DOT/FAA/TC-23/7

Federal Aviation Administration
William J. Hughes Technical Center
Aviation Research Division
Atlantic City International Airport
New Jersey 08405

Negative transfer of training: Simulator study into the effects of overruled pilot decision making

February 2023

Final report

U.S. Department of Transportation
Federal Aviation Administration

This report is available at the Federal Aviation Administration William J. Hughes Technical Center's Full-Text Technical Reports page: actlibrary.tc.faa.gov in Adobe Acrobat portable document format (PDF).

Technical Report Documentation Page

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/TC-23/7		

4. Title and Subtitle	5. Report Date
Negative transfer of training: Simulator study into the effects of overruled pilot decision making	February 2023
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Landman, A. Mol, D., Emmerik, M. L. van, Groen, E. L.	

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Netherlands Organization for Applied Scientific Research (TNO), Human Performance; Soesterberg, The Netherlands.	TNO 2023 R10152
	11. Contract or Grant No.

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
FAA Policy and Innovation Division Aircraft Certification Service Bldg. N243 Moffett Field, CA 94035	
	14. Sponsoring Agency Code

15. Supplementary Notes

This report was originally published by the Netherlands Organization for Applied Scientific Research (TNO), Human Performance; Soesterberg, The Netherlands as TNO 2023 R10152

16. Abstract

In this simulator study, we investigated the potential negative transfer of training that may result from situations where a correct decision made by the trainee is being overruled by the instructor to save training time. We hypothesized that discarding a correct decision in a training situation can introduce confusion when the trainee encounters a similar situation in operational practice. The experiment was performed in a fixed-base simulator of a twin-engine business jet. Two groups of 19 commercial pilots participated in a training session consisting of four training scenarios, immediately followed by a test session with four other test scenarios. Whereas in the “control” group the training scenarios were presented as briefed, the pilots in the experimental, or “overruled” group included some events (e.g., an enhanced ground proximity warning (EGPWS)) which required a response. However, the instructor directed the pilots to refrain from the response. It was expected that pilots in the overruled group were more likely to show delayed response or even ignore alarms in the test scenarios. The objective results showed no significant effect of the training manipulation on the performance of the pilots in the test scenarios involving the same type of events as the overruled group had encountered during the training. However, according to the subjective results, 32% of the participating pilots stated that they had been confused during the test, and that the training influenced their decision making. Within the limits of this study, these findings are an indication of negative transfer of training.

17. Key Words	18. Distribution Statement
Training – Simulation Training – Decision-making	This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actlibrary.tc.faa.gov.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	19. Security Classif. (of this report)
Unclassified	Unclassified	45	Unclassified

Contents

Figures

Tables

Acronyms

Acronym	Definition
ATC	Air Traffic Control
CAS	Calibrated air speed
DME	Distance measuring equipment
EGPWS	Enhanced Ground Proximity Warning System
EHWO	Woensdrecht Air Base, The Netherlands
FAA	Federal Aviation Administration
FD	Flight director
FI	Flight instructor
FO	First officer
GA	Go Around
HSI	Horizontal Situation Indicator
ICAO	International Civil Aviation Organization
ILS	Instrument landing system
IMC	Instrument meteorological conditions
LSGS	Sion Airport, Switzerland
MANOVA	Multivariate analysis of variance
MCP	Mode control panel
ND	Navigation display
OM-C	Operation Manual, Part C
PAPI	Precision approach path indicators
PFD	Primary flight display
SD	Standard deviation
SO	Second officer
TNO	Netherlands Organization for Applied Scientific Research
TRE	Type rating examiner
TRI	Type rating instructor
VMC	Visual meteorological conditions
VOR	VHF omnidirectional range

Executive summary

This report describes details of a flight simulator study investigating the potential negative transfer of training that may result from situations where a – principally correct – decision of a trainee is being overruled by the instructor to save training time. Prior to the study, participating pilots were told that its purpose was related to workload and decision making in go-around situations, and that the exact purpose would be explained afterwards. The study took place in a fixed-base flight simulator, running a flight model of a twin-engine business jet. In total, 38 commercial pilots participated on a voluntary basis, equally divided into two groups: an experimental, or 'overruled' group, and a control group. Each pilot received a briefing of 20 minutes, a simulator session of 80 minutes, followed by a debriefing of 20 minutes. Four pilots were tested each day.

The simulator session consisted of a familiarization part, a training part, and a test part. In the training part, all pilots practiced two go-arounds and landings in poor visibility on two different airfields: Woensdrecht Air Base (EHWO, The Netherlands) and Sion airport (LSGS, Switzerland). For the pilots in the control group, these training scenarios occurred uneventful. In contrast, the pilots in the overruled group experienced an event in two of the training scenarios (one at EHWO and the other one at LSGS) which should prompt them to abort the approach. However, the instructor told them to discard their decision, and to continue with the approach in order to practice the landing or go-around above the runway. In the manipulation scenario at EHWO, visual contact with the runway was lost while the aircraft was in the downwind leg. After an expression of confusion by the pilot, the instructor directed them to the runway and instructed them to land anyway. In the manipulation at LSGS, the pilots in the overruled group experienced a warning from the Enhanced Ground Proximity Warning System (EGPWS) with one repetition of the aural message "Terrain, terrain, pull up" when the participant was making a turn close by a mountain ridge. The instructor immediately said to this group, "whereas you should normally respond to this alarm, you can proceed for now in order to practice the response to the ATC instruction."

After completing the training part, all pilots performed the test part consisting of four test scenarios in which the pilots made an approach to LSGS starting at 5000 ft. Two of these test scenarios served as a transfer test: one to test the pilots' response to gradual loss of visual contact with the runway, which should prompt a go-around; and the other to test the pilots' response to an unexpected triggering of the EGPWS during the approach. To control for inter-individual differences, two baseline test scenarios were added in which the pilots were told in advance to perform a go-around, and an EGPWS escape maneuver, respectively.

The dependent variables comprised objective performance measures (for example, the time to the first response, time to reach 8000 ft), as well as subjective ratings on perceived stress and workload, etc. In addition, after each test scenario the pilots were asked to describe their decision-making process in free format.

For both tests (i.e., go-around and EGPWS warning), the objective performance measures significantly differed between the transfer test and the corresponding baseline test (which contained the same event as the transfer test, but pre-announced). For each scenario, there was no significant effect of group on the performance measures, nor on the subjective ratings. However, six (32%) of the 19 overruled pilots mentioned in the comments that the training session had influenced their decision making in the transfer test with the unexpected EGPWS alarm. They stated that the training manipulation had caused confusion and hesitation. In two pilots, this led to a substantial delay in response. These comments are an indication of negative transfer of training, which should be taken seriously, considering the controlled simulator environment in which the transfer test took place. This finding suggests that instructors should carefully consider how they overrule the trainees' decisions during training. If taking shortcuts is not avoidable, then the instructor should at least provide a clear disclaimer for why this is done, and explain that under normal circumstances, the pilot's own decision was appropriate.

1 Introduction

Negative transfer of training refers to situations in which learning in a training environment results in the degradation of performance in a different context, or in the real world (Alexander, Brunyé, Sidman, & Weil, 2005). On request of the Federal Aviation Administration (FAA), The Netherlands Organization for Applied Scientific Research (TNO) previously carried out a literature survey on the topic of negative transfer of training, including interviews with various training experts (Pennings, Oprins, Schoevers , & Groen, 2019). The findings of this survey led to two previously reported experimental studies. In the first study, we found evidence of negative transfer of training when the variability of practice with a procedural task was low or high (Landman, Pennings, Blankendaal, Bosch, & Groen, 2022). In the second study, we found no evidence for the experts' concern that deliberately going beyond alarms during stall recovery training leads to delayed responses in situations where a stall happens unexpectedly (Landman, Mol, Emmerik, & Groen, 2022).

In this third experimental study, we investigate the potential negative transfer of training that may result from situations where a correct decision made by the trainee is overruled by the instructor to save training time. This training practice was identified by several experts in the study of Pennings et al. (2019). One expert from this study stated: "Often choices are made in simulator scenarios which, because of limited time, are completely against what a pilot should do in practice." (Pennings, Oprins, Schoevers , & Groen, 2019, pp. 5, Appendix B). Another expert mentioned: "Due to limited simulator time, pilots are sometimes allowed to land from an unstable approach, or the instructor will stop the simulator, whereas in real life the pilot should initiate a go-around procedure." (Pennings, Oprins, Schoevers , & Groen, 2019, pp. 8, Appendix B). We hypothesize that discarding a correct decision during a training situation can possibly introduce confusion when the pilot encounters a similar situation in operational practice. This can be considered a negative training effect. This effect can be aggravated when the pilot becomes startled or surprised, which may impair one's ability to evaluate the context and make accurate considerations (Landman, Groen, Van Paassen, M. M., Bronkhorst, & Mulder, 2017).

Two flight conditions were selected for the current experiment, which seemed suitable to overrule a pilot's decision in a training setting. The first condition involves a pilot's decision to perform a go-around in response to deteriorating visibility on the runway during the approach. This scenario was inspired by a discussion with an instructor pilot of an airline company, who explained that during their simulator training in Instrument Meteorological Conditions (IMC), the visibility around the simulated airfield is sometimes set to Visual Meteorological Conditions (VMC) in order to avoid a time-consuming go-around, and to continue with the training. The second condition involves a pilot's decision to perform an escape maneuver at the activation of the Enhanced Ground Proximity Warning System (EGPWS). This scenario is interesting because some training organizations allow their pilots to ignore the EGPWS alarm in the training when it is activated during an approach in mountainous terrain while the pilots have the runway in sight. Furthermore, the EGPWS is also interesting because in our previous study on stall recovery training (Landman, Mol, Emmerik, & Groen, 2022), we coincidentally observed that pilots had difficulty with responding to an unexpected EGPWS warning that was part of one of the scenarios. Three out of 40 pilots performed the required procedure incorrectly, which led to flight into terrain for two pilots. This observation suggests that pilots may, at times, be confused about how to respond to the EGPWS, and we expected that instructing them to sometimes ignore the EGPWS may increase the confusion.

2 Method

2.1 Design

The study had a between-subject design: two groups of pilots were compared. Both groups received a simulator training. One group, the overruled group, was presented with two situations in which the conditions would warrant a response; however, the instructor directed them to refrain from responding. The other group, the control group, was confronted with the same situations in the training, although the conditions did not require a response. After the training, the performance of all pilots from both groups was measured in two test scenarios that included similar situations requiring a response. For each test scenario, a corresponding baseline test scenario was also included, in which the pilots were told in advance about the required response. The purpose of these baseline test scenarios was to better control for the within-subject variability in the pilots' responses. This was done by testing the effects of the training manipulation using the statistical two-way interaction effect between group (control, overruled) × test (test, baseline test). This way, we expected to find a manipulation effect when the difference in performance between the test (go-around, EGPWS) and the corresponding baseline test depends on the group.

2.2 Participants

Airline pilots (n = 38) were recruited in three different ways: 1) through news bulletins of their company; 2) through word-of-mouth; and 3) from a list of participants from previous experiments. Two groups were created that were balanced as much as possible based on the characteristics shown in Table 1. The control group had on average less flight hours, but this difference was not significant as indicated by an independent *t*-test, $p = 0.365$. The control group had more flight hours on smaller aircraft compared to the overruled group. Because flight experience was not normally distributed within the groups, a (non-parametric) Mann-Whitney U test was used to compare the groups. This revealed that pilots in the control group had nearly significantly more flight hours on small aircraft than those in the overruled group, $U = 115.50$, $p = 0.056$.

Table 1. Characteristics of the two participant groups

Criterion	**Overruled**	**Control**
Years of age: mean (SD)	39.0 (8.3)	40.9 (9.4)
Years of work experience: mean (SD)	15.0 (8.2)	17.3 (10.5)
Flight hours large: mean (SD)	8269 (5465)	7572 (4426)
Flight hours SEP/MEP: mean (SD)	285 (484)	689 (802)
Captains / FOs / SOs / Not employed yet	9 / 9 / 0 / 1	9 / 8 / 1/ 1
Pilots with instructor experience (TRI, TRE or FI)	7	10

2.3 Apparatus

The experiment was performed using a fixed-base flight simulator (Figure 1), developed by the simulator company Multisim B.V. It featured an aerodynamic model of a Cessna Citation S550 and was equipped with a primary flight display (PFD); navigation display (ND) with adjustable range; and a mode control panel (MCP): all based on those of a B737. Two radio frequencies could be set. The controls consisted of a yoke control column with pitch trim buttons, a throttle column, and rudder pedals. Configuration controls consisted of a gear lever with three gear indicator lights, and a Sion airport flaps lever with three settings: Up, Approach and Land. It did not include speed brakes. There were no autopilot functions, but guidance functions consisted of a flight director (FD) and the Horizontal Situation Indicator (HSI) in MAP mode (which shows airplane position relative to the flight route against a moving map background). Pilots could use the Instrument Landing System (ILS), Distance Measuring Equipment (DME), and VHF Omnidirectional Range (VOR) beacons. There was an aural “approaching minimums” and “minimums”. The EGPWS alarm was disconnected from actual proximity to the mountains, but it could be triggered when the scenario required it.

2.4 Procedure

The participating pilots performed the experimental tasks individually on a single day, guided by the instructor. The total duration of the visit was two hours: 20 minutes for briefing and questionnaires, 80 minutes for the simulator session, and 20 minutes for debriefing. Four pilots were tested each day, the first starting at 08:00 a.m., and the fourth ending at 05:30 p.m.

2.4.1 Briefing

After a welcome, the pilots received information about the experimental tasks, the aircraft model and the functions available in the simulator. They were told that the experiment was aimed at investigating pilots' decision making and perceived workload during various approaches and go-arounds. It was emphasized that there was a practice part and a testing part in the experiment, and that the instructor would help and guide them only in the practice part.

Figure 1. Fixed-base simulator with participant in the left seat, and instructor in the right seat

During the briefing, the procedure was explained to the pilots, as follows:

- The practice part would consist of two landings on Woensdrecht Air Base (EHWO in the nomenclature of International Civil Aviation Organization (ICAO), Netherlands (once with good visibility, and once with poor visibility), and two scenarios in Sion, Switzerland (LSGS): one landing, and one go-around on runway 25. LSGS is an airport in the mountains in Switzerland.

- The test part would start after the practice part. In this part, the pilot's performance would be measured during four approaches to LSGS. Performance measurements would consist of the pilot's responses to events, and their ability to fly at target speeds and headings. A video was shown of the approach to LSGS (https://youtu.be/DY-hrbpu5nM, starting at 5:36), and a schematic approach chart and checklist was handed out to the pilots to study (see, Appendix A). The same approach chart was also available in the cockpit, so no data needed to be memorized.

2.4.2 Training scenarios and manipulation

All training scenarios were flown manually and single-pilot, except that the participant could ask the instructor to perform actions such as setting gear, flaps, speed and heading bugs, or change the radio frequency. As mentioned above, the pilots had the schematic approach map and checklist available in the cockpit.

The training session started with a familiarization flight at 5000 ft. for about three minutes, including some changes in heading and flight level, so that pilots could familiarize themselves with the controls. At the end, pilots were instructed to make a maximum rate climb, while the instructor pointed out the maximum pitch angle and minimum calibrated airspeed (CAS; i.e., ca. 120 kt.) that could be flown. Then followed four training scenarios:

1. The first training scenario started paused at 1000 ft. and positioned downwind in a traffic pattern at EHWO. Pilots in both groups were instructed to finish the pattern and land. The instructor gave target speeds and instructed configuration changes. The purpose of this scenario was to let the pilots practice a landing.

2. The second training scenario was the same as the first, but now with low visibility (about 6 km) for both groups. Only in the Overruled group, the visibility decreased even further when the landing lights were switched off while they were in downwind, meaning that visual contact with the runway was lost. After an expression of confusion by the pilots of the overruled group, the instructor directed them to the runway and instructed them to land anyway. The purpose of this scenario was to manipulate a difference between the groups, intended to bring about different trainee behavior in a later transfer test when they would again be presented with a situation that involved losing visual contact with the runway.

3. The third training scenario took place at LSGS. Pilots in both groups made an approach and landing on runway 25 with good visibility. They started in the clouds at flight level (FL) 100 with the FD set to ILS/DME beacon "ISI" (see Appendix A), radial 244, and to

a 6 degrees glide slope. After passing the cloud layer, they would reach the decision altitude at 4490 ft. At, or below this altitude, visual contact with the runway and relevant surroundings is required to continue the landing. The glide slope changed to 4 degrees, also indicated by the precision approach path indicators (PAPIs). In this visual part of the approach, participants were to use landmarks to make a turn to the left, and back to the right, lining them up to the runway by flying ISI outbound radial 251. They were flying close by mountain ridges left and right. The required flight path was also indicated on the ND. The purpose of this scenario was to practice the approach and landing in the test area.

4. The fourth practice scenario also took place at LSGS, this time with poor but sufficient visibility. Before the scenario started, all of the participants were told that they would practice a go-around based on an (air traffic control) ATC message. On final approach, as they were close to the runway, an ATC message called their callsign and stated that the runway was blocked due to military activity, and that they should go-around. To fly the missed approach, they needed to climb heading R067 inbound of the LSGS VOR beacon, after which they would fly outbound R234. The scenario ended at 8000 ft. Only for the overruled group, the EGPWS triggered with one repetition of the aural message: "Terrain, terrain, pull up" when the participant was making the first turn close by a mountain ridge. The instructor immediately told the overruled group pilots "whereas you should normally respond to this alarm, you can proceed for now in order to practice the response to the ATC instruction."

After completion of the fourth practice scenario, the instructor declared that it was the end of the practice session, and that the participants would no longer receive training information from him, but he would answer any questions that the participants may have before the start of the test.

2.4.3 Test scenarios

The test consisted of two transfer test scenarios (1a and 2a) and two corresponding baseline test scenarios (1b and 2b), respectively. Test scenarios 1a and 1b involved a go-around maneuver, which was used to test if there was negative transfer induced by the second training scenario, in which pilots of the overruled group were instructed to continue the landing while visual contact with the runway was lost. Test scenarios 2a and 2b involved an EGPWS alarm, which was used to test if there was negative transfer induced by the fourth training scenario, in which pilots of the overruled group were instructed to discard the EGPWS alarm.

The order of test scenarios 1a and 2a was counterbalanced between participants, and both baseline test scenarios were always presented directly after the corresponding test (1b always followed 1a, and 2b always followed 2a). The visibility was always poor, starting at 6 km, and slowly decreasing so that the runway during the whole approach was always just barely visible. The test scenarios were as follows:

1. Go-around scenarios
    a. **Loss of visual contact**
    This first transfer test scenario (1a) consisted of a go-around that was triggered by faster than normal decreasing visibility and disappearing runway lights, just before the participant made the second turn to line up to the runway.
    b. **ATC instruction**
    This first baseline test scenario (1b) corresponded to transfer test 1a, and consisted of a go-around that was triggered by the ATC message. The same message was heard in the practice session, but this time it was heard before the final approach (i.e., just when the participant was lining up to the runway).
2. EGPWS scenarios
    a. **Unexpected EGPWS**
    In this second transfer test scenario (2a), the EGPWS alarm was triggered automatically at a certain distance from the runway, and continued to sound until the participant pitched up. The EGPWS was triggered after the decision altitude, just after the participant made their first turn. This scenario was used to test whether pilots in the overruled group would be more likely to discard, or to respond later to the EGPWS alarm than pilots in the control group. When the EGPWS triggers in this situation, the pilot could either choose to discard the alarm, or respond to it. Choosing to discard the alarm could be based on having visual contact with the runway and on one's certainty of the flight path. Choosing to respond to the alarm could be based on unfamiliarity with the approach, not being briefed that the alarm could be ignored at this location, or on uncertainty of the flight path. The correct response would be to start with an escape maneuver with maximum climb rate (i.e., 120 knots, practiced during familiarization), after which the go-around speed of 140-150 knots shown on the checklist is followed.

b. **Expected EGPWS**
   This second baseline test scenario (2b) corresponded to test scenario 2a, and was an exact repetition of that, but now the EGPWS trigger was announced beforehand, and the pilot was instructed to respond to it.

After each test scenario, participants filled in a questionnaire (see section 2.5 below, Dependent measures).

2.4.4 Debrief

After the simulator session, the pilots were told about the true objective of the experiment, and they were asked if they had recognized instances in their initial or recurrent training session where their own decision-making would be overruled in a possibly hindering manner. They received a present as thanks and were farewelled.

2.5 Dependent measures

The dependent measures are categorized as general performance measures, performance measures related to the EGPWS escape maneuver, and performance measures related to the standard go-around. The objective data were obtained from the recorded time histories of simulator data, which were first smoothened with a fourth-order Butterworth filter with cutoff frequency of 5 Hz.

2.5.1 Go-around performance

To perform the go-around, the pilot would need to set throttle to max, pitch up 15 degrees, raise the gear, if reaching a positive climb rate set flaps UP, and follow the headings required for go-around that are indicated on the approach chart.

Actions
In both go-around scenarios, we checked whether participants set flaps UP and raised the gear following the trigger (i.e., the disappearing runway lights, or the end of the ATC message with the go-around instruction).

Time to first response
This is the time in seconds from the trigger (i.e., disappearance of the runway lights, or the end of the ATC message with the go-around instruction) until the participant's first response. The first response could be any of the following: increasing power by at least 25%, increasing pitch by at least 10 degrees, changing flaps, or raising gear.

Time to set power 80%

This is the time from the trigger until the pilot adjusted power to at least 80%. This measure was used to indicate when a pilot made a true committed response to climb.

Procedure execution time

This is the time between the first response and the total execution of all procedural actions. The procedure was considered fully executed if all of the following conditions were met: power > 80%, pitch > 10 degrees up, flaps up, and gear up.

Percentage within speed bounds

In both go-around test scenarios (1a and 1b), the extent to which pilots managed their speed during the go-around according to the provided procedure (i.e., 140-150 kts, as indicated on the checklist) was measured. This was done by taking the percentage of time the pilot kept the airspeed between 135-155 kts from the moment of passing 135 kts. An extra margin of five kts was used to exclude the time during which pilots were flying slightly above or below the required speed.

Maximum speed excursion above upper bound

This measure is the number of knots with which the upper speed bound of 150 kts was exceeded during the go-around.

Maximum speed excursion below lower bound

This measure is the number of knots with which the lower speed bound of 140 kts was exceeded negatively during the go-around. It was counted from the first moment the pilot crossed 140 kts. If the pilot never reached 140 kts, the difference between the maximum speed reached during the go-around and 140 kts was taken for this measure.

2.5.2 EGPWS Escape maneuver performance

Actions

We checked whether participants responded to the EGPWS alarm by climbing and making a go-around, or not (i.e., a dichotomous variable “yes” or “no”). We also checked whether they set flaps *UP* and raised the gear following the start of the EGPWS alarm.

Time to first response

This is the time in seconds from the start of the EGPWS alarm until the participant’s first response. The first response could be any of the following: increasing power by at least 25%, increasing pitch by at least 10 degrees, changing flaps, or raising gear.

Time to set power 80%

This is the time from the trigger until the pilot adjusted power to at least 80%. This measure was used to indicate when a pilot made a true committed response to climb.

Procedure execution time

This is the time between the trigger and the total execution of all procedural actions. The procedure was considered fully executed if all of the following conditions were met: power > 80%, pitch > 10 degrees up, flaps up and gear up.

Time to reach 8000 ft.

This is the time in seconds from the moment participants responded to the EGPWS (see, Actions) until reaching 8000 ft. This measure was only obtained in the EGPWS scenarios, as it indicates whether participants performed the escape maneuver (maximum climb) instead of focusing more on the go-around parameters.

Maximum pitch angle

This variable was obtained in the EGPWS scenarios following the alarm, to check to what extent the participants executed the escape procedure, which requires a higher pitch angle, instead of the normal go-around procedure.

2.5.3 Subjective measures

Mental demand, time pressure

After each test scenario participants rated perceived mental demand and time pressure, with regard to their decision to make a go-around, or to continue. For this, the NASA Task Load Index (TLX) subscales of mental demand and time pressure were used (Hart & Staveland, 1988). These scales range from 0-100 with iterations of five.

Stress

After each scenario, the pilots also rated their perceived level of stress with regard to the moment of the decision and the response, using the Anxiety scale (Houtman & Bakker, 1989). This is an analogue scale (horizontal line) ranging from 0 to 10.

Surprise

To check if the unexpected EGPWS was indeed surprising as intended, we let the pilots rate how surprised they were after the unexpected EGPWS scenario as well as in the baseline EGPWS scenario. For the scale, we used the same format as the mental demand and time pressure scales.

Comments

Finally, as an open-ended question, pilots were asked to chronologically describe their thoughts and actions from the moment they noticed the trigger and during their response.

2.6 Data analysis

Dependent continuous performance variables were compared between the groups using two 2 × 2 mixed-models multivariate analysis of variance (MANOVA's): one for making the go-around (loss of visual contact scenario and ATC message scenario) and one for the EGPWS response (unexpected scenario and baseline scenario). They had the between-subject factor of Group (Overruled, Control) and the within-subject factor of scenario. Alpha is set at $p = 0.05$, but results with $p < 0.100$ will be reported as marginally significant.

2.7 Hypotheses

We hypothesized that for both scenarios (i.e., go-around and EGPWS), pilots in both groups would respond later to the triggering event in the transfer test compared to their response in the corresponding baseline, and that this difference would be larger for pilots in the overruled group than for pilots in the control group (i.e., a two-way interaction). Furthermore, we hypothesized that pilots in the overruled group were more likely to ignore the EGPWS alarm, and continue the approach. These hypotheses are based on the expectation that pilots in the overruled group would be more confused in their decision making due to the unnatural decision-making process that they experienced in the training sessions where the instructor advised them to continue against their own decision.

3 Results

3.1 Performance example

Figure 2 shows the inputs of a pilot in the control group who responded hesitantly to the unexpected EGPWS alarm. It can be seen that the pitch input was quite variable, making it problematic to use this input for determining the pilot's response time to the EGPWS alarm. Therefore, the "Time to first response" was identified as the moment where the throttle exceeded 25%, which is indicated with the black marker.

In this example, the pilots initially responded by temporarily increasing the throttle to approximately 70% and briefly pitching up. The black marker on the throttle graph indicates the point which is detected as the "first response" for the data analysis. After the initial increase in throttle, the throttle was momentarily reduced to 50%, after which it increased again to 100%. Because this example shows that the initial response (throttle > 25%) may differ from a committed unchanging response, we also obtained the moment at which throttle exceeded 80%. We checked the data to make sure that the 80% limit only included committed responses.

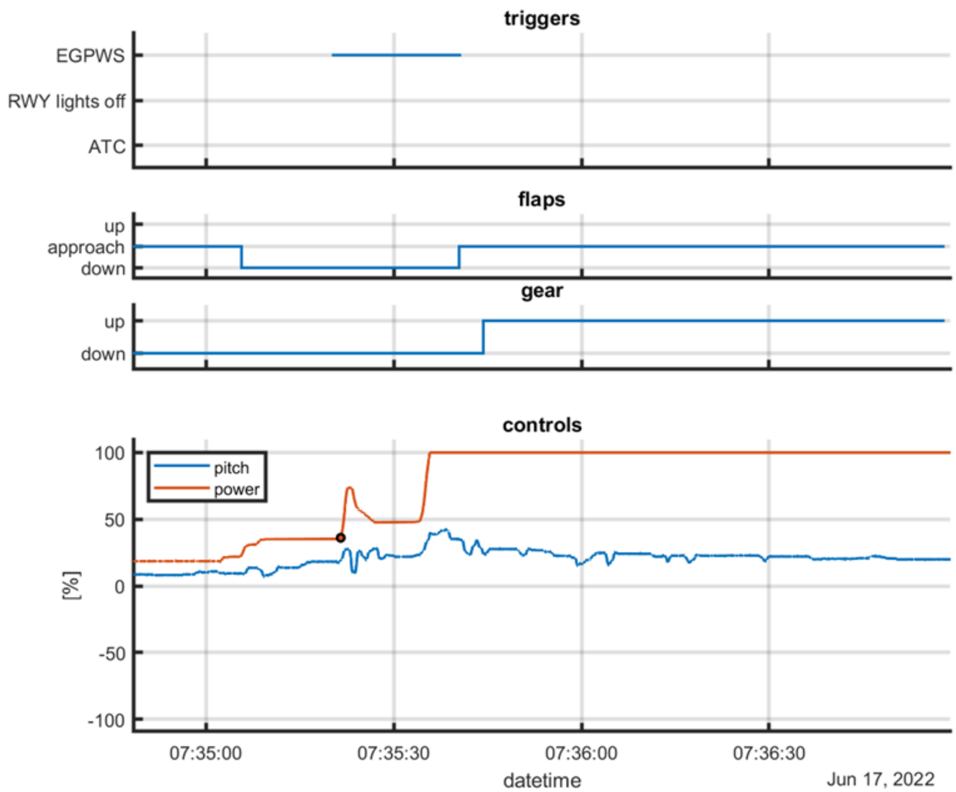

Figure 2. Time history of a pilot's response to unexpected EGPWS

Figure 3 shows the aircraft state resulting from the responses shown in Figure 2 (pilot's response to unexpected EGPWS). The dotted lines in the airspeed plot (CAS) show the boundaries of 140-150 kts that were set for the go-around performance. However, these were not used as a measure of performance in the EGWPS scenarios (which is shown here), because this requires an escape maneuver with a higher climb rate, and hence, a lower airspeed than a go-around maneuver.

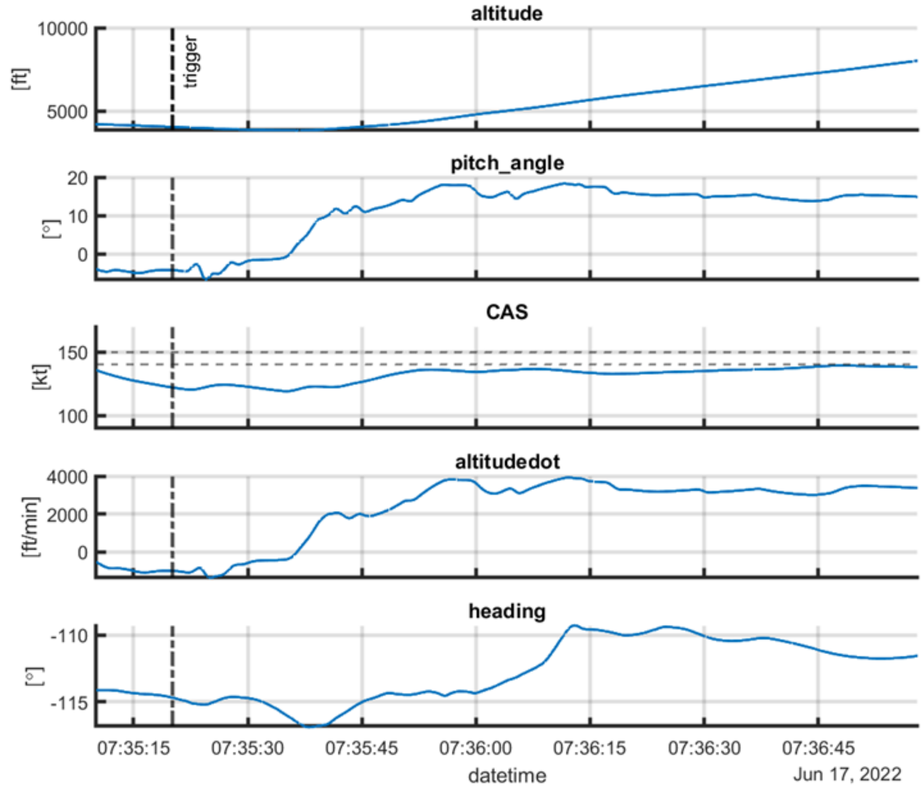

Figure 3. Aircraft state resulting from pilot responses to unexpected EGPWS

3.2 Performance of the go-around procedure

3.2.1 Actions

All pilots from both groups performed a go-around in both test scenarios (1a and 1b), and performed all required procedural actions.

3.2.2 Statistical analysis of performance measures

Table 2, Table 3 and Table 4 present the outcomes of the 2 × 2 MANOVA performed on the data of the go-around tests for the factors: scenario, group, and the two-way interaction, respectively. There was a main effect of scenario for all measures. Apart from the largest excursion above the upper speed bound, there was no significant or marginally significant main effect for group. The same variable was the only variable for which a significant Scenario × Group interaction effect was found. The next sections will provide more details on the size and direction of (marginally) significant results, as well as boxplots of the data.

Table 2. Main effects of scenario (MANOVA) for go-around test scenarios

Parameter	***F*(df, df_{error})**	***p***
Multivariate	$F(8,29) = 17.52$	< 0.001
Time to first response	$F(1,36) = 133.70$	< 0.001
Time to set throttle > 80%	$F(1,36) = 130.61$	< 0.001
Time to execute procedure	$F(1,36) = 4.23$	0.036
Percentage of time within the speed bounds	$F(1,36) = 21.83$	< 0.001
Largest excursion above upper speed bound	$F(1,36) = 5.23$	0.028
Largest excursion below lower speed bound	$F(1,36) = 4.41$	0.043
Stress	$F(1,36) = 31.86$	< 0.001
Mental demand	$F(1,36) = 14.54$	< 0.001
Time pressure	$F(1,36) = 12.47$	0.001

Table 3. Main effects of group (MANOVA) for go-around test scenarios

Parameter	***F*(df, df_{error})**	***p***
Multivariate	$F(8,29) = 1.48$	0.209
Time to first response	$F(1,36) = 0.22$	0.640
Time to execute procedure	$F(1,36) = 0.05$	0.818
Time to set throttle > 80%	$F(1,36) = 0.01$	0.908
Percentage of time within the speed bounds	$F(1,36) = 0.90$	0.349
Largest excursion above upper speed bound	$F(1,36) = 4.64$	**0.038**
Largest excursion below lower speed bound	$F(1,36) = 1.31$	0.261
Stress	$F(1,36) = 1.06$	0.309
Mental demand	$F(1,36) = 1.79$	0.190
Time pressure	$F(1,36) < 0.01$	0.958

Table 4. Scenario × group interaction effects (MANOVA) for go-around test scenarios

Parameter	***F*(df, df$_{error}$)**	***p***
Multivariate	F(8,29) = 1.44	0.223
Time to first response	F(1,36) = 0.76	0.388
Time to set throttle > 80%	F(1,36) = 0.01	0.942
Time to execute procedure	F(1,36) = 0.18	0.674
Percentage of time within the speed bounds	F(1,36) = 2.58	0.117
Largest excursion above upper speed bound	F(1,36) = 4.14	**0.049**
Largest excursion below lower speed bound	F(1,36) = 0.89	0.351
Stress	F(1,36) = 1.58	0.217
Mental demand	F(1,36) = 2.01	0.165
Time pressure	F(1,36) = 2.25	0.142

3.2.3 Time to first response

Two pilots in the transfer test scenario 1a (loss of visual contact) were coincidentally adjusting throttle at the moment the trigger started, leading to an apparent immediate reaction. For these pilots, the detection of a reaction in throttle was set to throttle exceeding 80%.

As can be seen in Table 2, there was a main effect of scenario for the time to the first response, which is not of interest as the triggers in the scenarios were different. The descriptive statistics of the time to the first response are shown in Figure 4. It can be seen that the decision to initiate the go-around varied widely within each group. There was no significant difference between the groups and no significant interaction.

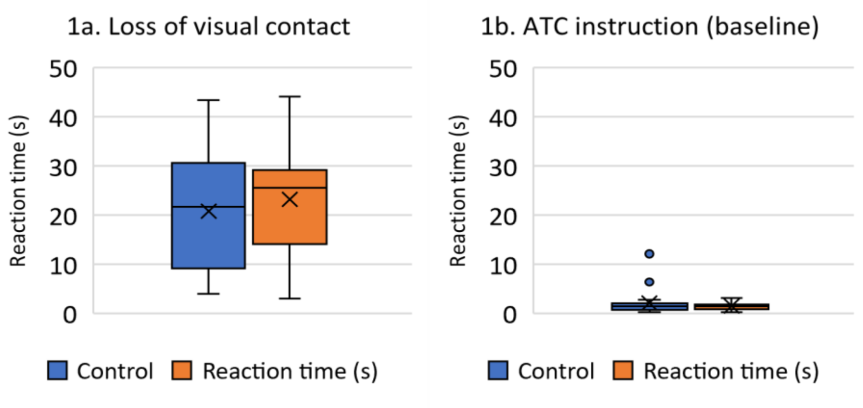

Figure 4. Tukey boxplots of time to first response in go-around test scenarios 1a and 1b

3.2.4 Time to set throttle > 80%

As can be seen in Table 2, there was a main effect of scenario for the time to set the throttle to 80%, which is not of interest as the triggers in the scenarios were different. The descriptive statistics of this performance measure are shown in Figure 5. It can be seen that the decision to initiate the go-around varied widely within each group. There was no significant difference between the groups and no significant interaction.

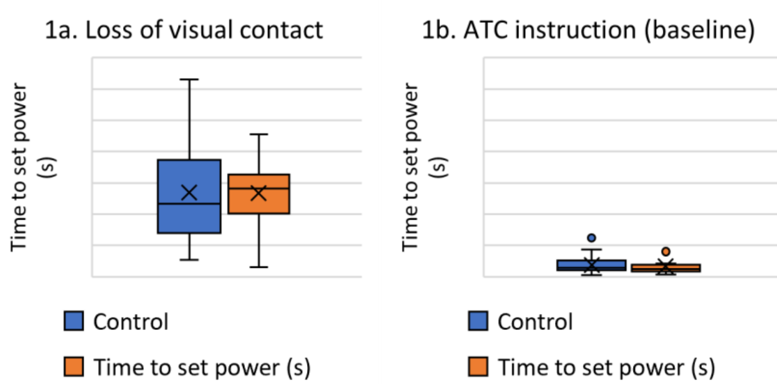

Figure 5. Tukey boxplots of time to set throttle to > 80% in go-around test scenarios

3.2.5 Time to execute procedure

As can be seen in Table 2 there was a main effect of scenario for the time to execute the full procedure from the moment the go-around was triggered. As shown in Figure 6, the time from the pilot's first reaction to full execution of the procedure was 2.97 seconds longer (SE = 1.37) in the loss of visual contact scenario than in the ATC instruction scenario. Some pilots responded by setting throttle at 60-70% at first, and later increased throttle to > 80%. There was no significant difference between the groups and no significant interaction.

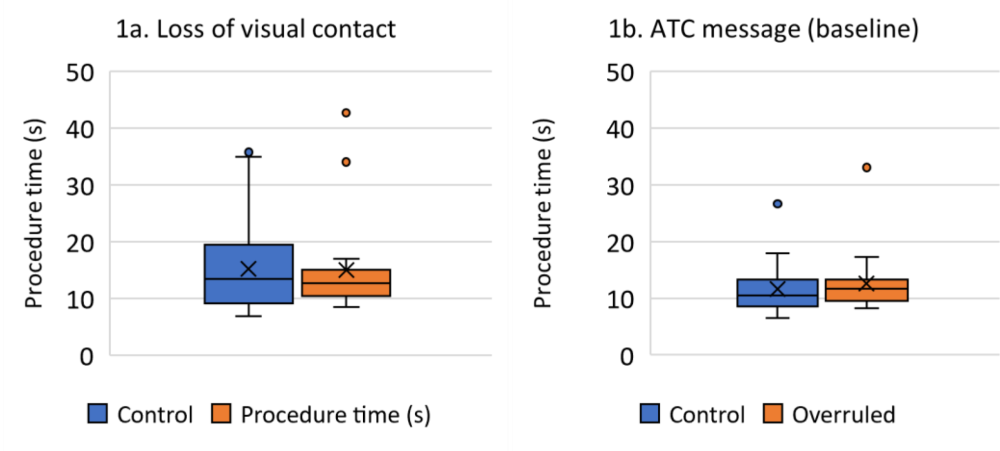

Figure 6. Time since the first response until full procedure was executed in go-around test scenarios 1a and 1b

3.2.6 Percentage of time within the speed bounds

As can be seen in Table 2, there was a main effect of scenario for the percentage of time within the speed bounds from the moment that the airspeed exceeded 135 kts depended only significantly on scenarios. It was on average 17.3 % lower in scenario 1a (loss of visual contact) than in scenario 1b (ATC instruction), as seen in Figure 7. It could be that the loss of visual reference in scenario 1a caused pilots to climb faster to escape terrain, which led to a lower CAS. There were two pilots in the control group who never reached 135 kts, meaning that they scored 0%. There was no significant difference between the groups and no significant interaction.

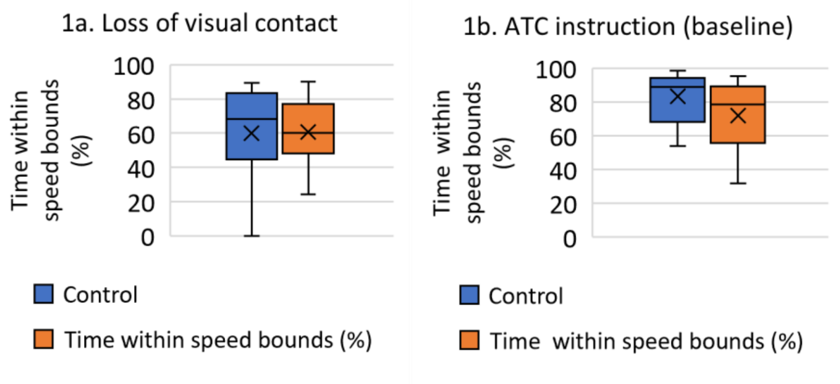

Figure 7. Percentage of time flown within required speed bounds in go-around scenarios

3.2.7 Largest excursion above the upper speed bound

The MANOVA revealed significant effects of scenario, group and of the interaction. However, this data was heavily skewed (see Figure 8), because most pilots did not exceed the upper bound of 150 kts at all in scenario 1a. Therefore, a non-parametric test was used as an alternative to the mixed-model MANOVA, namely a Generalized Linear Mixed Model for ordinal data, with group, test, and group × test as the predictors. This revealed only a significant effect of Group, $F(1,36) = 4.21$, $p = 0.048$. The control group exceeded the upper speed bound more than the overruled group in both posttest scenarios. This could be an effect of improper balancing of the groups.

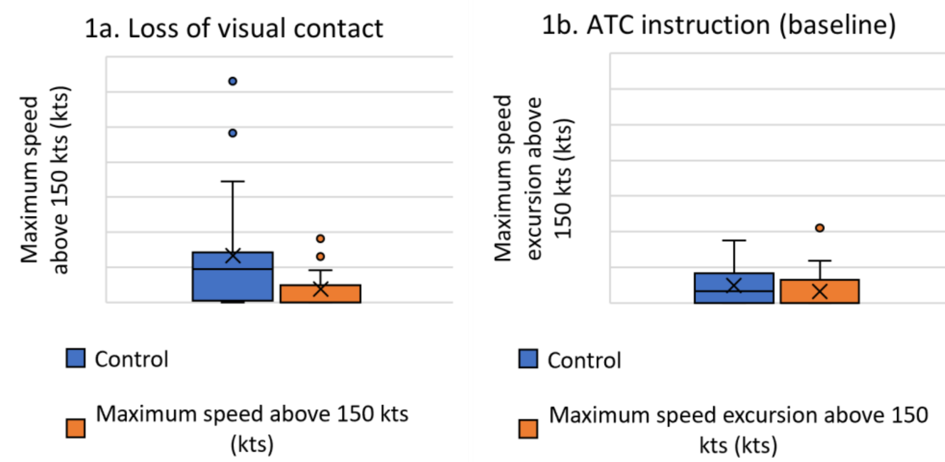

Figure 8. Maximum speed excursion above upper bound of 150 kts

3.2.8 Largest excursion below the lower speed bound

The MANOVA revealed a significant effect of scenario, with the scenario 1a (loss of visual contact) leading to larger excursions. However, similar to the excursion of the maximum speed bound, the data for this variable was heavily skewed because most pilots did not exceed the lower bound of 140 kts (see, Figure 9). Therefore, a non-parametric test was used as an alternative to the mixed-model ANOVA, namely a Generalized Linear Mixed Model for ordinal data, with group, test, and group × test as predictors. This showed no significant or marginal significant effects.

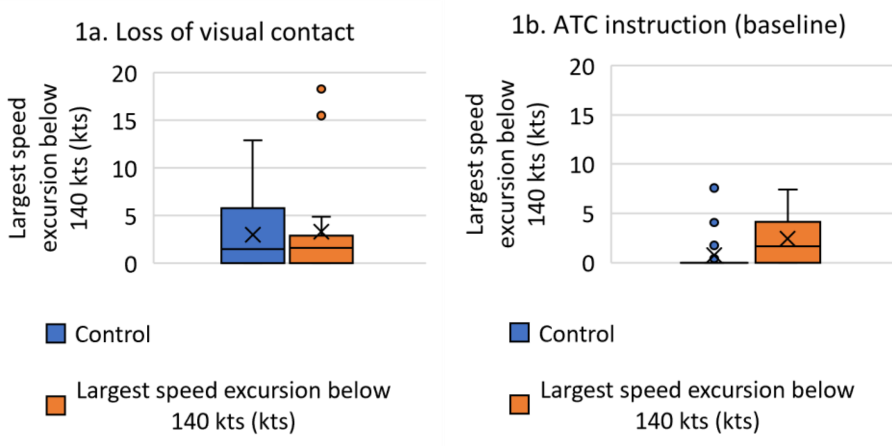

Figure 9. Largest speed excursion below lower bound of 140 kts, after first time 140 kts was exceeded

3.2.9 Subjective stress, mental demand and time pressure

The analyses of the subjective data revealed only significant effects of scenario. The loss of visual contact scenario (1a) was rated on average with 1.2 points (SE = 0.21) more stressful on a 0-10-point scale compared to the baseline test scenario (1b; see Figure 10). Mental demand and time pressure were also rated higher in the loss of visual contact scenario, with 12.6 points (SE = 3.3) and 10.1 points (SE = 2.9), respectively, both on a 0-100-point scale (see Figure 11 and Figure 12). The average scores in the transfer test scenario 1a (loss of visual contact) were somewhat on the low end of the scales, with only few pilots scoring higher than the midpoint of the scale (see Figure 10). This means that most pilots did not find the scenarios very challenging.

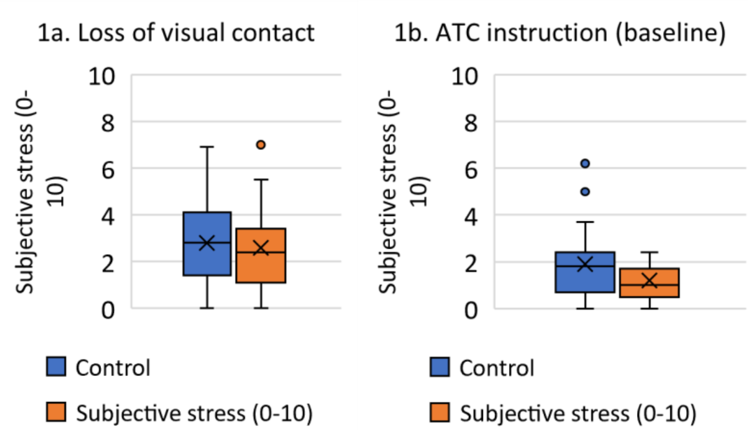

Figure 10. Tukey boxplots of subjective stress

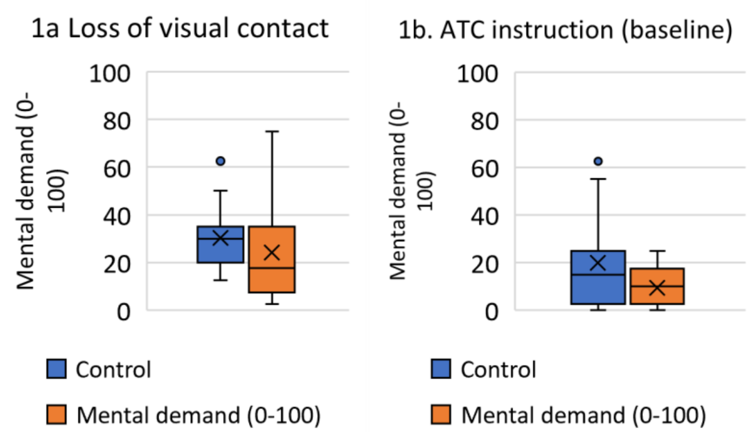

Figure 11. Tukey boxplots of subjective mental demand

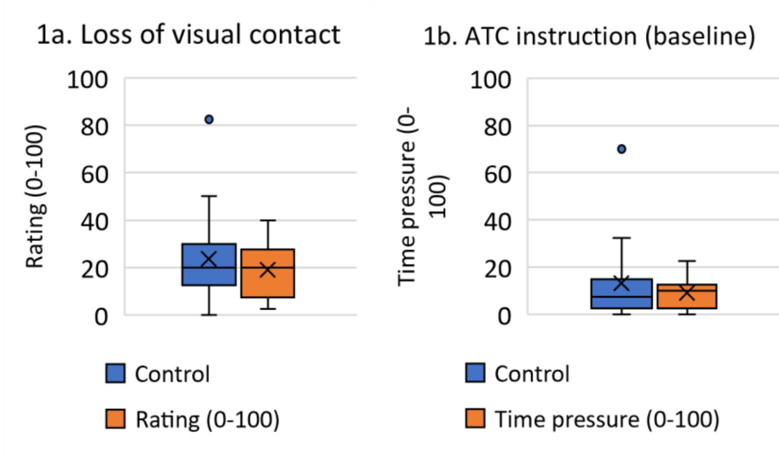

Figure 12. Tukey boxplots of subjective time pressure

3.3 Results of the EGPWS scenarios

3.3.1 Actions

3.3.1.1 Responding to the alarm

In the transfer test scenario 2a (unexpected EGPWS), 13 pilots in the overruled group responded to the alarm, compared to 12 pilots in the control group. In the baseline, all pilots responded. This was not significantly different between the groups, $p > 0.900$. Pilots who responded to the alarm generally commented that they were following the required procedure, or that they preferred to respond first and think later (e.g.: *"'Terrain, pull up' is not the moment to think (near Sion). Therefore, without doubt GA* [Go Around]*. It did astound me!"* (control group)). Pilots who continued with the approach reasoned that they were visual and suspected a faulty warning, (e.g., *"Fault with warning computer. I am visual - I am happy to continue - I am allowed."* (control group)). Two pilots, one in either group, mentioned that they suspected that the alarm was due to a simulation error.

3.3.1.2 Configuration change

Interestingly, two pilots, one in either group, did not change their configuration when responding to the unexpected EGPWS. The pilot in the overruled group commented: "*Sim malfunction? After no comment → executed escape maneuver."* The pilot in the control group commented: *"Escape maneuver procedure: follow correct SOP. Then think about terrain and GA procedure. As soon as clear of terrain: clean up."* It seems that these pilots were unsure at which altitude they would be clear of terrain and therefore waited with changing their configuration.

3.3.1.3 Escape maneuver

Of the pilots who responded to the alarm, some indicated that they performed the escape maneuver, whereas others did not. As one pilot in the control group indicated: *"Started go-around with 20 deg pitch and 120 kts for max climb. Changed config to appr/gear up. After visual separation from terrain on the left, accelerated to 140 kts.*" Some pilots, however, commented that not performing the escape procedure was a conscious decision: "*Because I didn't see terrain in front of me, I thought normal GA would be sufficient, instead of terrain escape with max climb."* (control group). Two pilots in the overruled group commented that they did not perform the escape procedure because they did not know the procedure for this aircraft, although one also mentioned performing only the go-around and no escape maneuver.

Of those who responded to the unexpected EGPWS alarm, a few more pilots in the control group (7/12 or 58%) than in the overruled group (5/13 or 38%) mentioned considering the escape maneuver. In the baseline EGPWS scenario, these numbers were 9/19 (47%) in the control group and 10/19 (53%) in the overruled group.

When inspecting the average climb rate in the first 40 seconds following the pilots' committed response (i.e., throttle exceeding 80%), it can be seen in Figure 13 that the control group indeed climbed more steeply in the first 15 seconds, which suggests better focus on performing the escape maneuver than the overruled group. However, as described in section 3.3.5, this did not lead to a statistical difference in time to complete the maneuver.

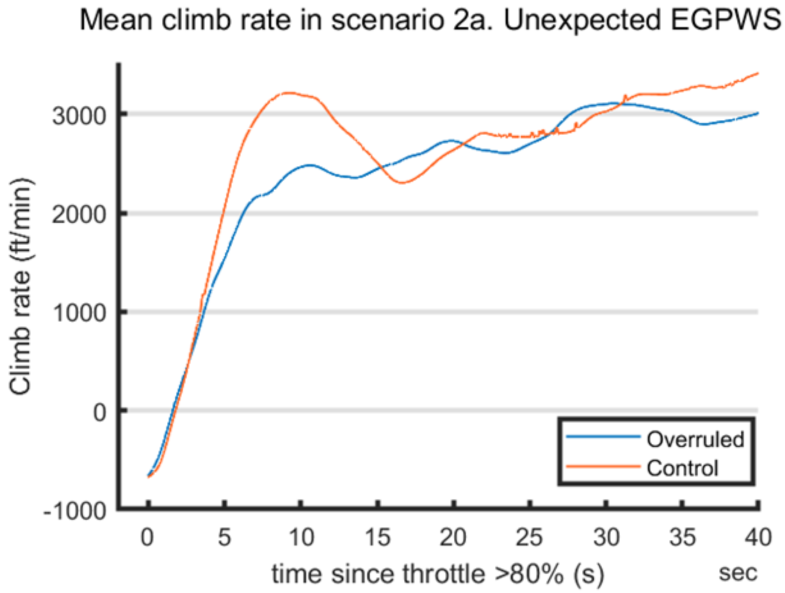

Figure 13. Mean climb rate per group following throttle exceeding 80% in unexpected EGPWS scenario

3.3.2 Statistical analysis of performance measures

The performance parameters were calculated only for the pilots who responded to the alarm. Table 5, Table 6 and Table 7 present the outcomes of the 2 × 2 MANOVA performed on the data of the EGPWS test scenarios, for scenario, group, and the interaction effects, respectively. Because there were many missing values for the performance measures due to pilots not responding to the alarm, separate MANOVAs were executed for performance measures and for subjective measures. There was a significant multivariate effect of scenario, which was based on several parameters being marginally different between the unexpected and baseline EGPWS scenario. The subjective measures were all significantly different between the scenarios. For the main effect of group, there was only a marginally significant multivariate effect. For the Scenario × Group interaction, there were no significant or marginally significant effects.

The next sections provide details on the size and direction of (marginally) significant results, as well as boxplots of the data.

Table 5. Main effects of scenario in 2 × 2 MANOVA of EGPWS test scenarios

Parameter	**F(df, df_{error})**	***p***
Multivariate (performance)	F(8,16) = 4.46	0.007
Time to first response	F(1,23) = 3.88	0.061
Time to set throttle > 80%	F(1,23) = 3.99	0.058
Time to execute procedure	F(1,23) = 1.02	0.324
Maximum pitch	F(1,23) = 0.19	0.664
Time to reach 8000 ft.	F(1,23) = 2.40	0.135
Multivariate (subjective measures)	F(3,34) = 23.59	< 0.001
Stress	F(1,36) = 35.74	< 0.001
Mental demand	F(1,36) = 35.88	< 0.001
Time pressure	F(1,36) = 24.96	< 0.001
Surprise	F(1,36) = 91.53	< 0.001

Table 6. Main effects of group in MANOVA of EGPWS test scenarios

Parameter	**F(df, df_{error})**	***p***
Multivariate (performance)	F(8,16) = 2.19	0.098
Time to first response	F(1,23) = 0.70	0.411
Time to set throttle > 80%	F(1,23) < 0.01	0.968
Time to execute procedure	F(1,23) = 0.04	0.841
Maximum pitch	F(1,23) = 2.83	0.106
Time to reach 8000 ft.	F(1,23) = 0.68	0.418
Multivariate (subjective measures)	F(3,34) = 2.08	0.106
Stress	F(1,36) = 0.02	0.901
Mental demand	F(1,36) = 0.71	0.406
Time pressure	F(1,36) = 0.05	0.820
Surprise	F(1,36) = 3.66	0.064

Table 7. Scenario × group interaction effects in MANOVA of EGPWS test scenarios

Parameter	**F(df, df_{error})**	***p***
Multivariate	F(8,16) = 0.59	0.706
Time to first response	F(1,23) = 0.08	0.782
Time to set throttle > 80%	F(1,23) = 0.11	0.741
Time to execute procedure	F(1,23) = 2.14	0.157
Maximum pitch	F(1,23) = 0.94	0.342
Time to reach 8000 ft.	F(1,23) = 0.62	0.438
Multivariate (subjective measures)	F(3,34) =`0.50	0.738
Stress	F(1,36) = 0.09	0.771
Mental demand	F(1,36) = 0.06	0.802
Time pressure	F(1,36) = 0.16	0.689
Surprise	F(1,36) = 0.63	0.431

3.3.3 Time to first response

Two pilots in test scenario 2a (unexpected EGPWS), and one pilot in baseline test scenario 2b (expected EGPWS) were coincidentally adjusting throttle at the moment the trigger started, leading to an apparent immediate reaction. For these pilots, the detection of a reaction in throttle was set to throttle exceeding 80%.

As shown in Table 5, there was only a marginally significant effect of scenario, which indicated that, on average, all pilots responded 0.87 seconds (SE = 0.44) more quickly to the anticipated EGPWS alarm than to the unexpected alarm (see, Figure 14). There was no main effect of group, nor an interaction effect of group x scenario.

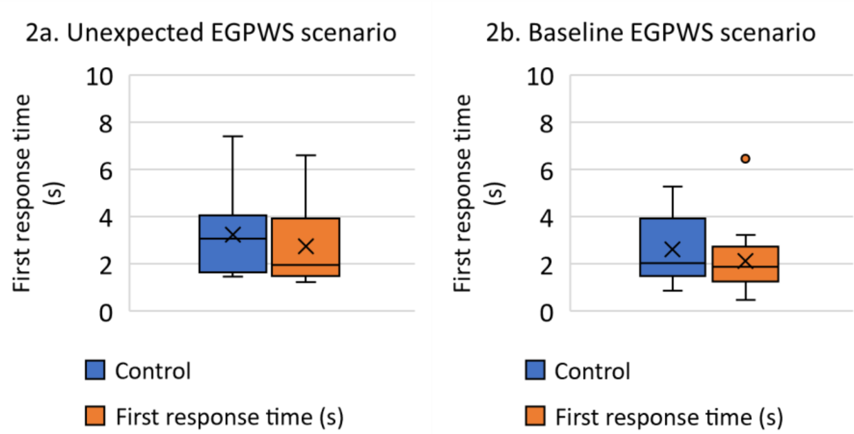

Figure 14. Tukey boxplots of time to first response in EGPWS test scenarios 2a and 2b

3.3.4 Time to set throttle > 80%

As shown in Table 5, there was only a marginally significant effect of scenario for the time to set throttle beyond 80%, indicating that all pilots on average reached 80% throttle 1.84 seconds (SE = 0.92) more quickly in the anticipated EGPWS scenario than in the unexpected one. As can be seen in Figure 15, in each group there was one outlier in the unexpected EGPWS scenario. The corresponding pilot in the control group noted: *"First I thought false warning. Outside I saw I was clear from terrain. Warning continued, so I chose GA. (...)."* The corresponding pilot in the overruled group noted: *"We trained for this approach and also received the GPWS. We were allowed to ignore that one. Now we were meant to respond? This negatively influenced my decision-making. I would otherwise not have shown this reaction."*

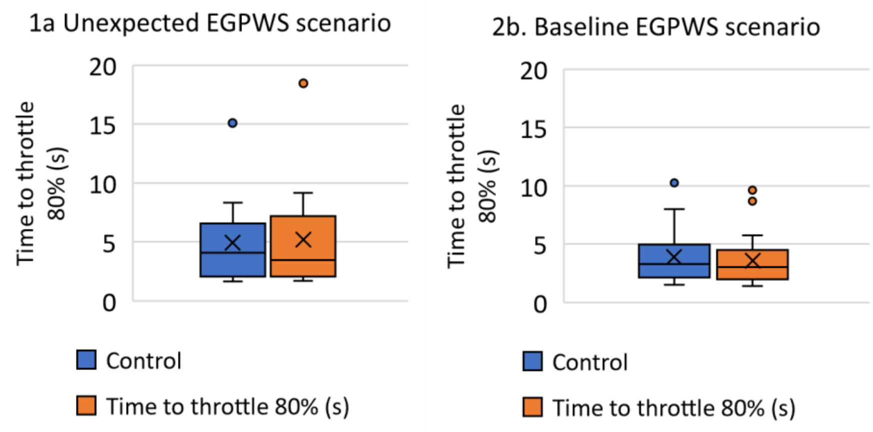

Figure 15. Tukey boxplots of time to set throttle to > 80% in test scenarios 2a and 2b

3.3.5 Time to execute procedure

For the total time to execute the whole procedure there were no significant or marginally significant results. The descriptive statistics of this performance measure are shown in Figure 16. The furthest outliers were caused by two pilots who configured very late in both EGPWS test scenarios.

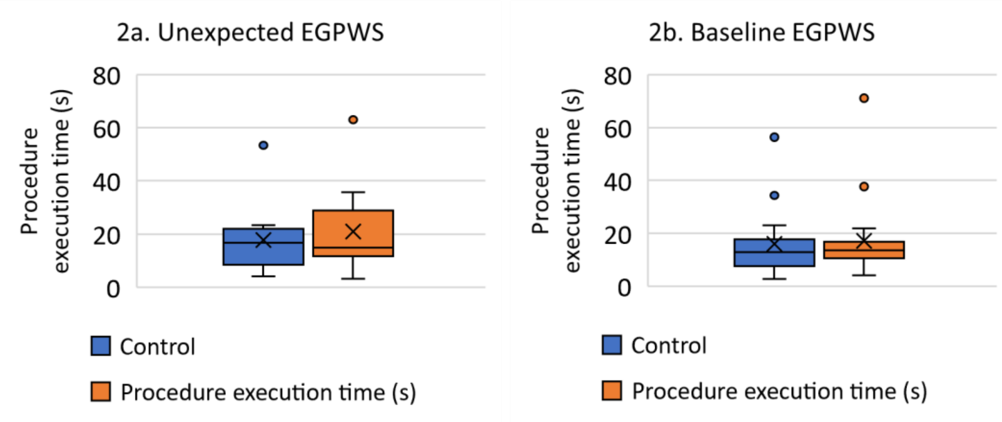

Figure 16. Tukey boxplots of time from first response until execution of full procedure in test scenarios 2a and 2b

3.3.6 Maximum pitch

There were no significant or marginally significant effects for maximum pitch. As can be seen in Figure 17, there were some outliers with excessive pitch, which were caused by different pilots in each scenario.

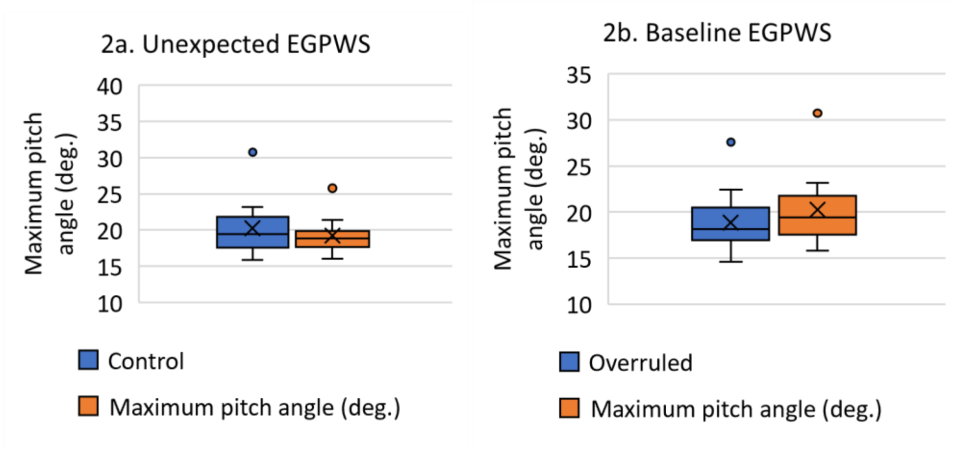

Figure 17. Tukey boxplots of maximum pitch angle in test scenarios 2a and 2b

3.3.7 Time to reach 8000 ft.

There were no significant or marginally significant effect for the time to reach 8000 ft. The descriptive statistics of this performance measure are shown in Figure 18.

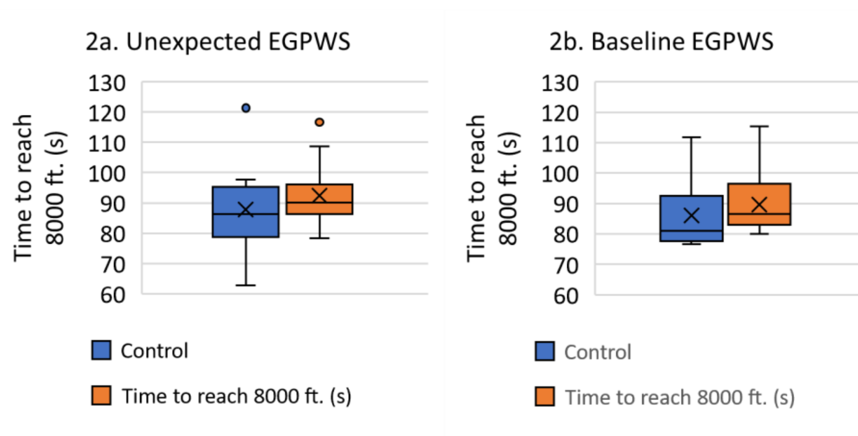

Figure 18. Time to reach 8000 ft. in test scenarios 2a and 2b

3.3.8 Subjective stress, mental demand and time pressure

The unexpected EGPWS scenario was rated as significantly more stressful, mentally demanding, and time pressure-inducing. The mean differences with the baseline scenario 2b amounted to 1.9 points (SE = 0.3) on a 0-10-point scale for stress (See Figure 19); 18.7 (SE = 3.1) on a 0-100-point scale for mental demand (See Figure 20); and 18.8 points (SE = 3.8) on a 0-100-point scale for time pressure (See Figure 21). The MANOVA revealed no significant main effect of group or scenario × group interaction effect. The scores in the unexpected EGPWS scenario were on average just below the midpoint of the scales (see Figure 19), indicating that pilots generally found the unexpected scenario somewhat but not very challenging.

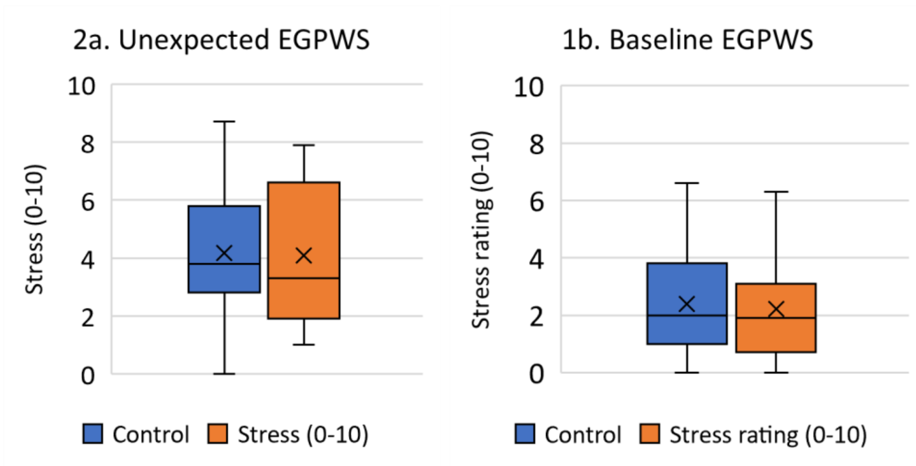

Figure 19. Tukey boxplots of subjective stress in test scenarios 2a and 2b

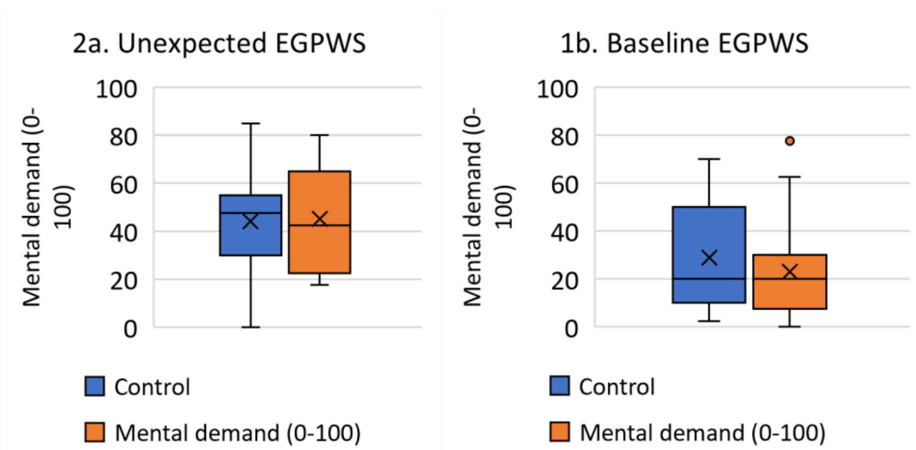

Figure 20. Tukey boxplots of subjective mental demand in test scenarios 2a and 2b

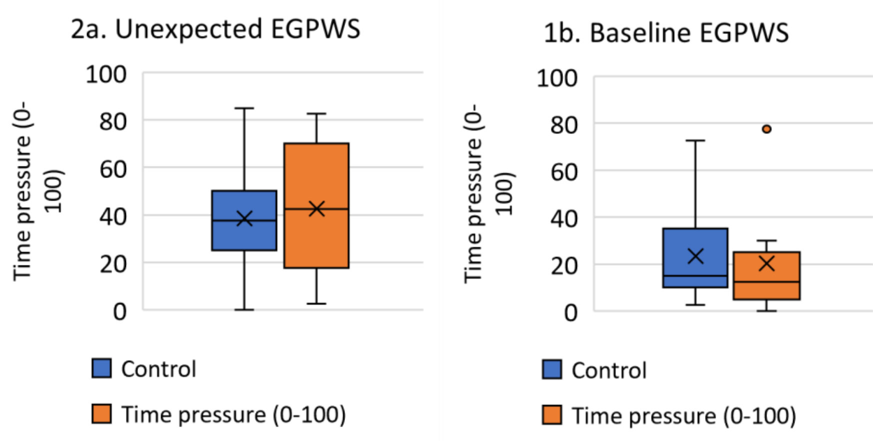

Figure 21. Tukey boxplots of subjective time pressure test scenarios 2a and 2b

3.3.9 Surprise manipulation check

The surprise ratings in the unexpected EGPWS scenario confirm that the alarm was indeed unexpected, as the average score was above the midpoint of the scale (see, Figure 22). Still, the variation in ratings was very high and some pilots scored zero or almost zero on surprise in the unexpected EGPWS.

The average values suggest that the pilots in the overruled group were less surprised by the alarm in both EGPWS scenarios. This would be logical because they had experienced the alarm in the training, while the pilots in the control group were only briefed that it could occur. However, according to a t-test, these differences were not significant, $p = 0.134$ (2a) and $p = 0.285$ (2b). There was also no significant interaction effect between scenario and group, $F(1,36) = 0.63$, $p = 0.431$, indicating that the groups did not differ in the increase in surprise ratings between the expected and the unexpected EGPWS scenario. One outlier in the control group in the baseline test perhaps indicates that the instruction may not have been clear.

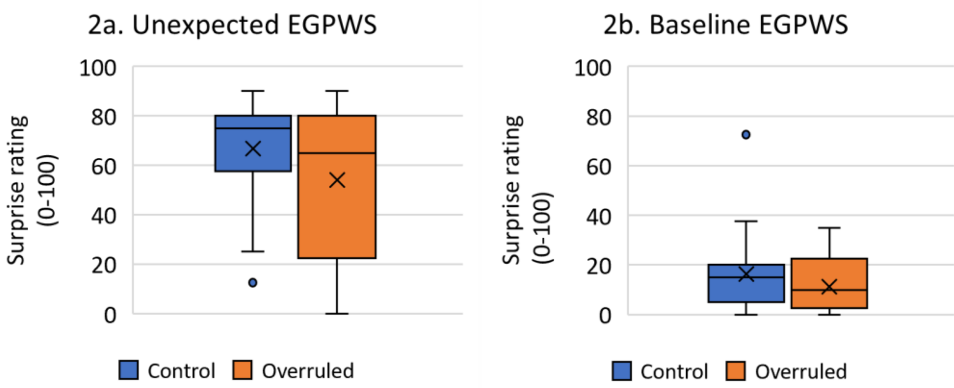

Figure 22. Tukey boxplots of surprise ratings in test scenarios 2a and 2b

3.4 Pilot's comments with regard to the training session

3.4.1 Go-around test scenarios

In the go-around test scenarios 1a and 1b, none of the pilots' comments referred to the training session.

3.4.2 EGPWS test scenarios

Regarding the EGPWS test scenarios, six pilots (32%) in the overruled group mentioned that the training session had influenced their decision making in the unexpected EGPWS test scenario. Two of these pilots responded to the EGPWS alarm in the transfer test, whereas the other four continued with the approach. Two of the six pilots responded relatively late, and indicated that the instructions during training had caused some confusion and hesitation:

1. *"We trained for this approach and also received the GPWS. We were allowed to ignore that one. Now we were meant to respond? This negatively influenced my decision-making. I would otherwise not have shown this reaction."* (Pilot responded to EGPWS alarm after 6.6 seconds, whereas the group mean was 2.7 seconds).
2. *"Terrain warning, in training generated but closer proximity to mountain. Discussed and continued. Now repeated initial action but decided to go-around."* (Pilot responded to EGPWS alarm after 4.7 seconds, whereas the group mean was 2.7 seconds).

The four other pilots continued with the approach, despite the EGPWS alarm, and mentioned that experiencing the warning in the training was a factor in their decision making:

3. *“It could trigger here. I am visual and am not going to hit a mountain. Was I allowed to ignore it here or not? I move away from the mountain a bit to be sure. It keeps calling, why? I'm definitely not hitting the mountain. It feels scary to continue.”*
4. *“Primed by instructor. EGPWS can trigger. Decision-making: are we visual? Yes. Are we safe in current airplane config and position to terrain? Yes --> continue.”*
5. *“Company policy is always terrain escape. Saw it during exercise, treated this as OM-C[1] instruction. Therefore, continued in VMC”.*
6. *“Had experienced it before in the practice, knew it could be triggered. Is also mountainous terrain, but was visual. -> continue.”*

4 Discussion

This simulator study was designed to investigate the possible negative transfer of training of overruling a pilot’s decision to perform a go-around in response to loss of visibility on the runway, or to perform an escape maneuver in response to the activation of the EGPWS alarm, respectively. The results of the transfer tests did not show significant differences in any of the objective performance measures between both groups, apart from an apparent difference in initial climb rate in response to the unexpected EGPWS warning in the transfer test. On average, the 13 pilots in the overruled group responding to this warning seemed to perform a go-around maneuver, whereas the 12 pilots in the control group responding to this warning seemed to perform an escape maneuver. However, this difference did not affect the time to reach 8000 ft. There were also no significant differences in subjective ratings between the groups. This means that we did not find a robust effect of the training manipulation of “overruling a pilot’s decision” on the dependent variables. However, the comments of six (32%) of the 19 pilots in the overruled group made clear that the training manipulation had confused them and influenced their decision making during the EGPWS transfer test. Two of them decided to abort the approach and initiated the evasive maneuver about 2-4 seconds later than the group’s average response time. The other four pilots explicitly decided to continue during the transfer test because they had been primed to do so by the instructor in the training.

In both test scenarios a comparable number of pilots in each group responded to the triggering events. In both go-around test scenarios, all pilots performed a go-around. In the EGPWS test scenarios, two-third of the pilots from both groups responded by an escape maneuver, or by a go around. However, it was difficult to distinguish between the two response types based on

---

[1] OM-C refers to Operation Manual, Part C, which contains route and aerodrome information.

objective performance measures. In this respect, the pilots' comments provided better insight in their response strategy.

This study had several limitations that made it a challenge to find statistically significant performance differences between both groups. First, the training manipulation, consisting of two instances where the instructor told the pilots in the overruled group to continue the approach, may not have been strong enough to produce a demonstrable effect. Second, all participants were professional line pilots, which made it difficult to induce a change in their behavior in a relatively brief practice session which contained only two rather subtle manipulations from the instructor. Third, although we carefully designed the simulator scenarios with the help of instructor pilots, the two test moments were created in a rather artificial way. In particular, the EGPWS was triggered at a certain distance before the runway, irrespective of the actual position relative to the mountains. Some pilots noted that they were confused about the EGPWS alarm sounding while they were sure that they were clear of terrain. This limitation is typical for simulator scenarios, where events are often initiated by the instructor, and not necessarily because the simulator crew is deviating from the flight plan.

By definition it is difficult to reproduce workload and risks perceived in real-life in a simulator environment. Therefore, the pilots' physical and cognitive responses to the events may have been more contained than would be the case in real flight. In the view of these limitations, the subjective findings observed in this study are particularly interesting because they may reflect uncertainties in the pilots which may become aggravated in more stressful real-life conditions. The finding that six pilots mentioned that the instruction to ignore the EGPWS warning during the training had influenced their decision making in the transfer test suggests that instructors should carefully consider how they overrule the trainees' decisions during training. If taking shortcuts is not avoidable, then the instructor should at least provide a clear disclaimer for why this is done, and explain that under normal circumstances the pilot's own decision was appropriate.

5 Conclusions

Although the objective performance measures did not show significant effects of the pilot's own decision being overruled by the instructor, 32% of the overruled pilots mentioned that the training session had influenced their decision making in the transfer test involving the unexpected EGPWS warning. They stated that the training manipulation had caused confusion and hesitation. In two pilots, this led to a substantial delay in response. These findings are an indication of negative transfer of training.

6 References

Alexander, A., Brunyé, T., Sidman, J., & Weil, S. (2005). *From gaming to training: A review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in PC-based simulations and games.* DARWARS Training Impact Group. doi:10.1016/j.athoracsur.2004.02.012

Hart, S., & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology, 52*, 139-183. doi:https://doi.org/10.1016/S0166-4115(08)62386-9

Houtman, I. L., & Bakker, F. C. (1989). The anxiety thermometer: a validation study. *Journal of personality assessment*, *53*(3), 575-582. doi: https://doi.org/10.1207/s15327752jpa5303_14

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

Landman, A., Mol, D., Emmerik, M., & Groen, E. (2022). *Negative transfer of training: Simulator study into effects of going beyond alarms during stall recovery training.* DOT/FAA/TC-22/10. doi:https://doi.org/10.21949/1524474

Landman, A., Pennings, H., Blankendaal, R., Bosch, K., & Groen, E. (2022). *Negative Transfer of Training of Suboptimal Degrees of Variability in the Training of Procedures.* DOT/FAA/TC-22/11. doi:https://doi.org/10.21949/1524477

Pennings, H., Oprins, E., Schoevers , E., & Groen, E. (2019). *Current insights in negative transfer of training (R11747). Soesterberg: TNO.* DOT/FAA/TC-22/10 , FAA.

A Approach chart and checklist

The following schematic approach chart of Sion airport (LSGS) was shown and explained during the briefing, and was also present in the cockpit during the experiment. It was based on the actual approach chart which we had available, but which was not in a format known to participating commercial pilots. The format was therefore adapted and it was simplified to ensure that all pilots read the essential information.

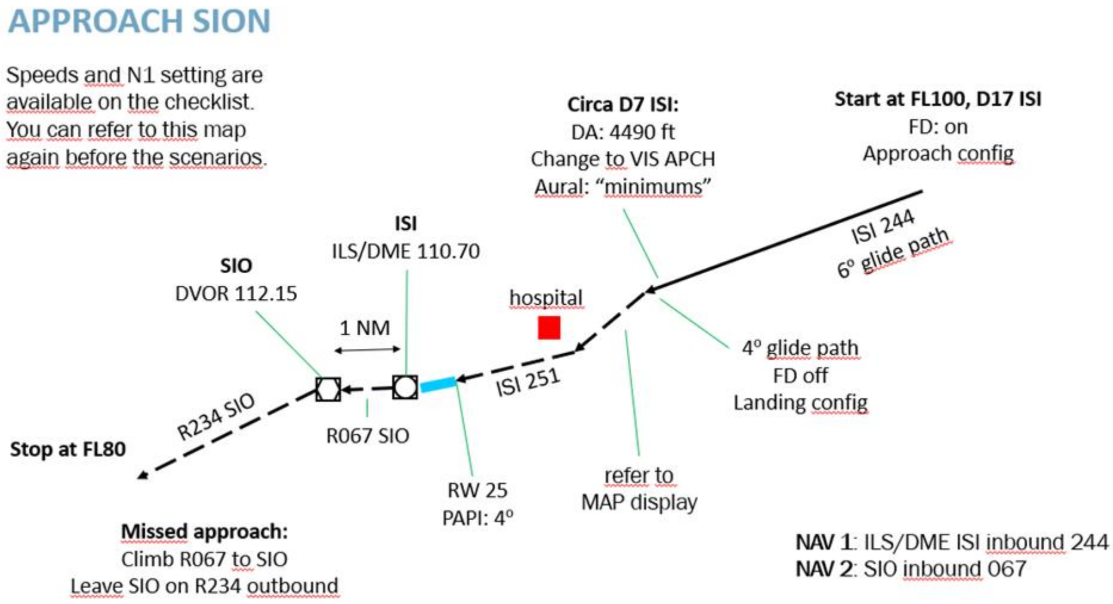

Figure A-1. Approach chart of Sion airport (LSGS)

The following instructions and checklist were explained during the briefing and present in the cockpit during the experiment:

1. Follow the IFR missed approach procedure when a go-around is required during the VIS APCH.
2. MISSED APPROACH
    a. Callout: "Go Around"
    b. N1: 95%
    c. FLAPS: TO/APRCH
    d. Positive rate: GEAR UP
    e. V: 140-150 KIAS (on checklist)
3. Scenario stops at 8000 ft.