This project has received funding from the European Union's Horizon Europe Programme

EMCO SIPO EASA.2022.C17

D-2.6 NOMINAL AND FAILURE CONDITION SIMULATOR STUDIES – EXPERIMENTAL RESULTS

eMCO-SiPO – Extended Minimum Crew Operations - Single Pilot Operations

An Agency of the European Union

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Union Aviation Safety Agency (EASA). Neither the European Union nor EASA can be held responsible for them.

This deliverable has been carried out for EASA by an external organisation and expresses the opinion of the organisation undertaking this deliverable. It is provided for information purposes. Consequently, it should not be relied upon as a statement, as any form of warranty, representation, undertaking, contractual, or other commitment binding in law upon the EASA.

DELIVERABLE NUMBER AND TITLE: D-2.6 Nominal and Failure Condition Simulator Studies – Experimental Results
CONTRACT NUMBER: EASA.2022.C17
CONTRACTOR / AUTHOR: NLR / DLR
IPR OWNER: European Union Aviation Safety Agency
DISTRIBUTION: Public

APPROVED BY:	AUTHOR	REVIEWER	MANAGING DEPARTMENT
A.D.J. Rutten	D. Niedermeier J.-P. Buch	A.L.C. Roelen	AOSH

DATE: 29-04-2025

SUMMARY

This report contains the evaluation of the Task 2 and Task 3 simulator experiments of the EASA research activity towards ”extended Minimum Crew Operations – Single Pilot Operations- – safety risk assessment framework” EASA.2022.C17, assessing the effects of the introduction of the eMCO concept of operations (ConOps) on various characteristic human factors (HF) values like workload, situational awareness and decision making in normal operations and in case of failure management. The study was conducted in September 2024 in an Airbus A350-900 Full-flight Simulator at the premises of Lufthansa Aviation Training in Munich. The experiment results should reveal potential hazards and provide a qualitative assessment of the reachable level of safety. The report describes the general setup of the simulator experiments, the selection of experiment participants, the applied measurement methods, and finally the experimental results with the focus on workload, situational awareness and decision-making.

Problem area

Recent advancements in automation, technology, and unmanned autonomous aircrafts sparked interest in and a desire to investigate if it is possible to run commercial air transport with fewer flight crew in large aircraft. The European Union Aviation Safety Agency (EASA) needs additional information about how this new concept of operations will affect safety with a focus on the analysis of human factors issues considering both normal operations and specific issues, such as sleep inertia, fatigue risk management, or pilot incapacitation.

Description of work

The objective of the Task 2 and Task 3 experiments is to analyse the impact of eMCO on HF values during nominal and abnormal (i.e. in case of system failures) operations. The study should qualitatively provide an indication of potential problem areas. It is not meant to identify the probability of the occurrence of particular hazards, as the number of participants will not have any statistical relevance.

Application

This document will feed into the final reports D-8 and D-9, which are generated in Task 8 and Task 9 of the project.

CONTENTS

ABBREVIATIONS

ACRONYM	DESCRIPTION
A/C	Aircraft
AP	Autopilot
ATC	Air Traffic Control
CAT	Commercial Air Transport
CM1	Crew Member 1, Pilot-in-Command
CM2	Crew Member 2
CONOPs	Concept of operations
CPDLC	Controller Pilot Data Link Communications
CPT	Captain
D	Demand (main dimension of 10D-SART)
DEC	Decrease
EASA	European Union Aviation Safety Agency
ECAM	Electronic Centralised Aircraft Monitor
EDDF	Frankfurt Airport ICAO code
EDDM	Munich Airport ICAO code
EFB	Electronic Flight Bag
EFIS	Electronic Flight Instrument System
ETP	Equal time points
EZE	Buenos Aires Airport (IATA code)
eMCO	Extended Minimum Crew Operation
FFS	Full-flight Simulator
FL	Flight Level
FLD	Factored landing distance (required landing distance incl. safety factor)
FMS	Flight Management System
FO	First Officer
FORDEC	Facts, Options, Risks & Benefits, Decision and Check (decision-making method)
FRA	Frankfurt (IATA code)
FSTD	Flight Simulation Training Devices
FTL	Flight Time Limitations
GVAC	Amílcar Cabral International Airport (ICAO Code)
HF	Human Factors
HI	High
HYD	Hydraulic system
ICAO	International Civil Aviation Organisation
ILD	In-Flight Landing Distance
INC	Increase
ISOS	International SOS (service hotline in case of medical issue onboard)

KCCU	Kansas City International Airport (ICAO code)
KMER	Castle Airport (ICAO code)
KSFO	San Francisco Airport (ICAO code)
KSMF	Sacramento International Airport (ICAO code)
kts	knots
LAND ASAP	Land as soon as possible (as ECAM message)
LAT	Lufthansa Aviation Training
LH	Lufthansa
LPPT	Lisbon Airport ICAO code
LOFT	Line-oriented Flight Training
MEDIAN	MED
Min	Minutes
NASA	National Aeronautics and Space Administration
NAT	Airspace of the North Atlantic
NCO	Normal Crew Operations (conventional role distribution of PF and PM according to SOPs)
NM	Nautical Mile
NOTAM	Notice to Airmen
OEM	Original Equipment Manufacturer
PF	Pilot Flying
PIC	Pilot in Command
PM	Pilot Monitoring
PR	Pilot Resting
QTY	Quantity
RCL	Recall
RSVR	Reservoir
S	Supply (main dimension of 10D-SART)
SA	Situational Awareness
SAEZ	Buenos Aires Airport ICAO code.
SART	Situation Awareness Rating Technique
SD	System Display (as part of the ECAM)
SFO	Senior First Officer
SiPO	Single-Pilot Operations
SO	Second Officer
SOP	Standard Operating Procedure
STS	Status
TEM	Threat and Error Management
TEMP	Temperature
TLX	Task Load Index
T0	Fixed time reference for all scenario events (passing of particular waypoint)
U	Understanding (main dimension of 10D-SART)

1. Context

1.1 Background

Due to the ongoing developments in technology, automation and autonomous unmanned aircraft, there is an interest and desire to explore whether it is feasible to operate commercial air transport (CAT) with reduced flight crews in large aeroplanes. This feasibility is considered from both the safety as well as efficiency perspectives.

EASA was approached by aircraft manufacturers regarding the regulatory and safety aspects of such new concept of operations (CONOPs). Two specific CONOPs were identified:

- Extended Minimum-Crew Operations (eMCO) are defined as operations where the flight time is extended by means of rest in flight with the minimum flight crew. It is achieved by allowing operations with one pilot at the controls during the cruise flight phase; however, offering an equivalent overall level of safety through compensation means (e.g. ground assistance, advanced cockpit design with workload alleviation means, pilot incapacitation detection, etc.). It is, in particular, relevant to large aeroplanes operated in CAT operations, for which no fewer than two flight crew members are currently required as per the Air Operations Regulation.
- Single-Pilot Operations (SiPO) are defined as end-to-end single-pilot operations. Annex III (PART-ORO) "Organisation requirements for air operations" to the Air Operations Regulation already foresees conditions and limitations under which these types of operations are allowed. In the future, it is expected that these conditions and limitations will need to evolve in order to extend single-pilot operations to large aeroplanes, provided that compensation means (e.g. ground assistance, advanced cockpit design with workload alleviation means, capability to cope with pilot incapacitation, etc.) are in place in order to provide for an overall level of safety equivalent to today's two-pilot operations.

1.2 Scope of the document

This report includes a detailed description of the results obtained from the simulator experiments conducted in Task 2 "Nominal Operations" and Task 3 "Failure Condition Management". The setup of the experiment is described in detail in [1]. This deliverable presents the selection of the participants, the measurement means applied to the experiments and finally the experimental results leading to a conclusion with the focus on the impact of eMCO on workload, situational awareness and decision-making.

It should be noted that the scope of this document is limited to a qualitative assessment of potential hazards. Considering the limited number of samples, it is not meant to make any quantitative statement on the likelihood of the occurrence of certain hazards during eMCO operation.

1.3 Definitions

Note that for the entire document the role distribution of the crew remains unchanged. For all experiments the Crew Member 1 (CM1) was Pilot Resting (PR) during the eMCO segment and Pilot Monitoring (PM) during the non-eMCO segment. The Crew Member 2 (CM2) was always PF. During the eMCO segment he was PF alone. During the non-eMCO segment he was still PFs but shared his tasks with the PM according to the Standard Operating Procedures (SOPs). The eMCO segment also includes the sleep inertia recovery time (15 min) defined in the eMCO procedures used for the experiments (see Section 2.1.2.2), i.e. the CM1 is still designated as PR during the sleep inertia recovery time even if he returned already to his cockpit seat.

2. Experiment Description

2.1 Objective of the Simulator Trials

The general objective of the simulator trials is a qualitative assessment of the impact of eMCO on essential human factor values, such as workload, situational awareness and decision-making. Two scenarios were flown in the simulator, both leading to an unplanned abortion of the eMCO segment.

The first objective is to assess the increase in workload for the Pilot Flying alone during the eMCO segment, who has to perform all tasks in the cockpit alone until the Pilot Resting is back on duty after the abortion of the segment.

In addition to that the impact of the separation of the crew for the duration of the eMCO segment on situational awareness and decision-making was identified as a potential hazard during nominal as well as during abnormal operation in the deliverables D2.1 [2] and D3 [3]. During the eMCO segment the crew is not able to build up a shared mental model of the situation, which also might have an impact of the final decision-making after the PR is back on duty. Furthermore, the PF might make some decisions alone before the PR returns to the cockpit.

The second objective of the trials is to evaluate the decision-making process, such as the consideration of different options and associated risks and benefits, and assess the impact of eMCO on the quality of the decision-making process.

A specific decision that has to be made during eMCO is the decision to abort the eMCO segment which has to be performed by the PF alone. As a third objective a specific focus in put on the timing of this decision and the related decision-making process.

2.1 Experimental Procedure and Test Execution

The experimental procedure and the execution of the simulator runs were fixed for all nine test days in order to guarantee the same condition for all crew. This relates to the briefing presentation content, the briefing documents as well as the execution of each simulator run and the scripting for every role.

2.1.1 Briefing

The briefing of the crews was divided into three different parts:

- One general briefing on the research project and some basic eMCO information, such as the general operational concept, prerequisites for eMCO and the specific contents of the transition briefing.
- The second part included details on the specific experimental setup: the procedure for the simulated resting (see Section 2.1.2.1) , available contacts during the experiment (see Section 2.1.2) and specific procedures for weather updates (see Section 2.1.2).
- The final part of the briefing included the dedicated operational briefing for the two scenarios including the planned routing and the fuel planning. A complete Operational Flight Plan (OFP) for each scenario was handed out to the crew. As part of the simulator study design generic eMCO procedures had to be developed based on information provided within the eMCO-SiPO project. A detailed description of those procedures can be found in Section 2.1.2.2. The procedures were provided to the crew as paper checklists.

After the briefing presentation the crew received sufficient time to prepare the flight using the OFP and specific briefing items listed on the eMCO procedure checklist.

2.1.2 Experimental Setup

For the experiment certain otherwise free parameters had to be fixed in order to have a consistent experimental setup. The Crew Member 2 (CM2) was assigned the PF and the CM1 the PR and PM roles in both scenarios. The participants were not allowed to switch roles during the experiment. This decision was made to avoid that the captain makes decisions during the eMCO phase on his or her own without involving the CM2 and to observe a greater variation of the abortion decision timing.

To increase the operational relevance, the scenarios were set in an open world in which the crews could interact freely. However, to ensure comparability over the trials, all available roles participating in the experiment were scripted in advance. The roles the crews could interact with where: Air Traffic Control (ATC), medical hotline, company/dispatch/maintenance, and the cabin. ATC services were emulated by the simulator operator, who also was a former Air Traffic Controller. All other services were taken over by the two participating researchers.

The crew carried their briefing packages from the briefing room into the cockpit as initial flight information (e.g. for the weather) and were subsequently updated with additional weather information either as printed handouts or on crew request by ATC.

During non-eMCO phases the simulator was occupied by CM1, CM2, simulator operator/ATCO, and the two researchers.

2.1.2.1 Simulated eMCO Resting Period

Based on project-internal information the maximum eMCO shift duration was set to 150 minutes from the time the PR leaves the flight deck until his or her return, so including the time required for waking up, recovering from sleep inertia, and performing the mandatory eMCO end briefing. The participants were aware that the eMCO segments were to be aborted earlier due to the experienced events in the scenarios and that they won't have the opportunity for real rest during the eMCO segment. After each eMCO start briefing was performed, the PR was led out of the simulator by one of the two researchers and either guided back into the briefing room or to the simulator building's coffee bar. The PRs were advised to stay in the respective place in order to guarantee comparable return times in case of the eMCO abortion.

The PFs were briefed before the experiment, that they abort the eMCO segment by pressing the flight crew rest call button in the overhead panel. This was observed by the researchers who brought back the PRs. The return was timed by the collecting researcher so that the PR was back in the flight deck approximately two minutes after the PF aborted the eMCO segment in order to have a comparable situation over all experiments.

2.1.2.2 EMCO Procedures Applied in the Experiments

The design of the experiment required the application of special eMCO procedures. As certified eMCO procedures are not existing yet, a generic procedure was developed for the study including the following items:

- Guidelines for the contents of the eMCO start transition briefing and for the eMCO end transition briefing
- A checklist for the eMCO prerequisites
- A list of the eMCO abortion criteria
- The sleep inertia recovery phase defined by the 15 min time period in which the PR was not allowed to take any action or participate in the decision-making process

The contents of the transition briefing were based on transition briefings already applied to current long-range operations in case of crew member changes. The complete procedure was provided a paper checklist to both pilots (see Annex A).

As, especially, the sleep inertia recovery procedure was unusual for the crews, it was explained in detail to prevent too much confusion during the experiment. The timeline presented to the crews during the briefing is shown in Figure 1.

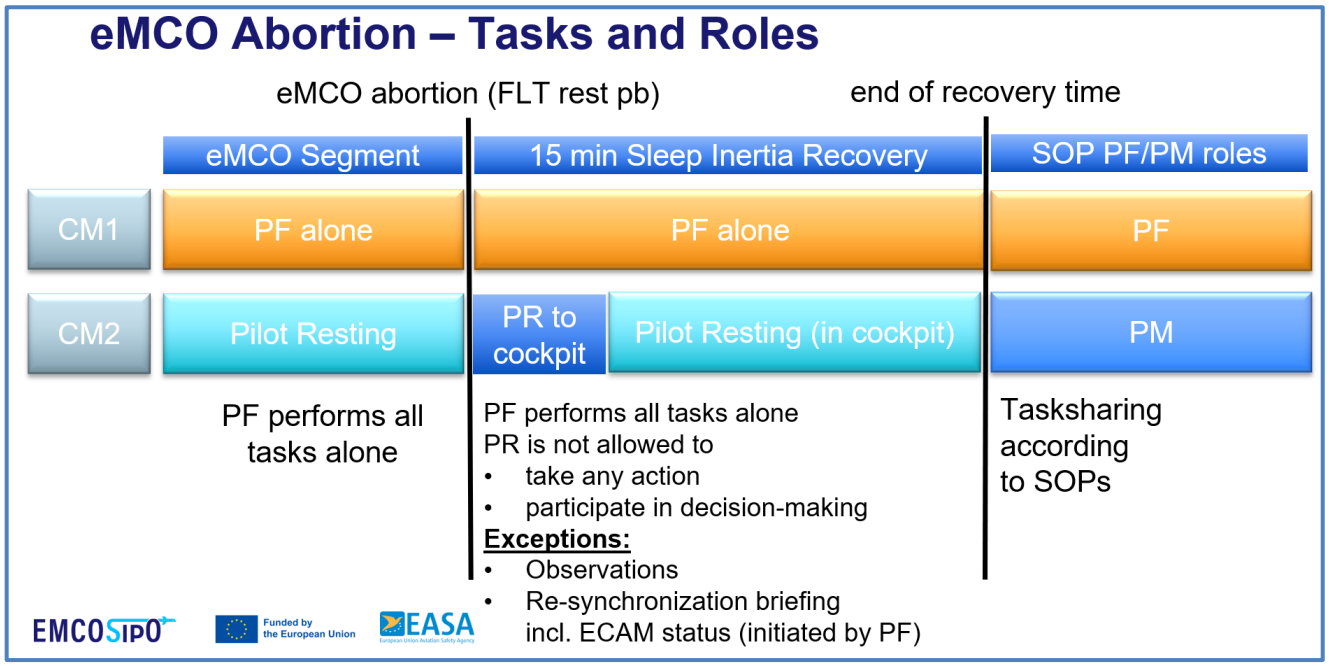


Figure 1 Briefing presentation for eMCO abortion timeline

2.2 Scenario Description

For each project task, Task 2 and 3, a specific scenario was designed. The basic events considered for the design of the scenario were pre-selected in [2] and [3]. The nominal scenario included a medical issue with a passenger aggravating over the time. This scenario was designed such that the timing of the eMCO abortion was not clearly defined by the occurring events. In that case the decision to abort the segment was not covered by the procedure but was based on the subjective assessment of the situation of the PF.

The abnormal failure condition scenario included two issues with the hydraulic system finally leading to double hydraulic failure. This scenario was selected due to the relatively high amount of ECAM actions and status entries resulting from the two system failures [4].

For both scenarios the position of the aircraft at the occurrence of the main scenario events as well as the weather situation at the potential alternates was defined such that an extensive decision-making process was required.

Both scenarios begin in cruise flight. In both scenarios the PR leaves the cockpit after the crew has performed the eMCO start briefing. The simulator flight freeze took as long as needed to perform the eMCO start briefing. This freeze was used to guarantee that after the PR left the cockpit, the run started always at exactly the same aircraft position independent of the duration of that briefing.

In order to counteract any familiarization effects to the eMCO procedure, the order in which the participating crews experienced the scenarios was changed with half of the crews beginning with the nominal operation scenario and the other half beginning the experiment with the failure condition management scenario, which are both described in detail in the following subsections.

2.2.1 Nominal Operation Scenario

The nominal operation scenario represents a Line-Oriented Flight Training (LOFT) scenario simulating a flight from Frankfurt (EDDF) to Buenos Aires (SAEZ). The scenario point begins just before the waypoint ASMAR which is approximately 20 min abeam Lisbon airport (see Figure 2).

Figure 2 Route planning and scenario start point of the nominal operation scenario

The cabin calls as part of the scenario are designed such that the medical condition of the passenger is not clear completely clear to the crew making also the decision to abort eMCO diffuse for the PF.

The timeline of the scenario is presented in Table 2.

Table 1 Timeline of the nominal operation scenario

When	Event	Notes
	Crew enters cockpit	
T0 – 20	Sim will be set up (10 min): take-off performed by crew, afterwards repositioning, camera setup, data recording	
T0 - 6	Crew performs eMCO briefing (5 min): at repos waypoint	10 NM inbound ASMAR
	PR leaves the cockpit	
	Sim is released when pilot has left the cockpit	
T0	**Passing waypoint ASMAR**	
	PF in eMCO, possibly checking surrounding airports and weather	Be prepared for information requests (weather on printouts)
T0 + 10	Call from cabin: *"Hello, here is the cabin with a brief information. We currently have a passenger in Economy, about 45 years old, who is complaining of dizziness and poor circulation for about 10 minutes now. He is responsive and seems to be oriented. We're continuing to look after him and I'll get back to you in 10 minutes with an update."*	
	Lisbon (LPPT) weather gets worse: • One-side-only approach • 10 kts tailwind • Windspeed 25G35	Handout weather & NOTAM (Notice to Airmen)
T0 + 20	Call from cabin: *„ Hello, here is the cabin with an update on our passenger with the circulatory problems. He is now unconscious but has a pulse and breathing. We have moved him to the rear galley. He is now in the recovery position, CardioSecur is connected. We're calling for a doctor right now and have begun to fill out the medical sheet. We will bring it to you in a few minutes."*	*Good crew reaction would be to• call ISOS• abort eMCO*
	If eMCO was aborted: PR has to wait 15 minutes sitting on his hands to counter the effects of possible sleep inertia / simulate gradual return	Expected PR return time from alarming approx. 5 minutes (controlled by experimental crew)
	If ISOS is contacted: They advise to continue	

	PR is back in the cockpit,	
T0 + 30	*Call from cabin: "Here is the cabin with an update on our medical. The passenger has regained consciousness and we have a doctor aboard who is with him. The doctor says the condition is stable. The passenger is still being monitored by the doctor and is currently being given oxygen, something to drink and eat. If his condition remains stable, the doctor says he can return to his seat in about half an hour. We have completed the medical sheet and will hand it to you right away. "*	*Maybe enroute holding because of unclear situation (not depart too much from possible alternate)*
	Hand prepared medical sheet to the cockpit	
	***If ISOS is contacted**: T*hey advise to continue	
	Call from cabin*: "Hello, here's the cabin again. The doctor says that the passenger is in a critical condition. The doctor recommends landing as quickly as possible for further medical treatment and transfer to a hospital."*	
	If ISOS is contacted (again): ISOS advises ASAP diversion, but cannot yet generate an operational decision / network problem	
	Lisbon closes due to weather, 25G37 kts cross wind out of limits	
approx.T0 + 45	*Crew performs decision making*	

The first call from the cabin is initiated by the experimental crew slightly more than 10 min after the PR has left the cockpit (and 10 min after ASMAR has been passed). This call just states that there is a medical issue in the cabin and that there will be an update soon. After 10 min a second call reports a worse medical condition of the passenger, already increasing the likelihood of a diversion including the abortion of the eMCO segment. At the same time the PF should set up a contact to the medical hotline in order to get medical support from ground. The purpose of the third call is to create a diffuse picture of the passenger's medical condition by reporting that the passenger's condition has improved. Finally, the last call from the cabin generates a medical emergency situation putting time pressure on the crew to increase workload.

To complicate the decision-making process, the weather of Lisbon, being the most likely alternate airport, is deteriorating. Finally, the weather situation results in a complete closure of Lisbon airport. The according NOTAM is handed out to the PF during the eMCO segment so that the PR is not immediately aware of this constraint requiring a good transition and re-synchronisation briefing to guarantee a sound decision performed by both crew members.

Due to the diffuse medical condition of the passenger the eMCO abortion decision is not straightforward for the PF. The resulting workload for the PF alone is very dependent on the timing of the abortion decision, because the PF might still be alone, when the final cabin call informs him on the medical emergency immediately triggering a diversion.

2.2.2 Failure Condition Management Scenario

The failure condition management scenario is designed as LOFT scenario simulating a flight from Munich (EDDM) to San Francisco (KSFO). The scenario starting point is just before the waypoint BALIX being the entry point to the airspace of the North Atlantic (NAT). The planned track is shown in Figure 3.

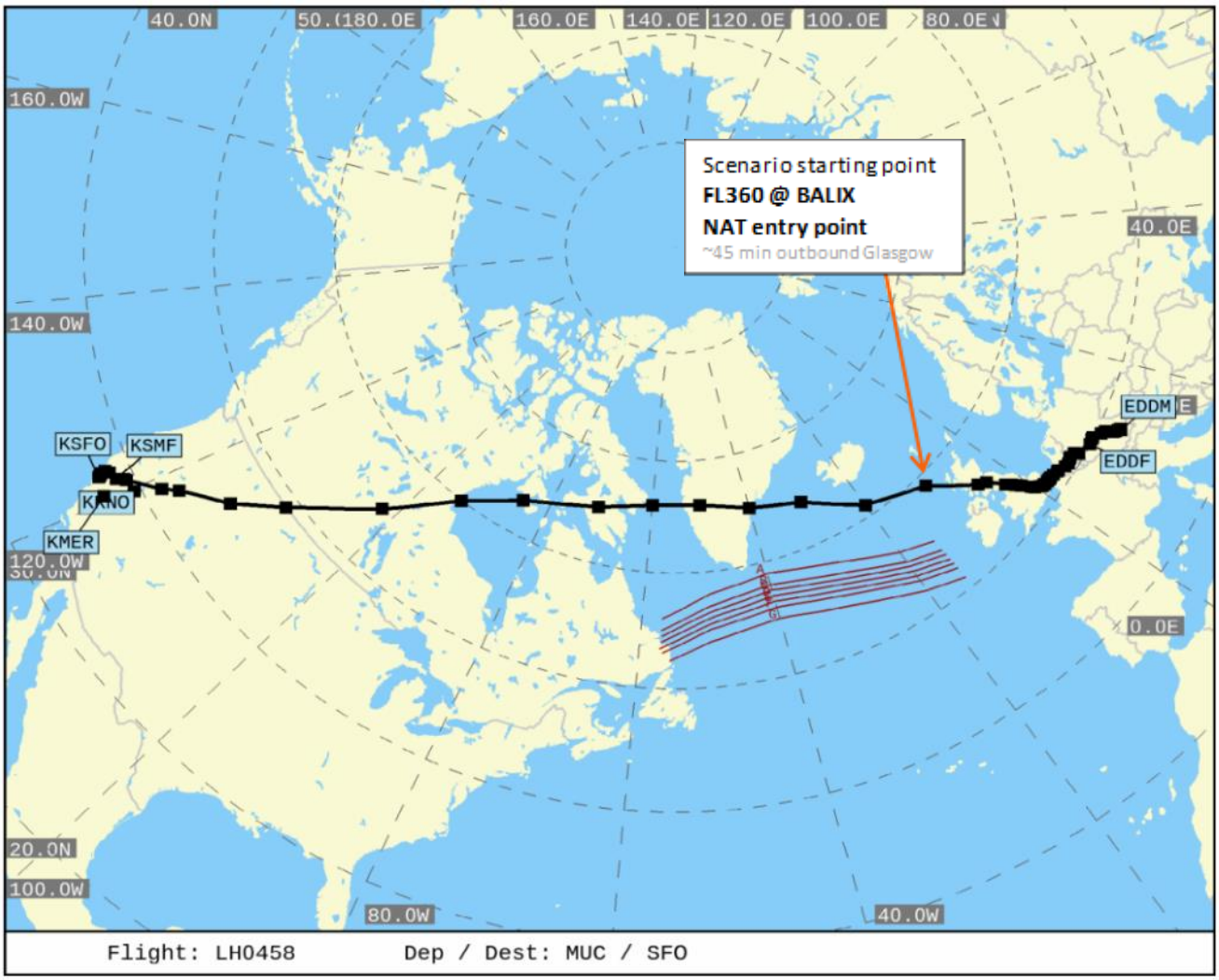


Figure 3 Route planning and scenario starting point for the failure condition management scenario

Table 2 presents the planned timeline of the main scenario events:

Table 2 Timeline for the failure condition management scenario

When	Event	Info
	Crew enters cockpit	
T0 - 20	Sim will be set up (10 min): take-off performed by crew, afterwards repositioning, camera setup, data recording	
T0 -6	Crew performs eMCO briefing (5 min)	10 NM inbound BALIX
	PR leaves the cockpit after briefing	Guided to briefing room, Just before passing BALIX
	Sim motion is released when pilot has left the cockpit	
T0	**Passing waypoint BALIX**	
T0 + 10	Trigger **HYD Y SYSTEM TEMP HI**	HYD Y causes slats half speed
	PR is requested by PF to return as per eMCO abortion criteria (single chime, amber caution)	
approx. **T0 + 15**	PR returns to the flight deck	PR needs to wait 15 minutes before being allowed any action
	Trigger **HYD G RSVR LOW LEVEL (FLUID QTY INC/DEC)**	Set quantity level to 9 Litres to trigger fault.
		Resulting ECAM warning HYD G + Y SYS PRESS LO
approx. **T0 + 25**	ECAM Actions completed	
T0 + 45	Crew performs decision-making	

The first failure is triggered slightly more than 10 minutes after the PR has left the cockpit (and 10 minutes after BALIX has been passed). The failure is a high temperature of the yellow hydraulic system requesting ECAM actions from the PF that lead to a de-activation of the yellow hydraulic pumps. This results in a temporary loss of the yellow hydraulic system which has a slight impact on the performance of the secondary flight control surfaces. In the scenario this failure is used as a trigger to abort the eMCO segment, because it causes an amber caution with a single chime that is clearly listed as abortion criterion in the eMCO checklist previously provided to the crew.

The PF aborts the eMCO segment by pushing the flight crew rest pushbutton in the overhead panel. Exactly after 3 min (controlled by the experimental crew) the PR returns to the cockpit, which is the trigger for a second system failure. The fluid level of the green hydraulic reservoir is reduced so that the related warning is triggered. In combination with the yellow hydraulic system failure, this results in low pressure in both hydraulic system and virtually in a loss of the complete hydraulics, only relying on the electrical backup system of the aircraft.

Both failures trigger a significant amount of ECAM actions that have to be handled by the PF alone which is an essential scenario characteristic for the workload assessment.

In addition to that the landing performance of the aircraft is significantly affected resulting in additional challenges in terms of workload, situational awareness and decision-making.

2.3 Simulator Description

For the simulator study the FT72 Full Flight Simulator (FFS) Airbus A350-900 with Rolls-Royce XWB-84 engines was used. It is operated by Lufthansa Flight Aviation Training (LAT) in Munich and qualified according to EASA CS-FSTD (A) issue 2, Level D. The device was manufactured by TRU Simulation and Training and modified by CAE. The aircraft systems software version is published on the website of Lufthansa Aviation Training [5].

The simulator includes a Lufthansa Electronic Flight Bag (EFB) using the integrated EFB displays of the A350 cockpit. It has to be noted that it was not possible to modify the EFB and that the crews could not use their dedicated airline EFB.

Figure 4: Outside view of the Airbus A350 simulator FT72 from Lufthansa Aviation Training in Munich [5]

2.4 Measurement Methods

2.4.1 Video Recording

Provided that the participants gave their consent, video recordings of the scenario flights were captured. For this purpose, a GoPro Hero 11 action camera was fixed to the upper backend of the overhead panel from which it captured an overview of the whole cockpit. The GoPro was also equipped with the Media Mod, an external microphone for increased voice quality. The video recordings were used in the analysis and the analysed data went into this report.

2.4.1 Simulator Data Recording

A very limited set of signals could be logged in the simulator, including position and speed data. For the analysis presented in this report, the aircraft position at time of the eMCO abortion was determined through a combination of simulator logging data and video recording data. Both sources were manually synchronized using the moment the first waypoint was sequenced (video recording) and the smallest great circle distance to the first waypoint. Once the synchronization was performed, the position of the aircraft at eMCO abortion was read from the logging data using the eMCO abortion time offset from the video recording.

2.4.2 Workload Assessment Using NASA-TLX Raw

The NASA TLX (Task Load Index) in its raw version was used to collect subjective ratings of perceived task and mental workload in the Task 2 and Task 3 simulator experiments. The NASA TLX is a multidimensional rating scale of six workload-related factors that are combined to derive a sensitive and reliable estimate of workload. The scale that was used had two extreme poles ranging from zero (low) to ten (high) (see Annex B).

A separate rating was performed for two parts of the experimental run. For both experiments, normal operation as well as failure condition management, the phases were defined as follows:

1. Time period from scenario start until PR is back at the controls (incl. sleep inertia recovery phase)
2. Time period after PR is back at the controls with standard PF/PM roles until the end of the scenario

The raw NASA-TLX rating uses the following questions from Table 3:

Table 3: NASA raw TLX rating categories and questions

Category	Question to rate
Mental demand	How much mental demand and perceptual activity was required (i.e. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical demand	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal demand	How much time pressure did you feel due to the rate or pace at which the scenario elements occurred? Was the pace slow and leisurely, or rapid and frantic?
Performance	How successful did you think you were in accomplishing the goals of the scenario? How satisfied were you with your performance in accomplishing these goals?
Effort	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration level	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

2.4.3 Situation Awareness Rating Technique

The Situation Awareness Rating Technique (SART) is a well-known and thoroughly tested subjective metric ([6], [7]) which was used to capture the PR's/PM's and PF's levels of Situation Awareness (SA). The ten-dimension SART scale (10D-SART) was used to collect ratings regarding instability of situation, variability of situation, complexity of situation, arousal, spare mental capacity, concentration, division of attention, information quantity, information quality and familiarity in the Task 2 and Task 3 experiments. The scale of the questionnaire was anchored by word descriptors at each end and ranged from zero (low) to ten (high).
As for the NASA-TLX ratings the 10D-SART questionnaires were separated into two parts: one part for the first phase in the scenario, which lasted from the beginning of scenario until both pilots were back at the controls, and a second part for the second phase of the scenario considering the time period from "both pilots back at the controls" until the end of the scenario.
In contrast to the NASA-TLX ratings the PM/PR had to answer the 10D-SART for both phases in order to rate his SA during the sleep inertia recovery time.

The 10D-SART questionnaire includes the following dimensions and questions:

Main Dimensions	Basic Dimensions	Questions
Demand (D)	Complexity of Situation	How complicated was the situation? Was it complex with many interrelated components (high) or was it simple and straightforward (low)?
	Instability of Situation	How changeable was the situation? Was the situation highly unstable and likely to change suddenly (high), or was it very stable and straightforward (low)?
	Variability of Situation	How many variables were changing in the situation? Were there a large number of factors varying (high) or were there very few variables changing (low)?
Supply (S)	Arousal	How aroused were you in the situation? Were you alert and ready for activity (high) or did you have a low degree of alertness (low)?
	Concentration of Attention	How much were you concentrating on the situation? Were you bringing all your thoughts to bear (high) or was your attention elsewhere (low)?
	Division of Attention	How much was your attention divided in the situation? Were you concentrating on many aspects of the situation (high) or focused on only one (low)?
	Spare Mental Capacity	How much mental capacity did you have to spare in the situation? Did you have sufficient capacity to attend to many variables (high) or nothing to spare at all (low)?
Understanding (U)	Information Quantity	How much information did you gain about the situation? Did you receive and understood a great deal of knowledge (high) or very little (low)?
	Information Quality	How good was the information you have gained about the situation? Was the knowledge communicated very useful (high) or is it a new situation (low)?
	Familiarity with Situation	How familiar were you with the situation? Did you have a great deal of relevant experience (high) or was it a new situation (low)?

The overall SA score is than determined by:

$$SA = U - (D - S)$$

For the evaluation of the main dimensions U, D and S ratings the median of the corresponding ratings of the basic dimensions was used. Other references also use the sum of all basic dimension ratings to determine U, D and S.

2.5 Demographic Data, Crew Composition and Roles

For the recruiting of the study participants pilots with a valid Airbus A330/350 type rating were contacted all over the world via e-mail, mostly through databases available through the national aviation authorities. The invitation offered travel expenses as well as an allowance fee of 500 € for the participation of each pilot. However, due to the combined Airbus A330/350 type rating, a lot of interested pilots had to be screened out as they did not have actual A350 flying experience, which was a prerequisite for participation in the experiments. Airlines were not specifically selected. However, pilots currently flying for Original Equipment Manufacturers (OEMs) were excluded.

In total, nine crews participated in each scenario. For three out of the nine crews the CM1 and Pilot in Command (PIC) and the CM2 were from the same airline (see Table 4). The majority of the crews was composed of a Captain (CPT) and Senior First Officer (SFO), for one crew two CPTs were combined and, finally, one crew was composed of a CPT and a Second Officer (SO) (see Table 4).

All participants were male. Fifteen out of the 18 pilots were flying for an airline, one pilot had a military background currently flying A350 in a government fleet. Two pilots retired recently.

All except of one airline pilots were working for Western-European airlines.

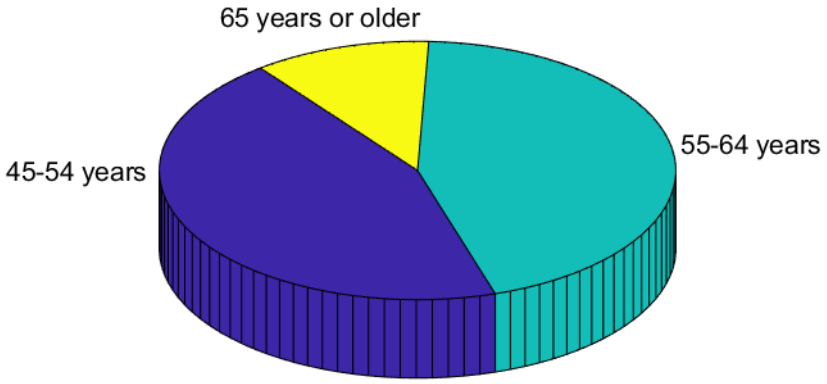


Figure 5 Age distribution of all participating Captains (N=9)

Figure 5 shows the age distribution of all Captains that participated in the experiment, whereas Figure 6 presents the age distribution among the Senior First Officers and First Officers. The age distributions are representative for airline operations.

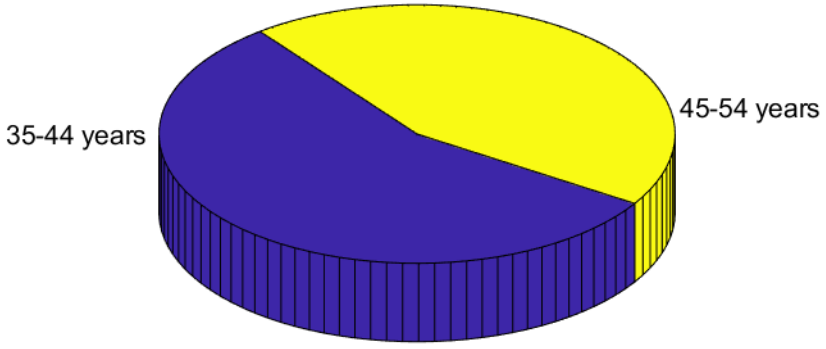


Figure 6 Age distribution of all participating Senior First Officers and First Officers

Table 4 Crew composition including rank and flight hours for all crews (N=9)

Crew Number	CM1		CM2 (PF)		Same Airline
	Rank	Flight Hours	Rank	Flight Hours	
1	CPT	15500	CPT[1]	3200	NO
2	CPT	15600	SFO	11000	NO
3	CPT	17000	SO	10000	YES
4	CPT	22500	SFO	6000	NO
5	CPT	20000	SFO	12000	YES
6[2]					
7	CPT	16500	FO	7000	NO
8	CPT	14000	SFO	6000	NO
9	CPT	18700	SFO	15000	YES
10	CPT	11500	SFO	10000	NO

All runs were conducted with fixed roles for the PF and the PR/PM to reduce the number of variables. It was decided that the CM1 participates in the experiment as PR/PM whereas the CM2 took part as PF.

The distribution of flight hours for the PR/PM and those of the PF are presented in Figure 7.

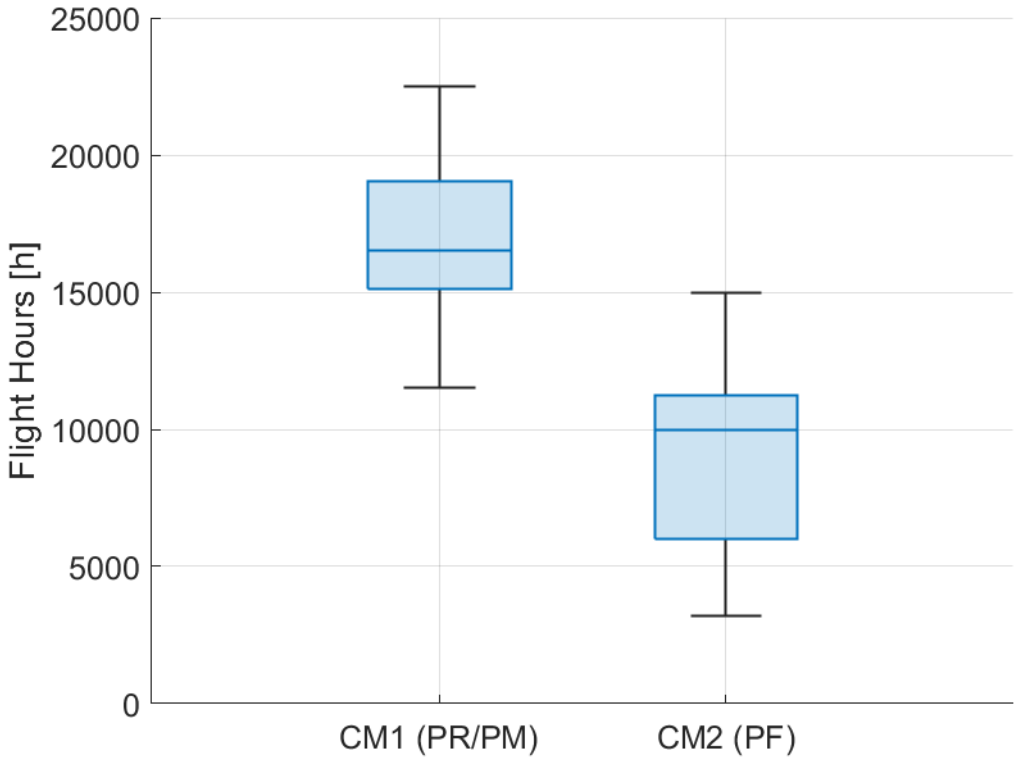


Figure 7 Flight hour distribution for CM1 and CM2

[1] The Captain was just before completing the Captain's course at the time of experiment.
[2] Crew 6 cancelled the experiment. However, the internal numbering for this report remains unchanged.

3. Nominal Operations Experiments

3.1 Crew Stats/Quick Facts/Complement for nominal operation scenario

Fact	Crew 1	Crew 2	Crew 3	Crew 4	Crew 5	Crew 7	Crew 8	Crew 9	Crew 10
Same Airline	no	no	yes	no	yes	no	no	yes	no
PR woken up T0+[mm:ss]	12:30	15:00	35:25	35:35	35:02	23:31	36:00	38:12	12:31 (22:31)[3]
PR back on seat T0+[mm:ss]	13:55	16:09	39:00	38:22	38:17	26:38	39:00	40:27	15:48 (25:48)
eMCO end briefing performed at T0+	22:42	26:18[4]	44:29	50:44	Started at 43:30, but never finished	31:13	No	46:23	No
PR → PM transition allowed T0+[mm:ss]	27:30	30:00	50:25	50:35	50:02	38:31	51:00	53:12	27:31
Actual PR → PM transition T0+[mm:ss]	15:26	17:30	41:51	46:51	51:45	38:25	49:52	52:39	25:24
PR active before recovery time[5] (+ performed actions, if applicable)	yes: EFIS CP, KCCU, decision-making	yes: EFIS CP, KCCU, EFB, decision-making	yes: KCCU, EFB, decision-making	yes: EFIS CP, KCCU, EFB, decision-making	no	no (within grace period)	no (within grace period)	no (within grace period)	yes: EFIS CP, KCCU, ISOS call, decision-making
CM2 eMCO abortion reason	Diversion might be required	Diversion might be required	Required diversion	Required diversion	Required diversion	Required diversion	Required diversion	Required diversion	Diversion might be required

[3] Due to simulator problems at the beginning the available time was limited. Therefore, the timing of the cabin calls was modified. The second cabin call (passenger lost consciousness) was moved from T0+20 to T0+10. The time in brackets is the corrected time consistent with the time of the other crews.

[4] Crew 2 wanted to start with the eMCO end briefing at T0+18 min, but asked experimental lead whether they should wait until end of sleep inertia recovery. This was confirmed by the experimental lead and then the crew waited for until T0+26.

[5] According to the procedures the PR was only allowed to observe or to take part in the re-synchronization briefing. For this analysis "observe" was defined as follows:

1. ok to switch onside displays in order to gather information without disturbing or interacting with the "PF alone"
2. Not ok: enter information into aircraft systems or EFB, e.g. FMS modifications, ATC COM, performance calculation on EFB

Initial diversion decision CM2	-	-	Seville	Seville (after calling med. Hotline)	Lisbon (didn't notice closure)	-	Prepares Gran Canaria (but waits for CM1)	-	-
Initial CM1 diversion decision	Lisbon (NOTAM not noticed)	Portugal or Morocco	-	-	-	-	-	-	-
Final CM1 diversion decision	Rabat	Seville	Seville	Seville	Faro	Porto Santo	Gran Canaria	Gran Canaria	Sevilla
Company was informed	Yes	Yes	Yes	Yes + ISOS	No	Yes + ISOS	Yes + ISOS	Yes + ISOS	Yes + ISOS
Declared PAN PAN	Yes, 36:14	No	No	No	35:18	Yes, 42:48	No	No	No
Declared Mayday	-		No	No	No	48:44	No	No	No
LAPA performed for	Rabat, FLD = 2287 m	Seville, FLD = 2400 m	Seville, 2338 m @ conf 3, do a conf full ldg	Seville	Faro	-	Gran Canaria	Gran Canaria	Seville
NITS performed	No	No	No	No	Yes, 55:26	No	Yes	Yes, 53:25	Yes, 32:52

3.2 Workload

As described in Section 2.4.2 the NASA-TLX rating in its raw version was answered by the PR/PM and the PF after each scenario. It was separated in two phases (see Section 2.4.2):

1. Time period from scenario start until PR is back at the controls (incl. sleep inertia recovery phase)
2. Time period after PR is back at the controls with standard PF/PM roles until the end of the scenario

All ratings of the pilots that participated as PF (N=9) are presented in Figure 8 and Figure 9. Each boxplot shows the results for every NASA-TLX category. Figure 9 illustrates the ratings for the 1st Phase of the Nominal Operation scenario.

The median of the ratings for mental demand is higher than for physical demand which could be expected as the amount of physical pilot actions was low for this scenario. Even for mental demand the median reaches only a rating of MED = 3.0 (SD = 2.2). The majority of the pilots stated that they accomplished the goals of the scenario and that they were satisfied with their performance (Median = 9.0, SD = 1.5). Basically, workload seems not be an issue for the majority of the PFs when operating the aircraft alone. However, it has to be noted that the range of all ratings is large for the first phase of the scenario.

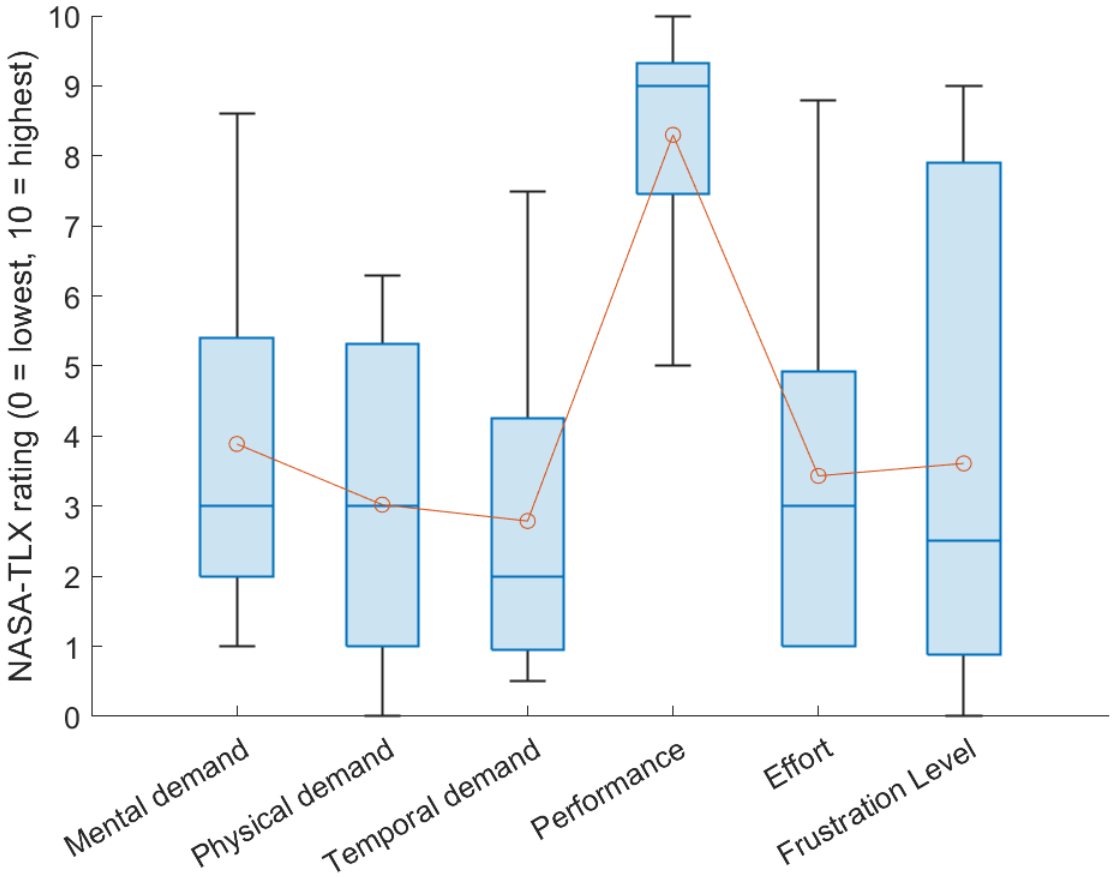


Figure 8 NASA-TLX rating of the PF for 1st phase of the nominal operation scenario (N=9)

The PF of Crew 7 rated his mental demand with 8.5, his effort with 8.8 and his frustration level even with 9.0. This crew had an increased workload, because they selected a diversion airport in close proximity. On the other hand, the PF of Crew 9 provided very low ratings for the mental, physical and temporal demand as well as for his effort and his frustration level (between 0 and 1) while assessing his performance with a rating of 10 (see Table 5 NASA-TLX ratings of each PF for both phases of).

Figure 9 illustrates the ratings of the PF during the 2nd phase of the scenario, in which the PR was back at the controls (i.e. after the sleep inertia recovery time). Unexpectedly, the boxplot shows a slight tendency of higher mental workload ratings (median for mental workload MED = 4.0 with SD = 4.0) for the PF compared to the 1st phase, whereas all other categories are lower, which would be expected. However, a comparison between both phases and also between each PF is difficult due to the differences in the abortion time (see Section 3.4). PFs who aborted the eMCO segment early in comparison to other PFs were exposed to lower workload, as the entire decision-making and the re-planning of the flight could be shared between both crew members.

This is also represented by the significant variation in ratings between the crews resulting in high standard deviations for almost all rating categories. On one hand, it has to be noted that the number of test samples is low, which might be the explanation for the high SD values. On the other hand, it could be observed that the workload was indeed very different between the crews depending on different factors, such as the abortion timing and the selection of the alternate airport.

For the PF of Crew 7 the workload ratings are constantly high, even after the return of the PR to the controls. However, the frustration level is significantly lower in the 2nd phase of the scenario (4.2 during the 2nd phase vs. 9 during the 1st phase). On the other hand, the PF of Crew 9 constantly provided workload ratings close to 0.

Figure 9 also shows an outlier for the performance rating of the Crew 9 PF during phase 2. It is very likely that this was a mistake during the completion of the questionnaire.

Table 5 NASA-TLX ratings of each PF for both phases of the nominal operation scenario.

	Mental demand		Physical demand		Temporal demand		Performance		Effort		Frustration level	
	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2
PF Crew 1	2	8	0	1.0	1.0	1.0	5.0	6.0	1.0	4.0	1.0	1.5
PF Crew 2	2.0	5.0	1.0	0.5	4.0	7.5	9.0	8.6	1.0	2.0	0.5	1.0
PF Crew 3	3	4	6	2.5	2	2.5	9	9	3	5.0	2.5	3.0
PF Crew 4	4.2	0.4	3	0.5	0.5	0.4	10	10	4.9	1.6	7.8	0.3
PF Crew 5	6	5	6.3	2.0	5	2.0	7	9	5	5	1	0
PF Crew 7	8.5	8.2	3.5	8.7	7.5	7.4	8	8.2	8.8	8.7	9	4.2
PF Crew 8	5.2	3.7	5.1	2.9	2.7	2.5	7.6	8.0	3	1.7	8.2	1.0
PF Crew 9	1	0	1	0	0.8	0	10	0	1	0	0	0
PF Crew 10	3	2.5	1.3	1.4	1.6	2.5	9.1	9.7	3.2	1.8	2.5	1.6

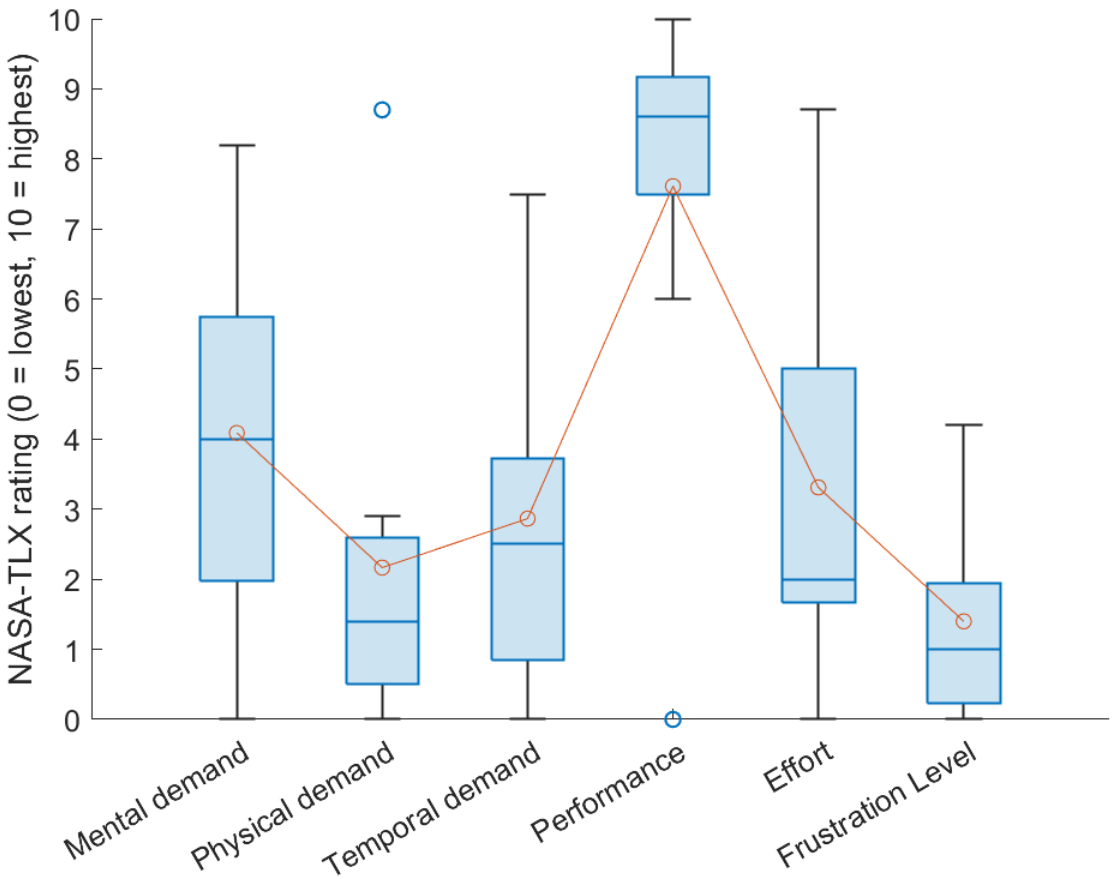


Figure 9 NASA-TLX rating of the PF for 2nd Phase of the nominal operation scenario (N=9)

Figure 10 presents the ratings provided by the PM for the 2nd phase of the nominal operation scenario (N=9). As the PM was PR during the first phase, TLX-ratings were not collected for this part of the scenario.

Basically, the PM ratings are significantly lower in workload and slightly higher in performance compared to the PF ratings. Only the PMs of Crew 3 and Crew 5 show high ratings for their mental workload. In addition, the PM

of Crew 5 assessed his performance only with 5.8 which is the lowest performance ratings of all participating pilots. This low performance rating is related to the initial diversion airport selection of the PM requesting radar vectors to Lisbon which was already closed at that time (see Section 3.4).

Table 6 NASA-TLX ratings of each PR/PM during the 2nd phase of the nominal operation scenario.

	Mental demand	Physical demand	Temporal demand	Performance	Effort	Frustration level
PM Crew 1	0.1	0.1	0.2	10	0.5	0.1
PM Crew 2	1.9	1.5	1.4	9.5	0.8	0.3
PM Crew 3	7.9	6.1	5.0	8.4	9.1	8.0
PM Crew 4	2.2	1.6	1.8	9.0	1.6	0.9
PM Crew 5	7.9	1.5	1.7	5.8	4.2	4.1
PM Crew 7	5.0	1.0	5.0	9.0	5.0	1.0
PM Crew 8	3.0	0.7	1.0	9.5	1.5	1.1
PM Crew 9	1.7	0.9	1.6	8.0	3.2	1.5
PM Crew 10	0.6	1.3	0.9	9.3	1.8	4.5

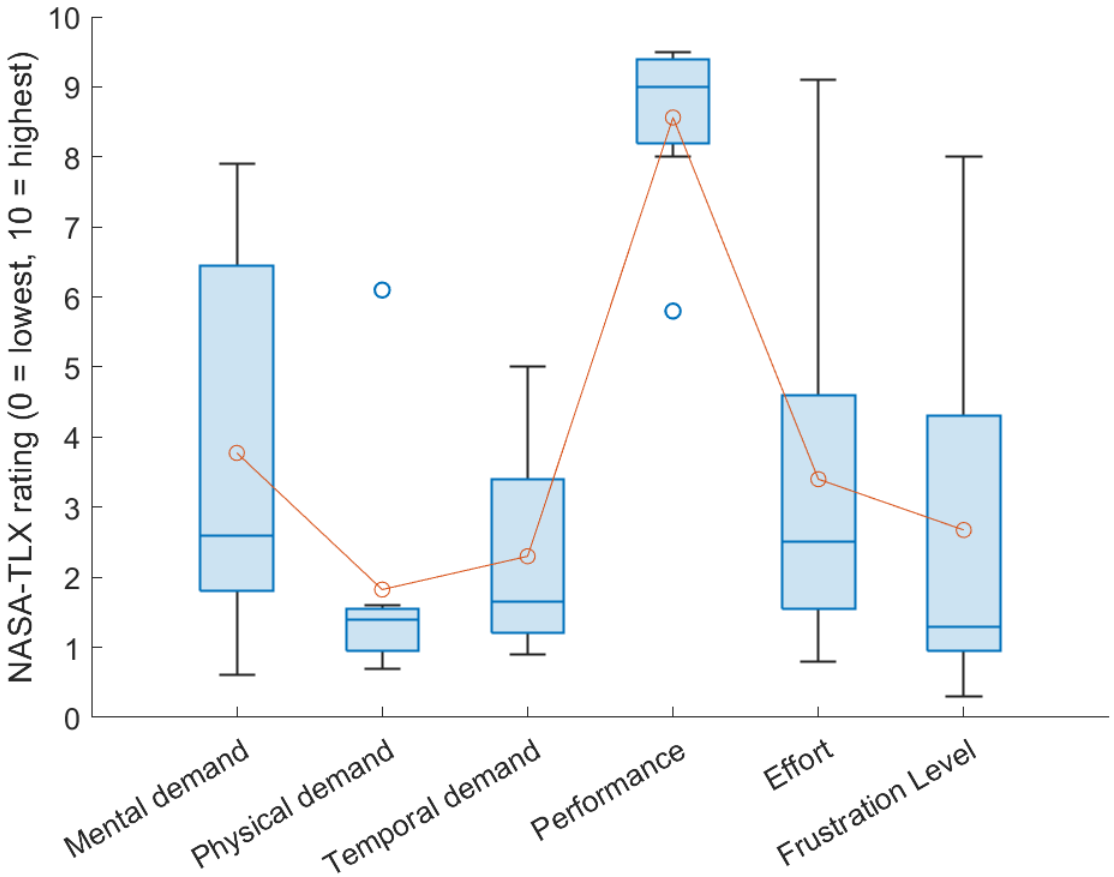


Figure 10 NASA-TLX ratings of PM during 2nd phase of the nominal operation scenario (N=9)

3.3 Situational Awareness

The 10D-SART ratings provided by the PF during the first and second phase of the nominal operation scenario are presented in Figure 11. The average situational awareness score of all PFs (N=9) during phase 1 was MED = 39.0 with SD =11.1. Note that the highest possible SA score is 70 whereas the lowest score is 4.

For phase 2 the ratings increased to an average of M = 44.7 with SD = 11.0.

The lowest PF overall situational awareness score for phase 1 was 20.0 provided by the PF of Crew 5. The same PF rated the second phase as 14.1. This low SA overall rating of the Crew 5 PF corresponds to the analysis of the decision-making process based on the video recordings. In the second phase of the scenario, during the re-synchronization briefing, the PF realized that his selection of the diversion airport was wrong, because he over-looked that the airport was already closed due to severe crosswinds (see Section 3.4).

The PM of Crew 5 rated the first phase of the scenario (for him this was time between entering the cockpit after the abortion up to the point in time in which he was allowed to take over the role as PM) as 14.1 which was the lowest SA score of all pilots. In the second phase his SA score increased to 28.4.

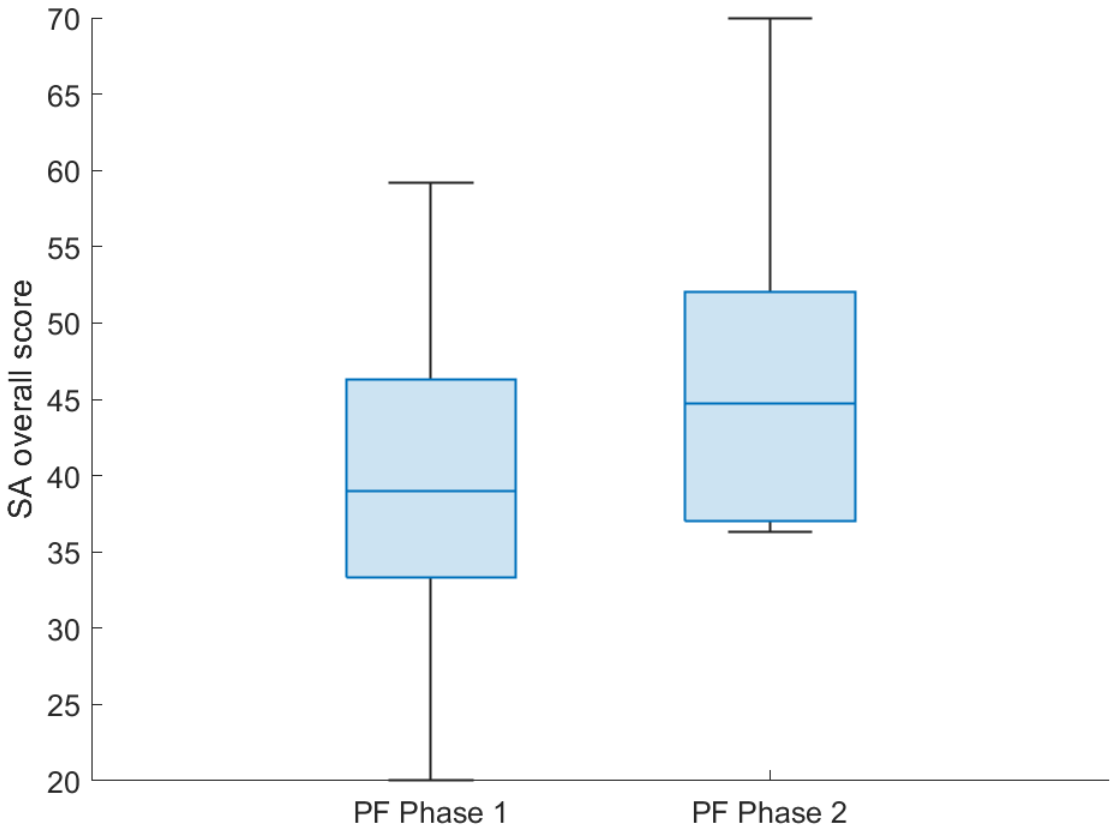


Figure 11 10D-SART Overall SA score for the PF (N=9, lowest score possible = 4, highest score possible = 70)

The PM's average rating in Figure 12 is MED = 49.6 (SD = 16.5) for phase 1 and M = 48.3 (SD = 13.7) for phase 1 (N=8). Note that the ratings of the Crew 3 PM had to be excluded, because they failed the validity check.

In average, the PMs rated their situational awareness significantly higher than the PFs did. This is in contrast to the expectation that the PRs/PMs might have a lower level of situational awareness due to their absence during the eMCO segment.

In addition, in case of the PFs there is a clear tendency for higher ratings in phase 2 compared to phase 1. As expected, the PF thus rated the phase in which the crew was acting in their usual role distribution as PF and PM with a higher SA overall score. The difference in the PM ratings between phase 1 and phase 2 is only small, but surprisingly the rating slightly decreases in phase 2.

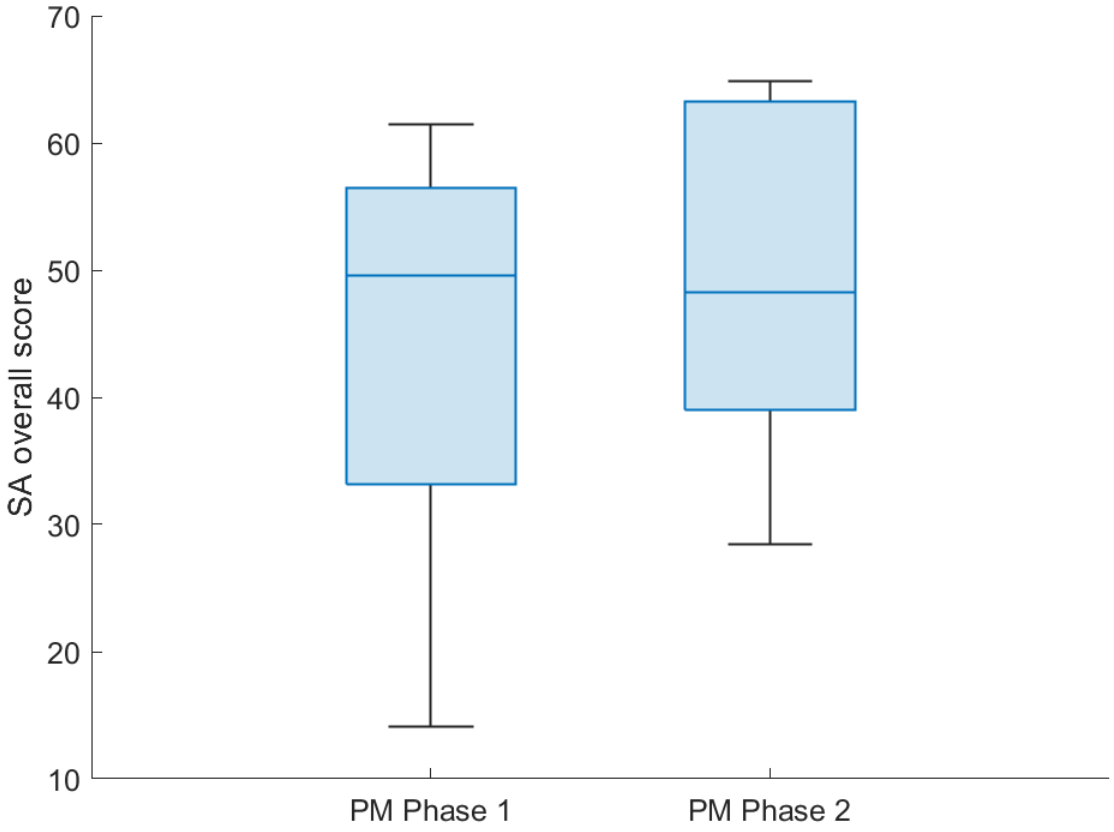


Figure 12 10D-SART overall SA score for the PM/PR (N=8, lowest score possible = 4, highest score possible = 70)

3.4 Decision-Making

3.4.1 eMCO Abortion Decision – nominal operation scenario

Figure 13 shows the track logs of all crews for the *non-normal* nominal operation scenario together with the characteristic scenario points (cabin calls). The content of the cabin calls was: 1.) first rise of a medical problem with a passenger, 2.) passenger unconscious, doctor will be called out, 3.) passenger conscious and condition stable, doctor was found, 4.) heart attack, doctor calls for immediate diversion. The calls were timed in reference to the eMCO entry point and occurred 10, 20, 30, and 35 minutes after its crossing. The only exception is Crew 10, where due to time reasons caused by a mandatory reset of the flight simulator before the experiment, the cabin calls were timed earlier to stay in the allotted simulator slot (6, 13, 21, 30).

In the eMCO abortion decision two groups could be identified. Out of the nine PF's, four aborted the eMCO segment after the first (N = 2) or second (N = 2) call, whereas the remaining five PF's aborted after the final cabin call (heart attack, doctor calls for immediate diversion). Unsurprisingly, no one aborted between the third and fourth call, were the medical situation improved.

The reason for the group of early eMCO aborters was that they anticipated a rising workload due to the passenger's medical condition and wanted to bring back the PR early so that there was more time to build up a shared analysis and mental model of the situation. The early eMCO end also tended to reduce the PF workload (see section 3.2), as consequently tasks could be shared earlier between the two crew members, the PR already experienced later cabin calls himself, and the necessary transition briefing would potentially contain fewer points of discussion. However, it must be considered that for a decision to abort the eMCO segment in the simulator, operational pressure was not a factor. This lack of consequences in the simulator experiment is in contrast to the operational reality, where an interruption of the eMCO segment results in at least a rescheduling of the flights subsequent eMCO segments. In the extreme case, the abortion of an eMCO segment could also lead to a required diversion in itself, if the flight was planned using all in the future possibly granted flight time limitation benefits to the maximum. From the second crew onwards, the participants were briefed that they should consider the operational pressure when deciding whether to abort the eMCO segment. The feedback of the participants was twofold: some stated that in reality, they *"probably would have waited longer [until the abortion]"* and that the abortion decision would also have more impact on their perceived workload, while others stated the consideration of *"…the decision to wake up the PR or to continue with a very sick passenger took quite a lot of mental capacity… [and]… the operational pressure was the main factor to not wake up …[the PR]… right away"*.

The group of the five late aborting PF's aborted only after the fourth cabin call, when the diversion was inevitable. In the debriefing, the late aborting crews stated that they see a correlation of operational flight experience with late abortion of the eMCO segment, especially experience with medical emergencies. To their experience there would likely be an abortion order from early to late with First Officers aborting early, then Senior First Officers, and then Captains aborting at the latest possible point in time. Although the sample size in the experiment was too low for a sound statistical evidence, this tendency was confirmed by all pilot during the debriefing. The feedback of the test subjects seemed to affirm that the scenario was designed realistically and the decision to abort was not easy even for the very experienced crews due to the dynamics of the medical situation with a first rise, a following improvement of the condition, before it worsened again. The highly experienced CM2's put more focus on the operational impact and often stated that their main priority is *"…to get to the destination and not to abort a flight on the first rise of a medical problem"*.

3.4.2 Diversion Decision – Nominal Operation Scenario

In this scenario the aircraft was fully intact and would have been able to land on at least all airports shown in Figure 13, except for Lisbon (ICAO code LPPT) which was the only airport that closed due to severe crosswinds over the course of the scenario. After the patient's medical condition had been varying up and down for 35 minutes, finally the diversion was mandatory after the fourth cabin call (see Figure 13), when the onboard doctor requested an immediate landing.

In the group of the late eMCO aborters, three PFs already decided to divert from the flight planned route without awaiting the return of the PR. In two cases the initial CM2 diversion decision was confirmed by the returned CM1 during the FORDEC[6]. In one case, the CM2 initially diverted to Lisbon overlooking the airport closure (due to the airport's prevailing weather conditions). The error was later noticed when the crew performed their the FORDEC and the diversion subsequently revised to Faro (see crew 5 in Figure 13). Although in this scenario the change of the diversion airport only resulted in a few minutes more flight time, it could have resulted in higher consequences like an increased workload, e.g. if the error was only noticed during the approach.

The majority of crews consulted additional simulated services like ATC, the medical hotline, or the company dispatch in their decision-making. However, the interaction with these services varied from announcing that

[6] The acronym FORDEC stands for Facts, Options, Risks & Benefits, Decision and Check and it is a method providing a framework for systematic thinking and action in decision-making processes. FORDEC is used by several major airlines. Other airlines use simulator methods such as DECIDE or DODAR.

the crew would in reality contact them, over sending fictional messages, to actually talking to the simulated services. It is not possible to assess how seriously the participants took this interaction in the experiments. In one of the nominal operation scenario runs a crew got under high workload with their decision to divert to a very close airport and performed a steep descend with speed brakes. This crew sent a fictional message to the med-link hotline about the medical infrastructure of the airport, but did not check for a reply during the self-induced high workload situation. This situation would potentially not have been encountered, if there had not been nearby airports at the point of the decision-making. This was also a point that came up several times in the debriefing interviews, where participants remarked that they would feel a higher workload when they would have to choose one of many nearby airports compared to choosing one of more far away airports and with no nearby airport. As a mitigation, emphasis could be laid on the avoidance of self-induced high workload situations due to time pressure in case of an eMCO abortion.

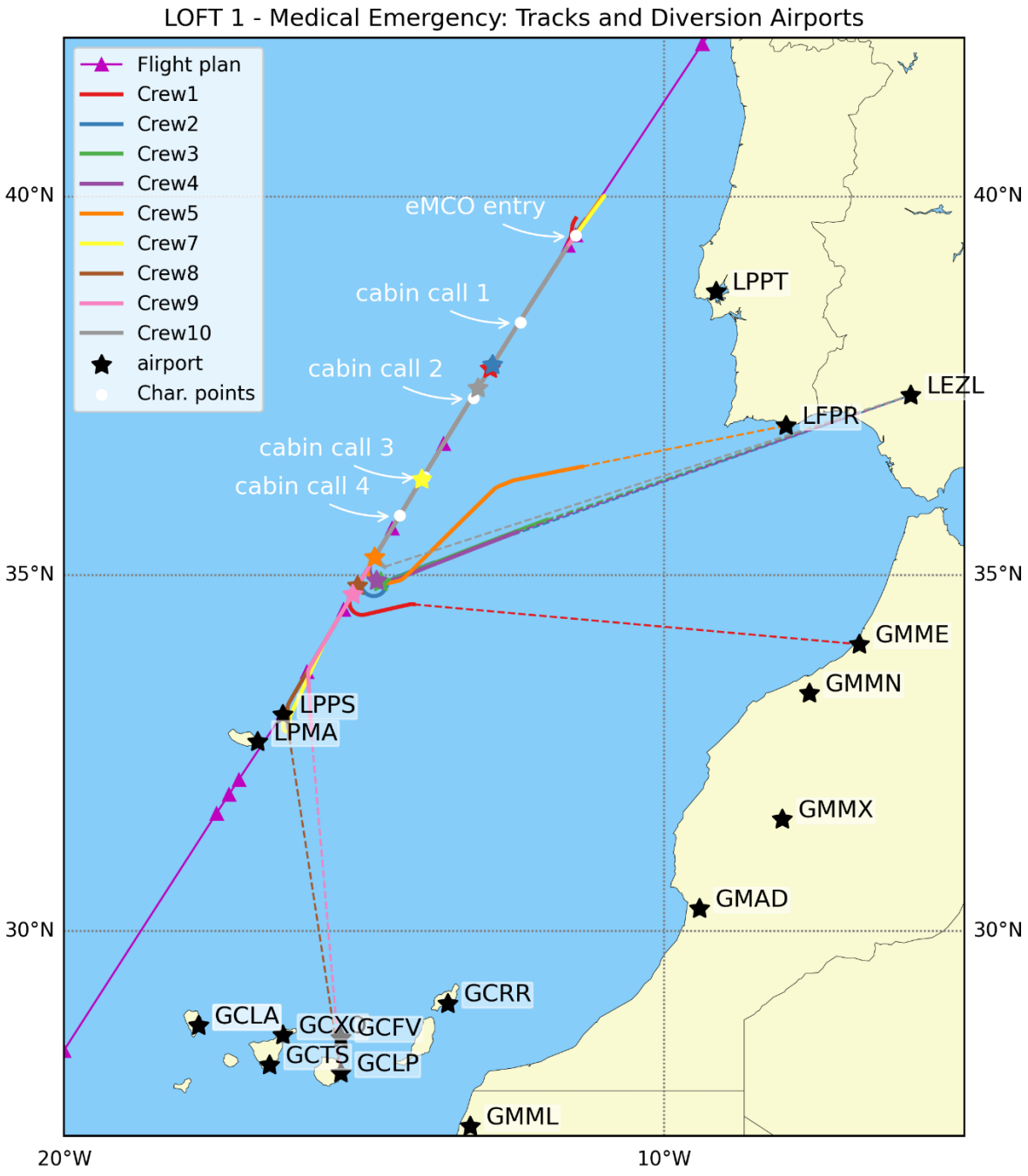


Figure 13: Nominal operation scenario track log showing the planned route, the crews' tracks and chosen diversion airports (dotted lines). The beginning of the dotted line marks the end of each experiment. The coloured star-shaped icons represent the location where the PR was back at his seat (but still in the sleep inertia recovery time of 15 minutes from waking up). The characteristic scenario points for crew 10 were triggered earlier in the flight due to simulator issues and time pressure. Crew 10 actually aborted after the 2nd cabin call.

4. Failure Condition Management Experiment

4.1 Crew Stats/Quick Facts/Complement the failure condition management scenario

Fact	Crew 1	Crew 2	Crew 3	Crew 4	Crew 5	Crew 7	Crew 8	Crew 9	Crew 10
Same Airline	No	No	Yes	No	Yes	No	No	Yes	No
PR woken up T0+[mm:ss]	10:02	10:20	10:17	10:12	10:23	11:11	13:34	13:23	12:25
CM2 diversion before CM1 return	yes	no	no	no	yes	no	no	no	yes
PR back on seat T0+[mm:ss]	12:02	12:03	13:23	13:51	13:20	14:23	16:35	15:23	15:16
eMCO end briefing performed at T0+[mm:ss]	27:46	-	-	24:51	-	23:45		-	-
PR → PM transition allowed from T0+[mm:ss]	25:02	25:20	25:17	25:12	25:23	26:11	28:34	28:23	27:25
Actual PR → PM transition T0+[mm:ss]	23:28	Hard to say	19:39	24:51	Hard to say	26:59	28:43	23:35	28:25
PR active before recovery time[7] (+ performed actions, if applicable)	yes: EFB, Decision-making	yes: EFB, FMS, Decision-making	yes: EFB Decision-making	no	yes: EFIS CP, EFB, FMS	no	yes: FMS, KCCU	yes: EFB Decision-making	no
CM2 eMCO abortion reason	Amber caution, single chime								
Initial diversion decision CM2	Prestwick	-	-	Shannon (CM1 already in cockpit)	-	-	-	Keflavik	Prestwick (CM 1 already in cockpit)
Initial CM1 diversion decision	-	-	-		Keflavik	-	-	Keflavik	-

[7] According to the procedures the PR was only allowed to observe or to take part in the re-synchronization briefing. For this analysis "observe" was defined as follows:

1. ok to switch onside displays in order to gather information without disturbing or interacting with the "PF alone"
2. Not ok: enter information into aircraft systems or EFB, e.g. FMS modifications, ATC COM, performance calculation on EFB

Final CM1 diversion decision	Shannon	Keflavik	Prest-wick	Shannon	Keflavik	Keflavik	Keflavik	London	London (Prestwick too short)
Company was in-formed	Yes, 37:33	Yes, 31:08	Yes, 21:20	No	Yes	Planned, but never played	Yes	Yes	Yes
Declared PAN PAN	-	-	-	-	-	-	-	-	-
Declared Mayday	Yes, 13:15	Yes, 19:19	Yes, 18:24	Yes, 23:03	Yes, 14:04	Yes, 20:30	Yes, 19:28	Yes, 12:36	Yes, 19:50
LAPA per-formed for	Prest-wick, Shannon	SOS, Keflavik	Prest-wick (2625 m, autbrk med 2286 m)	Shannon (max. rev.)	Keflavik	Keflavik (2660 m, autobrk med 2300 m)	Keflavik	Keflavik, Man-chester, London, ILD > 3,5 km	Prest-wick, London
NITS per-formed	-	-	Yes	Yes	Yes	Yes	Yes	Yes	Yes

4.2 Workload

As for the nominal operation scenario the NASA-TLX raw was separated into two phases exactly corresponding to the phases as defined in Section 2.4.2.

The ratings of the PF for the first phase of the scenario are presented in Figure 14. As expected the highest workload ratings are provided for mental workload. Similar to nominal operations the range of the ratings between the different crews is widespread. The median of the mental workload ratings corresponds to MED = 7.9 (SD = 3.6), which is far higher rating than the median rating of the nominal operation scenario. For the second phase of the failure condition management scenario (see Figure 15) the mental workload rating decreases slightly to 6.8 (SD = 2.7). A decrease in workload could be expected for phase 2, as this is the phase in which the PR re-joins the crew.

Although the PF performed all ECAM actions alone, the rating for his physical demand is low in comparison to the mental workload and only slightly higher than the median of the physical demand during phase 1 of the nominal operation scenario.

The ratings for the performance are high which is comparable to the nominal operation scenario ratings. The frustration is relatively low for both phases.

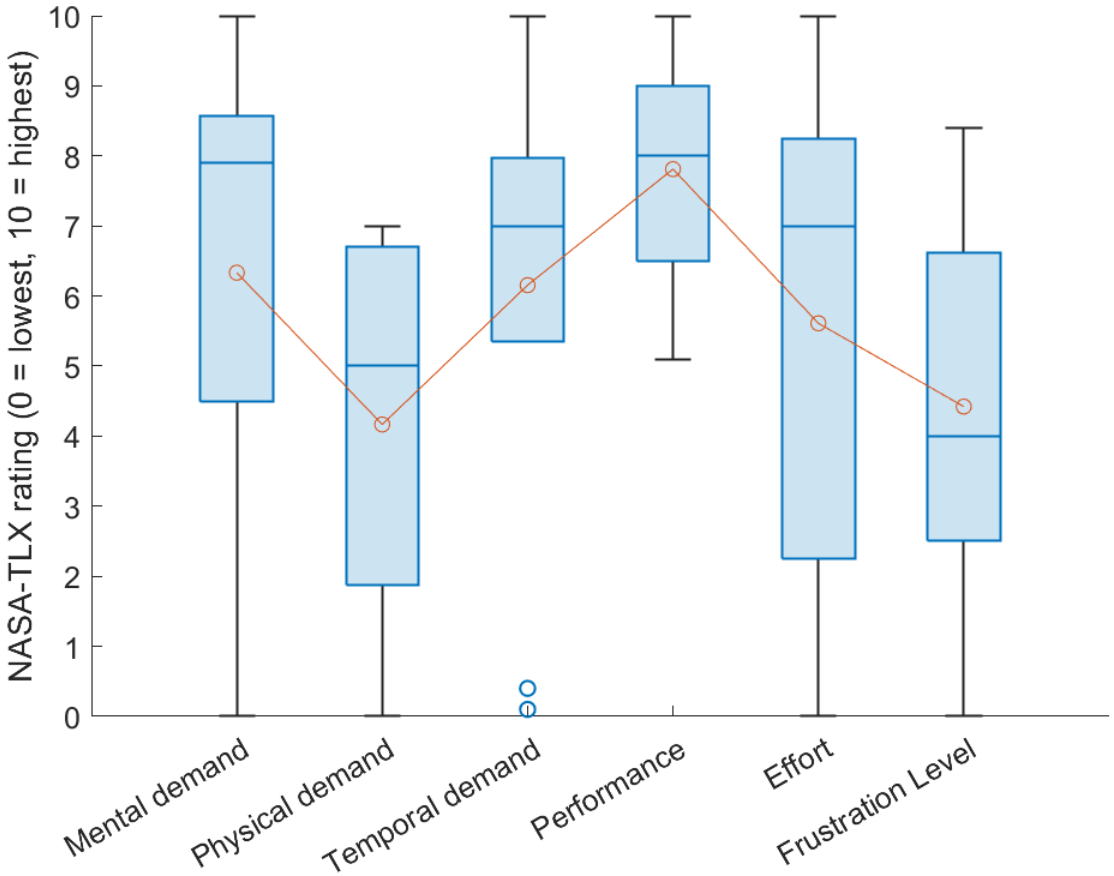


Figure 14 NASA-TLX rating of the PF for 1st Phase of the failure condition management scenario (N=9)

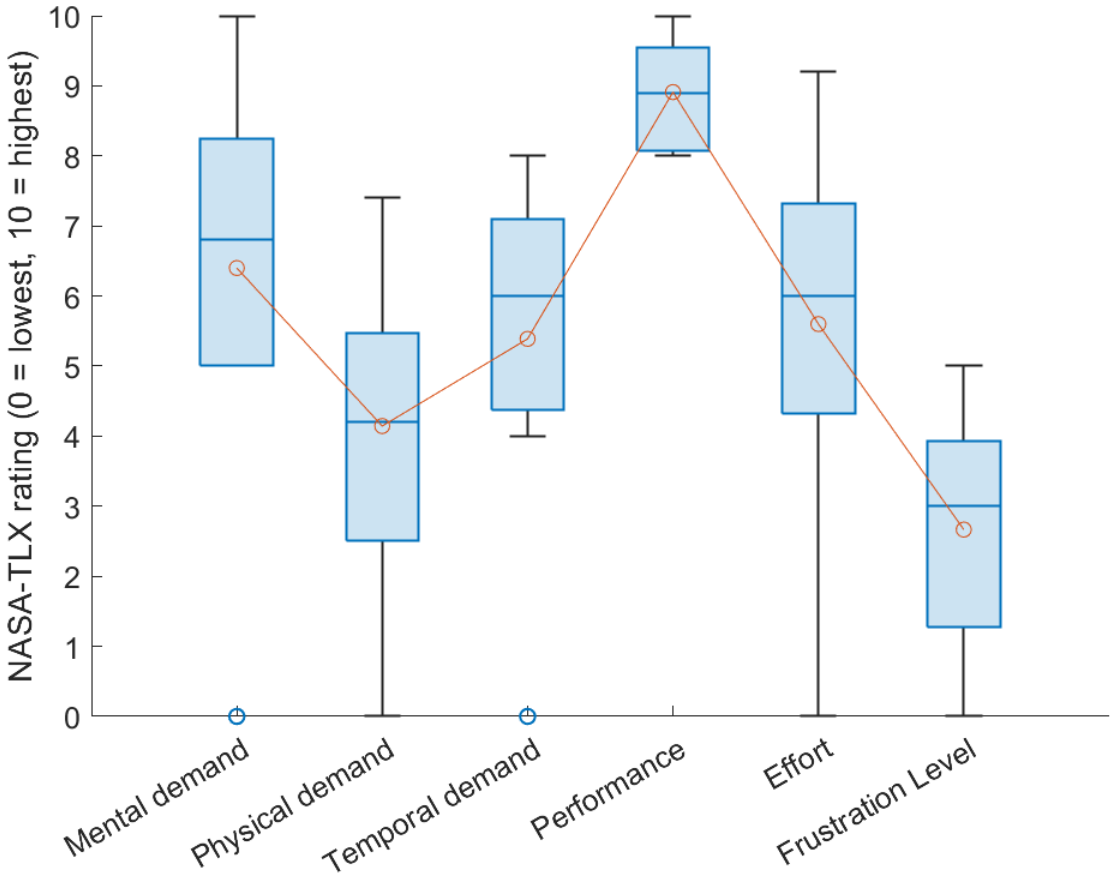


Figure 15 NASA-TLX rating of the PF for 2nd Phase of failure condition management scenario (N=9)

Table 7 presents all PF ratings for both phases of the failure condition management scenario. Extreme cases are highlighted in yellow. Two crews (Crew 1 and Crew 4) reached the maximum level of mental workload, at least during the first phase of the scenario in which the PF had to perform all ECAM actions alone while also starting to re-plan the route, e.g. by calculating the landing performance on the EFB.

On the other end of the extreme cases the PF of Crew 9 rated his workload close to zero for all workload categories.

Crew 7 is the only crew that provided comparable high workload ratings for both LOFT scenarios which is most probably due to their alternate decision in the nominal operation scenario (see Section 3.4).

Table 7 NASA-TLX ratings of each PF for both phases of the failure condition management scenario

	Mental demand		Physical demand		Temporal demand		Performance		Effort		Frustration level	
	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2
PF Crew 1	10	10.0	7.0	5.1	10.0	8.0	5.1	8.0	10.0	8.0	4.0	4.0
PF Crew 2	6.0	8.0	2.5	2.5	7.0	7.0	9.0	10.0	3.0	7.0	4.0	1.5
PF Crew 3	7.0	5.0	7.0	5.0	7.0	5.0	8.0	8.0	6.5	6.0	6.5	5.0
PF Crew 4	10.0	9.0	3.4	7.4	7.0	7.4	6.2	9.4	9.0	9.2	7.0	2.7
PF Crew 5	0	5.0	0	2.5	0.4	4.5	9.0	8.9	0.0	2.3	1.0	0.6
PF Crew 7	8.1	6.8	6.6	6.6	9.4	6.0	8.4	8.1	7.0	5.8	8.4	3.3
PF Crew 8	7.9	6.8	5.0	4.2	7.5	6.0	6.6	8.8	7.0	7.1	5.9	3.9
PF Crew 9	0	0	0	0	0.1	0.0	10.0	10.0	0.0	0	0	0.0
PF Crew 10	8.0	7.0	6.0	4.0	7.0	4.0	8.0	9.0	8.0	5.0	3.0	3.0

The mental workload ratings of the PM result to a median value of MED = 6.8 (SD = 2.6). As expected, the value is lower than that of the PF. However, it is only a small difference. The PMs rate their performance level similar to that of the PF.

Considering the box plots it should be noted that, corresponding to the nominal operation scenario, the variation of the ratings is very high. This is especially the case for the PF ratings during the first phase of the failure condition management scenario. The mental workload ratings range from 0 to 10 with a significant standard deviation.

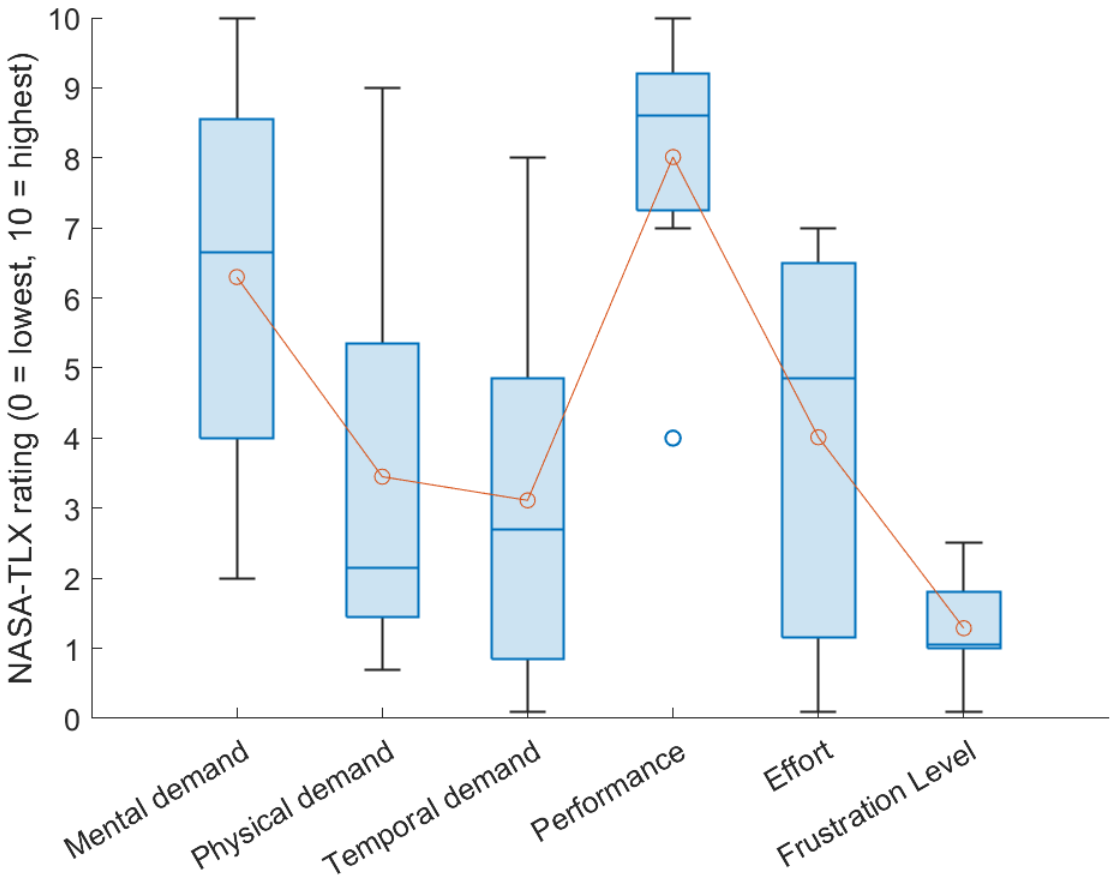


Figure 16 NASA-TLX rating of the PM for 2nd Phase of the failure condition management scenario (N=9)

Table 8 shows the NASA-TLX ratings provided by the PM for the 2nd phase of the failure condition management scenario. Corresponding to the ratings of the PF the ratings of the Crew 1 and Crew 7 PMs' represent the extreme cases in comparison to the other PMs. The lowest ratings in terms of workload are provided by the PM of Crew 8.

It is striking that the ratings of the Crew 9 PM are significantly higher than those of the PF. This was not the case for the nominal operation scenario.

Table 8 NASA-TLX ratings of each PR/PM during the 2nd phase of the failure condition management scenario

	Mental demand	Physical demand	Temporal demand	Performance	Effort	Frustration level
PM Crew 1	10.0	8.0	0.1	10	0.1	0.1
PM Crew 2	2.0	0.7	0.7	9.4	1.3	1.0
PM Crew 3	7.1	5.0	5.7	7.1	8.2	5.1
PM Crew 4	6.0	1.0	2.0	9.0	5.0	1.0
PM Crew 5	6.4	2.7	4.7	8.2	6.0	2.5
PM Crew 7	9.0	9.0	8.0	7.0	7.0	2.1
PM Crew 8	2.0	1.9	1.0	9.0	1.0	1.0
PM Crew 9	6.9	1.9	3.4	7.5	4.7	1.1
PM Crew 10	8.1	2.4	5.0	4.0	7.0	1.5

4.3 Situational Awareness

The 10D-SART ratings provided by the PF during the first and second phase of the failure condition management scenario are presented in Figure 17. The median value of the overall SA score provided by all PFs during phase 1 results to MED = 32.8 with a standard deviation of SD = 9.4. It is thus lower than the overall SA score for the first phase of the nominal operation scenario.

Equivalent to the nominal operation scenario the median of the overall SA score of the PM is higher than the score of the PF (MED = 46.9 with SD = 16.7), although the PM enters the cockpit after the first failure has already been triggered and just before second failure is triggered. However, the relatively high standard deviation shows that there is a significant variation between the crews.

For the PF the overall score slightly increases for the second phase of the failure condition management scenario, once both crew members are back on duty (MED = 37.4 with a relatively small SD of 8.6, see Figure 17). The same tendency can be observed for the nominal operation scenario.

Similar to the nominal operation scenario the PM's SA score is lower for phase 2. During the nominal operation scenario the PM's overall score decreases only slightly, whereas for the failure condition management scenario it is reduced significantly (see Figure 18). The median SA overall score for the PM is MED = 46.9 (SD = 16.7) for phase 1 and MED = 38.6 (SD = 13.2) for phase 2 (N=8). Note that the ratings of the Crew 3 PM had to be excluded, because they failed the validity check.

During the first phase the PM's overall SA score is significantly higher than that of the PF The overall SA score is even higher for the PM during phase 1. This indicates that the PM seems to quickly regain SA after his entry into the cockpit while observing the PF's actions and analyzing the ECAM SD as well as the Status Page. This topic was also addressed during the de-briefing (see Section 5.1).

The lowest overall situational awareness score for phase 1 was 15.0 provided by the PF of Crew 4. The same PF rated the second phase as 34.4. The PM of Crew 4 rated the first phase of the scenario (time between entering the cockpit after the abortion up to the point in time in which he was allowed to take over the role as PM) as 33.7 and the second phase as 38.2.

The lowest overall situational awareness score provided by a PM for the first phase was 14.0. This was the PM of Crew 7 which was the only crew that had high workload ratings for both LOFTs. In the second phase of the scenario the PM re-gained an overall SA score of 36.0.

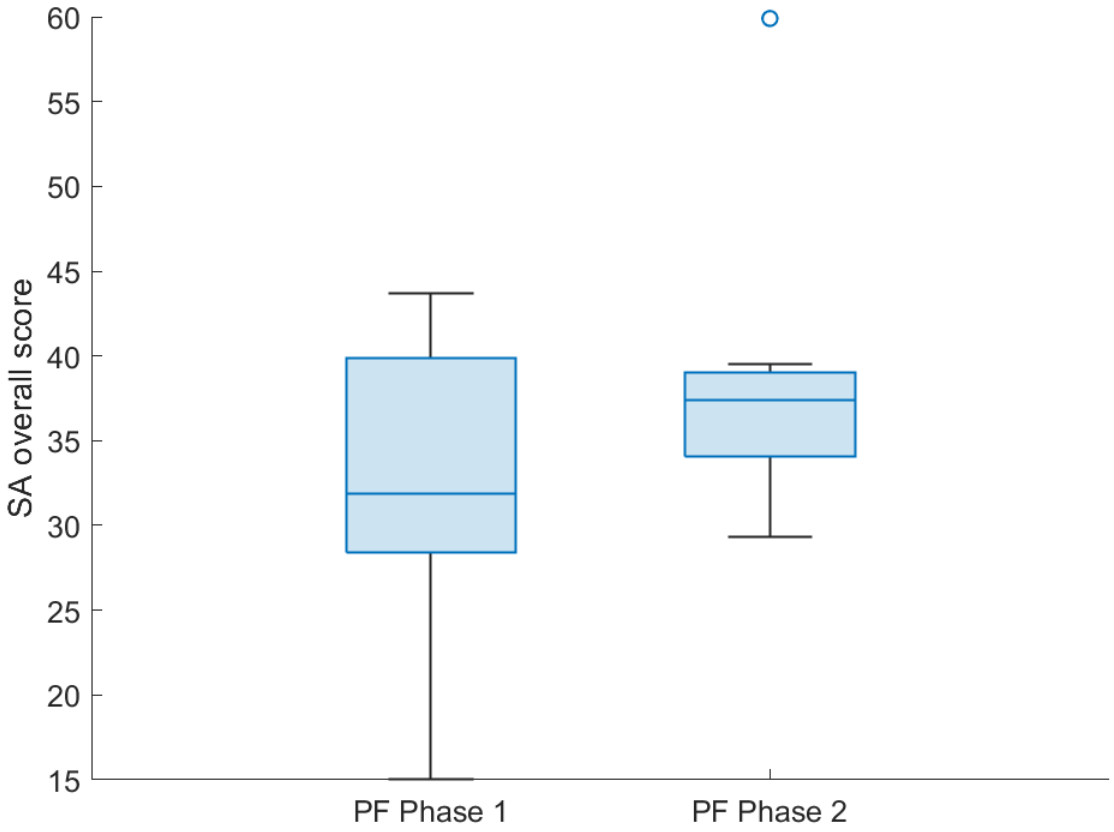


Figure 17 10D-SART Overall SA score for the PF (N=9, lowest score possible = 4, highest score possible = 70)

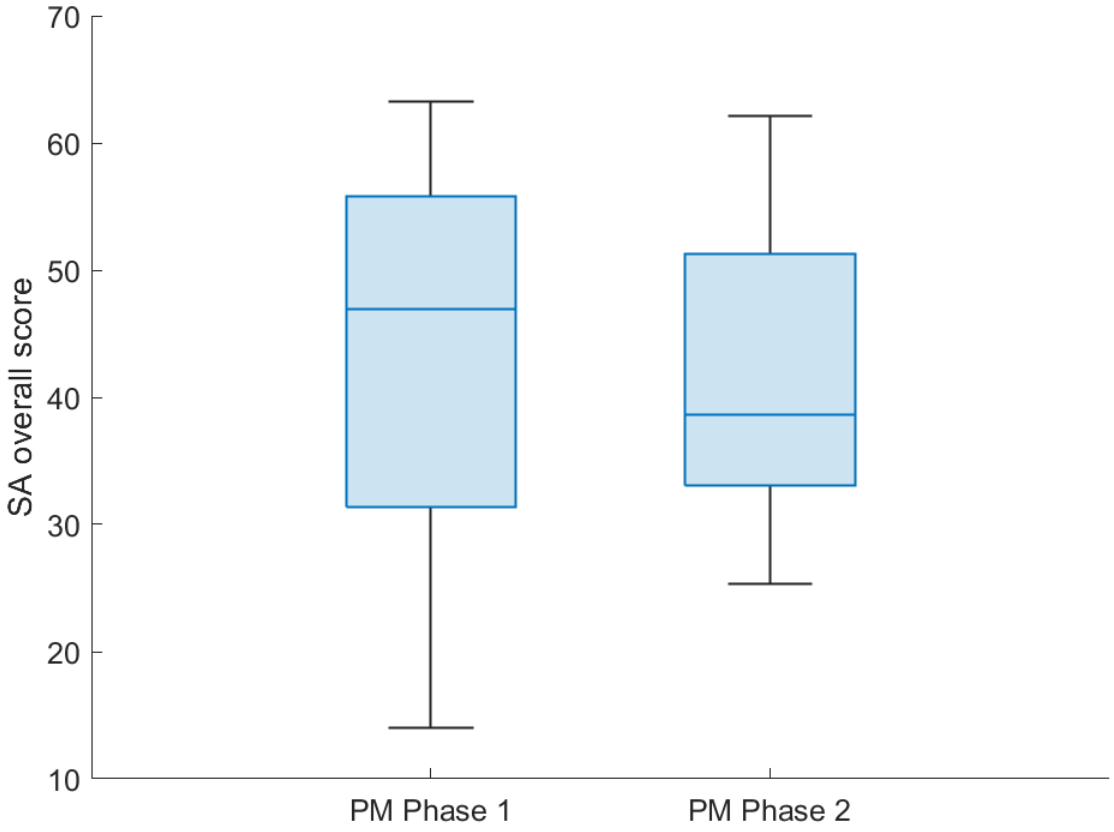


Figure 18 10D-SART Overall SA score for the PM (N=8, lowest score possible = 4, highest score possible = 70)

4.4 eMCO Abortion Decision – Failure Condition Management Scenario

Figure 19 shows the track logs of all crews for the failure condition management scenario in relation to the flight planned route and characteristic scenario points. The design goal of this scenario was to have a clear abortion criterion for the eMCO segment in order to create a situation that focusses more on the effect of eMCO on the failure management in abnormal situations. It is therefore not surprising, that all PFs aborted the eMCO segment at the occurrence of the first amber caution/single chime, which was in line with the previously briefed abortion criteria. The PFs woke the PRs in a time span ranging from one to three minutes.

The three characteristic scenario points where (see Section 2.2.2):

1. The eMCO entry point, that also acted as time reference T_0, where the simulated flight began. The aircraft was released from flight freeze and the timer was started after the PR left his or her seat.
2. The point at which the first hydraulic system was failed at $T_0 + 10\ min$. The Hydraulic Yellow System Temperature High is an amber caution with a single chime, which fulfilled the eMCO abortion criteria. Note that the abortion criteria where chosen for this experiment with the focus on being able to study the effects of an eMCO abortion and that the definition of abortion criteria for real life operations would require further research.
3. The point at which the second hydraulic system was failed. The Hydraulic Green Reservoir Low Level failure was triggered at the moment the returning PR was back on his or her seat. The location of this point differs slightly between the crews as also the time it took the PRs to return to the cockpit differed

slightly (around 30 seconds between slowest and quickest return). However, due to the general location of the aircraft this variation is negligible for the analysis as even a few minutes flight time are not likely to change the diversion decision.

Looking at Figure 19, the participating crews' decisions were indifferent with four out of nine crews continuing in the general flight direction towards Iceland while the other five crews turned back to Europe. It is interesting to note that in three out of nine runs the PF performed the initial decision to divert before the PR had returned to his seat. In only one case the crew revised the initial decision of the PF and diverted to a different airport due to a longer available landing distance and better weather.

4.5 Diversion Decision – Failure Condition Management Scenario

The scenario was designed so that the dual hydraulic failure emerged shortly after the equidistant time point between Scotland and Iceland. The waiting time until the end of the sleep inertia recovery time places the continuation to Keflavik as alternate in favour. Weather-wise neither Keflavik nor the airports of Great Britain have ideal conditions. Keflavik is more isolated and the only airport coming into question in Iceland while Great Britain offers more alternates, albeit these are further away at the time a crew decision can be made again according to the eMCO procedure provided to the crews.

The simulated dual hydraulic failure resulted in an approximate required In-Flight-Landing-Distance (ILD) ranging between 2500 m and 3800 m across the available airports. The associated system failure to be entered in the EFB (the built-in EFB was not linked to the aircraft's sensed failures) was HYD G+Y SYS PRESS – LO (INR FLAPS < 1). The factors with the highest influence on the ILD were the reverser setting, the autobrake setting, and the runway condition. Three crews decided to calculate without reversers, thereby reducing the number of possible alternates to one, with London being the only airport with an available landing distance in excess of 3400 m. However, diverting to London would result in an additional hour of flight time, which, when viewed in conjunction with the dual hydraulic failure, also poses an additional risk by itself. Given the operational integrity of the reversers even in the context of the simulated system failure, the risk of calculating the landing performance with reversers (i.e. not using the reverser margin as additional stop margin) has to be weighed against the increased risk of a prolonged flight due to not using them in this case of a dual failure. Whether the above crews forgot to select the reversers for landing or used the increased resulting ILD could not be resolved in the debriefing.

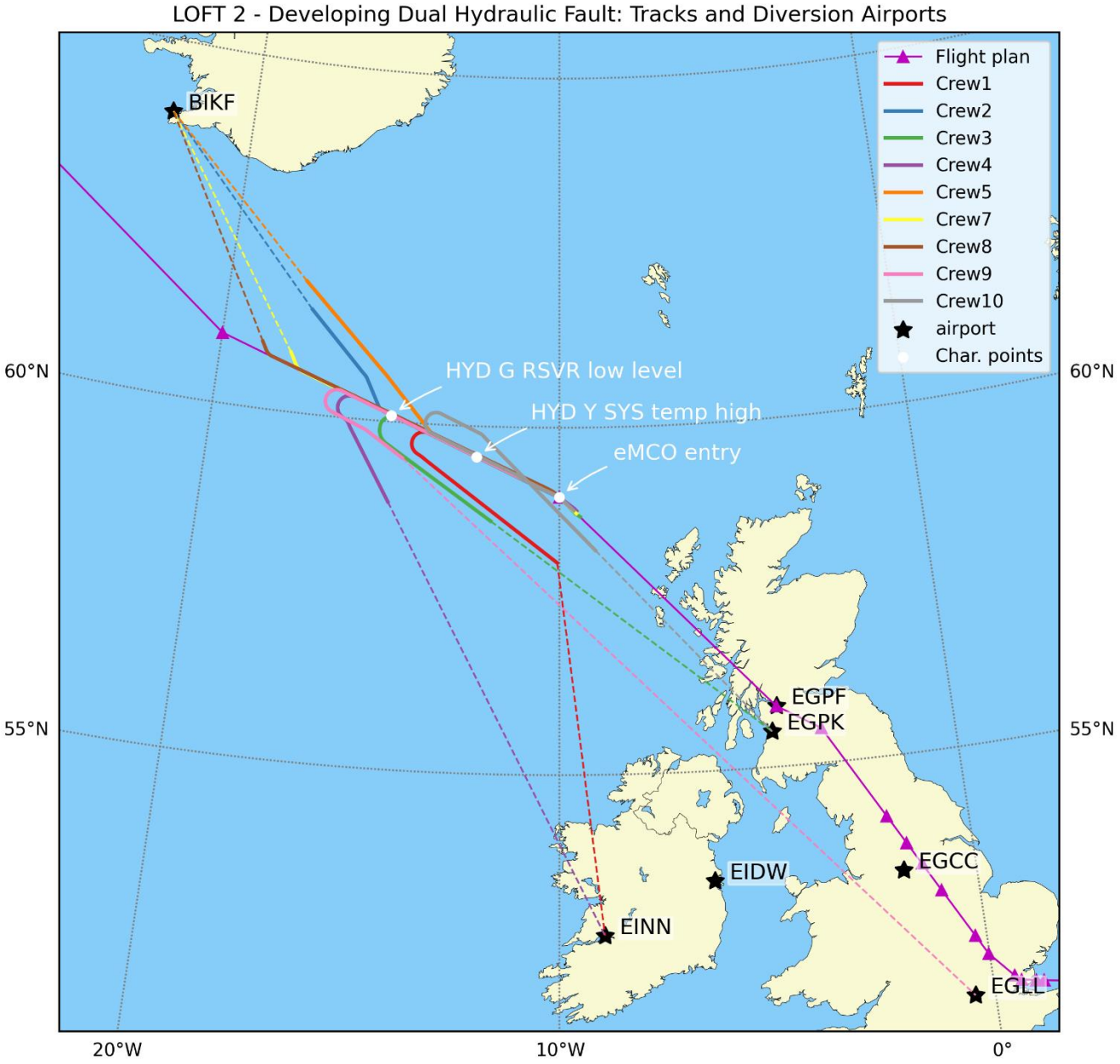


Figure 19: The failure condition management scenario track log showing the planned route, the crews' tracks and chosen diversion airports (dotted lines). The beginning of each dotted line marks the end of each experiment. The experiment was ended after the final diversion decision had been made by the crew. Because all PFs aborted the eMCO segment as previously briefed after the first appearance of a failure, the eMCO segment was aborted at the same position for all crews.

5. Experimental Results Summary

In D-2.1 [2] and D-3 [3] potential hazards arising from eMCO during nominal and failure management conditions were described in detail. This section shall discuss the results described in Section 3 and Section 4 with respect to those potential hazards and describe the associated potential mitigations mainly derived from experimental observation and the de-briefing sessions. However, it must be noted that the simulator study is only able to provide qualitative insight into potential problems and hazards. Due to the limited number of participants it is not meant to identify and quantify the associated risks.

5.1 Situational Awareness and Shared Mental Model

The 10D-SART ratings show that the SA overall score of the PR is not lower than the score of the PF. This indicates that the PR was able to regain SA quickly after his return to the cockpit. However, it should be noted that the variation of the ratings between the different crews is significantly high making a general conclusion difficult.

In addition to the SA of the PR, who misses the history of a particular event, the SA of the PF is also affected due to the missing exchange with the PM who might have different views and additional knowledge.

This is supported by the statement of one PF during the de-briefing which pointed out that he missed the exchange of mental models with the captain, who has a different view of the situation.

A good example is the lack of situational awareness of the Crew 5 PF leading to a non-optimal alternate airport decision. The lack of situational awareness is also supported by the low 10D-SART rating provided by this participant. Before the PR returned to the cockpit the PF decided to divert to Lisbon, although the landing there would have to be performed with a significant tailwind component due to the closure of the opposite runway direction published in a NOTAM. In a later weather update Lisbon airport was even closed due to severe wind conditions. Only during the eMCO end briefing performed after the sleep-inertia recovery phase of the PR the PF recognized his error, as CM1 the PR/PM then moderated the decision-making process which led to adaptation of the alternate airport.

This shows that the re-synchronization briefing during the sleep-inertia recovery phase of the PR and the eMCO end briefing after the recovery phase are essential mitigations that help to build a shared mental model, to improve threat management und to exchange different views on the situation after the abortion of the eMCO segment.

Concerning the system status awareness during the failure condition management scenario, the PMs were asked during the de-briefing how easy it was for them to regain situational awareness. Most PMs stated that using the Recall (RCL)- and Status (STS)-buttons to analyse the failures and the resulting consequences on the System Displays (SD) and the Status Page made it easy for them to get back into the loop. Moreover, some PFs pointed out that it was easy to brief the PMs using the RCL- and STS buttons, already representing important mitigations in terms of SA during eMCO operation. As one CM1 stated during the de-briefing: "*There was not delay when I got back into the cockpit and saw the information*".

However, some PRs mentioned that, compared to current operations, "it was different to get into the situation, when you did not see/hear the first master caution" and that they saw the situation, but that they did not see the history of what CM2 had done before.

One Captain proposed that the PFD could present a box with a summary of the situation on the PR's side.

5.2 Failure Diagnosis Capabilities

Failure diagnosis during the failure condition management scenario did not seem to be an issue for the crews. Especially, the ECAM SD pages as well as the Status Page helped the crews to understand the situation. This is also indicated by the statement of the crews during the de-briefing (see Section 5.1).

Based on the video recordings it could be observed that the PF discussed the failure and resulting consequences loudly to increase his awareness similar to the so-called “rubber duck debugging” used in information technology [8]. After the abortion of the eMCO segment a very detailed discussion of the failure development enhanced the failure diagnosis capabilities of the crew. In most cases the crew used the ECAM SD and the Status Page to perform the re-synchronization briefing in the failure condition management scenario. For some crews the PR performed the re-synchronization by himself using the ECAM, in other cases the PF initiated the briefing and described the situation as well as the history of the failure also using the ECAM pages.

The latter approach bears the advantage that the PR learns more about the view of the PF on the situation and on the facts that the PF considered for his decision-making approach. On the other hand, the PR might be biased by the PF’s diagnostic results.

The failure condition management scenario was designed in such a way that the first failure led to a loss of the yellow hydraulic system that could potentially be reset for landing and approach. This failure was triggered during the eMCO segment. The knowledge that some of the landing performance limitations resulting from the loss of the yellow hydraulic system could be reset was essential for the decision-making of the crew. Every crew recognized and considered this factor. Again, also for this specific point, the Status Page of the ECAM was an essential mitigation in terms of failure diagnosis and decision-making, as the possibility to reset the system could be recalled by the PF or the PR after the eMCO abortion. As one CM1 said during the de-briefing: “

It has to be noted that for the failure condition management scenario the system failure was automatically detected by the aircraft systems thus leading to a clear diagnosis on the ECAM.

During the de-briefing some crews pointed out that diffuse failure scenarios might be more difficult in terms of failure diagnosis. Several crews pointed out that unreliable airspeed might be a critical scenario. However, those crews also stated that an unreliable airspeed case is very unlikely on the Airbus A350-900 due to its redundancies in the speed calculation.

5.3 Decision-Making and Problem-Solving

With respect to decision-making different approaches could be observed during the experiments. Three out of the nine PFs performed the diversion decision alone and changed the route on their own before they aborted the eMCO segment. During the de-briefing the PRs, being always CM1 during the experiment, stated that they would always want to be involved in the diversion decision as they take the full responsibility for the flight. However, they stated that even in current operations some of the Senior First Officers tend to decide for a diversion without waking up the Captain.

In eight of the 18 scenario runs the Senior First Officers performed a diversion without waiting for the Captain having returned to the cockpit and waiting for the 15 minutes sleep inertia recovery time defined by the SOPs. For four crews of the eight crews mentioned before the first decision for an alternate airport made by the PF alone was subsequently corrected after the CM1 had been back on duty. Some of the PFs delayed the decision-making process until the PR was back in the cockpit (some of them also considering the 15 min for sleep-inertia recovery; see also Section 5.5) instead to making the decision to divert on their own. For example, during the failure condition management scenario the PF of Crew 8 continued the flight inbound Iceland in order to wait for the CM1 to become PM again, thereby limited the distance to the other available options on the British Islands. Especially, in case of the double hydraulic failure, for which a LAND ASAP in red was triggered, this delay is questionable. But also, with respect to the medical emergency time might be a critical factor.

Basically, this results in a conflict for the PF who has to decide between making the decision on his own or to delay the decision until the PR is back on duty. This is especially critical, if the PR is the CM1.

One SFO gave a good summary on that issue during the de-briefing, also with respect to the sleep-inertia recovery time of 15 min that was part of the eMCO procedure applied in the experiment: “The diversion decision is not a SFO decision and waiting for 15 minutes is a long time. I would like to be allowed to initiate the decision. It was really difficult to wait for the FORDEC to start. The eMCO procedure feels artificial.” Other pilots stated during the de-briefing that the main difference between current operations and eMCO regarding the failure condition management scenario is that the decision would have been made earlier. One pilot remarked: “Decision-making was affected and there was no double-checking as when together. This retards the decision-making and is a bad thing for safety.”

One mitigation for the issue of missing double-checks during the decision-making process is a detailed re-synchronization and/or eMCO end briefing and, especially, a repetition of the decision-making process including a possible revision of the decision once the PR is back on duty. This should include the application of a structured decision-making scheme, such as FORDEC or DECIDE which was not always performed by all crews during the experiment.

During the experimental runs it could be observed that some PFs described the course of actions that led to the situation as well as their thoughts and considerations that led to their decision, if they already made one. This procedure proved to be very effective in creating a shared mental model und pick up the decision-making process in the conventional crew composition as PF and PM.

Especially, in the role distribution chosen for the Task 2 and Task 3 experiments with the CM1 being PR (as well as PIC), it is essential to perform the decision-making process again after the abortion of the eMCO segment, as the CM1 has the sole responsibility for the safety of the flight. This was pointed out by several Captains during the de-briefing.

With respect to the diversion decision-making process several pilots recommended a potential system mitigation that supports the pilot in selecting a suitable alternate airport. They proposed to have an automated system that recommends an alternate airport by considering the weather and the landing performance in order to support the PF with a second opinion and further options. For example, one CM1 said: “Yes, I would call him earlier than the third pilot when flying with three pilots. FORDEC doesn't work as good when you are alone. Support systems that propose alternate airports could help in getting a different input than one’s own."

In addition to that, one pilot stated there should be an EFB link to the aircraft warning system to automatically consider system failure. Such a functionality is basically already available but is not implemented for all airlines and aircraft types.

The decision for possible system reconfigurations and for alternate airports in case of a required diversion is supplemented by the decision to abort the eMCO segment solely made by the PF. The timing of the eMCO abortion might have a big impact on the safety of flight as well as on the operations (see Sections 5.4 and 5.6). The nominal operation scenario was designed so that the timing of the abortion decision was diffuse which had a significant impact on workload as well as on the timing of the diversion decision. According to all pilots the timing of the abortion is highly influenced by the experience of the PF. During the de-briefing the pilots stated that the checklist criteria for an abortion decision should be defined as clear as possible to minimize the probability of diffuse situations. In the failure condition management scenario, most pilots adhered to the provided eMCO procedure and called the PR immediately according to the “single chime/master caution” abortion criterion. However, one PF began to work on the ECAM actions before calling the PR. During the de-briefing he explained: "The initial actions on the ECAM are of higher importance than waking of the PR a few seconds earlier. After the initial actions I then called the PR back."

5.4 Workload Management

As expected, for both scenarios mental workload was predominant. Physical workload was not a factor. Based on the NASA-TLX raw ratings and on the observations during the experiment even mental workload was not an issue for the majority of the crews during the nominal operation scenario. Mental workload was rated with a median of MED = 4.0. However, due the diffuse abortion decision, the amount of workload was significantly affected by the timing of the abortion and, vice versa, often the abortion was triggered by the expectation of a high workload situation as part of the workload management. For example, the PF of Crew 1 aborted the eMCO segment very early, although he rated the mental workload during the first phase of the scenario to 2 which then increased in the second phase to 8. During the de-briefing he pointed out that he used the high workload abortion criterion listed on the eMCO checklist as reason to abort eMCO in the nominal operation scenario.

Basically, there are significant differences in the workload ratings between the crews. This applies to both scenarios. In the nominal operation scenario, in contrast to many crews, the PF of Crew 7 did not wait for the medical emergency call from the cabin until he woke up the PR. Consequently, the PR was already back at the controls when the emergency call came. However, the mental workload rating of the PF was the highest of all for pilots in the nominal operation scenario due to the crew's decision to land in close vicinity of the aircraft position without fully considering the non-optimal health care at that alternate. The crew even sent a fictional message to the company with a request for medical information, but afterwards never asked again. Due to the urgency of the landing preparations and route changes the workload ratings of the PF make it questionable whether he could have handled that workload alone during the sleep-inertia recovery phase.

Generally, also as expected, the workload in the failure condition management scenario was significantly higher based on the median of the NASA-TLX ratings. However, again there is a significant variation between the crews.

In contrast to the nominal operation scenario the timing of the abortion was completely clear leading to a comparable amount of workload for all crews. All PFs performed the ECAM actions alone. This is supported by the higher physical workload ratings in comparison to the nominal operation scenario. The main increase in workload can be, however, observed for the median of the mental workload of the PF trying to understand the situation and the consequences resulting from the system failures. Physical workload does not seem to be an issue. The mental workload of the PF is only slightly reduced in the second phase of the scenario and the median of the PM's mental workload ratings is comparable

During the de-briefing pilots stated that workload was not an issue for them while performing the ECAM actions due the well-structured checklists. At the same time, they pointed out that this scenario would have been completely different in terms of workload on other aircraft that do not provide a similar robustness in automation. This is supported by the statements made by different pilots during the de-briefing:

- CM2: "The DUAL HYD was not a big surprise. At the beginning, workload was not very high. I was not used to the eMCO procedure. It would be really hard if you had to fly manually."
- CM1: "Yes, with the A55O it's easy, with the A330/A340 it would have been really hard."
- CM1: "It would have been totally different in a less automated plane like the A320. In the 350 it's easy."

The availability of the autoflight system after the double hydraulic was a prerequisite for the PF to able to handle all actions until the end of the sleep-inertia recovery phase alone. Robust automation also in case of critical system failure is a key mitigation for workload management during eMCO.

However, it must be noted that even with the availability of the autopilot the PF of Crew 7 rated the mental workload with 8.1 and the PFs of Crew 1 and 4 with 10 indicating that they were close to their workload limit.

On the other hand, there is the extreme case of the Crew 9 PF who provided ratings close to zero for all workload categories in both scenarios.

When evaluating workload for the first phase of the scenario, in which the PF was supposed to perform all actions alone, it should be noted that for the nominal operation scenario five and for the failure condition management scenario six out of the nine PMs did not fully stick to the eMCO procedure not allowing any action or participation in the decision-making for 15 min. A significant number of PR/PMs started to support the PF with an EFB performance calculations reducing the PF's workload in that phase (see Section 5.5).

During the de-briefing the abortion decision was also mentioned in conjunction with workload. One pilot stated that in real operations, considering peer pressure (see Section 5.6), the "*consideration of pros and cons of the abortion would probably increase my workload*".

To conclude, workload management is significantly affected by the timing and the unambiguousness of the abortion decision. The pilots clearly stated during the de-briefing that the criteria to abort eMCO should be as clear and unambiguous as possible, to reduce workload during the abortion decision as well as for the upcoming tasks, such as the potential decision for an alternate airport.

5.5 Adherence to Procedures and Complacency

During the experiments it could be observed that all PFs flying alone stuck to the procedures. Especially while performing the ECAM actions in the failure condition management scenario the PFs spoke out loudly the checklist items and carefully checked the hydraulic system before pressing the correspond overhead switch to switch-off the pumps. Also, the Status Page was read aloud by the PF alone after having completed the ECAM actions.

The adherence to the eMCO procedure with respect to the sleep inertia recovery time was difficult for the PR. During the briefing both crew members were told that the PR is not allowed to take any action or to participate in the decision-making process for 15 min after the activation of the flight crew rest pushbutton by the PF. According to the eMCO procedure the PR was only allowed to take part in the re-synchronization briefing initiated by the PF or to re-synchronize himself, e.g. by observing the System Display or the Status Page in the ECAM. During the experimental design it was, however, decided that the experimental lead would not interfere in case the PR does not comply with the procedure.

Most of the PRs could not wait for the complete 15 min to support the PF with taking certain actions in the cockpit. During the de-briefing several pilots pointed out that such a procedure with a fixed recovery time feels artificial and is not feasible in real operations. Especially, the fact that the PR was CM1 made it difficult just to observe the actions of the PF. Some pilots stated during the de-briefing that they, in extreme cases with a high level of adrenaline, would be awake in a shorter time and that 15 min recovery time are too long. On the other hand, the PFs pointed out the advantage that the returning PR did not disturb or distract him while he was working on the ECAM actions. Three pilots stated that the procedure should not define a fixed sleep inertia recovery time. They pointed out that the degree of sleep inertia should be assessed by the PR him- or herself and that is significantly impacted by his or her adrenaline level.

Finally, it could be observer that some crews did not call out a decision-making scheme and performed the decision-making in a rather unstructured process. As mitigation a stricter adherence to a structured decision-making scheme, such as FORDEC, could be very helpful to make the final decision after the abortion of the eMCO segment. This gives the PF the opportunity to describe the history of the events, his thoughts, ideas and considerations and, finally, it helps to re-discuss options and associated risks and benefits leading to potential revision of the PF's decision.

5.6 Peer Pressure

During the briefing as well as during the de-briefing almost all pilots mentioned peer pressure in case of an eMCO abortion as hazard that can obviously not be emulated in a simulator study. This pressure might arise

from the “risk” that the PR is woken up without a good reason, e.g. a too early abortion in case of medical issue. Depending on the hierarchy in the cockpit this is even worse, if the PR is CM1, which was often stated by the FOs in the de-briefing.

Another reason for concern among the participants mentioned in the de-briefing was that, if future eMCO operations one day grant benefits on Flight Time Limitations (FTL) and these are not utilised to only reduce pilot fatigue but also to extend flights, situations might arise in which the flight depends on these eMCO segments. The participants feared, that in these cases the decision to abort the eMCO segment is also seals the decision to abort the whole flight, because the destination cannot be reached inside the FTL anymore. This even aggravates the situation explained in the above paragraph and was also mentioned as a reason of concern during the project’s pilot workshops.

5.7 Threat and Error Management

Almost all of the participants raised concerns about missing monitoring and thereby missing cross checks during the eMCO phase. During normal operations for example misheard ATC communications were mentioned, e.g., an ATC heading instruction was erroneously set in the auto flight systems but read back correct. Although this is addressed by modern CPDLC (Controller Pilot Data Link Communications) , this specific example remains a problem for areas without CPDLC coverage. In case of immediate abnormal or emergency situations the application of irreversible actions remains another problem to solved regarding the current cockpit and eMCO philosophy. As means of mitigation, either these actions are delayed by the sleep inertia recovery time until the crew is back to normal crew operations (NCO), the cross check is performed via an automated system that checks or grants the action first, or the system automation is increased, so that these actions are also automated. An example of this could be the automatic engine shutdown and fire extinction in case of an engine fire.

Two observations could be made during the experiments. The first observation concerning monitoring could be made during one experiment run where the PF during the eMCO segment accidentally deactivated the autopilot seconds after the first hydraulic failure was triggered. During the de-briefing the PF linked this pilot error to his high workload. Only because this triggered a second master caution the PF noticed his error and promptly reengaged the autopilot. However, depending on the situation, it could be imagined that there is a possibility that flying with deactivated autopilot remains unnoticed for a short duration.

The second observation could be made during the performance of the landing distance calculation for the Failure Condition Management scenario with the dual hydraulic fault. All crews considered the failure for the calculation (the EFB configuration required a manual insertion of the failure), but three crews did not consider the usage of the reversers. This, in combination with the respective runway conditions, led to required landing distances in excess of 3400 m, which in turn rendered all nearby airports unusable. To solve the situation, these crews had to fly to London which was the nearest airport with available landing distances of more than 3400 m, which extended the flight time with by approximately one hour. It remains a point of discussion whether it is safer to increase flight time with a dual failure and associated LAND ASAP for using the gained thrust reverser margin on a longer runway, or to land sooner with the planned application of the reversers (that were not affected by the system failure). However, the *As Soon As Possible* is linked to the thrust reverser usage decision, thereby possibly creating a slack in the definition.

It could not be verified from the experiments with certainty whether the crews that did not use the reverser for their landing distance calculation did this on purpose or if the non-selection could be related to human factors. However, workload and error potential could be reduced by an EFB configuration in which sensed aircraft system failures are transferred to the EFB automatically, which was not the case in the simulator that was used for the trials.

6. Conclusion

In September 2024 simulator experiments assessing the effects of eMCO on various characteristic human factor values were performed. These experiments addressed one nominal operation scenario as well as one abnormal operations scenario (called failure management scenario according to the naming of the Task 3 work package). This report contains the evaluation of both experiments performed in a certified Airbus A350-900 FFS. As means of measurement questionnaires, video recording and a structured de-briefing were used to get an insight into potential hazards related to eMCO and to collect potential mitigation means in the fields of procedures, training and aircraft systems. It should be noted that, due to the limited number of participants, a quantitative risk assessment is out of the scope of this report.

6.1 Experimental Findings and Potential Mitigations

Workload measurements using NASA-TLX raw showed the medical emergency was not an issue for most crews in terms of mental workload. However, due to the vicinity of and distance to the selected alternate airport and thus also due to the timing of the decision-making, it could be observed that workload can quickly rise. This changes for the failure condition management scenario, where the double hydraulic failure created significant mental workload, especially for the PF. Three PFs had a mental workload rating between 8 and 10 showing that they were approaching or reaching the workload limits.

In terms of SA the 10D-SART ratings are not completely meaningful. As a general trend for the PF, it can be observed the SA overall score of the PF increases in the second phase of both scenarios when the crew is back to NCO. However, there are only slight differences. An impact of the resting period on the SA of the PR cannot be identified. A very interesting SA observation is based on the video recordings showing the Crew 7 PF's wrong alternate airport decision. At that point the PF alone was not aware of the closure of the airport. This corresponds to his SART rating being the lowest of all PFs during the nominal operation scenario. This error was identified and corrected by the PF himself during the eMCO end transition briefing when he explained the previous events and his decision to the PM.

The decision-making processes of the crews were evaluated using the video recordings. Interesting observations could be made in terms of the landing distance calculation. Three of the nine crews did not consider to use the thrust reverser for the landing performance calculation and also accepted an extension of the remaining flight time by roughly one hour, although the ECAM showed a LAND ASAP in red. The reason for their decision is not completely clear. During the experiment it could be, however, observed that a structured decision-making process was not always followed by the crews after they were back to normal operation.

The abortion decision is a decision specific to the eMCO operation and it might cause specific hazards, as the timing of the abortion has a significant impact on the current workload, the upcoming workload and the decision made by the crew. The abortion decision might be influenced by peer pressure due to the possible operational consequences resulting from an abortion.

Section 5 describes all potential hazards and mitigations observed during the experiment or collected in the debriefing that was performed with each crew. Based on those contents the following table provides a short overview of potential mitigations, of which some of them already exist in modern passenger aircraft (see Table 9).

Table 9 Potential mitigations identified based on the Experiments and the de-briefing contents

Field	Mitigation	Description
Procedures	Transition briefing	The transition briefing at the beginning of the eMCO should include a detailed discussion on potential alternates and related weather during the eMCO segment. The pre-preparation of eMCO diversion airport is an effective mitigation in terms of decision-making. This could be a prerequisite for allowing to activate a diversion alone even if the CM1 is PR. During this briefing specific abortion criteria could be agreed upon. The transition briefing after the eMCO abortion increases the shared mental model and situational awareness.
	Clear abortion criteria	The abortion criteria should be as clear and unambiguous as possible to reduce peer pressure and workload.
	Structured decision-making process after eMCO abortion	The structured decision-making process after the eMCO abortion should be performed between PF and PM (or CM1 and CM2). This approach questions the decision of the PF, possibly made during the eMCO segment, and gives the PM the opportunity to understand the considerations of the PF. Especially, if the CM1 was PR, this is essential, as he has the responsibility for the safety of the flight.
	Reading checklist and status aloud	Some PFs read the checklist and the aircraft status aloud when working alone on the ECAM actions increasing the adherence to procedures and their SA.
	"Rubber duck" approach	By explaining himself/herself the problem aloud (or a rubber duck sitting on the glareshield), the event and the decision, SA and Threat and Error Management (TEM) are improved [8]. As part of the transition briefing this approach helps also to create a shared-mental model after the eMCO abortion.
	Test eMCO in defined airspace	During the de-briefing one pilot remarked that it could be a good test to introduce eMCO only in a defined airspace that is covered by CPDLC and with only professional traffic, e.g. North Atlantic Tracks (NAT).
Training	Specific eMCO training	LOFT scenarios dealing with eMCO and the unplanned eMCO abortion, such as applied in this experiment, should be part of the regular training.
Systems	Alternate airport decision support system	An alternate airport decision support system could automatically assess the weather at potential alternate airports and evaluate the landing performance at those airports by considering the wind situation, runway surface condition and possible system failures. This would serve as a "second" opinion during the decision-making process of the PF alone and would also support the decision-making process, when the crew is back to NCO.
	Exchange of aircraft system failure information with the EFB	The integrated EFB used for the experiments did not receive information on system failures. The crews had to enter the failure manually to perform the landing performance calculation and thus the diversion decision-making process. This mitigation was mentioned during the de-briefing to decrease workload and reduce the likelihood of crew errors.
	Robust autoflight system	The NASA-TLX ratings show that some PFs already had to handle high mental workload situation although the autoflight system was fully

Systems		available. Based on those rating and the information provided in the de-briefing the availability of the AP was essential for safety. An autoflight system robust against system failures but also against exceedance of the envelope is an essential means of mitigation.
	Sophisticated warning system and status page	The design of the ECAM SD and the Status Page was an important means of mitigation to support the re-synchronization of the PR and the generation of a shared mental mode.
	Automated cross-checks	All PFs in the failure condition management scenario performed the ECAM actions correctly, e.g. all PFs switched off the correct pumps. During the briefing the crews mentioned that they missed the cross-checks by the second crew member. This applies also crew actions during normal operation. Automated cross-checks performed by the system, especially before irreversible actions are taken, would mitigate the risk of crew errors.

6.2 Limitations and Future Work

This section discusses the main limitations of the conducted experiments and open points for future research work. The most essential limitation of the experiments is the limited number of crews that participated in the simulator study making statistically relevant statements impossible.

Considering the numerous design possibilities when generating normal and abnormal operation simulator scenarios the selection and the design of the scenarios is always debatable. For that reason the crews were asked during the de-briefing what kind of scenario elements they would generally consider as most critical. In terms of workload resulting from system failures all pilots stated that the Airbus A350 is not comparable to any other Airbus aircraft. According to the pilots, critical system issues such as a fuel leak or unreliable airspeed either generate far less workload or their occurrence is very unlikely.

However, some crews proposed more diffuse scenarios for future experiments. In addition to that they suggested a higher time criticality to increase workload either by a greater proximity to suitable diversion airports or by a critical amount of available fuel on board.

For future studies realistic eMCO segments, including a real rest of the PR, should be considered. It is expected that the duration for which the PR is taken out of the loop has a significant impact on the crew's shared SA after the abortion of the eMCO segment.

Finally, as part of an internally funded project, DLR will perform further studies on the effects of sleep inertia on the crew's performance as well as on the evaluation of potential active sleep inertia mitigations.

Bibliography

References

[1] eMCO-SiPO, “D-2.2 Detailed Research and Test Activity Plan,” 2023.

[2] eMCO-SiPO, “D-2.1 Report on nominal operations,” 2023.

[3] eMCO-SiPO, “D-3 Report on failure condition management,” 2023.

[4] Airbus S.A.S., “A350XWB - Airbus Standard A350 Training Devices Std. 1.2.0 Malfunctions Description: Simulator Document,” G05 - V00604901, 2020.

[5] Lufthansa Aviation Training, “Full-Flight Simulator AIRBUS A350-900 (FT72): Specifications and Product Sheet,” https://www.lufthansa-aviation-training.com/de/-/ffs-airbus-a350-900-ft72.

[6] Salas, E., and Dietz, A. S., *Situational Awareness*, Routledge, 2017.

[7] Endsley, M. R., Bolte, B., and Jones, D. G., *Designing for Situation Awareness*, CRC Press, 2003.

[8] Byrd, N., Joseph, B., Gongora, G., and Sirota, M., “Tell Us What You Really Think: A Think Aloud Protocol Analysis of the Verbal Cognitive Reflection Test,” *Journal of Intelligence*; Vol. 11, No. 4, 2023. doi: 10.3390/jintelligence11040076.

Annex A eMCO Checklist

EMCO SIPO FLIGHT CREW OPERATING MANUAL	STANDARD OPERATING PROCECURES	3.03.15	P1
	EMCO CRUISE	SEQ 100	REV 3

EMCO PREREQUISITES & ABORTION CRITERIA

eMCO may be performed under the following prerequisites that must be fulfilled anytime. Violation of any of these requirements leads to a required abortion of the active eMCO segment.

- **Flight rules** . . . . . **IFR**
- **Valid cruise flight phase** . . . . . **after ToC & 45 min before ToD**
- **Altitude** . . . . . **above FL 200**
- **AP and A/THR** . . . . . **ON**
- **Shift duration alone** . . . . . **150 min**
- **Pilot flying** . . . . . **at controls & seat belts fastened**
- **No crew member incapacitated**

Beyond the conditions mentioned above, an **eMCO segment is also aborted on**

- **Amber caution, single chime** or higher failure severity
- **PF incapacitation**
- **Physiological break** that cannot be delayed any further
- **Required diversion**
- **High workload**, such as oceanic clearances or adverse enroute weather

TASKS & ROLES DURING AND AFTER EMCO ABORTION

- **Pilot flying (PF) immediately wakes up Pilot Resting (PR)**
- **PF remains PF**
- **PR remains PR for 15 minutes**, the only allowed tasks are
  - **Observe** cockpit and PF actions
  - **Resynchronize** with the flight progress and aircraft status
- **After 15 minutes have passed: PR becomes PM**

THESE PROCEDURES ARE SPECIFIC TO THE SIMULATOR AND ARE FOR USE IN EMCO-SIPO SIMULATOR TRIALS ONLY. THEY ARE NOT INTENDED FOR REAL WORLD USAGE

14-2-2025

EMCO START BRIEFING

Before the start of the eMCO segment review and discuss:

- **Status at the start of the eMCO segment**
  - Time
  - Position and FL
  - Fuel
  - ATC
  - Aircraft systems and autoflight settings
  - Cabin
  - Crew fitness
  - Physiological needs

- **Expected conditions and activities during eMCO segment**
  - Route
  - Terrain and safety altitudes
  - Weather
  - ATC and communication
  - Diversion airports
  - Expected PF actions

- **Conditions that require unscheduled end of eMCO segment**
  - See ***EMCO PREREQUISITES & ABORTION CRITERIA***

- **Tasks & Roles in case of unscheduled end of eMCO segment**
  - Pilot flying (PF) immediately wakes up Pilot Resting (PR)
  - PF remains PF
  - PR becomes PM

14-2-2025

EMCO END BRIEFING

At the end of the eMCO segment review and discuss:

- **Status at the end of eMCO segment**
  - Time
  - Position and FL
  - Fuel
  - ATC
  - Aircraft systems and autoflight settings
  - Cabin
- **Past conditions and activities**
  - Route deviations
  - ATC restrictions
  - Aircraft systems
  - Cabin
  - Dispatch
- **Destination – in case of significant changes**
  - Expected time of arrival
  - Expected weather at destination
  - Expected fuel at arrival

14-2-2025

Annex B NASA-TLX Rating Scale

Phase 1 - PF alone / PR not in cockpit or in sleep inertia recovery (from begin of eMCO segement until 15 min after eMCO abortion)

Please use the slider inside the scale to indicate your answer

Mental demand

How much mental demand and perceptual activity was required (i.e. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

Only numbers may be entered in these fields.
Your answer must be between 0 and 10

low high

Physical demand

How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Only numbers may be entered in these fields.
Your answer must be between 0 and 10

low high

Temporal demand

How much time pressure did you feel due to the rate or pace at which the scenario elements occurred? Was the pace slow and leisurely, or rapid and frantic?

Only numbers may be entered in these fields.
Your answer must be between 0 and 10

low high

Performance

How successful did you think you were in accomplishing the goals of the scenario? How satisfied were you with your performance in accomplishing these goals?

Only numbers may be entered in these fields.
Your answer must be between 0 and 10

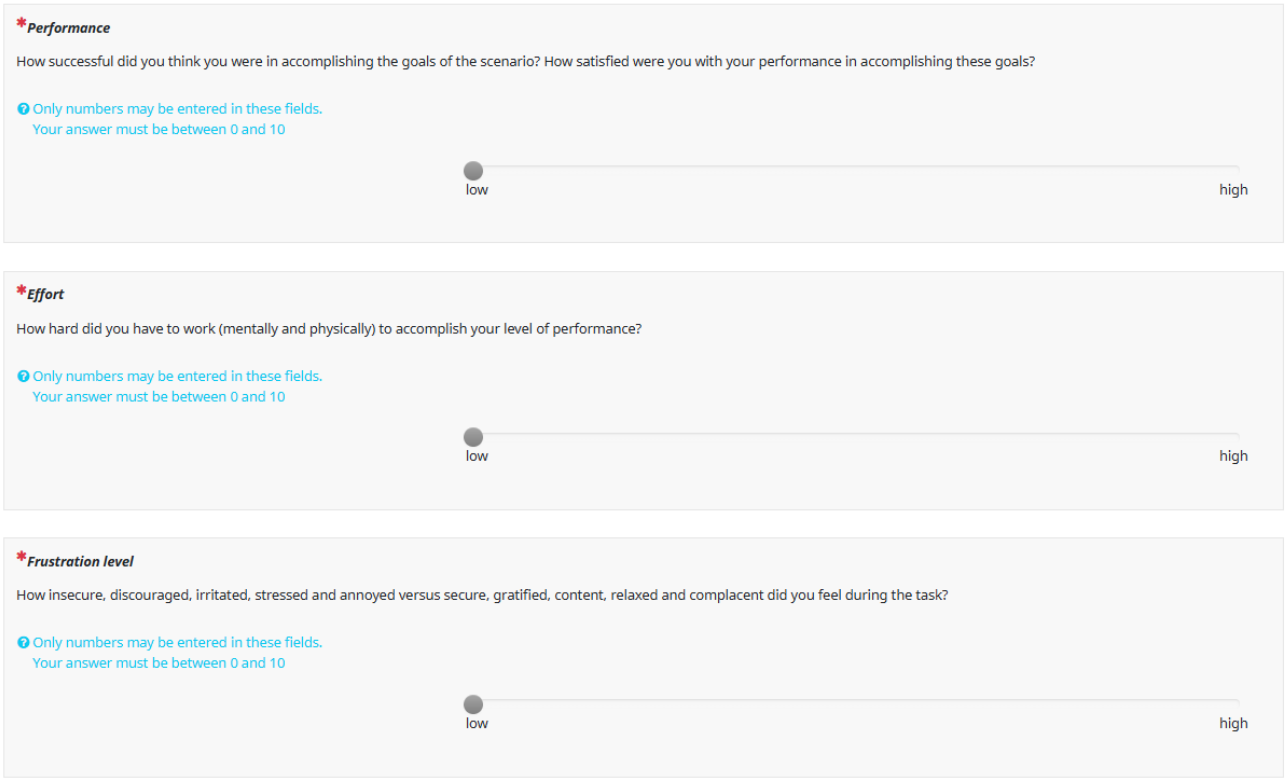
low high

Effort

How hard did you have to work (mentally and physically) to accomplish your level of performance?

Only numbers may be entered in these fields.
Your answer must be between 0 and 10

low high


*Performance
How successful did you think you were in accomplishing the goals of the scenario? How satisfied were you with your performance in accomplishing these goals?
Only numbers may be entered in these fields.
Your answer must be between 0 and 10
low
high
*Effort
How hard did you have to work (mentally and physically) to accomplish your level of performance?
Only numbers may be entered in these fields.
Your answer must be between 0 and 10
low
high
*Frustration level
How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?
Only numbers may be entered in these fields.
Your answer must be between 0 and 10
low
high

European Union Aviation Safety Agency

Konrad-Adenauer-Ufer 3
50668 Cologne
Germany

Mail EASA.research@easa.europa.eu
Web www.easa.europa.eu

An Agency of the European Union