pilots were asked to fly manually and complete a secondary task at the same time as shown in Figure 3.2. This was found to effectively increase workload for the pilots. From personal interviews conducted with active pilots, it was hypothesised that a similar idea can be implemented in this experiment with the goal of providing participants with effective motivation to try and reduce their workload by using autopilot as from previous experiments it was found that simply asking them to do so did not often result in participants feeling motivated to do so.

This will also simulate what a pilot experiences when issues arise in flight as they are often required to complete checklists which distract from flying the aircraft. Since the task used by **Brouwer** is not well suited for this experiment, the secondary task that the pilots will be asked to do will be determined after testing as the secondary task should not be so demanding in workload that pilots feel overloaded yet it should provide enough visual and mental distraction from the goal of navigation that pilots feel the need to use automation to aid them.

4
Experiment Plan

In this chapter, the experiment plan is laid out firstly stating the hypotheses in Section 4.1 followed by the equipment that will be used in the experiment in Section 4.2. The experiment design and participants that will be invited are then explained in Section 4.3 and Section 4.4. The control, independent and dependent variables are listed in Section 4.5. Finally, the experiment execution is described in Section 4.6 followed by the expected behaviour of the participants in Section 4.7.

4.1. Hypotheses

Based on the literature review conducted in Chapter 2 a number of hypotheses for the experiment can be posed these are shown below.

1. Pilots who were under pressure in the training phase will receive a lower quality of training
2. Pilots who were under pressure in the training phase will be less likely to use the optimal autopilot mode in the test scenarios

4.2. Equipment

The experiment will be conducted using the SIMONA Research Simulator (SRS) at the faculty of Aerospace Engineering at TU Delft shown in Figure 4.2. The SRS is a full flight simulator with a six-degrees-of-freedom hydraulics system. On the interior is a modular cockpit that can be adapted to simulate a variety of aircraft and the visual display system provides a collimated $180^\circ \times 40^\circ$ field of view image (Stroosma et al., 2003). This provides a very similar simulator experience to what a pilot may utilise in recurrent training which makes the SRS suitable for the experiment.

The model that will be used in the SRS is the Piper Seneca III fitted with glass cockpit avionics resembling the Garmin G1000 and the accompanying Garmin GFC700 autopilot for which the primary flight display (PFD) is shown in Figure 4.1. The Piper Seneca III is a twin propeller aircraft popular among general aviation enthusiasts and flight schools. The Seneca was chosen specifically since a well developed and accurate model of it was already used in multiple experiments within the section of Control and Simulation thus avoiding the need to develop a new model. Although the models flown in recurrent training would be based on commercial airliners; for the purposes of this experiment, continuity on previous research was prioritised hence the choice of the Seneca model. Moreover, using a model of an airliner would require a two-pilot crew making it less suitable for this project.

The simulation is based on the extensions written by van Leeuwen (2020) on top of the model used by van Oorschot (2017). The simulation is written in C++ using the Delft University Environment for Communication and Activation (DUECA) framework. DUECA allows for custom scenarios to be written with simulated failures occurring at preset times within the scenario. For the secondary task, a touchscreen tablet computer will be mounted within the cockpit of the simulator with which the participants can complete the secondary task.

4.3. Experiment Design

Similar to previous experiments on pilot training in the section of Control and Simulation, this experiment will also take the form of a between-subjects experiment with two groups rather than within-subjects. The

Figure 4.1: Screenshot showing the simulated PFD

Figure 4.2: The SIMONA Research Simulator (Stroosma et al., 2003)

Figure 4.3: The Piper Seneca V (Popovic, 2016)

control group will undergo training with no pressure and the experimental group will instead be pressured using a strategy derived from literature and from interviews with pilots which is elaborated on in Subsection 4.5.2. A between-subjects design was ultimately chosen since that would minimise carryover pressure effects between different runs in the simulator. Furthermore, there would be a confounding effect between a plateau in the participants' learning curve and the potentially detrimental effect of the pressure on their learning which would ultimately make understanding the effects of pressure on learning unclear. Finally, although a within-subjects experiment would require fewer participants. It would also require more scenarios to be devised which is ultimately the more limiting factor than the number of participants that can be found. Hence, a between-subjects experiment design was used.

4.4. Participants

For this experiment, twenty participants will be invited to participate in the experiment. The participants will ideally be active airline pilots with experience flying aircraft with glass cockpits and autopilot but no experience using the Garmin G1000. At a minimum, all participating pilots should hold airline transport pilot licences (ATPLs) and be type rated for a large multi engine aircraft. The decision was made to use active airline pilots as it was hypothesised based on the interviews conducted with pilots that new pilots would lack the experience diagnosing unknown failures thus defeating the purpose of the experiment. It was also decided to use pilots with autopilot experience and no experience with the G1000 to allow the ground school phase to be shorter as basic autopilot concepts would not have to be taught yet the pilots would not be completely familiar with the autopilot thus providing an opportunity to study the quality of learning they would receive.

4.5. Variables

4.5.1. Control Variables

The control variables in this experiment are:

- **Ground School:** Both groups will receive the exact same ground school training. This is to ensure that all participants begin the simulator training with the same amount of knowledge of the autopilot in the simulator.
- **Scenarios:** Both groups will undergo the same scenarios in the training and test in the same order. This will avoid either group undergoing any unexpected learning effects between scenarios.
- **Training time:** Both groups will undergo the same number of training scenarios and efforts will be made to ensure that all participants will have spent approximately the same amount of time training.
- **Aircraft:** Both groups will use the same aircraft in the same configuration.
- **Instruction:** Except for the pressure manipulation, both groups will receive the same instructions when undergoing the experiment in the simulator. I.e. the same experiment script will be used for all participants.

4.5.2. Independent Variables

The independent variable in this experiment will be the performance pressure applied on the participants in the training phase. The participants will be divided into two groups with the control group undergoing the training scenarios with no performance pressure and the experimental group having to do so whilst under performance pressure. Following personal interviews with active line pilots to determine what sources they typically experience pressure in training from, a pressure manipulation strategy was devised which is shown below:

- Pilots will be informed that all the data from the simulator is being recorded and will be analysed by other pilots afterwards
- Pilots will be told that the number of training scenarios depends on their performance and that "better" participants will have fewer scenarios

Since it is impossible to replicate the true stresses and pressures of recurrent training in real life. It was decided that the strategies to apply pressure in the experiment should utilise the specific strengths of the context of the experiment. Firstly, since the experiment is conducted in an academic context; participants in the experimental group will be told that for the purposes of the research all of the data from their runs in the simulator will be anonymously recorded and possibly analysed later by other pilots. Participants will therefore feel under pressure to perform to an adequate level for not only themselves but their peers as well hence inducing performance pressure.

Subsequently, participants in the experimental group will also be told that the number of training scenarios are dependent on their performance. This is similar to real recurrent training where underperforming pilots must undergo further training. In the context of the experiment it can be explained as all participants requiring a certain level of competency to be able to undergo the test scenarios. Once again, participants will experience performance pressure from their desire to prove their skills and pass the training scenarios in as few attempts as possible.

Finally, To ensure that the experiment only focuses on the core question (the effect of pressure on pilots in training), there will only be one independent variable. Furthermore, since pressure is subjective it would not be practical to try and vary the level of pressure hence only two groups (pressure and no pressure) will be used.

4.5.3. Dependent Measures

To evaluate the participants' knowledge and performance in the final test scenarios, a number of dependent measures must be defined these are listed below:

- **State anxiety:** Participants will be asked to evaluate their level of anxiety during the experiment using the state-trait anxiety inventory developed by Spielberger et al. (1983). This will be to ensure that the pressure manipulation was in fact successful and that participants in the experimental group experienced a higher level of anxiety during training.
- **Time spent out of the optimal mode:** Participants will be judged based on their selections of autopilot modes during the scenarios and how much time they spend in either too high or too low of an automation mode.
- **Choice of final automation mode:** Ahead of the secondary task in each scenario, participants will be motivated to choose as high a level of automation as possible to reduce their workload for the secondary task. Participants will be evaluated based on what they choose as their final automation mode.
- **Workload:** Participants will be asked to measure their workload using the rating scale mental effort (RSME) from Zijlstra and van Doorn (1985) after every run in the simulator.
- **Motivation of actions:** Participants will be asked to explain their actions during the scenario and will be evaluated based on whether they performed actions randomly or based on their knowledge of the autopilot systems.
- **Paper test scores:** Participants will be asked to complete a multiple choice paper test at the end of the experiment to evaluate their theoretical knowledge of the autopilot that they gained during the experiment.

4.6. Experiment Execution

The experiment plan can be seen in Figure 4.4. The experiment will be split into three phases: the preparation phase, experiment phase and post-experiment phase. The preparation phase involves sorting the participants into the two groups and balancing them as well as the basic ground school instruction. From there the experiment phase begins where the participants undergo simulator training and testing as well as the paper test. Finally, in the post-experiment phase the data is processed and analysed.

4.6.1. Group Balancing

Since the experiment is a between subjects experiment, in the preparation phase the groups will have to be balanced to ensure minimal differences between the groups. This will be done on the basis of a few different metrics: age, flight hours and susceptibility to pressure. The final metric will be measured using a predictive model based on that created by Clarke et al. (2020) where correlation between personality traits and "paradoxical performance" was measured. This model can be seen in Figure 4.5.

In this case paradoxical performance is defined by Baumeister and Showers as: "*the occurrence of inferior performance despite striving and incentives for superior performance*" (Baumeister & Showers, 1986, p. 362) which provides an accurate definition for what is to be explored in this experiment.

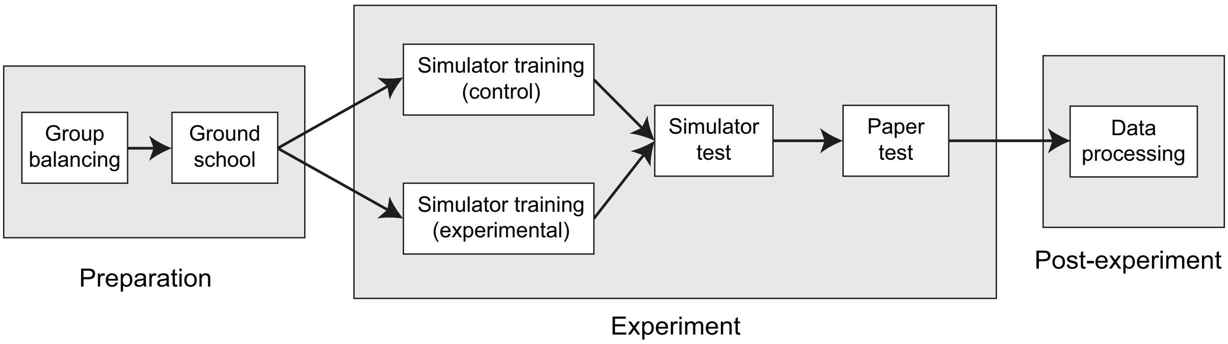


Figure 4.4: Diagram showing the experiment plan

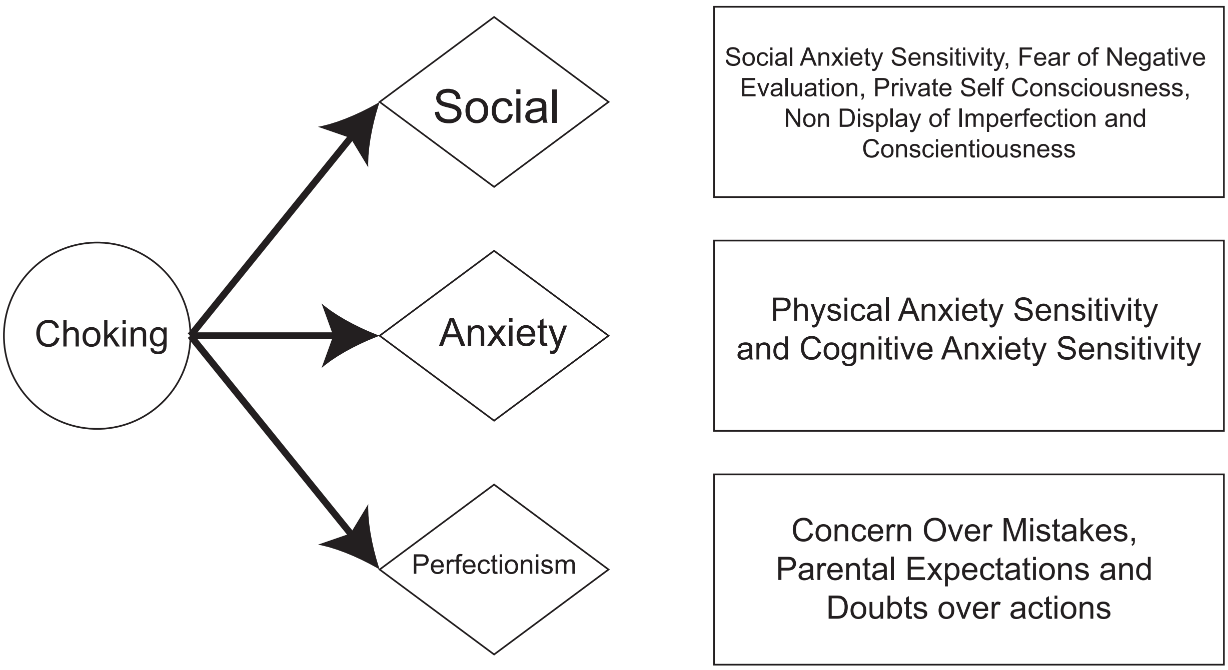


Figure 4.5: The “choking predictive model” (Clarke et al., 2020)

4.6.2. Ground School

Since the participants will have no experience with a Garmin G1000. The first step of the experiment will be to provide them with basic instructions on the operation of the autopilot. This will be done in the form of a ground school similar to what pilots will undergo during recurrent training. Instruction during the ground school phase will be kept to a minimum as to ensure that most of the learning by the participants on the autopilot will take place in the simulator. The training will therefore explain the basic autopilot modes available to them in the simulator as well as how to select and fly with them. However, no information on how the subsystems interact will be given. These training slides will be based on the Piper Seneca's G1000 guide an example of which is shown in Figure 4.6. Moreover, participants will also be instructed on a problem solving strategy based on that devised by van Leeuwen (2020).

This problem solving strategy is shown below:

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 available, or
5. Switch to lower level automation

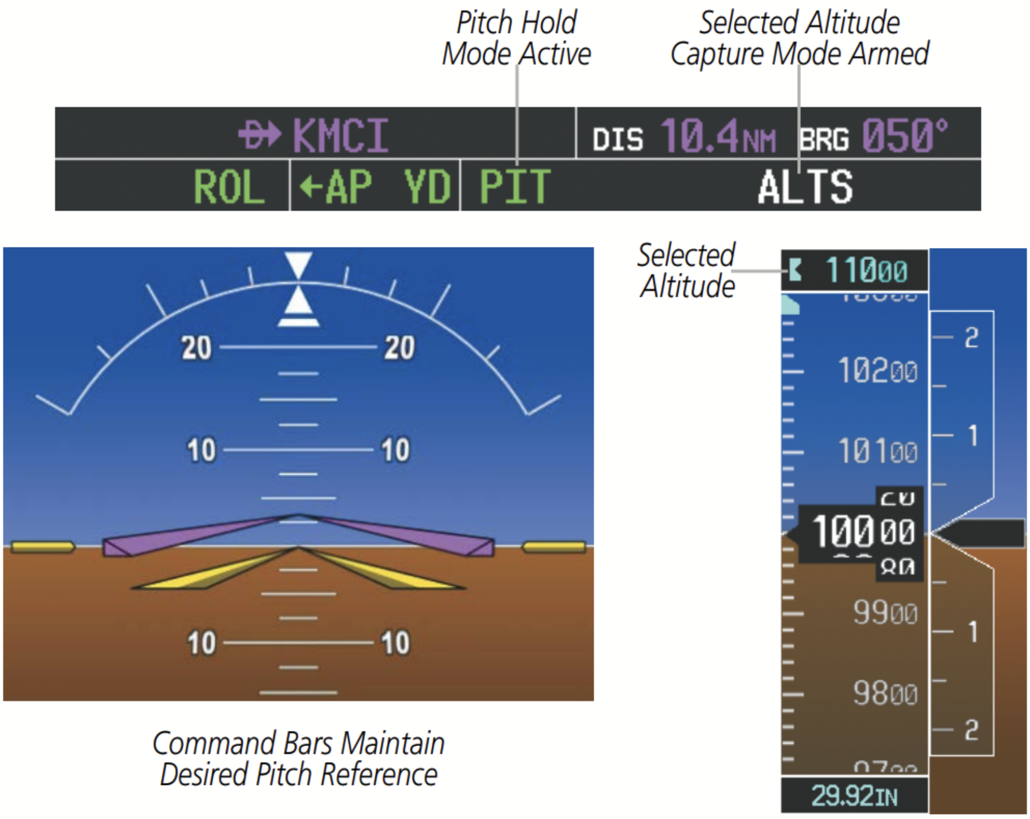


Figure 4.6: Pitch hold mode as shown in the G1000 pilot's guide (Garmin, 2014)

4.6.3. Simulator Training

After the ground school instruction, the participants will then move onto the simulator training phase of the experiment. Here both groups will undergo the nominal flight and instrument failure training scenarios detailed in Section 3.3. The goal of the nominal flight scenarios will be to allow the pilots time to familiarise themselves with the aircraft should the need arise to have to switch to manual control and to be able to explore the autopilot functions of the G1000 under no extra workload due to any instrument failures. The instrument failure training scenarios will then force the pilots to use the problem solving strategy taught in the ground school and the previous knowledge gained to resolve the instrument failures as effectively as possible. Pilots will be told that their motivation is to keep their workload as low as possible to complete the secondary task adequately, which can be achieved by using the highest level of automation possible. Moreover, after every training scenario pilots will be asked to rate their workload using the RSME (Zijlstra & van Doorn, 1985).

The control group will perform the scenarios with no extra instructions from the control room whereas the experimental group will perform the scenarios with pressure applied as detailed in Subsection 4.5.2.

4.6.4. Simulator Test

After having completed the training scenarios in the simulator, the participants will move onto the test scenarios detailed in Section 3.3. Here both groups will be put under the same amount of pressure to ensure that there is no confounding pressure effect in the test. Furthermore, participants will be told that unlike the training scenarios they should prioritise the safety of the aircraft first and then subsequently attempt to lower their workload using automation. This is provides a motivation to the pilot more akin to one they might have in the context of recurrent training as the first priority of a pilot should always be to fly the plane safely. After each test scenario, pilots will be asked to rate their perceived effort during the test scenarios using the RSME and to verbally explain their motivations for their actions during the test scenarios. This will provide an opportunity to qualitatively understand whether or not their performance was influenced due to the quality of the training received in the training phase and whether or not the pressure manipulation did in fact lead to a lower system knowledge.

4.6.5. Paper Test

Finally, after having completed all the scenarios in the simulator. Participants will be asked to complete a multiple choice test. This will be in the form of a 15 question multiple choice test based partially on the failure scenarios but also on other features not explicitly covered in the ground school. The purpose of this test will be to ascertain the technical knowledge of all the participants and whether or not the pressure manipulation in training had affected the quality of the participants' training. Participants will complete the paper test after having completed training and testing in the simulator since the questions would be multiple choice and therefore there could be a possibility that participants could gain some extra knowledge from the paper test that they did not have from the simulator training.

4.7. Expected Behaviour

The aim of this experiment is to determine whether there are detrimental effects resulting from performance pressure induced on pilots during recurrent training. It is thus useful in this section to discuss how the design of the scenarios might aid in proving or disproving the hypothesis. Firstly, the performance pressure may have the least effect in the nominal flight scenarios as pilots will be allowed to fly however they desire. However, the control group may utilise a wider variety of modes or switch more often between modes resulting in a better understanding compared to the experimental group who might be more incentivised to simply reach the end of the scenario.

In the training scenarios it is likely that both groups may instinctually decide to disconnect the autopilot upon noticing the failure. However, it will be expected that since the control group is not under pressure; the participants in the control group would eventually utilise trial and error or their problem solving skills to increase the level of automation to the highest level to complete the scenario. The experimental group on the other hand would simply attempt to complete the scenario using any means necessary. In the test scenario both groups will be under performance pressure and once again it would be expected that both groups would first disconnect the autopilot. However, it would then be likely that the control group would more confidently increase the level of automation using the experience gained from a non-pressured training. Finally, in the paper test it is possible that there will be no significant differences between the groups since as was found in the literature study; exploratory learning does not necessarily translate to technical knowledge.

5 Conclusion

From the literature review and personal interviews with airline pilots, it is clear that there is a need for further research into recurrent pilot training practices and the effects of pressure during pilot training. This was shown in this preliminary thesis report in the literature review in Chapter 2 from which an experiment testing the effects on learning during recurrent pilot training can be proposed. The scenarios for the experiment are explained in Chapter 3 followed by the experiment plan in Chapter 4.

Moreover, the research objective was defined in Chapter 1 as:

What are the effects of performance pressure during recurrent pilot training?

From this primary research objective a number of sub-questions can be defined some of which can be answered following the research done for this preliminary thesis report:

What are current practices in recurrent pilot training and what deficiencies are there?

Currently, pilots undergo recurrent training every six to twelve months to be assessed and maintain their skills such that they are allowed to keep on flying. This is essential since allowing pilots who under perform in training to continue flying can potentially lead to air accidents. During recurrent pilot training, pilots will split their time between learning in a classroom environment and in a flight simulator for practical training and assessment. One identified issue in recurrent training is the manner in which pilots are assessed. Traditionally, pilots underwent task-based training where they were assessed based on how well they completed SOPs and while these SOPs are useful in training and predictable scenarios; pilots struggle to adhere to SOPs and exhibit unpredictable behaviour in real life emergency situations. Moreover, despite pilots relying more and more on flight automation nowadays. There is still a lack of training for pilots that explicitly covers diagnosis and problem solving skills. This again leads to pilots being more prone to errors when confronted with an emergency situation in flight.

Through what sources do pilots experience performance pressure from during training?

There are a number of sources identified from which pilots experience performance pressure during training. Firstly, there is the fact that a pilot's progression in their career can depend on their performance in training. Pilots who under-perform in recurrent training can be assigned to remedial training, have promotions withheld and in extreme cases be terminated. Therefore, pilots who wish to continue flying and advance in their career are under pressure to pass recurrent training. Moreover, another source of pressure for pilots comes from the inadequacy of current training regulations. The lack of competent assessment from CAs leads to training being more of a box checking exercise rather than demonstrating technical knowledge or skills. This leads to pressure being put on pilots to simply memorise SOPs which have been proven to be less effective in flight.

To what extent does performance pressure inhibit a pilot's exploratory behaviour during unexpected and varied training?

To answer the final research question, an experiment will be conducted to investigate the effects of performance pressure on pilots' learning. Since pilots learn better with more exploratory learning during unexpected and varied training. The experiment will be designed to investigate how performance pressure affects pilots' exploratory behaviour and consequently how that affects their technical knowledge.

To do so, 30 pilots will be invited to participate in a simulator experiment where they will learn to operate the Garmin G1000 avionics and more specifically, the autopilot. Half of the pilots will undergo training with no pressure and the other half will be put under an artificial high pressure situation. The pilots will undergo 5 training scenarios where they can familiarise themselves with the avionics and utilise the problem-solving strategy that will be taught to them. Subsequently, they will have 2 test scenarios where their performances will be evaluated. This will be followed finally by a paper test where their theoretical knowledge will be tested. It is hypothesised that the group under no pressure will receive a higher quality training in the training phase and thus be able to more effectively lower their workload in the test scenarios.

References

Allsop, J., & Gray, R. (2014). Flying under pressure: Effects of anxiety on attention and gaze behavior in aviation. *Journal of Applied Research in Memory and Cognition*, *3*, 63–71. https://doi.org/10.1016/j.jarmac.2014.04.010

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*, 775–779. https://doi.org/10.1016/0005-1098(83)90046-8

Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology*, *16*, 361–383. https://doi.org/10.1002/ejsp.2420160405

Brouwer, D., Grootendorst, F., Bedoya, C., van Paassen, M., & Mulder, M. (2004). Experimental evaluation of a safety augmentation system for commercial aircraft. *AIAA Guidance, Navigation, and Control Conference and Exhibit*. https://doi.org/10.2514/6.2004-4996

Burratto, F., & Graef, R. (2021). Training pilots for resilience. *The Airbus Safety Magazine*. https://safetyfirst.airbus.com/training-pilots-for-resilience/

Cameron, M., Dahlstrom, N., & Kennedy, R. J. (2022). Competency scoring in simulation for experiential training (set). https://www.researchgate.net/publication/366216805_Competency_Scoring_in_Simulation_for_Experiential_Training_SET_A_white_paper_from_CAE_and_Emirates

Carolan, T. F., Hutchins, S. D., Wickens, C. D., & Cumming, J. M. (2014). Costs and benefits of more learner freedom. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *56*, 999–1014. https://doi.org/10.1177/0018720813517710

Casner, S. M., Geven, R. W., & Williams, K. T. (2013). The effectiveness of airline pilot training for abnormal events. *Human Factors*, *55*, 477–485. https://doi.org/10.1177/0018720812466893

Cassell, V. E., Beattie, S. J., & Lawrence, G. P. (2018). Changing performance pressure between training and competition influences action planning because of a reduction in the efficiency of action execution. *Anxiety, Stress and Coping*, *31*, 107–120. https://doi.org/10.1080/10615806.2017.1373389

Childs, J. M., & Spears, W. D. (1986). Flight-skill decay and recurrent training. *Perceptual and Motor Skills*, 235–242. https://doi.org/https://doi.org/10.2466/pms.1986.62.1.235

Clarke, P., Sheffield, D., & Akehurst, S. (2020). Personality predictors of yips and choking susceptibility. *Frontiers in Psychology*, *10*, 479670. https://doi.org/10.3389/fpsyg.2019.02784

Dahlstrom, N., Cameron, M., & Kennedy, R. J. (2022). Simulation for experiential training (set) as an enabler for evidence based training (ebt). https://www.researchgate.net/publication/363376160_Simulation_for_Experiential_Training_SET_as_an_enabler_for_Evidence_Based_Training_EBT_A_white_paper_from_CAE_and_Emirates

Dahlstrom, N., & Kennedy, R. J. (2022). A future learning ecosystem for pilot training. https://www.researchgate.net/publication/360715020_AVIATION_TRAINING_A_white_paper_from_CAE_and_Emirates_Airlines_A_future_learning_ecosystem_for_pilot_training

Demerouti, E., Veldhuis, W., Coombes, C., & Hunter, R. (2019). Burnout among pilots: Psychosocial factors related to happiness and performance at simulator training. *Ergonomics*, *62*, 233–245. https://doi.org/10.1080/00140139.2018.1464667

EASA. (2015a, September). *Annex ii to ed decision 2015/022/r*. https://www.easa.europa.eu/en/document-library/agency-decisions/ed-decision-2015022r

EASA. (2015b, December). *Ed decision 2015/027/r*. https://www.easa.europa.eu/document-library/agency-decisions/ed-decision-2015027r

EASA. (2016). *Part-fcl*. https://www.easa.europa.eu/sites/default/files/dfu/Part-FCL.pdf

EASA. (2021). *Flight examiner manual*. https://www.easa.europa.eu/en/downloads/135918/en

Eisenberger, R., & Aselage, J. (2009). Incremental effects of reward on experienced performance pressure: Positive outcomes for intrinsic interest and creativity. *Journal of Organizational Behavior*, *30*, 95–117. https://doi.org/10.1002/job.543

European Commission. (2019). *Report from the commission to the european parliament, the council, the european economic and social committee and the committee of the regions - aviation strategy for europe: Maintaining and promoting high social standards*. https://ec.europa.eu/transport/modes/air/25years-eu-aviation_en

Fletcher, G., & Bisset, G. (2017). *Pilot training review-interim report: Literature review cap 1581b*. https://publicapps.caa.co.uk/docs/33/CAP1581bLiteratureReview.pdf

Garmin. (2014). *Garmin g1000 pilot's guide for the piper pa-34-220t seneca*. http://static.garmin.com/pumac/190-01404-00_0A_web.pdf

Hinze, S. R., & Rapp, D. N. (2014). Retrieval (sometimes) enhances learning: Performance pressure reduces the benefits of retrieval practice. *Applied Cognitive Psychology*, *28*, 597–606. https://doi.org/10.1002/acp.3032

Holder, B. (2013). *Airline pilot perceptions of training effectiveness*. Boeing.

Huijbrechts, E.-J., & van Paassen, M. M. (2023). Flight deck procedures for a new generation of pilots. https://www.researchgate.net/publication/371721442

ICAO. (2013). *Manual of evidence-based training*. https://skybrary.aero/sites/default/files/bookshelf/3177.pdf

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

Landman, A., van Oorschot, P., van Paassen, M. M., 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*, *60*, 793–805. https://doi.org/10.1177/0018720818779928

Mavin, T. J. (2016). *Supporting learning across working life* (Vol. 16). Springer Nature. https://doi.org/10.1007/978-3-319-29019-5_9

Nikolic, 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*, 553–563. https://doi.org/10.1518/001872007X215647

Noon, H., & Murphy, M. (1984). Innovative approaches to recurrent training. https://ntrs.nasa.gov/api/citations/19850009719/downloads/19850009719.pdf

NTSB. (2010, February). *Aircraft accident report: Loss of control on approach, colgan air, inc. clarence center, new york, february 12, 2009*. https://www.ntsb.gov/investigations/accidentreports/reports/aar1001.pdf

NTSB. (2022, July). *Aviation investigation final report: Commutair flight 4933*. https://data.ntsb.gov/carol-repgen/api/Aviation/ReportMain/GenerateNewestReport/99050/pdf

Popovic, S. (2016, October). *Piper pa-34-220t seneca v*. https://upload.wikimedia.org/wikipedia/commons/8/85/Piper_PA-34-220T_Seneca_V_YU-VGI_RV_i_PVO_VS.jpg

Ranieri, A., Gregorini, S., Moreno, F., & Cano, M. (2019). *Evaluation report: Evaluation of the applicable rules for initial and recurrent pilot training, testing and checking*. https://www.easa.europa.eu/sites/default/files/dfu/evaluation_report_on_the_applicable_rules_for_initial_and_recurrent_pilot_training_testing_and_checking.pdf

Spielberger, C. D., Gorssuch, R. L., Lushene, P. R., Vagg, P. R., & Jacobs, G. A. (1983). Manual for the state-trait anxiety inventory. *Consulting Psychologists Press*.

Strauch, B. (2018). Ironies of automation: Still unresolved after all these years. *IEEE Transactions on Human-Machine Systems*, *48*, 419–433. https://doi.org/10.1109/THMS.2017.2732506

Stroosma, O., van Paassen, M. M., & Mulder, M. (2003). Using the simona research simulator for human-machine interaction research. *AIAA Modeling and Simulation Technologies Conference and Exhibit*. https://doi.org/10.2514/6.2003-5525

van Leeuwen, J. K. (2020). *The effectiveness of simulating active problem solving in pilot training to deal with automation surprises*.

van Oorschot, P. (2017). *Using unpredictability and variety in pilot training to improve performance in surprise situations*. http://resolver.tudelft.nl/uuid:537c5257-f785-4121-827c-8d5bbfeaeb92

Zijlstra, F. R. H., & van Doorn, L. (1985). *The construction of a scale to measure perceived effort*.

Part III

Appendices

A
Full Experiment Results

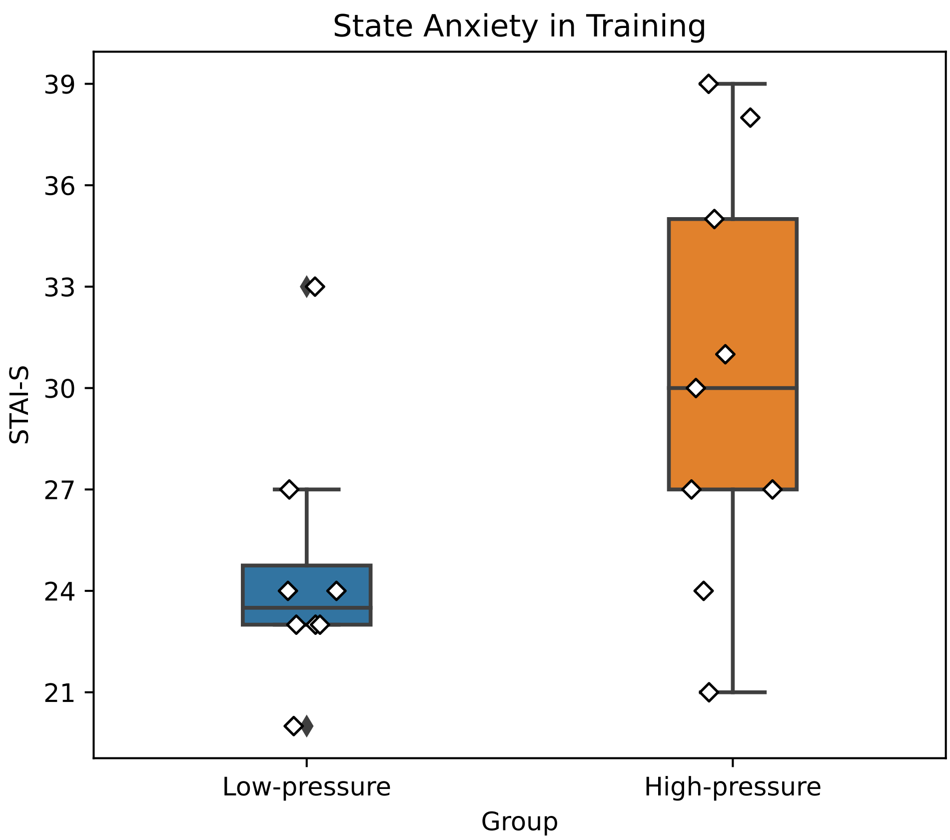


Figure A.1: State anxiety in training

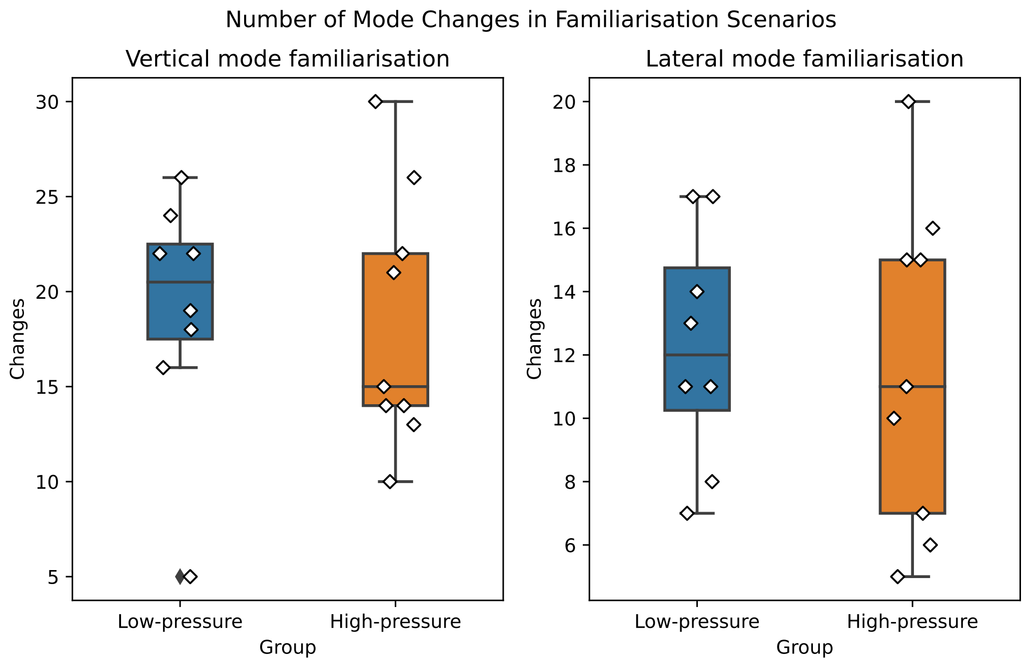


Figure A.2: Number of mode changes in the familiarisation training scenarios

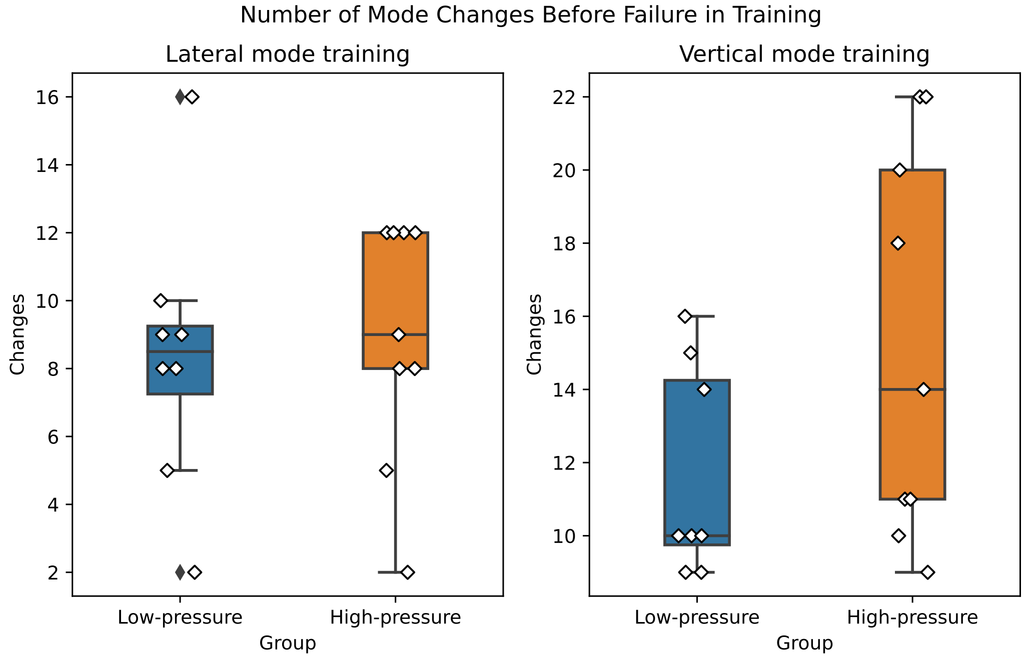


Figure A.3: Number of mode changes before the instrument failure in the training scenarios

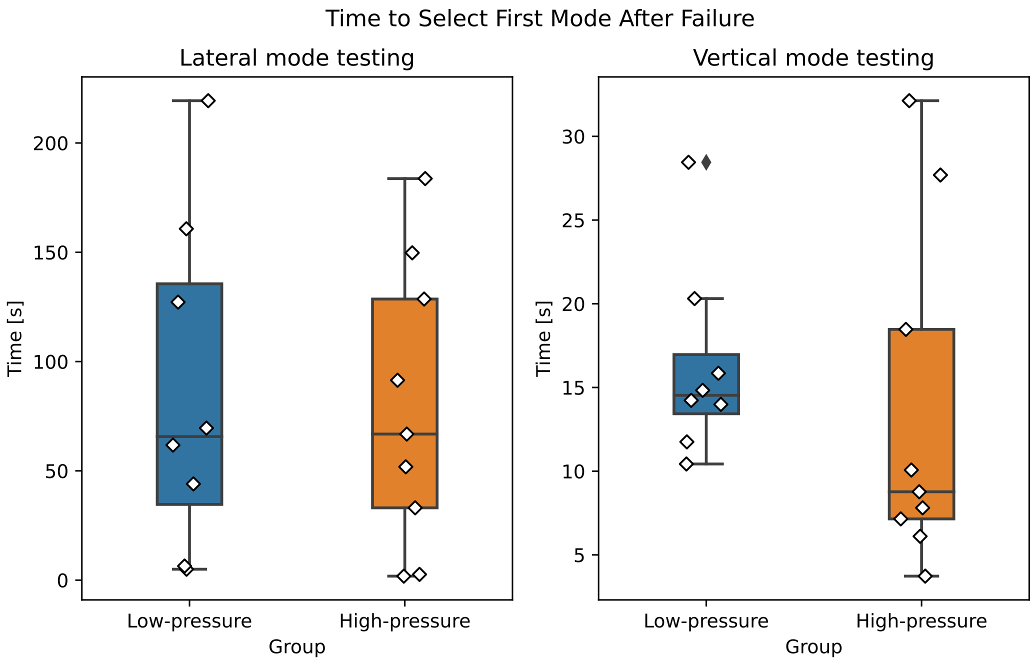


Figure A.4: Time taken to select the first mode after the instrument failure in the testing scenarios

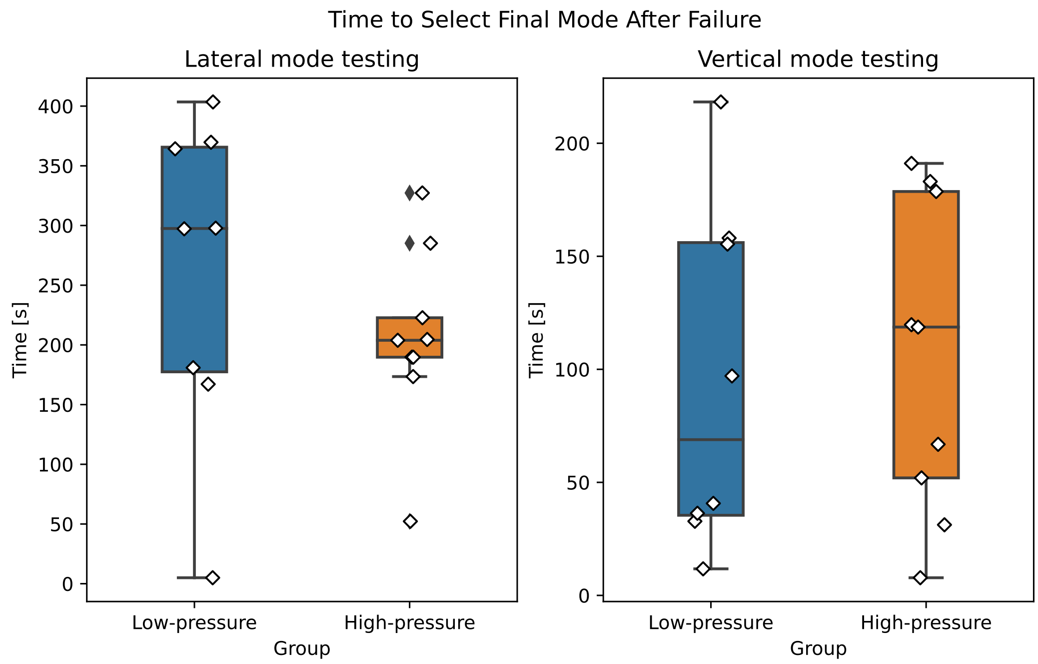


Figure A.5: Time taken to select the final mode after the instrument failure in the testing scenarios

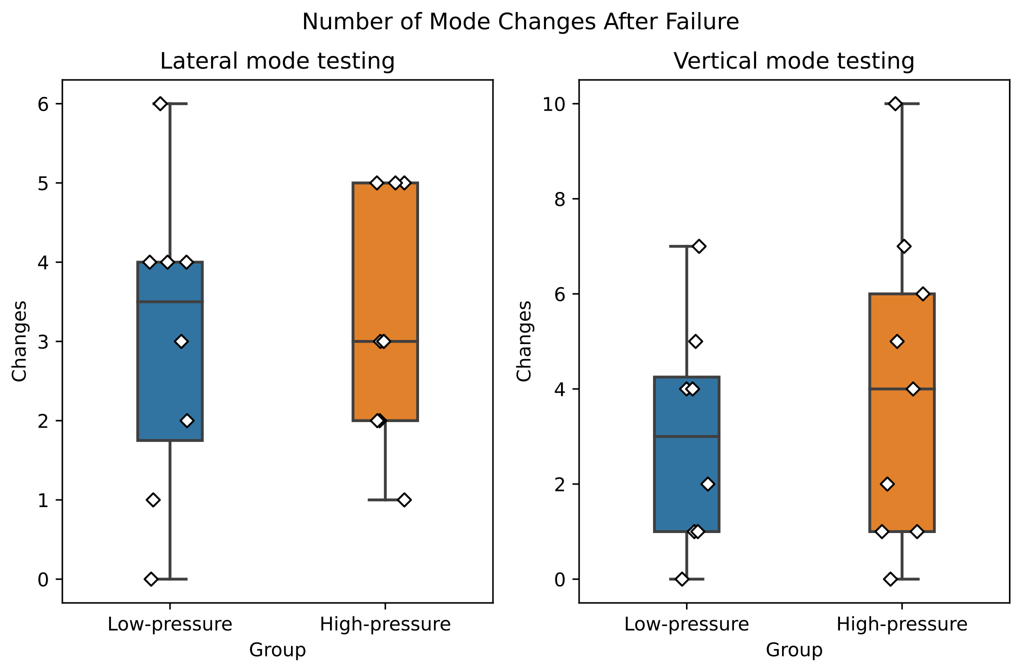


Figure A.6: Number of mode changes after the instrument failure in the testing scenarios

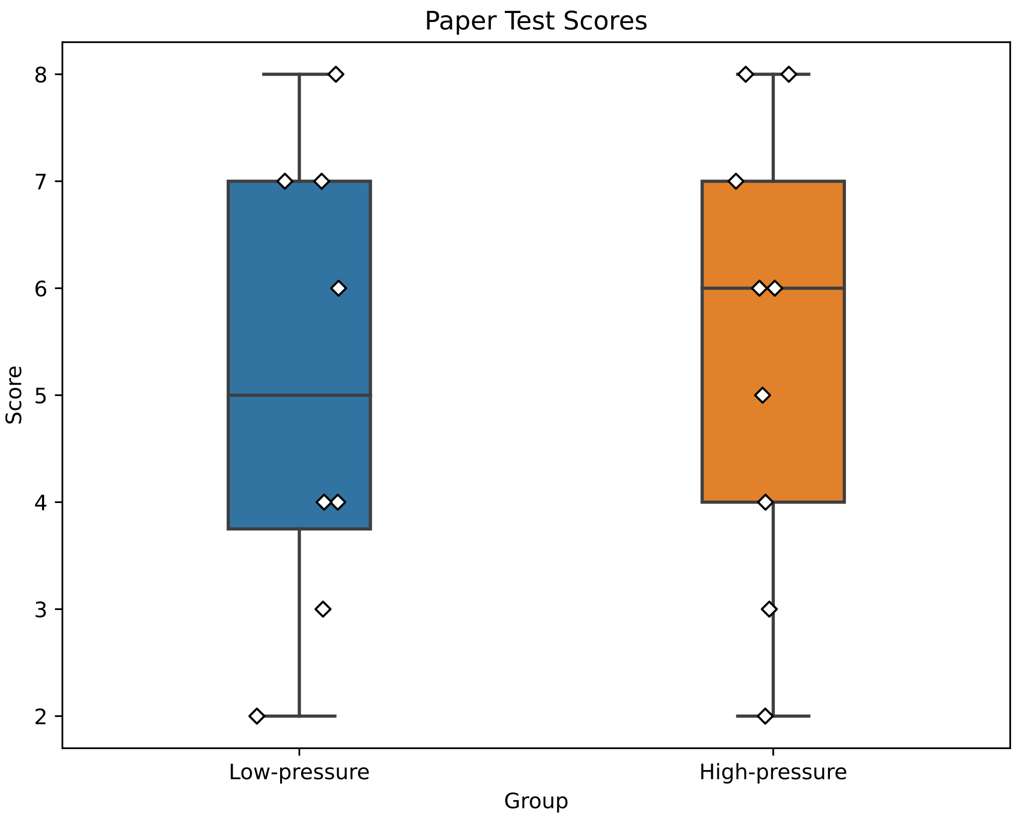


Figure A.7: Paper test scores

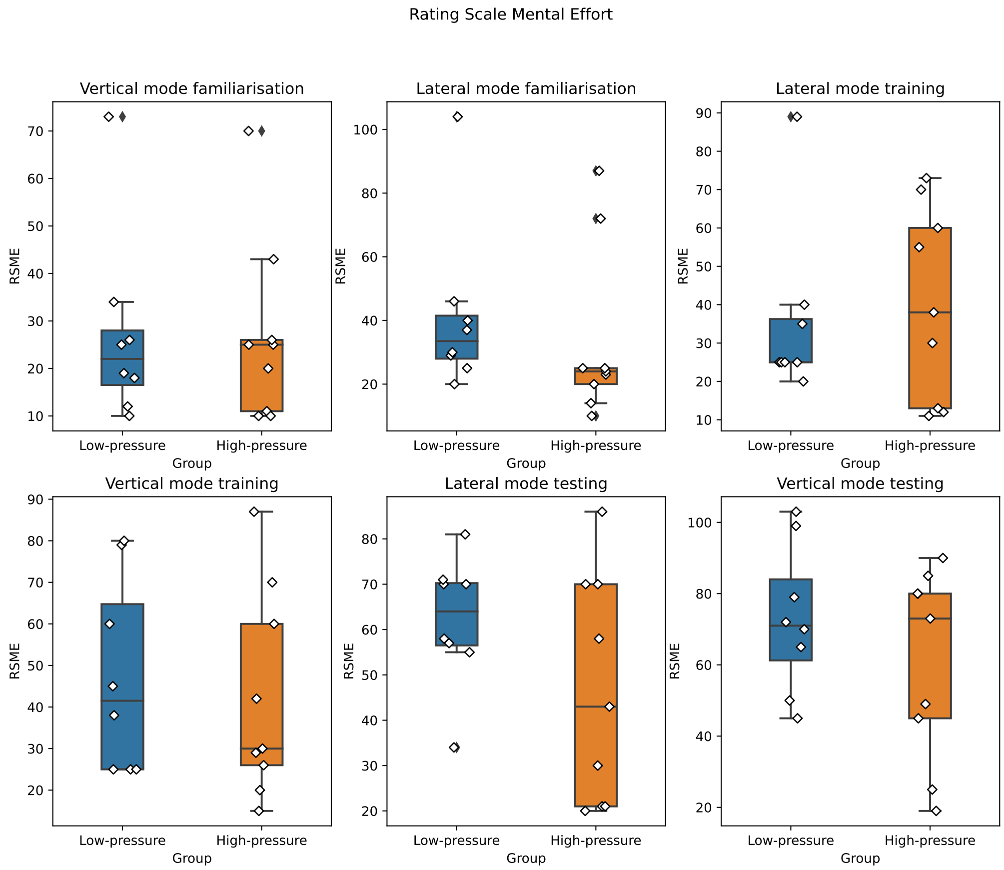


Figure A.8: Mental effort using the rating scale mental effort

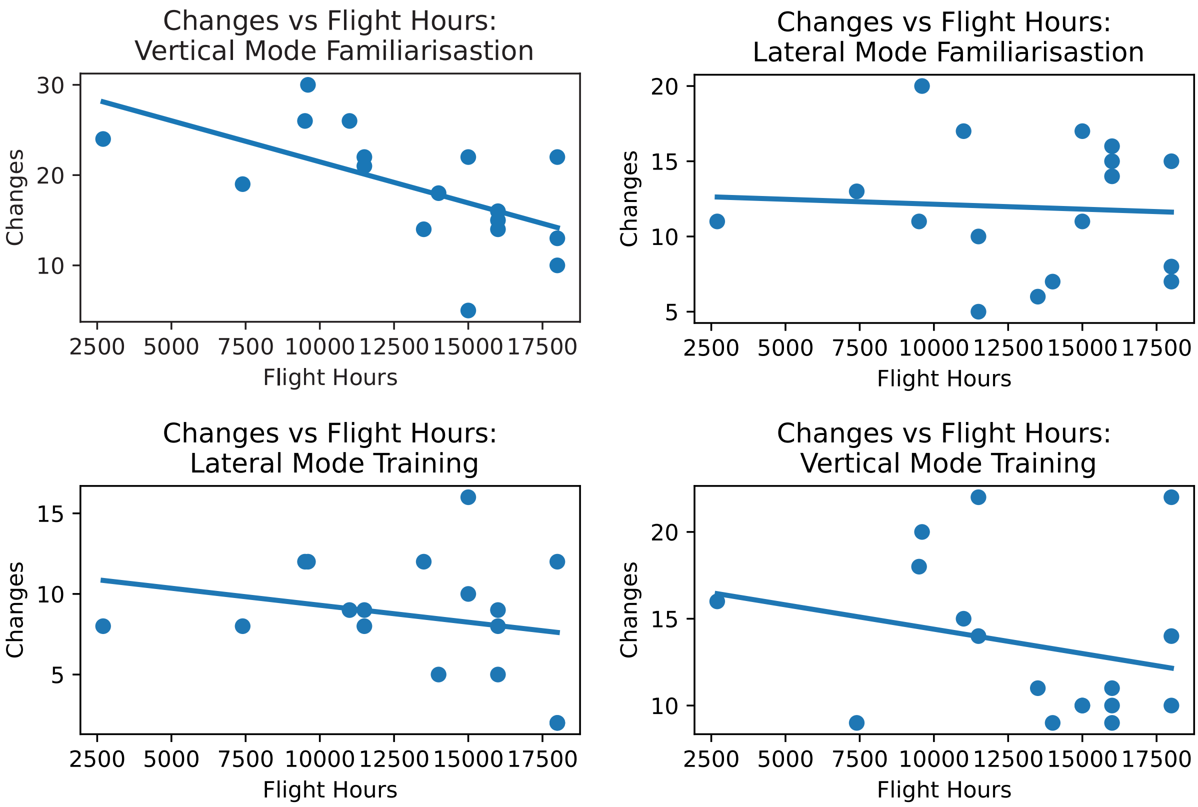


Figure A.9: Number of mode changes vs flight hours per scenario

B

Changes to SenecaAutomationTraining

B.1. VOR Navigation

The ability to utilise the standby frequencies was added enabling the pilot to navigate using four unique VOR beacons. These can be added in the `.nav` files by adding the desired frequencies to be displayed in `nav1stby` and `nav2stby`, four booleans can be used in `use_vors` to choose which VORs are active, and the extra locations for the beacons in `vor1stby` and `vor2stby`.

B.2. GPS Navigation

Completing the programmed GPS route will then remove the flight plan from the ”FPL” screen and the MFD. If in the navigation mode, the autopilot will then revert to the roll hold mode.

B.3. Position Failure Triggers

A new method for triggering failures was added in the `.sce` files. A trigger called `eventposition_x` and `eventposition_y` can be used which will trigger a failure based on the x and y position of the aircraft.

B.4. Outside Air Temperature Display

A new property to the `.sce` files was added called `base_temperature` which set the temperature shown on the PFD in the simulator. This enabled a more realistic clogged pitot tube scenario where the temperature on the display would drop showing a freezing temperature as the failure was occurring.

B.5. New Flight Level Change Controller

A first order lag filter was added to the airspeed reference used by the controller for the flight level change autopilot mode, this reduced the rapid oscillations in the controller due to the changes in airspeed from the turbulence model in the simulation.

C

Flight Plans

Flight Plan Scenario 612

	Identifier	Frequency	Course	Distance	Altitude
			[deg]	[NM]	[ft]
	SY411		018	2.6	**6000**
	SY803		018	3.6	**5000**
	SY416		042	6.1	**4000**
	SY417		033	4.6	3000
	LUSUL		033	5.8	2000

Flight Plan Scenario 613

	Identifier	Frequency	Course	Distance	Altitude
			[deg]	[NM]	[ft]
	VABAT		141	4.3	5000
	UMIRU		118	6.7	5000
	GINOK		065	5.0	5000
	LAW	115.80	116	5.1	5000
	CMP	114.50	193	6.3	5000

Flight Plan Scenario 621

	Identifier	Frequency	Course	Distance	Altitude
			[deg]	[NM]	[ft]
	MOL	116.60	145	5.2	4000
	WAR	114.90	184	5.3	4000
	OKC	113.45	116	4.9	4000
	SIE	114.70	091	20.0	4000

Flight Plan Scenario 622

	Identifier	Frequency	Course	Distance	Altitude
			[deg]	[NM]	[ft]
	GD902		113	2.5	2000
	IPLAM		113	4.0	3000
	GD903		113	5.0	4000
	GD904		113	5.0	5000
	BERNU		166	5.0	6000

Flight Plan Scenario 631

	Identifier	Frequency	Course	Distance	Altitude
			[deg]	**[NM]**	**[ft]**
	ABRAX		277	3.2	5000
	SUKAV		265	4.5	5000
	EKNEV		296	4.9	5000
	VLM	114.30	250	5.3	5000
	VOZ	116.95	216	6.1	5000

Flight Plan Scenario 632

	Identifier	Frequency	Course	Distance	Altitude
			[deg]	**[NM]**	**[ft]**
	XALUR		245	4.6	**5000**
	LL523		247	6.5	**4000**
	LL522		245	4.0	**3000**
	D066I		241	4.8	**2000**
	EPLL		244	8.2	**600**

D

Briefing

12/4/2024

1

Experiment Goal

- Investigate the effectiveness of exploratory learning during automation training

TU Delft

09-10-2020

2

Experiment Tasks

- Ground school
- Fly a variety of training and testing scenarios
- Each scenario has a corresponding flight plan with waypoints, course, distance and altitude
- Follow the flight plan as closely as possible whilst using the highest possible mode of automation
- Some scenarios may contain instrument failures
- Resolve instrument failures using a problem-solving strategy and subsequently fly with the highest possible mode of automation
- Evaluation based on variety of factors i.e. deviation from flight plan, use of highest automation mode etc…

TU Delft

3

Informed Consent

- This experiment will involve collecting data
- This data will possibly be published
- The informed consent form is the only link between your name and your subject ID

TU Delft

4

Safety

- Please refer to the SIMONA safety briefing video
- If you feel motion sick during or want to stop at any point in the experiment, please say so
- The control loading actuators can exert large forces
- Keep clear during startup and calibration
- https://www.youtube.com/watch?v=PXijsyJ3hro

TU Delft

5

E
Ground School Slides

12/4/2024

1

Primary Flight Display (PFD)

TU Delft

2

Primary Flight Display (PFD)

Functional buttons to interact with AFCS

TU Delft

3

3

Primary Flight Display (PFD)

Active lateral and vertical modes (green)

Armed vertical mode (white)

TU Delft

Primary Flight Display (PFD)

Mode
selectors

TU Delft

5

Primary Flight Display (PFD)

Setting altitude and heading reference

TUDelft

6

6

Primary Flight Display (PFD)

TU Delft

Primary Flight Display (PFD)

Change map
zoom on
MFD

Vertical Modes

Vertical Mode	Button	Annunciation
Pitch Hold[1]	**AP**	PIT
Altitude Hold[2]	**ALT**	ALT nnnn ft
Vertical Speed[1]	**VS**	VS nnn fpm
Flight Level Change[1]	**FLC**	FLC nnn kts
Vertical Path Tracking[3]	**VNV**	VPTH

[1] Will automatically arm Selected Altitude Capture mode (**ALTS**)

[2] Will hold current altitude unless transitioned to automatically from **ALTS**

[3] Will automatically arm VNV Target Altitude Capture (**ALTV**) and capture when in range

TU Delft

9

9

Lateral Modes

Lateral Mode	Button	Annunciation
Roll Hold	**AP**	ROL
Heading Select	**HDG**	HDG
Navigation (GPS/VOR)	**NAV**	GPS/VOR

TU Delft

10

10

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** switch to lower-level automation

TU Delft

11

11

F

Paper Test

1. What happens when the CDI softkey is pressed while in Navigation (**NAV**) mode?
    a. The autopilot reverts to Roll Hold mode and the CDI cycles to the next navigation source
    b. The CDI cycles to the next navigation source
    c. The autopilot reverts to Roll Hold mode and the CDI softkey must be pressed again to cycle navigation source
    d. The autopilot reverts to Roll Hold mode

2. What is the highest possible mode of automation that can be used with an unreliable airspeed indication?
    a. Pitch Hold
    b. Vertical Speed
    c. Flight Level Change
    d. Autopilot cannot be used

3. The plane is flying at 4000 feet, the autopilot is on and climbing using Vertical Speed (**VS**) mode at a rate of +300 feet per minute. The altitude bug is set at 5000 feet. What happens when the ALT button is pressed?
    a. The autopilot will arm Altitude Hold (**ALT**) mode and continue climbing until the aircraft reaches 5000 feet
    b. The autopilot will switch to Selected Altitude Capture (**ALTS**) mode with Altitude Hold (**ALT**) mode armed and continue climbing until the aircraft reaches 5000 feet
    c. The aircraft will switch to Altitude Hold (**ALT**) mode and hold an altitude of 4000 feet
    d. The aircraft will switch to Altitude Hold (**ALT**) mode and arm Selected Altitude Capture (**ALTS**) mode and hold an altitude of 4000 feet

4. The autopilot is in Roll Hold (**ROL**) mode, the current heading is 200 degrees. The Horizontal Situation Indicator (HSI) is on VOR1. The current information is available: the heading bug is set to 10 degrees, the selected course is 100 degrees, the desired track in the flight plan is 180 degrees. The NAV key is subsequently pressed, what heading reference will the autopilot try and follow?
    a. 200 degrees
    b. 10 degrees
    c. 100 degrees
    d. 180 degrees

5. Which statements are true? (1) The DME information will still be available after a VOR failure. (2) When in VOR navigation, the DME information window will automatically cycle to correspond with the active VOR source.
    a. Statement (1) is true
    b. Statement (2) is true
    c. Both statements are true
    d. Neither statement is true

6. The following is shown in the status box. What statements are true? (1) Roll Hold is the active lateral mode. (2) The autopilot is not engaged. (3) Selected Altitude Capture mode (**ALTS**) is armed.

   a. Only statement 1 is true
   b. Statements 1 and 3 are true
   c. All the statements are true
   d. Only statement 3 is true

7. What is an indication that the autopilot is correctly moded to capture the current vertical navigation profile?
   a. The VNV key has been pressed
   b. The flight plan indicates an altitude for every waypoint
   c. The altitude bug is set to the desired altitude
   d. VPTH is annunciated in the status box

8. Which information gives the course to the active waypoint from the current position?
   a. Desired Track (DTK) on the MFD
   b. Track (TRK) on the MFD
   c. Bearing (BRG) on the PFD
   d. This information is only shown on the flight plan in the MFD

9. This following is shown in the status box. What will pressing the AP button do?

   a. The lateral and vertical modes will revert to Roll Hold (**ROL**) and Pitch Hold (**PIT**) mode
   b. The autopilot will turn off
   c. The autopilot will turn on
   d. The flight director will turn off

10. Which statements are true? (1) VOR1 and VOR2 use the same course selected by the CRS knob. (2) Vertical Speed mode will try and maintain the reference vertical speed regardless of the airspeed. (3) The pitch reference in Pitch Hold mode cannot be changed using the Nose Up and Nose Down buttons.
   a. Statements 1 and 2 are true
   b. Only statement 1 is true
   c. Only statement 2 is true
   d. Statements 2 and 3 are true

G Experiment Script

EXP-1 Participants enter simulator

- Please keep your hands and feet out the way while I calibrate the controls

Calibrate controls

- You may encounter some instrument failures scenarios in this experiment so please recall the problem solving strategy that was presented in the ground school
- Notice autopilot behaviour is off nominal
- Identify which sensor or system is faulty
- Identify implications on autopilot
- Switch to alternative information source if possible **OR** switch to lower-level automation

EXP-2 Manual flight

[LOAD 611_IntroductionManualFlight.sce]

- We'll now begin with the first training scenario
- We'll go through a small manual flight scenario so you can familiarise yourself with the plane
- Make sure flaps and gear are up and set the throttle to cruise (around 60%)
- I'll give some instructions on how to fly
- Are you ready? 3, 2, 1. Your controls

[INSTRUCT PARTICIPANT]

- Fly at 100 knots and trim
- Turn right to heading 030
- Decrease power and descend to 5000 ft of altitude
- Increase power to 120 knots and trim
- My controls

End manual flight scenario

- For reference the scenarios have been designed to be flown between 115-120 knots

EXP-4 Vertical mode familiarisation

- We'll now move on to the next scenario
- Please take the flight plan marked "612" from the pack marked "training" and study it briefly so you're familiar with what route you'll have to fly

[LOAD 612_IntroductionVerticalModes.sce]

- Please remember your objective will be to fly the route as closely as you can whilst in the highest possible mode of automation

No pressure	**Pressure variant**
Use this scenario to learn the automatic flight control system to the best of your ability regardless of safety or performance. There are no consequences if you fly poorly or crash the plane.	The number of scenarios you'll fly is dependent on how well perform on these training scenarios. Since we're interested in how well you can learn, your data will be recorded on this run for analysis.

- Are you ready? 3, 2, 1. Your controls

[PARTICIPANT FLIES SCENARIO 612]

Stop simulation after reaching LUSUL

- You've now reached the end of this scenario

EXP-3 Measure workload

- Please take a new form from the pack marked "RSME" and indicate your participant number, the flight plan you just flew and compile the form

EXP-5 Lateral mode familiarisation

[LOAD 613_IntroductionLateralModes.sce]

- We'll now move on to the next scenario
- Please take the flight plan marked "613" from the pack marked "training" and study it briefly so you're familiar with what route you'll have to fly
- Please remember your objective will be to fly the route as closely as you can whilst in the highest possible mode of automation

No pressure	Pressure variant
Use this scenario to learn the automatic flight control system to the best of your ability regardless of safety or performance. There are no consequences if you fly poorly or crash the plane.	The number of scenarios you'll fly is dependent on how well perform on these training scenarios. Since we're interested in how well you can learn, your data will be recorded on this run for analysis.

- Are you ready? 3, 2, 1. Your controls

[PARTICIPANT FLIES SCENARIO 613]

Stop simulation after reaching CMP

- You've now reached the end of this scenario

EXP-3 Measure workload

- Please take a new form from the pack marked "RSME" and indicate your participant number, the flight plan you just flew and compile the form

EXP-6 End familiarisation scenarios

- We're going to take a quick break so I'm going to let you step out of the simulator now

Give participant a break

EXP-7 Brief training scenarios

No pressure	Pressure variant
Please take the flight plan marked “621” from the pack marked “training” and study it briefly so you’re familiar with the route you’ll have to fly.	I think we’re going to do a few extra training scenarios. Please take the flight plan marked “621” from the pack marked “extra” and study it briefly so you’re familiar with the route you’ll have to fly.

EXP-8 VOR receiver failure

[LOAD 621_VORReceiverFailure.sce]

- Please remember your objective will be to fly the route as closely as you can whilst in the highest possible mode of automation

No pressure	Pressure variant
Use this scenario to learn the automatic flight control system to the best of your ability regardless of safety or performance. There are no consequences if you fly poorly or crash the plane.	The number of scenarios you’ll fly is dependent on how well perform on these training scenarios. Since we’re interested in how well you can learn, your data will be recorded on this run for analysis.

- Are you ready? 3, 2, 1. Your controls

[PARTICIPANT FLIES SCENARIO 621]

Stop simulation after passing OKC

- You’ve now reached the end of this scenario

EXP-3 Measure workload

- Please take a new form from the pack marked “RSME” and indicate your participant number, the flight plan you just flew and compile the form

EXP-9 Blocked static port

[LOAD 622_BlockedStaticPort.sce]

No pressure	Pressure variant
Please take the flight plan marked “622” from the pack marked “training” and study it briefly so you’re familiar with the route you’ll have to fly.	I think we’re going to do one more scenario. Please take the flight plan marked “622” from the pack marked “extra” and study it briefly so you’re familiar with the route you’ll have to fly.

- Please fly the flight plan to the best of your ability whilst using the highest mode of automation possible

No pressure	Pressure variant
Use this scenario to learn the automatic flight control system to the best of your ability regardless of safety or performance. There are	Since we’re interested in how well you can learn, your data will be recorded on this run for analysis.

no consequences if you fly poorly or crash the plane.	

- Are you ready? 3, 2, 1. Your controls

[PARTICIPANT FLIES SCENARIO 622]

Stop simulation after passing GD904

- You've now reached the end of this scenario

EXP-3 Measure workload and anxiety

- Please take a new form from the pack marked "RSME" and indicate your participant number, the flight plan you just flew and compile the form
- Now take a form marked "STAI-S", please think back to how you felt at the beginning of the training and fill in the form

EXP-10 End training scenarios

Explain end of training scenarios

- We've now reached the end of the training scenarios and we're going to move onto the testing scenarios

EXP-11 Brief test scenarios

- In these testing scenarios your aim is now to firstly fly the aircraft as safely as possible and then subsequently if possible, to reduce your workload using as high a level of automation as you think is safe to fly with
- All the data from your performance during these scenarios will be recorded and used to analyse your performance
- Your performance will be scored based on how well you flew the flight plan and how much you use the highest level of automation

EXP-12 GPS Failure

[LOAD 631_GPSFailure.sce]

- Please take the flight plan marked "631" from the pack marked "testing" and study it briefly so you're familiar with the route you'll have to fly
- Please remember your task is to fly safely and then reduce workload through automation
- Are you ready? 3, 2, 1. Your controls

[PARTICIPANT FLIES SCENARIO 631]

Stop simulation after passing VLM

- You've now reached the end of this scenario

EXP-3 Measure workload

- Please take a new form from the pack marked "RSME" and indicate your participant number, the flight plan you just flew and compile the form

EXP-13 Ask to motivate actions

- In this testing scenario, after the instrument failure you switched to ____
- Why did you pick this mode?
- Do you think this was the highest and safest level of automation you could have used?

EXP-14 GPS failure

[LOAD 632_BlockedStaticPort.sce]

- Please take the flight plan marked "632" from the pack marked "testing" and study it briefly so you're familiar with the route you'll have to fly
- Please remember your task is to fly safely and then reduce workload through automation
- Are you ready? 3, 2, 1. Your controls

[PARTICIPANT FLIES SCENARIO 632]

Stop simulation after passing D066I

- You've now reached the end of this scenario

EXP-3 Measure workload and anxiety

- Please take a new form from the pack marked "RSME" and indicate your participant number, the flight plan you just flew and compile the form
- Now take a form marked "STAI-S", please think back to how you felt at the beginning of the testing scenarios and fill in the form

EXP-13 Ask to motivate actions

- In this testing scenario, after the instrument failure you switched to ____
- Why did you pick this mode?
- Do you think this was the highest and safest level of automation you could have used?
- Thank you for doing that, I'm now going to stop the simulator and allow you back out

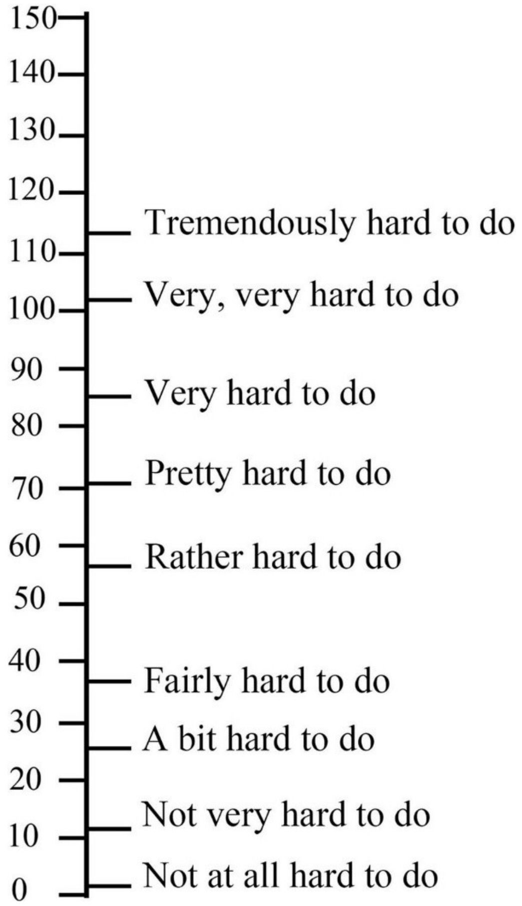
Explain the experiment to the pressured group

H

RSME and STAI-S

RSME

Scenario: ____________ How much mental effort did the scenario require? (Place a cross on the line)



STAI-S

Scenario: __________

Some statements that people have used to describe their feelings are given below. Read each statement and then circle the response option to the right to indicate how you feel right now, that is, at this moment. Do not spend too much time on any one statement but give the answer which seems to describe your present feelings best.

1 – Not at all

2 – Somewhat

3 – Moderately

4 – Very much

I feel calm	1	2	3	4
I feel secure	1	2	3	4
I am tense	1	2	3	4
I am regretful	1	2	3	4
I feel at ease	1	2	3	4
I feel upset	1	2	3	4
I am currently worried about possible misfortunes	1	2	3	4
I feel rested	1	2	3	4
I feel anxious	1	2	3	4
I feel comfortable	1	2	3	4
I feel self-confident	1	2	3	4
I feel nervous	1	2	3	4
I am jittery	1	2	3	4
I feel “high-strung”	1	2	3	4
I am relaxed	1	2	3	4
I feel content	1	2	3	4
I am worried	1	2	3	4
I feel overexcited and rattled	1	2	3	4
I feel joyful	1	2	3	4
I feel fine	1	2	3	4