Evaluation of Pilot Vigilance during Cruise towards the Implementation of Reduced Crew Operations

Bewertung der Vigilanz von Piloten im Reiseflug zur Implementierung von Reduced Crew Operations

Stefan Manuel Neis M. Sc.

Dissertation D17

Darmstadt 2020

TECHNISCHE UNIVERSITÄT DARMSTADT

FSR

Evaluation of Pilot Vigilance during Cruise towards the Implementation of Reduced Crew Operations

Bewertung der Vigilanz von Piloten im Reiseflug zur Implementierung von Reduced Crew Operations

Vom Fachbereich Maschinenbau
an der Technischen Universität Darmstadt
zur
Erlangung des Grades eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte

Dissertation

vorgelegt von

Stefan Manuel Neis M. Sc.

aus Bensheim

Berichterstatter: Prof. Dr.-Ing. Uwe Klingauf
Mitberichterstatter: Prof. Dr.-Ing. Ralph Bruder

Tag der Einreichung: 10. Oktober 2019
Tag der mündlichen Prüfung: 03. Dezember 2019

Darmstadt 2020

D 17

There are only two emotions in a plane: boredom and terror.

Orson Welles

Abstract

Airlines operate in a highly competitive economic environment and hence seek to reduce cost. This thesis examines effects of a proposed reduction of flight crew to one pilot during cruise flight. It is hypothesized that the barrier to implementing Single Pilot Operations is the failure to consider the socio-technical system. Research to date focuses on high workload phases. A literature research reveals that no concept exists that addresses flight phases of low workload and their related challenges, such as reduced vigilance.

A task analysis was conducted to identify those tasks that pilots execute during the cruise phase. It was assessed that under normal operations, workload is minimal, and pilots keep themselves engaged and thus vigilant through operations-unrelated tasks. An experiment was then designed to estimate vigilance levels during a simulated cruise flight in a realistic, non-laboratory environment. 10 engineering students acted as pilots and executed a 4 hour cruise flight under realistic conditions including communication and check tasks. For comparison, both Single and Dual Pilot Operation conditions were simulated. Subject's vigilance was estimated through Psychomotor Vigilance Tasks, subjective assessments, and changes in physiological parameters over time. These include Engagement Indices obtained through Electroencephalogram, concentration of oxygenated hemoglobin and heart rate through near infrared spectroscopy, and eye blink frequencies and durations.

Results were inconsistent, as vigilance is very dependent on personal characteristics. Nevertheless, based on 3 physiological parameters, the experiment confirmed that vigilance decreased significantly when no critical events occurred. An objective performance decrement was not detected. With the onset of a simulated critical event, vigilance increased significantly. No significant differences were found in the vigilance decrement between operating regime conditions. It was concluded that not the crew complement is the underlying cause of the vigilance decrement, but the nature of the cruise phase and lacking opportunities for meaningful engagement. To close the research gap, a new human centric single pilot concept of operations was developed. An on-board mission manager is assigned mission management tasks. Mission planning and airline operations support functions keep the mission manager engaged and vigilant during the flight.

Following this thesis, the new concept of operations should be implemented and validated with regards to vigilance levels.

Kurzfassung

Fluggesellschaften operieren in einem wettbewerbsintensiven wirtschaftlichen Umfeld und versuchen, Kosten zu senken. Diese Arbeit untersucht Auswirkungen der vorgeschlagenen Reduzierung der Crew auf einen Piloten im Reiseflug. Es wird vermutet, dass die Hürde zur Einführung von Single Pilot Operations darin liegt, dass das sozio-technische System nicht berücksichtigt wird. Bisherige Forschung konzentriert sich auf Phasen mit hoher Arbeitsbelastung. Eine Literaturrecherche zeigt, dass kein Konzept existiert, das Flugphasen mit geringer Arbeitsbelastung und damit verbundene Herausforderungen, z.B. verminderte Vigilanz, adressiert.

Eine Aufgabenanalyse wurde zur Identifikation aller Aufgaben durchgeführt, die Piloten während des Reiseflugs ausführen. Die Arbeitsbelastung im Normalbetrieb ist dabei minimal und Piloten beschäftigen sich während des Flugs mit anderen Tätigkeiten. Ein Experiment wurde entwickelt, um die Vigilanz während eines simulierten Reisefluges in einer realistischen Umgebung abzuschätzen. 10 Ingenieursstudenten agierten als Piloten und führten einen vierstündigen Reiseflug unter realistischen Bedingungen einschließlich Kommunikations- und Checkaufgaben durch. Zum Vergleich wurden sowohl Single als auch Dual Pilot Operations simuliert. Die Vigilanz der Probanden wurde anhand Psychomotor Vigilance Tasks, subjektiver Beurteilungen und Änderungen physiologischer Parameter im Zeitverlauf geschätzt. Dazu gehören Engagement Indizes mittels Elektroenzephalogramm, Konzentration von oxygeniertem Hämoglobin und Herzfrequenz mittels Nahinfrarotspektroskopie sowie Augenblinzelfrequenzen und -dauern.

Die Ergebnisse waren inkonsistent, da Vigilanz stark von persönlichen Eigenschaften abhängt. Basierend auf 3 physiologischen Parametern bestätigte das Experiment jedoch, dass die Vigilanz signifikant abnahm, wenn keine kritischen Ereignisse eintraten. Ein objektiver Leistungsabfall wurde nicht gemessen. Mit einem simulierten kritischen Ereignis nahm die Vigilanz signifikant zu. Es wurden keine signifikanten Unterschiede in der Vigilanzabnahme zwischen den operationellen Systemen festgestellt. Daraus wurde geschlossen, dass nicht die Anzahl der Piloten sondern die geringe Einbeziehung des Menschen hauptursächlich für Vigilanzabnahme ist. Um die Forschungslücke zu schließen, wurde ein neues menschzentriertes Single Pilot Konzept entwickelt. Einem Missionsmanager werden Aufgaben des Missionsmanagements zugewiesen. Missionsplanungs- und Flugbetriebsunterstützungsfunktionen sorgen dafür, dass der Missionsmanager aufmerksam bleibt. Im Anschluss an diese Arbeit sollte das neue Konzept validiert werden.

Acknowledgments

This dissertation was written during my time as research associate at the Institute of Flight Systems and Automatic Control (FSR) at Technische Universität Darmstadt (TUDA) and in close cooperation with Boeing Global Services. During and after that time, colleagues, superiors, students, friends, and family have supported and contributed to this thesis. To all of these, I extend my sincerest gratitude:

First and foremost, I thank Prof. Dr.-Ing. Uwe Klingauf, head of the Institute of Flight Systems and Automatic Control for the chance and the freedom to perform the research presented in this dissertation. Furthermore, he allowed me to contribute to Boeing Global Services' Reduced Crew Operations (RCO) Research Thrust besides my teaching assignments at TUDA. This ultimately enabled me to perform research on the chosen topic. I thank Prof. Dr.-Ing. Ralph Bruder from the Institute of Ergonomics at TUDA for acting as second examiner for this dissertation.

For mentoring me as a person and this dissertation, I extend a sincere "thank you!" to Dr.-Ing. Jens Schiefele, Director Research & Rapid Development at Boeing Global Services. He has had a great influence on the direction of this thesis and success of this dissertation.

My colleagues and peers both at FSR and at Boeing Global Services have contributed in many different ways. I thank all for the unforgettable time together, and their continuous motivation and support with regards to this dissertation. In particular, I thank Dr.-Ing. Millie Sterling and Sebastian Sprengart for the fruitful discussions on RCO, and the fun we had in designing, developing, building, and using the RCO simulator at TUDA. I thank my roommate Jonas Schulze for introducing me to the D-AERO research simulator, the technical discussions, and his continuous support. Furthermore, I thank the current RCO-crew Michelle Wenzel, Eric Sprenger, and Pascal Menner, for their support in finishing this dissertation. I thank Miriam Cornel for providing her expertise in medical questions, and Sebastian Stern for always offering a helping hand with regards to the research simulator.

I am obliged to thank my students, who have contributed to this dissertation through the supporting research they have conducted as my Bachelor or Master students or student assistants: Maximilian Funck, Caroline Schott, Matthias Eiden, Lucas Gräff, and Pia Lenhardt.

Last, I would also like to express my gratitude to those colleagues at FSR who have taken some of the teaching assignments off of my shoulders: Torben Bernatzky, Martin Michel, Frederik Prochazka, Sebastian Baumann, and Jennifer Hoffmann.

Of course, all students who participated in my simulator experiments have my gratitude. Unfortunately, but for obvious reasons, they shall remain anonymous. Only with their time was it possible to gather the data presented in this work.

Not all the support I have received was technical or professional in nature. I thank my parents Brigitte and Kurt Neis for their support in so many different ways, for encouraging me on this journey, and ultimately for fueling the interest for aviation from early on. I thank my brother Mario Neis for his continuous support and encouragement. Last, but not least, I thank my wife Anja Neis, for her love, support, and patience during the execution of this dissertation.

Bensheim, Germany, in February 2020

Stefan Neis

Contents

Nomenclature

Symbols

Notation	Description	Unit
A	Age of subject	$years$
DPF	Differential path-length factor	—
E_V	Illuminance	$cd \cdot sr \cdot m^{-2}$
EAR	Eye Aspect Ratio	—
EAR_b	Eye Aspect Ratio baseline	—
EAR_{sq}	Squared difference of Eye Aspect Ratio	—
EBD	Eye blink duration	s
EBF	Eye blink frequency	$1/min$
EI	Engagement Index	—
G	Geometry factor	—
HR	Heart Rate	$1/min$
I	Intensity of detected light	W/sr
I_0	Intensity of emitted light	W/sr
I_P	Current measured by photodiode	A
OD	Optical density	—
S	Spectral sensitivity	nA/lx
a_x	Acceleration along x-axis (horizontal)	m/s^2
a_y	Acceleration along y-axis (side)	m/s^2
a_z	Acceleration along z-axis (vertical)	m/s^2
c	Sample concentration	—
c^{HbO2}	Virtual concentration of oxygenated hemoglobin	—
c^{HbR}	Virtual concentration of deoxygenated hemoglobin	—
d	Distance between light source and photo detector	cm
h	Relation between horizontal facial distances	—
k	Multiplication factor for blink identification	—
m	Slope coefficient of linear regression solution for †	$\dagger/s$ with $[\dagger] = 1, s, 1/min$

Continued on next page

Continued from previous page

Notation	Description	
t	Time	s
t_0	Experiment start time	s
v	Relation between vertical facial distances	—
α	Absolute power of alpha frequency band	μV^2
β	Absolute power of beta frequency band	μV^2
ε_λ	Molar extinction coefficient for given λ	$L \cdot cm^{-1} \cdot mol^{-1}$
λ	Wave length	nm
θ	Absolute power of theta frequency band	μV^2

Acronyms

ANOVA Analysis of Variance

AS Autonomous Systems

ATC Air Traffic Control

BPMN Business Process Model and Notation

COH Concentration of Oxygenated Hemoglobin

ConOps Concept of Operations

COTS Commercial-Off-The-Shelf

CPDLC Controller-Pilot Data-Link Communication

DPO Dual Pilot Operations

EAR Eye Aspect Ratio

EBD Eye Blink Duration

EBF Eye Blink Frequency

EEG Electroencephalogram

EI Engagement Index

FAA Federal Aviation Administration

fNIRS Functional Near InfraRed Spectroscopy

FCOM Flight Crew Operating Manual

FMAI Full Mission Awareness and Involvement

FMS Flight Management System

GO Ground Operator

HbO2 oxygenated hemoglobin

HbR de-oxygenated hemoglobin

HR Heart Rate

HRV Heart Rate Variability

HTA Hierarchical Task Analysis

ICAO	International Civil Aviation Organization
LED	Light-Emitting Diode
LoA	Level of Automation
MCDU	Multipurpose Control and Display Unit
MM	Mission Manager
MMAI	Minimal Mission Awareness and Involvement
MoO	Mode of Operation
MWL	Mental Workload
NIR	Near InfraRed
NOTAMs	Notices to Airmen
OCC	Operations Control Center
PERCLOS	Percentage of Eye Closure
PF	Pilot Flying
PM	Pilot Monitoring
PVT	Psychomotor Vigilance Task
RCO	Reduced Crew Operations
RMAI	Reduced Mission Awareness and Involvement
SA	Situation Awareness
SOP	Standard Operating Procedure
SPO	Single Pilot Operations
TMM	Total Mission Management
ToD	Top of Descent
TUDA	Technische Universität Darmstadt
UDP	User Datagram Protocol
USB	Universal Serial Bus

1 Introduction

This first chapter serves as an introduction to the topic at hand, commercial Reduced Crew Operations (RCO) and Single Pilot Operations (SPO), as well as to the need for the research conducted and described in this thesis. In particular, the motivation to evaluate pilot vigilance under RCO/SPO, the goals and boundaries of the herein described research, and the chosen approach are detailed.

1.1 Motivation

Commercial airlines today operate in a highly competitive and challenging environment [BOB09, Cen09]. The industry is characterized by low profit margins, high fixed-cost, and a high dependency on external factors such as economic conditions [WBM11, BOB09]. Evidence of this challenging environment are the numerous airline mergers, acquisitions, and bankruptcies of the past years, the formation of airline alliances and cooperation, the emergence and rise of so-called Low Cost Carriers, and the restructuring of legacy carriers [Han07, BOB09].

According to the INTERNATIONAL AIR TRANSPORT ASSOCIATION (IATA), flight deck crew cost account for 6.1% of total airline cost on a world-wide average; with up to 31% in the United States [IAT15, CBL+13, CFG+17]. Cost reductions have been achieved through alternative employment models, lower wages, and more efficient crew scheduling [KK04, BOB09].

Research by UBS suggests total world-wide cost savings of 15$bn through the reduction of flight crew to one single pilot in commercial aviation [CFG+17], which is the next logical step following the de-crewing trend from five to two crew over the past 60 years [Har07, JLF+12]. These projected savings come from decreasing direct crew, training, and accommodation cost [CFG+17]. Besides, SPO could counter an imminent pilot shortage [Lou13, Nor07, LFM+15, Boe16, MG16]. BOEING forecasts that 804,000 new civil aviation pilots are required in the next 20 years and points out that pilot labor supply remains constrained [Boe19]. SPO could also add flexibility to an airline's operations, as it reduces crew scheduling complexity.

While the terms *Reduced Crew Operations (RCO)* and *Single Pilot Operations (SPO)* are often used interchangeably throughout literature and media, there is a notable difference. While SPO refers to operations by a single pilot for the complete flight, RCO means temporary SPO. It is also referred to as Single Pilot Cruise (SPC).

In fact, SPO is standard for most general aviation and military aircraft today [LBB+14, Har07, CFG+17]. Thus infrastructure, knowledge, and experience for SPO exist, cf. [MG16]. The implementation of RCO or SPO in commercial operations, however, is constrained by regulatory authorities and public perception. While the de-crewing trend to two crew members has not resulted in a decrease in safety [Har11, Har07, BS92, MDL81], the transition from Dual Pilot Operations (DPO) to SPO while achieving at least today's level of safety is a more complex undertaking [DP05, BKK+17].

Existing RCO/SPO concepts all face inherent dilemmas that have not yet been solved. At the root often lies the failure to take the whole socio-technical system into account [SHS14]. Studies on SPO reveal that anything less than two pilots requires increased levels of automation, in particular during phases of high workload [EKB+16]. Adding (intelligent) automation will relieve the remaining pilot [EKB+16], however, it will not solve some of the SPO-innate human factors challenges such as loss of human redundancy, pilot incapacitation, or new forms of error introduced by automation [Sch15, Har07, Dek02, SWB97]. Simple solutions, such as deadman circuits used in trains, are not effective [CCMF93]. Furthermore, (intelligent) automation on future flight decks will most likely move safety-critical tasks away from the pilot [NSK18, FSWL98]. Left on the flight deck then is a potentially bored, out-of-the-loop, and left-alone operator, tasked with monitoring systems [Sch15]. This non-desirable state of reduced *vigilance* may become a new problem on long-haul flights. The thesis at hand aims at quantifying pilot vigilance under real-world conditions to answer the questions *how bad is it?* and *what can be done?*

Under these conditions, the main challenge for SPO environments is defining the human operator's role [Boy14]. A certifiable RCO or SPO concept for commercial aviation must fundamentally change the nature of work undertaken on the flight deck [HSS15, WG15, DH99, LDRDD12, DP05, CFG+17], the flight deck itself, and the role of the human operator [HSS15, NSK18]. SPO means redefining the operation and thus presents a paradigm shift impacting all areas of aviation including regulation, insurance, and society [CFG+17]. This shift opens up the path for additional benefits besides cost reduction. Two examples are the reduction of errors resulting from poor monitoring of automated systems and crew resource management, cf. [Civ13, LBB+17, CBL+13, CFG+17], and a more effective utilization of pilot's workforce during phases of low workload [LBB+17, LBM+14], such as the cruise phase. This also serves to relieve the monotony of long monitoring tasks, cf. [Sce01].

1.2 Goals and Boundaries

This thesis shall contribute to the ongoing discussion of the feasibility of RCO and SPO in commercial operations in general by providing additional insight into the benefits, challenges, and many unknowns of operations executed by a single operator. The focus lies on the cruise phase of commercial operations.

The goal is to determine pilots' vigilance levels during long-haul cruise flights under RCO/SPO and compare them to their vigilance levels under DPO. As vigilance is typically only investigated in laboratory settings to date, it is further the goal to investigate and implement methods for human vigilance estimation in real-world applications.

Based on the particular findings and following COMERFORD ET AL. [CBL+13] and STANTON, HARRIS AND STARR [SHS16], a new concept shall be developed as part of this thesis, to progress research on SPO. In particular, the whole socio-technical system shall be taken into account adequately by ensuring satisfactory vigilance and human engagement, as well as mission and situation awareness throughout the mission. Following HANCOCK [Han13] and SCERBO [Sce01], this shall be achieved through reconsidering pilot tasks and the flight deck environment.

1.3 Approach and Thesis Structure

To begin, chapter 2 introduces the current state of the art regarding RCO and SPO in commercial operations. As commercial SPO are not yet realized, an overview of existing literature concepts follows. The concepts are then categorized and evaluated. Subsequently, selected aspects of aviation psychology are discussed. In particular, the current research body on vigilance theory, application, and measurement is presented in detail.

To better understand today's DPO and relevant pilots' tasks during the flight phase of interest, a comprehensive pilot task analysis was conducted. The methods, outcomes, as well as an assessment of flight crew workload are summarized in chapter 3. Based on this knowledge and the outcomes of the task analysis, chapter 3 closes with the definition of the research gap for this thesis; research hypotheses are formulated. The need for an evaluation of pilots' vigilance levels under current DPO and future SPO is highlighted.

To research the hypotheses, an experiment was conducted in a realistic flight simulator environment. The following chapter 4 details the experiment setup, methods, and procedures designed to answer the research question on pilot vigilance during RCO/SPO and DPO. Chapter 5 reports the major findings from the experiments, discusses the results, and draws conclusions.

Based on literature knowledge and the experiment findings, a new Concept of Operations (ConOps) for future commercial SPO is developed based on a human-centric approach. The process and results are described in chapter 6. Assumptions and goals are stated, requirements are derived and justified. Each element of the ConOps is defined, although the focus lays on the human agent. Tasks of the future human operator are designed with the experiment findings in mind. The methodology and derivation results are presented. The future flight deck environment is presented briefly. An operational scenario is introduced to illustrate the new ConOps.

The closing chapter 7 is used to summarize the findings of this thesis and to draw conclusions. It provides an outlook for future developments.

Throughout this thesis, only the cruise phase of commercial flight operations is considered. While the results apply to both states of single pilot operation, RCO and SPO, usually the term SPO will be used. Although the experiment described in chapter 4 resembled today's DPO and possible future RCO more than future SPO, the term Single Pilot Operations (SPO) will be used to highlight the two states of the independent variable, crew complement, in the experiments.

2 State of the Art

This chapter provides a brief overview of single and dual pilot operations in civil aviation today (section 2.1). Commercial Single Pilot Operations (SPO) / Reduced Crew Operations (RCO) concepts in literature are categorized and reviewed in section 2.2. Research gaps, in particular towards the cruise phase of commercial operations, are identified as the failure to take the socio-technical system into account. In this phase, reduced vigilance may be an issue of interest for future long-haul SPO. To fully understand and then to close the selected research gaps through the design of new Concept of Operations (ConOps) for commercial SPO, an understanding of underlying aspects of aviation psychology is necessary. Situation Awareness (SA) and Mental Workload (MWL) will be reviewed in section 2.3, automation and relevant design considerations in section 2.4, and vigilance as the main focus in section 2.5.

Vigilance decrement theories and the physiological bases of vigilance will be presented. As performance-based measures are not applicable in real flight deck environment operations, the focus lies on physiological parameters to estimate operator vigilance. Theoretical principles and underlying physiological processes, applications, and explanatory powers of Electroencephalogram (EEG), Concentration of Oxygenated Hemoglobin (COH), Heart Rate (HR), Eye Blink Frequency (EBF), and Eye Blink Duration (EBD) will be presented in detail.

2.1 Dual and Single Pilot Operations Today

Commercial aviation is strictly regulated. Flight phases consist of a predetermined sequence of events. Provisions for certifying and operating an aircraft differ by the type of aircraft, by operation, and by regulatory agency. Considered for this thesis are the U.S. airworthiness regulations (certification specifications, 14 U.S. CFR Part 25) for "large aeroplanes" as well as "scheduled air carrier" operations (14 U.S. CFR Part 121). These are the strictest regulations.

This section briefly summarizes pilot roles in multi-pilot crews, regulations applying to SPO, as well as a critical evaluation of single and multi-pilot crews on the flight deck.

Pilot roles

Traditional pilot tasks may be summarized under aviating, navigating, communicating, and managing systems (ANC+S) [FAA16]; detailed pilot tasks are described in chapter 3. All tasks on the flight deck are split between two pilots irrespective of actual rank [Har07]: one pilot acts as Pilot Flying (PF) and the other as Pilot Monitoring (PM) [DP05]. The PM supports the PF, handles communication to Air Traffic Control (ATC), and serves as a monitoring instance [DP05, Har07]. Independent of the tasks carried out during a flight, one of the two pilots must be designated pilot in command (14 CFR §121.385). The pilot in command is responsible for the operation of the aircraft, whether actively manipulating the controls or not. The pilot in command also has the final authority. [ICA05]

Regulations and Exemptions

Today, commercial airline operations must be operated by at least two pilots (14 CFR §25.1523 and §121.385c, similar provisions apply in the European Union, see EU-OPS 1). This provision is based on workload and pilot incapacitation considerations (FAA AC 25.1523-1). Most general aviation aircraft (Part 91: general operating aircraft, Part 135: commuter and on demand operations), in contrast, are operated with a single pilot only. Notwithstanding the two pilot requirement for commercial aviation, §25.1523-1 requires all aircraft to be capable of operation by one pilot only from either seat. Several examples (see incident reports [AAI17, AAI15, Com12]) demonstrate that commercial aircraft can be safely operated by a single pilot, albeit under increased workload. Operating and flight crew member duty and rest requirements (14 CFR Part 117) detail when additional crew members must complement the two pilots during long-haul flights.

Temporary exemptions to the two-person rule are permitted e.g. due to physiological needs of a pilot (§121.543b) and controlled rest, cf. [ICA12]. Whenever one pilot leaves their station, a second authorized person is required on the flight deck [FAA15]. Additionally, temporary single pilot operations are further constrained by the requirement of wearing an oxygen mask when flying above a certain altitude, see §121.333, §91.211, and §135.89; EU-OPS 1.770 does not contain such provisions.

Critical Evaluation of Dual and Single Pilot Operations

Despite the distribution of workload in Dual Pilot Operations (DPO) and thus an overall reduction of workload per person, DPO comes with an inherent additional workload associated to crew coordination, communication, and management [Har07]. By nature, mis-communication and poor crew resource management may lead to errors [Har07, Civ13], see also [Eid17]. Harris [Har07] questions the effectiveness of the second pilot as an "error-checker", cf. [WD95]. Pritchett [CBL+13] supports this hypothesis, arguing that both pilots are vulnerable to the same frailties and distractions, cf. [LBB+17]. Skitka et al. [SMBR00] found that the presence of a second crew member does not guard against automation bias errors. In this regard, SPO might come with safety benefits [LBB+17, Har07] through less distractions and false senses of safety.

On the other hand, similar to roadside transportation, the presence of a second person may also be engaging through talking and reassuring in case of need, and thus have positive effects, cf. [BNP18, ES01, RM01]. Several critical incidents could only be resolved with two or more pilots. Examples include Lion Air flight JT43 on October 28th, 2018 [LS19], and Qantas Airlines flight QF32 on November 4th, 2010 [Aus13]. Additionally, a second human can identify declines in cognitive abilities of the other pilot [Air19], which is an automation challenge not yet solved (see Appendix A.2.2). Literature suggests that today's pilots are critical of commercial SPO [BEH+09]: Pilots list SPO as the third most important risk. In particular, the negatively connotated vision of a "lone fighter pilot", missing team-work, and the omission of the four-eyes principle are mentioned. [BEH+09]

2.2 Existing Commercial Single Pilot Operation Concepts

Based on an extensive literature research, this section elaborates on existing commercial SPO concepts. Two design philosophies towards SPO were identified. A classification of SPO concepts is derived, and a category review is conducted. The comprehensive analysis and review is given in Neis, Schiefele and Klingauf [NSK18]. Research gaps pertaining to existing concepts are elaborated.

2.2.1 Concept Design Philosophies

A simple "remove the second pilot" concept will not lead to safe SPO [NSK18, GHLT14]. Bottlenecks are situations of heavy workload, pilot indisposition (e.g. biological breaks), or pilot incapacitation. Concepts and frameworks have been developed to find solutions for these bottlenecks [NSK18]. While concepts and frameworks found in literature differ greatly, SPO concepts follow two logical steps (design philosophies) from today's DPO: remove the second pilot and replace them with automation (aircraft-centric system), or relocate the second pilot (respectively their functions) to the ground (distributed system), see Figure 2.1. [NSK18, LBB+14, SHS16, SRM+15, SK17, MSV+17] Combinations due to overall complexity are also found. Also, neither design approach offers the ultimate perfect solution.

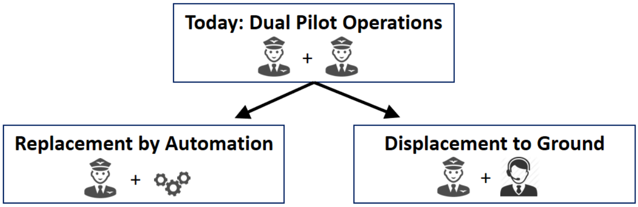

Figure 2.1.: Design approaches to SPO

Replacement design approach (aircraft-centric system)

Early approaches have widely focused on sophisticated technological solutions to replace the second pilot using physiological monitoring of human performance, intelligent knowledge-based systems, and adaptive automation [Har07, SHS16]. These can be seen as assistant, associate, or coach systems for the remaining single pilot [BGS08]. This design approach makes use of technological advancements, and does not require major changes to today's operational concept or air transportation system in general.

Displacement design approach (distributed system)

It may be argued that required on-board automation for the replacement design approach is difficult to develop and challenging to certify [HSS15]. Large-scale automation on the flight deck might increase the risk for human error [Sch15, Haa07]. New designs using existing technology were proposed i.a. by Harris [Har07]. Focusing on real-time distribution of tasks between flight deck and ground stations, they displace the second pilot to the ground [SHS14, HSS15].

2.2.2 Concept Classification

Existing SPO concepts in literature were identified, and a classification scheme has been developed. It bases on the two design approaches introduced in the previous section, and the number and nature of agents (entities, human or machine, which may act independently, cf. [CB09]) involved. In total, seven categories (A - G) have been identified (Figure 2.2). Both displacement and replacement approaches are clearly visible in their various specifications. All identified concepts come with unique requirements and challenges, which are summarized in Appendix A.2. A summary of each category follows; detailed descriptions can be found in [NSK18].

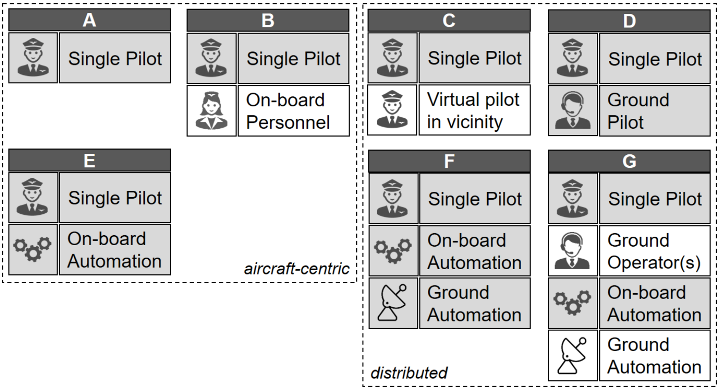

Figure 2.2.: Classification of existing SPO concepts. Grey color indicates agents that are always online; white indicates agents acting on request. [NSK18]

(A) Removal of second pilot

In the most simple concept category, discussed in [CBL+13, SHS14, GHLT14], the second pilot is removed without any change in the operating concept. Ultimately, this results in undesired states such as unacceptable heavy workload at times, inadequately low workload at other times, and leaves unanswered questions (how to deal with pilot incapacitation?). All other concept categories B - G exist to overcome these challenges by using either one or both design philosophies.

(B) Removal of second pilot, capable person for relief when required

Capable persons (e.g. flight attendants, commuting pilots, or any other untrained persons on-board the aircraft) are granted access to the flight deck and assist the single pilot during phases of high workload and whenever required. Concepts are discussed in [CBL+13, GOW+14, MG16].

(C) Virtual pilot in an aircraft in the vicinity

A pilot in the vicinity is connected to a pilot requesting assistance. In case of single pilot incapacitation, a support aircraft with a pilot in a virtual cockpit could rendezvous and take over control. Details are found in [CBL+13, GOW+14].

(D) Displacement of the second pilot to the ground.

Today's second pilot is displaced to the ground, cf. [GOW+14, DRPD17]. Both pilots continue to have the same tasks as today. A remote copilot comes with advantages over on-board personnel, such as independence from depressurization, G forces, smoke, etc. [LL07]. If communication fails, the single pilot would operate according to category A or B concepts. Additional equipment, such as crew resource management indicators, might be required, cf. [LFM+15, LBB+14, SRM+15].

(E) Replacement of the second pilot with on-board automation.

Replacing the second pilot by advanced on-board automation is the next logical step in the de-crewing trend described in section 1.1. Many research projects, not only in the civil aircraft domain, were conducted; Appendix A.3 lists and describes some of the most prominent. An important contribution to this category, due to transformed roles of the human operator, is the *Naturalistic Flight Deck* concept by Schutte et al. [SGC+07] and Flemisch et al. [FAC+03].

(F) Replacement of the second pilot with on-board and ground automation.

Category F concepts, discussed in [SHS14, SK17] and as part of ACROSS [GOW+14], represent an extension to the previous category (E). Additional automation on the ground act as safety net ("automation mirror" [SHS14]).

(G) Displacement of the second pilot with one or multiple ground operator(s), additional on-board and ground automation.

It is argued that SPO-related challenges may only be overcome and benefits harvested when changes on a broader scale are implemented. Concepts in this category often require a transformation of pilot roles and profound changes in pilot tasks. It is envisioned that the single pilot on board no longer is a traditional aviator or systems manager; instead they may be tasked with managing the mission, risk, and resources [LBB+17, SBL+16a, MSV+17], while (autonomous) machines micromanage systems (e.g. trajectory adherence, engines).

Four entities (on-board pilot, on-board automation, Ground Operators (GOs) and ground automation) provide resiliency, so that one entity may fail. Concepts are discussed i.a. in [SK17, A4A16, SHS16, BLBJ15, WG15, BJS14, LFM+15, MG16, BB09]. NASA gained extensive knowledge on remote crew interaction and change of tasks, distributed crew with enhanced collaboration tools, and the role of GOs, cf. [DKC+15, CBL+13, LBB+14, LBM+14, BLBJ15]. GOs will likely be part of existing airline Operations Control Centers (OCCs) with expanded functions [BJS14]. Second pilot functions might be transferred to traditional dispatchers, who already perform flight monitoring and decision-support functions [LBB+17].

Four GO roles, or organizational structures, are imaginable: remote controller, harbor pilot (as in maritime operations, cf. [BJS14, KRS+15, MSV+17]), hybrid operator, and specialist operator. All base on requesting dedicated assistance by the on-board single pilot, cf. [BJS14, WG15, BLBJ15], and are not mutually exclusive. Irrespective of the organizational structure, GO's responsibilities include monitoring single pilots, decision-making support to the on-board pilot, and taking over second pilot tasks under certain conditions [LGR+16, SK17].

2.2.3 Review of Concept Categories

In general, all existing concepts face similar challenges with regards to human factors. Little progress, however, has been achieved over the past decade to resolve these challenges. Each category has unique advantages and challenges associated to it, which makes determining the "most feasible" category difficult. Furthermore, the feasibility analysis is a multi-dimensional decision, as safety, and social, economical, and legal considerations must be taken into account, cf. [NSK18]. A general statement, as to whether SPO or RCO is feasible with a given category, cannot be given; concepts must always be reviewed individually. A review of the seven categories is constrained to the criteria named above. The only exception is category A: safety considerations and high chances for undesired states make such concepts not feasible, cf. [GHLT14].

While category B and C concepts are generally seen as feasible, the flexibility gained through SPO might be offset by the requirement to have the aforementioned crew members on board [CBL+13]. Flight deck controls must be simple in a way that capable persons can use them intuitively [CBL+13, NSK18]. Besides, such concepts are critical with regards to security, as they invert the current trend to limit access to the flight deck to anyone but authorized flight deck crew.

Concepts envisaging to control an aircraft from external sources (categories C, D, F, G) will likely require high initial development and implementation cost of providing reliable and secure means of communication [MG16, CBL+13, NSK18]. This initial investment cost and only minimal personnel cost savings could make such concepts economically unfeasible, cf. [MG16, CBL+13]. Authority and responsibility allocation and delegation protocols must be developed, which may, at the end of the day, require discussion on the ethical level. Such concepts could, however, be an intermediate step [NSK18].

SPO in commercial aviation is likely to be implemented with a category G concept, cf. [HSS15]. Although complex in nature, the combination of added automation and the existence of human ground operators can best deal with both standard and unforeseen situations [HSS15]. Such concepts will require a new operational concept [HSS15, NSK18]. Although such concepts deal with some of the human factors issues (e.g. MWL, time off), they still leave many questions unanswered. Challenges include, among others, resource and authority delegation management, communications reliability, integrity, and security issues, error management, human decision support, and vigilance. Due to the required complexity, category G concepts present a fundamental change, most likely associated with high investment costs.

2.2.4 Research Gaps with Regards to Existing Concepts

Two observations from the literature review are of relevance for this thesis, which pertain to the methodology of concept development: concept consistency and concept detailedness in literature (refer to Neis, Schiefele and Klingauf [NSK18] for a detailed review). Both observations apply to concepts from researchers regardless of size or significance of their institution. Firstly, most researchers try to master the challenge of developing SPO concepts by looking at the two extreme sides, normal but workload-heavy ("as planned") operations and emergency operations. Few concepts in literature accommodate for the uneventful, typical long-haul flight and associated human factors issues. Secondly, many concepts integrate enclosed, small, workaround-like solutions into one complex concept to solve the multifaceted challenge of RCO, cf. [NSK18]. This approach carries the risk of inconsistency, in particular with regards to the allocation of responsibility and authority [SNS18]. This applies in particular to present-day automation paradigms, which are "not sufficient for SPO" [BKK+17].

The evolution and introduction of automation on the flight deck has allowed to de-crew from a five person cockpit in the 1950s to the two person cockpit used today [Har07, CBL+13]. With the introduction of complex automation came a shift in the pilots' role from the traditional aviator to a systems manager [LDRDD12, HSS15]. Typically, pilots today only spend about 3-7 minutes per flight actually "flying" the aircraft (that is touching the controls), the majority of the flight time is left to systems and automation monitoring [CSC15]. Monitoring is a task humans are not good at [CS15]. Against this background it is important to understand underlying concepts of automation and aviation psychology such as Situation Awareness (SA) and Mental Workload (MWL). They are introduced briefly in section 2.3. The degree and nature of automation and function allocation on the flight deck directly influence operator MWL and SA during the cruise phase. Effects and design considerations to minimize these effects are discussed in the following section 2.4. To further understand why and how automation principles, function allocation, and the resulting task profile influence the human's ability to uphold SA during cruise, vigilance as the describing parameter for the required and desired state of the human operator is discussed in detail in section 2.5. This knowledge combined with experiment findings will later shape a new category G-like ConOps for SPO.

2.3 Situation Awareness and Mental Workload

Situation Awareness (SA) is essential for human decision-making [SSWG06] as it "represents a continuous diagnosis of the state of a dynamic world" [PSW08]. In literature, Endsley's three-level model ([End88]), Smith and Hancock's perceptual cycle model (see [SH95]), and the activity theory model by Bedny and Meister (see [BM99]) dominate to describe SA [SSWG06, SSJ+07]. In this thesis, Endsley's three-level model is used, as it is widely accepted and seen as the most useful for informing system design and evaluation [SSJ+07]. Endsley defines SA as

> "*the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.*" *[End88]*

Mental Workload (MWL) is multidimensional [YBWH15, HM88]. It may be compared to the two components, stress and strain, of physical workload [YBWH15], cf. EN ISO 10075 (2000), and available resources. Young and Stanton define:

> "*The mental workload of a task represents the level of attentional resources required to meet both objective and subjective performance criteria, which may be mediated by task demands, external support, and past experience.*" *[YS05]*

Besides the number of tasks to be executed, MWL also incorporates task saturation [LDB09]. Both SA and MWL may effect task performance [YBWH15, Hen95]. Suboptimal MWL, either overload or underload (see [BD01]), may lead to errors [YBWH15], which is known as the *Yerkes-Dodson-Law* (Figure 2.3).

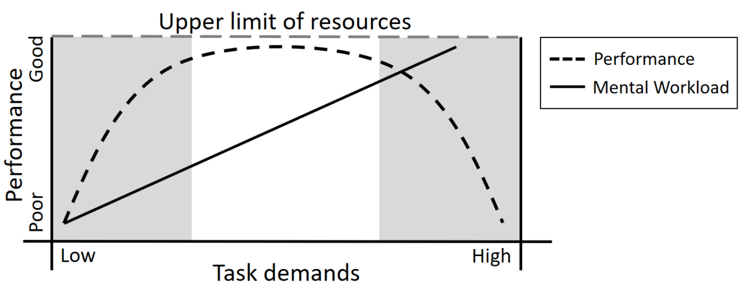

Figure 2.3.: Relationship between MWL (task demands) and performance, adapted from [deW96], cf. [YBWH15]. Mismatches between demand and capability represent mental over- or underload [Csi90, YD08, YBWH15].

For the purpose of this thesis, underload in particular is relevant for a ConOps for the cruise phase. It is a state in which the operator may not be vigilant, which may lead to reduced alertness and lowered attention. [YBWH15]

2.4 Automation and Design Considerations for the Cruise Phase under Single Pilot Operations

Automation is one of the major trends of the past century [End96] and "implies operating or acting, or self-regulating, independently, without human intervention" [Nof09]. It played an important role in the evolution of the flight deck [Ros89, CBL+13] and will play in the design of SPO, cf. [Com14]. Automation comes with the cost of managing it [BPR+13, Har13]. While designed to have beneficial effects on human performance, ironically, automation has often resulted in the opposite [Edw77, EJ12, PR97] and led to new forms of error [DP05, WC80, Bil96, SWB97]. Limitations of the human perceptual and cognitive system make it impossible to completely understand complex automation [DKC+15]. Hence, humans have to rely on automation guided by trust [DKC+15, LS11], which requires time [LS04]. Higher transparency and predictability, influenced by information about the system and the human's ability to understand this information, lead to higher trust in automation [SBH+16, MFBK17].

Function Allocation Strategies in Complex Systems

Parasuraman, Sheridan and Wickens state that "automation design is not an exact science" [PSW00]. In general, four aspects of a task can be automated: monitoring, generating options, selecting options, and implementing options [EK97]. Complex systems in aviation today require humans to interact and cooperate with increasingly autonomous systems [SGC+07, Hoc00, Bai83, BMRW75]. They are known as Joint Cognitive Systems [Har13]. Static task allocation strategies cannot achieve satisfactory SA [SDL16, SBL+16a], hence context-dependent function allocation strategies have been developed: adaptive automation [PSW00, Rou88, JFM15], "Levels of Automation" approaches [SV78, EK97, EJ12], and design patterns for Human-Autonomy Teaming, cf. [SBL+16a, SBL+16b, BLRS18]. Further strategies include "complemation" (technology designed to enhance human skills and abilities [Sch99, Sch00]), by taking into account the human's unique abilities of troubleshooting, abstraction, and adapting to new situations, cf. [Bil96].

Automation-Induced Design Considerations for SPO

Endsley and Jones [EJ12] list out-of-the-loop syndrome, lost mode awareness, and the decision support dilemma as challenges that automation may have on operator SA. Parasuraman adds *complacency* as "sub-optimal monitoring of automation or its information inputs" [PMS93, KJP14]. It is described as a state of self-satisfaction, over-confidence, non-vigilance, and inappropriate reliance on automation, cf. [Dek13, PMS93, BLF+76, End96, NTS14]. Lee [Lee06] discusses further automation-induced problems. All elements of the joint cognitive system must operate transparently, communicate and cooperate with each other to maintain a shared mental model, and adequately address autonomously changing dynamic situations, cf. [MFBK17, BLRS18]. Degrees of freedom must be maintained in man-machine interaction in SPO to allow both human and machine to adapt in real-time to unforeseen contingencies [Hoc00].

2.5 Vigilance and Human Engagement

This section introduces the phenomenon of vigilance. Definitions and explanatory theories will be presented, and design considerations for future flight decks will be summarized. Vigilance estimation methods are reviewed in detail.

2.5.1 Vigilance Theory

This section serves to shed light onto the phenomenon of vigilance and establishes a working definition for this thesis. In light of increasing automation, the flight deck becomes a highly automated supervisory control environment [CGT16] characterized by a low MWL profile and under-arousal. Inevitably, such environments lead to undesired operator mental states including boredom, task-unrelated thought (mind-wandering), and mental and physical fatigue [ENB+16, SS06, CGT16]. These mental states are partly correlated with each other and influence vigilance [CGT16]. *Boredom* may be defined as a "state of low arousal and dissatisfaction caused by a lack of interest in an inadequately stimulating environment" [CGT16]. Boredom resulting from lack of interest acts as a driver for shifting away attention from the primary task [CGT16]. *Task-unrelated thought* draws away resources from the primary task, but can also have positive effects, as it stimulates creativity [Fun18]. *Vigilance* (sustained attention), in general, refers to the ability to sustain attention to a task for a prolonged period of time [DP82, OSE06, WPM08] [MDWS00, Par98, Mac57]. Vigilance may also be defined as a state or a degree of readiness to react to stimuli in the environment [KSZ15]. This definition may be seen with sight to flight decks.

Debates on the definition of vigilance as a psychological phenomenon and implications on research and estimation methods is still ongoing, cf. [Han13, OSE06]. The term and definition of vigilance still remain to be somewhat blurry today

[OSE06] to a degree that Oken et al. [OSE06] recommend to not use the term vigilance at all. Particularly, vigilance research is not so much located in natural science, but in social and behavioral science [Fun18]. Current understandings of vigilance are most likely too simplistic [Sat93].

Vigilance Decrement

Mackworth [Mac64] first described the phenomenon of *vigilance decrement*, the decline in attention-requiring performance over an extended period of time. Vigilance decrement may arise irrespective of the task complexity [Par87, EJ12], both under over- and underload [WCF84, Don01, SGC+07]. Besides task characteristics (duration, MWL profile, signal characteristics) and the environment, personal characteristics (intrinsic motivation, sleep deprivation, circadian rhythms) influence vigilance decrement [Don01].

Two theories as to why vigilance decrements dominate in literature [TBS15]. The *resource depletion* (overload) theory hypothesizes that vigilance tasks are effortful [WDH96] and result in mental fatigue [HR15]. Due to the depletion of information processing resources over time, vigilance decrements. The *mind-wandering* (underload) theory postulates that vigilance tasks are monotonous and understimulating. Hence, the human mind tends toward self-generated, task-unrelated thought (mind-wandering) [SS06]. Both theories cannot fully explain all findings on vigilance decrement. Thomson, Besner and Smilek [TBS15] therefore introduce the *resource-control theory* which combines the two previous theories. They argue, that executive control over the distribution of attentional resources wanes over time (e.g. due to motivation decrease) [TBS15]. Mind-wandering results in performance cost for the primary task (see Figure 2.4).

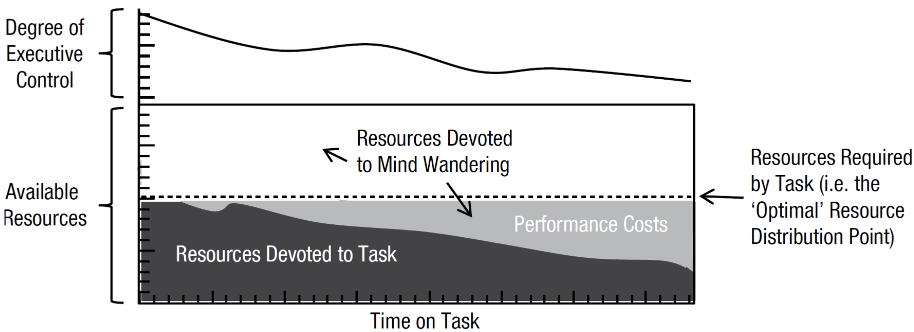

Figure 2.4.: Resource-Control Theory by Thomson, Besner and Smilek, from: [TBS15]

2.5.2 Physiological Bases of Vigilance

Vigilance, sustained, focused attention, and alertness are complex and somewhat abstract constructs. Being cognitive processes, they build on multiple underlying and interconnected physiological (physical, chemical, and biochemical) processes and mechanisms in the human body, in particular in the brain, and on psychological constructs. Some of these processes and constructs are not yet fully understood, cf. [OSE06]. In general, the capability to process information is modulated by activation states of the cerebral cortex of the brain. It is further influenced by sleep-wake states, by neurotransmitter systems, and by influencing parameters such as motivation, stress, and habituation. [OSE06] Oken et al. [OSE06] summarize and discuss the important underlying physiological and psychological mechanisms and physiology to vigilance, and critique related thereto.

The activation state of certain regions in the human brain is directly related to the human's vigilance state [OSE06]. By measuring brain activation and brain activity, a vigilance statement can be deduced [OSE06]. Dependent processes linked to brain activation are also of interest, such as the neuro-vascular coupling: local neural activity and subsequent changes in cerebral blood flow are related. Vascular-based functional brain imaging techniques, along with the measurement of dependent cardiac and dermatological activity can thus also be used to infer vigilance, its decrements and replenishments over time, or mental states leading to reduced vigilance [OSE06]. The operator's current cognitive state, amongst others inferred through physiological measurements of distinct "bio-behavioral signatures" [LCH+14] (see subsection 2.5.4), will play an important role on future flight decks [CBL+13]. It may also be used to detect pilot incapacitation, and to allocate tasks dynamically. Such methods may provide a continuous reflection of vigilance levels without unintentionally changing these levels through the measurement itself, cf. [KrK+07, SSS+11]. Besides, such measurements can be obtained continuously, and they do not require any events.

The relation between factors impacting cortical activation and performance is often U-shaped and follows the Yerkes-Dodson-law as described in section 2.3. Performance is best when arousal states are neither very low nor very high. [OSE06] Given the inherent complexity of the human physiology, a high number of known and unknown influencing factors, and large inter-individual differences, it is nearly impossible to make statements on the validity of certain measures at all.

As vigilance and physiological states and processes are linked, certain physiological preconditions may influence vigilance. Literature reports that females seem to react differently to monotony and boredom [SLFJ91, CGT16]. Handedness is known to influence brain waves [PC72, PPG+12]. Past and current neurological disorders or alcohol and drug abuse may also influence physiological processes [OBM07, SH92].

The following two sections present vigilance estimation methods: Often, vigilance is equalized to objective performance. Hence performance-based measures are commonly found in literature. They will be presented in the following section along with related critique. Physiology-based measurements derived from the knowledge gained in this section, their advantages, and underlying physiological processes are then presented in detail in subsection 2.5.4.

2.5.3 Vigilance Estimation through Performance Measures and Critique

Vigilance states and vigilance decrement are often inferred through objective performance measures. Sometimes, vigilance and performance (and their decrements) are even equalized. A vigilance decrement typically manifests itself through decreasing detection rate, increasing identification errors, slowed response times, and increasing levels of false alarms in psycho-motor vigilance tasks [Don01]. One problem of measuring performance to infer vigilance is known as circularity reasoning: when postulating that a decrement in vigilance will lead to a performance decrement, an actual decrease in performance cannot be explained by decreasing vigilance [WPM08], cf. [Fun18].

Besides, performance metrics are often obtained in controlled laboratory settings ("sit and stare" experiments). Typical cockpit monitoring tasks are frequently interrupted by secondary tasks (e.g. communications with external persons) [CS15], which laboratory studies do not reflect, cf. [Koe99, Han13, CS15]. Koelega [Koe99] criticizes laboratory experiments for unrealistic high signal rates. On the flight deck, few stimuli over several hours result in few observable events to measure performance and infer vigilance, cf. [SBC⁺14]. When faced with few stimuli, humans will apply coping strategies to counter boredom such as mind-wandering [CS15], which will influence vigilance levels. Along this line, artificially inserting events to measure operator vigilance during a cruise flight potentially falsifies measurements, as these events represent a stimulus, requiring the operator to engage, and thus potentially replenish vigilance levels, cf. [BWHA18, CS15, Buc66, Ada56].

While there certainly is a change in vigilance over time (which is deducted from literature findings, in particular the Yerkes-Dodson-Law, and common sense), the vigilance decrement as described by Mackworth may indeed be an "iatrogenical phenomenon" [Han13] not found in airline cockpits.

2.5.4 Vigilance Estimation through Physiological Measures

Gathering and interpreting physiological measures is not straightforward as performance measures. Multiple measures are required to infer vigilance as a single measure does not exist [OSE06, KrS+06, Sat93, Wil02]. Oken et al. [OSE06] recommend to control and measure as many physiological and behavioral parameters as possible due to the complexity of the underlying mechanisms. Second, such measures are relatively non-specific in nature [OSE06, Sat93]. Third, it is questionable, if physiological measures provide a vigilance statement or rather an "index of an unknown human behavioral parameter" [Sat93], which is due to potential simplistic definitions of vigilance. Non-laboratory conditions also effect physiological measurements [Sat93]. In particular when comparing physiological parameters, inter-individual differences must not be neglected.

Nevertheless, based on the cause - effect relations of cognitive, physiological, and psychological processes, despite that not all of them are fully understood, certain methods are commonly employed to deduce vigilance, alertness, and fatigue statements from physiological parameters and their change over time, respectively. The most promising physiological measures, determined through literature reviews (e.g. [BWHA18]) and own pre-testing are brain waves, blood oxygenation, eye blink rates, and heart rate. These will be discussed in detail in the following. All of these methods are non-invasive.

Neurological Metrics through Electroencephalogram (EEG)

Brain activation is correlated to an EEG-signal [OSE06]. An EEG shows spontaneous electrical activity (voltage fluctuations resulting from ionic current within neurons) of the brain over time, which is recorded using electrodes placed on the human scalp [SdS11, THBK07, BWHA18]. The difference in voltage between neighboring electrodes indicates cerebral activity (information processing processes of the cerebral cortex [SLH11]), referring to the synchronous activity of millions of similarly spatially oriented neurons. It is a direct measure of neuronal brain activity. [SdS11, THBK07, SLH11]

Besides clinical applications, EEG are also used in cognitive science and psychophysiological research. EEG is widely used for vigilance research. Example studies include SESAR's STRESS (Human Performance neurometrics toolbox for highly automated systems design) project [STR18] and other international studies, cf. [AKRP16, SSS+11, LCH+14, HTG+11, WM01, JM94, MI93, SI15, SSS17]. In this regard, EEG has been considered as the most sensitive, predictive, and reliable signal [LC05], cf. [BWHA18, SBC+14]. Oken et al. [OSE06] give a comprehensive overview of the reliability and validity of the EEG-signal. Of interest are both the detection of event-related potentials (a change in neuronal activity after a specific event) and the analysis of rhythmic activity (neural oscillations). Using Fast Fourier Transformation, such activity is divided into four frequency bands: alpha, beta, theta, and delta bands [Tat14], see Table 2.1. Note that in literature the number of frequency bands and their boundaries are arbitrary to a certain degree [McA12]. Although no physiological meaning is associated to the frequency bands, they are often linked to arousal states [McA12].

Table 2.1.: EEG frequency bands, taken from TATUM ET AL. [THBK07], and associated arousal states, taken from HAUS ET AL. [HHK+16] and SCHOLZ [Sch14], cf. [MBB+07].

Band	Frequencies	Arousal state
Alpha	8 Hz – 13 Hz	awake relaxation, reduced readiness to react
Beta	13 Hz – 25 Hz	focused attention, brain reactivation after sleep, increased alertness
Theta	4 Hz – 7 Hz	deep relaxation, sleep, low alertness, dreaming
Delta	< 4 Hz	deep sleep

EEG measurements are correlated to vigilance and valid for vigilance estimation [OSE06]. Being non-intrusive and relatively cheap compared to other brain function measurements, they offer a high temporal resolution for analysis [McA12]. On the downside, EEG signals contain noise [SSP02] and artifacts (signals not originating from the cerebral cortex). Such result i.a. from eye blinks, eye-movements, muscle activity, and motion [Luc14], as well as power supply interferences and electrical noise.

Based on the findings that neural activity in various frequency bands is linked to different states of arousal, and, in particular, that an increase in beta power along with a decrease in alpha and theta activity is related to increased brain activity [FMPS99, SBD+03, PCC09, SBC+14], POPE, BOGART AND BARTOLOME [PBB95] developed an *Engagement Index (EI)* of the form beta power/(alpha power + theta power) to reflect task engagement. Besides, two other indices of the form β/α and $1/\alpha$ were discussed [PBB95, CSB+15]. These indices were validated in multiple studies, i.a. by FREEMAN ET AL. [FMPS99] and COELLI ET AL. [CSB+15]. Another index of vigilance is the *alpha spindle rate*, which bases on brief spindle-shaped increases in the amplitude in the alpha band during states of reduced vigilance [SSS+09, TH01]. This index is robust against external noise and artifacts [SSS+09]. Furthermore, alpha synchronization (increased alpha power) is viewed as an indicator for hypovigilance in literature [NK06, FMPS99, MBB+07].

Different parts of the human brain take over different functions, hence electrode placement is important. Literature provides inconsistent findings on which brain areas control vigilance [KKI17]. Of the four cortical areas, frontal, parietal, occipital, and temporal cortex, most promising seems to be a frontal electrode montage. Evidence exists that the frontal cortex plays an important role in in sustained attention and vigilance, cf. [KKI17, SLH11, FMS+00, MMGM14, CSB+15, LRH+03, OK77, LSSO95]. Other studies focus more on occipital and parietal cortices, cf. [SBC+14, WM01].

Neuro-Vascular Metrics through Functional Near InfraRed Spectroscopy (fNIRS)

fNIRS is typically used in hospital applications, such as the monitoring of oxygen supply to the brain during heart operations. With regards to vigilance research, including ecologically valid environments, Functional Near InfraRed Spectroscopy (fNIRS) has also been successfully used before (Bogler et al. [BMSH14] provide a list of relevant studies). fNIRS uses near-infrared spectroscopy for the purpose of neuroimaging [KW14]. As neuronal activity is linked to related changes in localized cerebral blood flow (neuro-vascular coupling, see section 2.5.2), fNIRS measures brain activity through hemodynamic responses [LCLD12, DDP+14, SLH11]. For vigilance research, measurements on the parietal and prefrontal cortex seem to be best suited [CS02, BMSH14, SLH11]. fNIRS is a non-invasive method that detects concentration changes of oxygenated hemoglobin (HbO2) and de-oxygenated hemoglobin (HbR) resolved from the measurement of Near InfraRed (NIR) light attenuation or temporal or phasic changes (see Figure 2.5). [KW14, AIB+07] COH is expected to increase after focal activation of the cortex due to higher blood flow resulting from higher metabolic demand, while concentration of HbR is expected to decrease as it gets washed out by the higher blood flow [KW14].

Human skin and tissue are mostly transparent to near infrared light at wavelengths of 700-900nm [IBI+07]. HbO2 and HbR, which are linked to tissue oxygenation and metabolism, are absorbers of light. Significant differences in their absorption spectra (see Figure 2.6) are the basis for detecting relative changes in hemoglobin concentration. Two or more emitters with different wavelengths are employed (one above and one below 810nm, at which both HbO2 and HbR have identical absorption coefficients). [IBI+07] Photons entering the head are either absorbed by HbO2 and HbR or scattered by inter- and intracellular boundaries of different layers of the head. A photo detector collects those photons not absorbed and those scattered. While the latter amount is assumed to be independent from cognitive activity, changes in the attenuation measured must result from variations in the changes in absorption, which in turn results from changes in concentrations of HbO2 and HbR in the brain tissue. [IBI+07]

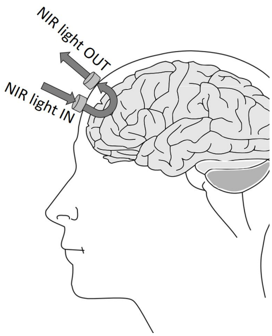

Figure 2.5.: fNIRS schematic after [AABM17]

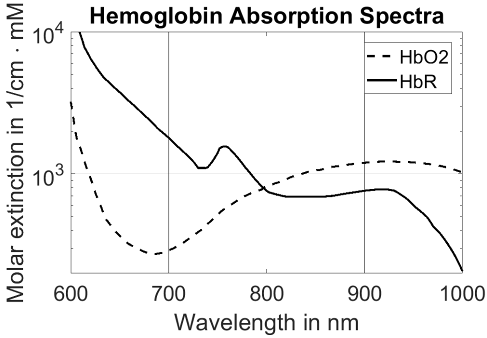

Figure 2.6.: Absorption spectra of HbO2 and HbR (taken from [Pra99]).

HbO2 and HbR concentration changes can be calculated as a function of total photon path length. The Beer-Lambert Law describes the attenuation of light when traveling through a non-scattering absorbing medium:

$$OD = -log\frac{I}{I_0} = \varepsilon_\lambda \cdot c \cdot d \tag{2.1}$$

where OD is the optical density or absorption, I the intensity of detected light, and I_0 the intensity of emitted light; ε_λ is the molar extinction coefficient (in $\frac{L}{cm \cdot mol}$), c the sample concentration, and d the distance between source and detector. To calculate the concentration of two constituents when extinction coefficients are known, measurements at two different wavelengths are required [Fat15, ASB$^+$12]:

$$\begin{aligned} OD_{\lambda 1} &= (\varepsilon_{1,\lambda 1} \cdot c_1 + \varepsilon_{2,\lambda 1} \cdot c_2) \cdot d \\ OD_{\lambda 2} &= (\varepsilon_{1,\lambda 2} \cdot c_1 + \varepsilon_{2,\lambda 2} \cdot c_2) \cdot d \end{aligned} \tag{2.2}$$

The Beer-Lambert Law is limited to non-scattering mediums. Biological tissue, however, is a turbid medium, and photons are scattered in random directions which increases the traveling distance between light source and detector [Fat15]. Therefore, the modified Beer-Lambert Law must be used for fNIRS applications:

$$OD = ln(10) \cdot \varepsilon_\lambda \cdot c \cdot d \cdot DPF + G \tag{2.3}$$

where DPF (differential path-length factor) is used to correct the distance factor. G is a geometry factor. While the modified Beer-Lambert-Law calculates absolute concentration values, in fNIRS applications DPF, G, and the number of chromophores are uncertain [ASB+12, Fat15]. In general, G and DPF are assumed to be constant [ASB+12], and can be eliminated if only relative optical density ΔOD is used. This leads to the following equation, in which ΔOD is again measured through the change in detected light intensity for the two relevant wavelengths:

$$\begin{bmatrix} \Delta OD_{\lambda 1} \\ \Delta OD_{\lambda 2} \end{bmatrix} = \begin{bmatrix} \varepsilon_{\lambda 1}^{HbR} \cdot d \cdot DPF_{\lambda 1} & \varepsilon_{\lambda 1}^{HbO_2} \cdot d \cdot DPF_{\lambda 1} \\ \varepsilon_{\lambda 2}^{HbR} \cdot d \cdot DPF_{\lambda 2} & \varepsilon_{\lambda 2}^{HbO_2} \cdot d \cdot DPF_{\lambda 2} \end{bmatrix} \begin{bmatrix} \Delta c^{HbR} \\ \Delta c^{HbO_2} \end{bmatrix} \tag{2.4}$$

This equation set can be solved for HbR and HbO2 concentrations when DPF-values are known. These are mostly dependent on wavelength and subject age [SW13]. An approximation formula was constructed based on work by Scholkmann and Wolf [SW13] and Kamran et al. [KMJ18], see [Sch19]):

$$DPF(\lambda, A) = 223.3 + 0.05624 A^{0.8493} - 5.723 \cdot 10^{-7} \lambda^3 + 0.001245 \lambda^2 - 0.9025 \lambda \tag{2.5}$$

with λ being the specific wavelength in nm and A subject age (in years).

Existing studies show that an increase in task difficulty results in a change in oxygen level [ASB+12], and a change in task reaction time correlates with HbO2 values [BMSH14]. Assuming that a decrease in reaction time indicates decreasing vigilance, fNIRS is a method to estimate vigilance, cf. [BMSH14]. However, fNIRS offers bad temporal resolution [DP17]. Additionally, it is notable that the time to the Δc^{HbO_2} peaks is 2 s to 15 s after activation onset [KLR+00].

Cardiac Metrics: Heart Rate Variability (HRV), Heart Rate (HR), and Respiration

HRV describes the variation in the time interval between subsequent heartbeats. The peaks in electrocardiogram readings are called R-waves, the intervals are called R-R intervals. As HRV allows conclusions on the activity of the sympathetic and the parasympathetic autonomic nervous system [DRST11], it is of interest for vigilance research. In general, physiological changes in respiration and the autonomic nervous system occur during the awake-sleep transition: HR decreases, respiration becomes more rhythmic, and tidal volume decreases [OSE06]. Activation of the sympathetic nervous system ("fight or flight" response [McC11]) is usually accompanied with low HRV, while activation of the parasympathetic nervous system ("rest and digest" response) is accompanied with high HRV [McC11]. Kaida et al. [KrK+07] predicted performance using HRV in a simple vigilance task, however, findings with regards to HRV and performance changes are not consistent in literature, cf. [Fun18]. Still, HRV is an indicator of changes in mental states and cognitive demands [AMM87, MRJBI15], and might be helpful in vigilance research.

Linked to HRV are HR and respiration. While Pattyn et al. [PNHS08] show that both HR and respiration decline over time in a simple vigilance task, both measures are rather unspecific [Fun18]. Against the background that typical simple vigilance tasks are laboratory applications with little validity towards real-world applications, see Hancock's [Han13] critique discussed in section 2.5.3, HR and respiration might offer even less explanatory power in the applications aimed at in this thesis. Brain activity, task difficulty, mental task load, stress, working time, task demand, muscular fatigue, and anxiety are all found to effect HRV and respiration [BAV+14], hence it is impossible to isolate vigilance effects.

Ocular Metrics: Eye Blink Frequency (EBF) and Eye Blink Duration (EBD)

Ocular metrics are widely and successfully used to detect car driver vigilance, fatigue, and drowsiness, cf. [BWHA18]. While the employed methods vary, they all include the measurement of the same parameters. These include, among others, Eye Blink Frequency (EBF), Eye Blink Duration (EBD), and Percentage of Eye Closure (PERCLOS). Eye blinking is a fast, semi-autonomic movement of the eyelid. Its frequency is dependent on external factors such as activity or environment [Dou11]. Blink duration, in particular, is influenced by drowsiness and fatigue [MGB+11, MMM+13, SSS+11, CEU03]. Several studies highlight the effect of visual or mental workload and task-induced fatigue on EBF and EBD, cf. [MMM+13, MGB+11, AFB06, SBS94]. Accordingly, it is assumed that vigilance and sustained attention states also influence eye metrics [MMM+13, AFB06]. In particular, EBF and EBD are found to increase with decreasing vigilance (performance) [BWHA18, MGB+11].

Another commonly used ocular parameter and accepted standard to estimate fatigue and alertness is PERCLOS, which describes the percentage of time the eyes are at least 80% closed, cf. [DG98]. PERCLOS correlates negatively with signal detection task performance [MGB+11], and correlates with increasing fatigue and declining performance in vigilance tasks [DG98, BWHA18]. Again, PERCLOS can be assumed an indicator of vigilance. Other relevant ocular parameters include pupil diameter, pupil eccentricity, pupil velocity, [MMM+13] and fixation [BNS+06]. These are more complicated to measure, or their relation to performance changes and vigilance is debatable.

The discussed parameters are usually obtained through normal or infrared video recording of the subject, along with an algorithm to identify facial landmarks (eyes, pupils and pupil size, eye aspect ratio). Other methods include fixed or head-mounted eye trackers, and filtering EEG towards artifacts also allows for blink detection.

2.5.5 Design for Vigilance

Knowledge about the undesired mental states and the resource-control theory indicates how tasks on a future flight deck must be designed to support operator vigilance. To counter boredom and underload, and thus reduce vigilance decrement and low vigilance levels in general, many authors, including [SGC+07, End17, AC11, PSKD12, Bou06, Haa07], promote human *engagement* (creating *alertness* through active involvement) in the task in a meaningful manner. In contrast to today's "hours of boredom, moments of terror" pilot workload profile, engagement serves to even out workload across the whole mission [SGC+07], and to keep vigilance at acceptable levels. Additionally, engagement prepares the human for back-up and troubleshooting functions in case of automation failure [SGC+07]. Creating a balance in task complexity and MWL along with diversifying tasks is important to maintain vigilance on the single pilot flight deck, cf. [Don01].

Hancock [Han13] and Scerbo [Sce01] promote the creation of environments and tasks that actively encourage individuals to engage in, and thus increase their performance. This includes stimulating intrinsic motivation of task engagement (in contrast to monitoring tasks dictated by someone else), but also displays which are specifically designed for the human [Han13, Sce01]. The relevance of motivation becomes visible when looking at video game players, whose performance does not decrement after long periods of sustained attention [Han13]. Generating (intrinsic) motivation through meaningful involvement and maintaining operator motivation through performance feedback alleviates vigilance decrement [HDWM99, Han13]. Meaningful involvement tasks require authority over the current mission, and allow self-determined scheduling and content of tasks.

Besides, Schmidt et al. [SSS+11] report a positive (although short) effect of verbal interaction on vigilance in a monotonous daytime driving task. Interaction with another human being might help reduce the vigilance decrement.

2.6 Summary

Chapter 2 first gave an overview of current legislation with regards to SPO in commercial aviation and briefly introduced pilot roles and a high-level evaluation of DPO and SPO.

SPO has already been under research for the past decades. This existing research body was summarized and reviewed in section 2.2 based on an extensive review by the author, cf. [NSK18]. Research gaps with regards to SPO were identified as the failure to take the whole socio-technical system into account, in particular during the uneventful cruise phase. In this phase, reduced vigilance might be an issue of critical interest for future commercial long-haul SPO. Following this insight, the current research body on vigilance was presented next, including vigilance decrement theories and the physiological bases of vigilance. As performance-based measures are not applicable in real flight deck environment operations, the focus lay on physiological parameters to estimate operator vigilance. Theoretical principles and underlying physiological processes, applications, and explanatory powers of EEG, COH, HR, EBF, and EBD were presented in detail. A further benefit of these methods is that physiological parameters can hardly be influenced by test subjects [BWHA18]. Finally, design considerations for high vigilance levels were summarized.

To better understand the operating environment and tasks that pilots must execute during the cruise phase today, the next chapter will present a flight crew task analysis. Based on this input and the herein given state-of-the-art overview, research hypotheses will be derived.

3 Analysis of Flight Crew Tasks during Cruise

The goal of this chapter is understanding what pilots must do during a typical cruise phase flight, and how they remain vigilant. The scope, methodology, assumptions, and findings of a task analysis performed on current Dual Pilot Operations (DPO) are presented. The task analysis is a process in which the tasks to achieve an objective are analyzed and documented, cf. [KA92]. The outcome is a model of the world (context) and how work is performed in it [Dia04]. Besides information on the tasks themselves, the analysis gives insight into the order and distribution of these tasks. This is helpful towards an understanding of operator vigilance under today's operational regime. The analysis reveals a generally low workload profile during the cruise phase, and the predominant types of operations as monitoring, communication, and planning. As such tasks are not engaging enough to remain vigilant over the long duration of the cruise phase, pilots typically engage in further activities unrelated to the flying task. The combination of the low workload profile, the nature of tasks itself, and the effects of a further reduction of flight crew are identified as the research gap to be investigated in this thesis. Corresponding research hypotheses will be derived.

With regards to a new Concept of Operations (ConOps) for future Single Pilot Operations (SPO), reducing the number of flight crew requires a systematic approach to the allocation of work [HSS15]. It is presumed that current DPO functions will remain the same in SPO, but the agent who carries out operations may change. The approach to function allocation for SPO must allocate extant tasks appropriately and dynamically between the different agents. Additionally, it must identify those functions and tasks that are no longer required and those that are new, cf. [HSS15]. The DPO task analysis will allow the construction of a SPO ConOps by eliminating the existing "second pilot" agent and reallocating tasks between remaining and new agents.

3.1 Scope and Sources

From a broader perspective, the flight crew task consists of five elements: flight management (aviate, navigate), the management of operational and environmental complexity, team management, reporting, and task management [Fun91]. As such, it is embedded in the complex air transport system, with information flowing between various entities and across system boundaries. Flight crews communicate with the outside world, in particular with their airline Operations Control Center (OCC), Air Traffic Control (ATC), cabin crew, and ground handling personnel. Information input to the flight crew is further provided by aircraft sensors and weather services. Flight crews and their tasks are influenced by national and international regulations, airline business model objectives and airline processes.

The herein described task analysis bases on the analysis of Boeing B787 and Airbus A320 Flight Crew Operating Manuals (FCOMs). It was complemented with flight deck video, various International Civil Aviation Organization (ICAO) and Federal Aviation Administration (FAA) material, and existing task analyses performed by Hankers [Han16] and Burian et al. [BCF+13]. It was refined and validated with the help of a subject matter expert. Purposely, the task analysis deviates from procedures described in the FCOM to depict procedures as they are executed in the everyday pilot life, and not as they should be according to theory.

Assumed was a typical commercial trans-Atlantic flight from Europe to the U.S. on a Boeing B787 aircraft under Extended-range Twin-engine Operational Performance Standards (ETOPS). It was presumed that no off-nominal events occurred. While the analysis was performed for all phases of flight, the herein reported results cover the cruise phase only.

3.2 Method and Results Representation

For the analysis of today's DPO the Hierarchical Task Analysis (HTA) method was chosen. HTA is an economical and flexible method, thus may be used to describe any system [Sta06]. It is widely accepted and used, allows to focus on crucial aspects of a task, and can be used to describe both team-work and non-human (i.e. automation) tasks [KA92, Sta06]. The underlying theory bases on goal-directed behavior, thus a HTA describes a system in terms of its goals [ADSG71, Sta06].

According to Kirwan and Ainsworth [KA92], the HTA results in a hierarchy of operations (things people must do to achieve goals) and plans (condition statements to undertake the operations to achieve goals) [KA92]. Goals are defined as desired states of a system under control. Advantageous is that the level of detail (stopping rule for description depth) can be set by the analyst as required. An HTA is commonly represented in hierarchical diagrams or tabular formats [ADSG71, KA92]. In order to use the HTA results for function allocation, tasks should be described as what should be done, and not how [MK05].

Generally, a pilot's overall goal may be derived from the goals of an airline, and may be stated as: *Operate a flight safely, efficiently and timely*, cf. [MHR04]. Efficiency may include economic (minimum cost), ecologic (minimum emission), and time factors (arrive before set time). The stated high-level goal may be achieved through a temporal sequence of operations. The operations may then be further split down into sub-operations with sub-plans until the desired level of detail is achieved.

Operations are described in the format of "do something". They may be seen as the smallest part of the overall task. Operations may be of different nature, they may include actions, decisions, information processing, and communication. An operation may be carried out by one or multiple agents. In total, seven agents have been identified as listed in Table 3.1. In the task analysis, no difference was made between different ATC controllers, ground-handling personnel, and airline OCC personnel. Plans provide the temporal and causal logic for carrying out operations. Operations can be executed either sequential or in parallel. An operation is usually executed once the execution of the previous operation is finished, or when meeting certain conditions.

Table 3.1.: Agents identified in the task analysis. Note that PF and PM are roles that both Captain and First Officer can take.

On-Board	Ground-based
Captain / Pilot Flying (PF)	ATC
First Officer / Pilot Monitoring (PM)	ground handling personnel
aircraft automation	airline OCC personnel
cabin crew	

Business Process Diagrams from the Business Process Model and Notation (BPMN) 2.0 (see [OMG11]) in Microsoft Visio 2013 were used to model and represent these relations graphically. BPMN allows to depict complex plans including condition-based events, events and conditions that interrupt operations, and parallel operations in a standardized way, and model the interaction between internal and external entities. BPMN models are a network of graphical objects made of four basic element categories. These include flow objects (e.g. events, activities: tasks and processes, gateways), connecting objects (sequence flow, message flow), swim lanes, and artifacts (data object, annotation). Swim lanes depict each actor within a system and all associated activities in a separate swim lane [Ber12], thus graphically distinguishing job sharing and responsibilities. [CT12, Rec10]

3.3 Plans and Associated Operations during the Cruise Phase

During the en-route climb phase, the pilots have configured aircraft systems for the cruise phase. Seats have been moved back and the shoulder harness has been released for comfort. This section now elaborates on the plans and associated operations in the cruise phase, which are represented in Figure 3.1 (high-level) and Appendix B (details). The cruise phase starts at the Top of Climb waypoint and ends at the Top of Descent (ToD) waypoint. Besides those in the following described tasks,

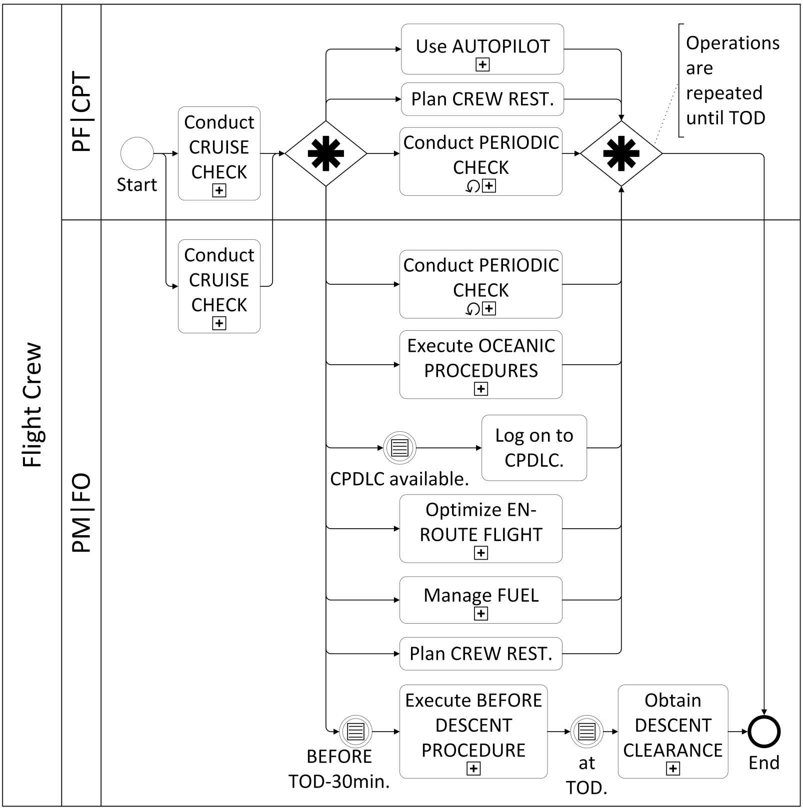

Figure 3.1.: Plan and associated operations of the cruise phase.

several operations must be executed over all flight phase (see Figure 3.2): Pilots must monitor communication channels at all times and communicate with external actors (ATC, OCC, cabin crew) when necessary. Communication with ATC largely includes sector hand-overs. Also, pilots need to monitor the aircraft's systems, and be vigilant and ready to respond to any non-normal and emergency situation at all times. Non-normal and emergency procedures are detailed in the FCOM, checklists, and Standard Operating Procedures (SOPs) material, which must be reviewed to execute the described procedures. Whenever required, pilots may temporary switch PM / PF roles ("handover control").

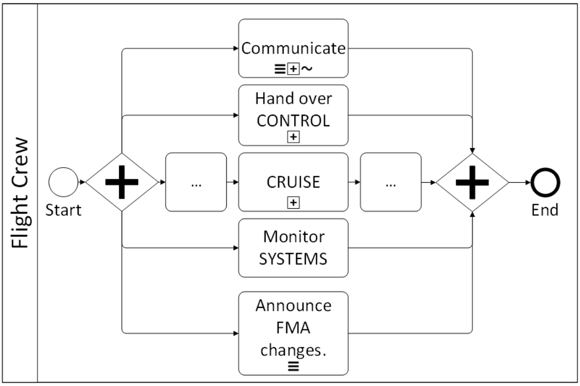

Figure 3.2.: Plan and associated operations over all flight phases.

Operations at the beginning of the cruise phase

When reaching the cruise level, a cruise check is completed (Figure B.1). It includes setting engine thrust to the pre-calculated cruise thrust level, crosschecking altimeter settings to be the standard setting, checking the cabin pressure altitude, as well as current fuel consumption and fuel used. If applicable during longer flights, the pilots schedule crew rest (including planned rest outside the cockpit in case of the presence of a relief pilot and controlled rest in the flight deck seat).

Regularly repeated operations during the cruise phase

A periodic check (Figure 3.3) shall be carried out regularly (every 30 minutes) during the course of the cruise phase to maintain Situation Awareness (SA). This check includes the following operations: All displays, panels, and controls shall be checked for any abnormal indications. The focus lies on the aircraft and engine status displays. Paper-based or electronic flight plan logs contained in the briefing package provided by dispatch must be updated. This allows for a comparison of actual values against the predictions (in particular amounts of fuel used and fuel remaining, position and time). Last, flight planning, in particular determining the nearest suitable airport and planning the respective route to be used in case of an emergency (e.g. a rapid decompression) must be conducted. This is particularly important in case of flying over a mountainous area. It includes getting and familiarizing with the latest weather reports for nearby, destination and alternate airports, and the joint decision, if the weather at those airports is within limits.

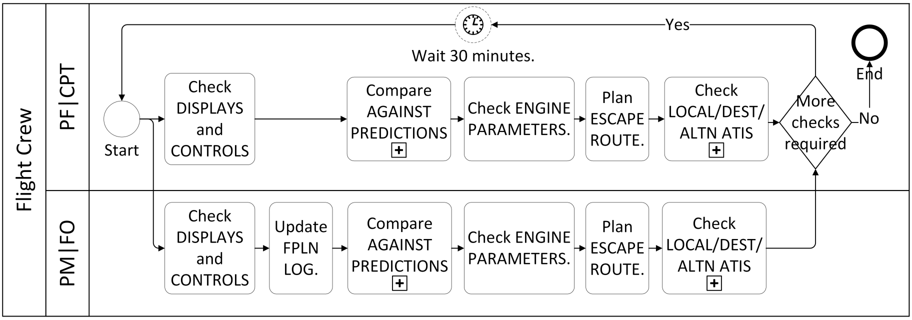

Figure 3.3.: Periodic checks: plan and associated operations.

Condition-based triggered operations during the cruise phase

If Controller-Pilot Data-Link Communication (CPDLC) is available in a certain airspace, ATC will request the pilots to log on and continue communication via data-link until further instructed. Pilots then will have to establish communication. Once data-link coverage is no longer provided, pilots will revert back to voice communication with ATC stations.

Triggered by the Flight Management System, by significant weather phenomena along the route, or as part of the periodic check, pilots may optimize the cruise flight (Figure B.2). It includes deviating from the planned track to avoid weather, changing the flight level to optimize fuel burn, to avoid or make use of winds, or adjusting cruise speed. Each operation requires performance checks (to determine the impact e.g. on fuel usage of the change and ensure if the operation is within aircraft performance limits), possibly communication with relevant airline stations, and a clearance request from ATC. Moreover, ATC may change the cleared trajectory. In these instances, pilots must respond and comply through manipulating the trajectory through the Flight Management System (FMS) or autopilot panels.

If the route includes crossing an ocean, pilots must execute trans-oceanic procedures before entering and while within oceanic airspace (Figure B.4). About 45 minutes before reaching the oceanic entry waypoint, pilots must obtain a crossing clearance. Before doing so, they must determine if the pre-filed flight level and speed for the oceanic crossing are still valid and optimal. While in oceanic airspace and if equipped appropriately, ATC will initiate a check of the selective calling radio system. The pilots need to ensure that the system is active. The PM will regularly submit a position report to ATC containing the latest position as well as the name of and estimated arrival time over the next flight planned waypoint.

Operations towards the end of the cruise phase

About 30 minutes before reaching the ToD waypoint, the flight crew will execute the Before Descent Procedure, again checking weather reports, reviewing aircraft status, configuring the aircraft's systems for landing, conducting an approach briefing, and executing the descent checklist (Figure B.5). If the descent clearance was not given by ATC by the time of passing the ToD waypoint, the PM must obtain the descent clearance from ATC.

3.4 Flight Crew Workload Assessment during Cruise

The predominant task types during cruise are monitoring, communication, and planning tasks. Under routine operations, there is very little, if any, activity related to the aviate and navigate tasks, as those tasks are mostly automated today. While at times tedious and not engaging, the purpose of the monitoring tasks is to enable the operator to regularly acquire enough SA to be able to quickly and safely intervene in case of non-normal events. These operations are non-critical, in particular non-time-critical, and may be executed in any order, although pilots usually develop a task flow. Most operations are triggered by time, the environment, position, or external actors. By reason of the duration of the cruise phase in comparison to take-off, taxi, or approach phases (cf. [FAA05]), by the number and nature of operations, and by the variability in execution order, it is derived that as long as no non-normal states occur, Mental Workload (MWL) of the two pilots during cruise is minimal, cf. [NCT07, Wei13].

Due to the low mission-, flight-, and aircraft-related workload for a great part of the flight [Wei13], pilots need to actively keep themselves busy to maintain vigilance as well as psycho-physiological fitness to allow for timely and accurate response in case of need. An online survey conducted with 18 commercial pilots by Eiden [Eid17] reveals activities pilots today spend time on during cruise to stay awake and bridge phases of low workload: tactical planning, reading company briefing or training material, and non-flying-related activities such as chatting, eating, and reading newspapers. Of course, these engagements mean a shift of mental resources away from the primary task (monitoring systems), still, they are required as the primary task is not engaging enough to maintain vigilance, cf. [CS15, Wei13].

The introduction of SPO likely eliminates the possibility to chat with each other, unless communications are established with ground personnel. Deprived of human-human communications, and left with a non-engaging, repetitive, non-rewarding primary task, the future single operator is prone for reduced vigilance.

3.5 Research Gap

The thesis at hand looks at those human factors that are not or only insufficiently addressed in existing concepts for future commercial SPO and Reduced Crew Operations (RCO). This chapter has demonstrated that both MWL and engagement are minimal during the cruise phase (as long as no off-nominal events occur). The relation of MWL, human engagement, and vigilance has been elaborated in the previous chapter. In particular, the influence of mental underload (see subsection 2.5.1) on a vigilance decrement was shown. The task profile on the flight deck does not create optimal engagement, as humans are not good at monitoring tasks, cf. [CS15]. Instead, they disengage eventually and develop coping strategies, cf. [Wei13]. Combining today's low MWL profile during a two-pilot cruise, the further reduction of engagement opportunities through the removal of the second human from the flight deck, the vigilance decrement due to underload theory, and the effects that current monitoring tasks have on vigilance, negative effects on operator vigilance are hypothesized. This hypothesis will be further evaluated in this thesis.

The influence of removing the second pilot from the flight deck on the vigilance of the remaining pilot is not investigated to date. It is not surprising that no dedicated SPO concepts exist that address human factors issues related to low workload such as vigilance decrement and boredom during the cruise phase. In fact, most SPO concepts concentrate on flight phases characterized by high MWL such as ground maneuvering, take-off, initial climb, approach, and landing, and emergency situations. Although highly critical, during long-haul flight these phases account for only a small portion of the total flight time. Low vigilance is the main human factor implicated in fatal accidents [YHYH09]. Further adding more automation, as suggested by many concepts, will aggravate this problem and increase the probability of undesired vigilance levels by leaving a potentially bored, left-alone operator on the flight deck with a monitoring function [Sch15]. Besides decreasing performance, social acceptance due to perceived boredom, loneliness, complacency, monotony, motivation, and fatigue becomes a non-negligible barrier to the implementation of SPO / RCO in commercial operations, cf. [LDRDD12].

This thesis aims at quantifying vigilance levels of single pilots in commercial airline operations, and the comparison of those levels to DPO. As only very few events typically occur during long-haul cruise flights, traditional performance-based vigilance estimation methods cannot be applied to the problem at hand.

Regardless of the outcome, future SPO ConOps must create optimal levels of operator MWL at all times during the cruise phase, and thus allow for adequate levels of vigilance through providing engaging and meaningful tasks on the flight deck. Such a ConOps must be consistent, and detailed enough to be evaluable with regards to operator vigilance to close the research gap.

3.6 Research Hypotheses

It is the goal of this thesis to answer the question "*how bad is it?*" with regards to vigilance under SPO. To determine vigilance, associated mental states, and the change of vigilance over time in flight deck operations under SPO, two global hypotheses are derived:

Global Hypothesis 1: Pilot vigilance under SPO/RCO decreases significantly with time en-route when no critical events occur.

This hypothesis is derived from the pilot task profile and the removal of opportunities for engagement through the removal of the second pilot and today's flight deck operations in general. No direct measurement of vigilance exists, it must be inferred through other parameters. The following relations are expected based on the literature findings (see sections 2.5.2 and 2.5.4): Objective performance decreases with time en-route. Brain activity, measured directly through the Engagement Index (EI) and indirectly through Concentration of Oxygenated Hemoglobin (COH), decreases with time, alike Heart Rate (HR). Both Eye Blink Frequency (EBF) and Eye Blink Duration (EBD) are hypothesized to increase due to increasing fatigue and decreasing vigilance.

Potentials for meaningful human engagement, and more so time- or safety-critical events, hold the potential to increase human vigilance. The second global hypothesis validates this:

Global Hypothesis 2: Pilot vigilance under SPO/RCO increases significantly with the onset of a critical event.

Again, the following expectations aim at the individual estimation methods of vigilance: Brain activity (EI and COH) and related cardiac activity (HR) increase with the onset of a critical event. Likewise, EBF and EBD decrease.

This thesis further aims at validating and quantifying reduced vigilance during cruise flights under SPO / RCO compared to DPO. The respective hypothesis, based on the findings of this thesis, would read as: Pilot vigilance under RCO is significantly lower than under DPO. Comparing absolute physiological parameters, even intra-subject parameters, is not necessarily meaningful. In particular when obtained on different days, many unknown variables come into play and might falsify results. The evaluation will be limited to the relative changes in vigilance:

Global Hypothesis 3: Pilot vigilance under SPO/RCO decreases significantly faster than under DPO.

This statement will be evaluated using physiological measurements. According to literature, it is expected that EI, COH, and HR decrease faster under SPO than under DPO, while EBF and EBD both increase faster.

3.7 Summary

This chapter detailed the motivation to conduct a task analysis. The main methods, HTA and result representation through BPMN along with their benefits, were explained. Then, all plans and associated operations of pilot tasks during the cruise phase were presented. An assessment of flight crew workload revealed a generally low workload profile and that predominant operations are monitoring, communication, and planning tasks. Such tasks do not engage pilots enough over the long duration of the cruise phase under slowly changing environment parameters; pilots typically engage in further activities not necessarily related to the current mission and flying tasks.

The combination of the low workload profile and the nature of tasks itself, as well as the effects of a further reduction of flight crew under the same workload and task profile were identified as the research gap to be investigated in this thesis. Corresponding research hypotheses were derived, which will be tested in a simulator experiment. The next chapter details this experiment.

4 Vigilance on the Flight Deck

The purpose of the following two chapters is to answer the question *how bad is it?* in relation to the hypothesized reduced vigilance under Reduced Crew Operations (RCO). The research gap and the hypotheses formulated in sections 3.5 and 3.6 form the basis for the experimental analysis of vigilance on the flight deck, and directly shaped the experiment's design as presented in this chapter.

In this experiment, 10 subjects completed two simulated 4 hour cruise flights, one under each of two operating conditions (independent variable): Dual Pilot Operations (DPO) and Single Pilot Operations (SPO). Subjects were engineering students, who acted as pilots and executed today's pilots' tasks. The experiment was conducted in TECHNISCHE UNIVERSITÄT DARMSTADT (TUDA)'s AIRBUS A320 research simulator, D-AERO. Dependent variables were objective performance, subjective self-assessments, and five physiological parameters: Engagement Index (EI), Concentration of Oxygenated Hemoglobin (COH), Heart Rate (HR), Eye Blink Frequency (EBF), and Eye Blink Duration (EBD). To record these parameters, low-cost Commercial-Off-The-Shelf (COTS) hardware was used: INTERAXON's Muse™ headband for Electroencephalogram (EEG), EBF, and EBD, and BIOSIGNALS PLUX' Functional Near InfraRed Spectroscopy (fNIRS) sensor for COH and HR. Additionally, two GOPRO HERO 4 cameras were used to measure ocular parameters.

The herein described experiment was designed in accordance with the provisions set by the Declaration of Helsinki in its latest revision, as well as in accordance with TUDA's ethical standards. Each participant signed informed consent. TUDA's Ethics Board approved the experiment without conditions (approval EK 01/2019).

4.1 Experiment Design

In section 3.5, three global hypotheses were derived. In order to test these, a simulator experiment was designed. First, it was hypothesized that pilot vigilance decreases with flight time en-route when no critical events occur. Accordingly, the null hypothesis $H_{0,1}$ reads as follows:

$H_{0,1}$: Flight time has no significant effect on pilot vigilance when no critical events occur.

The experiment consisted of a simulated long-haul cruise flight over mostly monotonous terrain, little (routine) interaction with Air Traffic Control (ATC), no abnormal events, and no deviations from the original flight plan. The independent variable was experiment duration (cruise flight time). The dependent variables represented those measurable parameters (i.e. pilot vigilance) that were hypothesized to change by reason of experiment duration (flight time). In lieu of vigilance, performance to the Psychomotor Vigilance Task (PVT), and EI, COH (local oxygen saturation), HR, EBF, and EBD, were used as dependent variables (see section 2.5.4). Due to the setup used, Percentage of Eye Closure (PERCLOS) was not used in this thesis. Seven null hypotheses were derived:

$H_{0,1.1}$: Self-reported vigilance levels do not decrease over time without critical events.
$H_{0,1.2}$: Reaction times to the PVT are not significantly different before and after the flight.
$H_{0,1.3}$: EI does not decrease significantly over time without critical events.
$H_{0,1.4}$: COH does not decrease significantly over time without critical events.
$H_{0,1.5}$: HR does not decrease significantly over time without critical events.
$H_{0,1.6}$: EBF does not increase significantly over time without critical events.
$H_{0,1.7}$: EBD does not increase significantly over time without critical events.

An artificial interrupt task was administered during the cruise flight. To study the effects of such a situation on pilot vigilance, the task created immediate time-pressure, stress, and the need to decide and act quickly, similar to an emergency situation. Regarding this interrupt task, it was hypothesized that vigilance increases during and after the task, hence the null hypothesis $H_{0,2}$ is:

$H_{0,2}$: The interrupt task has no significant effect on pilot vigilance.

Concurrent with literature it was also hypothesized, that vigilance levels drop again soon after the test. Still, non-directional null-hypotheses to $H_{0,2}$ are used to account for the variability of human physiology:

$H_{0,2.1}$: Subjective vigilance levels do not differ significantly with the onset of a critical event.
$H_{0,2.2}$: The EI does not differ significantly with the onset of a critical event.
$H_{0,2.3}$: COH does not differ significantly with the onset of a critical event.
$H_{0,2.4}$: HR does not differ significantly with the onset of a critical event.
$H_{0,2.5}$: EBF does not differ significantly with the onset of a critical event.
$H_{0,2.6}$: EBD does not differ significantly with the onset of a critical event.

Last, it was hypothesized that pilot vigilance, respectively the change in vigilance, differ between DPO and SPO due to very little engagement between periods of required tasks (through the removal of the second pilot). The according null hypothesis $H_{0,3}$ reads as

$H_{0,3}$: The operational regime has no significant effect on the change of pilot vigilance over time.

To answer this hypothesis, the experiment design was expanded to cater for a comparison between the operational regime (independent variable: number of pilots). The dependent variables again represent those measurable parameters that were hypothesized to change by reason of the scenario (EI, EBF, COH, HR, EBF, and EBD, respectively their trends). Again, sub null-hypotheses are:

$H_{0,3.1}$: The subjective vigilance level does not differ between operating regimes.
$H_{0,3.2}$: EI trends do not differ significantly between operating regimes.
$H_{0,3.3}$: COH trends do not differ significantly between operating regimes.
$H_{0,3.4}$: HR trends do not differ significantly between operating regimes.
$H_{0,3.5}$: EBF trends do not differ significantly between operating regimes.
$H_{0,3.6}$: EBD trends do not differ significantly between operating regimes.

The complete study consisted of 10 experiments in total. Each experiment consisted of two scenarios (crew complement conditions), which differed by the operational regime (DPO vs. SPO). The scenario included a simulated commercial flight (see section 4.5); mission and environment parameters (time, weather, flight plan) were not changed between the two conditions. Independently of the operating regime, subjects' tasks were the tasks of today's pilots (see chapter 3), although some of them were simplified. Under SPO, the single pilot executed both the captain's and first officer's tasks. The flight trajectory was loaded into the Flight Management System before the experiment, and the autopilot executed this given trajectory; manual control was not necessary. Both conditions had to be completed by the same subjects to allow for intra-subject comparison of physiological parameters. To account for psychological factors (a participant undergoing the DPO scenario first might not be very motivated to do the same again under SPO) and fatigue effects, the order of the two scenarios per experiment was alternated. Experiment starting times were the same for all participants and scenarios (13:30 o'clock ±30 min). Dependent variables were only recorded for left seat occupants.

4.2 Participants

Due to the long duration of the experiment (see section 4.1) and relating ethical considerations of constraining participants to a monotonous environment for multiple hours, the associated difficulty in finding volunteers, limited budget for compensation, and the limited availability of the research simulator, the number of participants was limited to 16 (age 22.3 ± 2.0 years). Ten participants (age 22.2 ± 2.2 years) completed both conditions, four participants took over the role of first officer twice, two participants took over the role of first officer once (no data was captured for participants in first officer role). Table 4.1 shows the assignment of participants to the experiments and their condition order: four participants started with the DPO condition as first scenario, and six started with the SPO condition. Participants were assigned to either order based on schedule and their availability.

Table 4.1.: Experiment design with allocation of participants (P-ID).

#	Condition	Left Seat	Right Seat	#	Condition	Left Seat	Right Seat
1	DPO SPO	P-01 P-01	P-02 -	2	SPO DPO	P-03 P-03	- P-02
3	DPO SPO	P-04 P-04	P-05 -	4	SPO DPO	P-06 P-06	- P-08
5	DPO SPO	P-07 P-07	P-05 -	6	SPO DPO	P-09 P-09	- P-16*
7	DPO SPO	P-10 P-10	P-11 -	8	SPO DPO	P-12 P-12	- P-11
				9	SPO DPO	P-13 P-13	- P-14
				10	SPO DPO	P-15 P-15	- P-14

*Due to unplanned unavailability of P-08 for the second part, a new participant, P-16, took over.

To exclude as many unknowns as possible, selection criteria were applied based on section 2.5.2 and for practical reasons. Only male subjects were chosen. Both handednesses were allowed due to a limited number of available subjects. Eight subjects were right-handed, two were left-handed (P-06 and P-15). No participants with a history of neurological disorders or alcohol or drug abuse were allowed. For comfort reasons, the maximum subject height was limited to 2 m. To increase data analysis performance (facial landmark algorithm, see subsection 4.4.3), wearer of glasses were excluded (contact lenses were accepted).

Pilots were the natural candidates for this experiment. The long duration of the experiment, however, made it difficult to find pilots willing to sacrifice two free days for the simulator experiment, and often, pilot schedules did not allow for simulator sessions in between actual flights to adhere to rest time rules. Besides, general reservations of the pilot community towards the general topic of SPO exist [Air19]. Instead, participants were registered (engineering) students at TUDA. The experiments were advertised in selected courses offered by TUDA. Any kind of pilot license was not required as the experiment only replicated the cruise phase with the autopilot executing the given trajectory. Further expertise in flight deck operations was not necessary, operations were simplified for the participant group. All required information was given to participants during a briefing.

All participants took part voluntarily. Risks and benefits as well as data handling were explained before the experiment. All participants signed an informed consent form. They received compensation of 100€ after successful completion of both parts of the experiment. Since P-08 and P-16 took part in only one part, they were compensated with 50€ each. All participants received a briefing package and the informed consent and data privacy forms beforehand to familiarize themselves with the experiment and their tasks.

4.3 Apparatus

For the experiments of this thesis, TUDA's fixed-base Airbus A320 research flight simulator D-AERO with collimated outside view was used (see Figure 4.1). The flight simulator uses six COTS servers as illustrated in Figure 4.2 and four COTS monitors; the displayed elements match those in a real aircraft. The core is composed of Laminar Research's X-Plane Professional 10.X with the QPAC A320 plugin (realistic implementation of flight mechanics and aircraft systems) on *XPLANE*. Its source code was available and adapted subsequently for the experiments. Aircraft and flight data, such as position, attitude, and system states were communicated to the other servers using TUDA's datapool, see [Eng01], hosted on *CONVERTER*. For projection, simulation data was transferred via native X-Plane functions to another instance of X-Plane running on *VISION*.

Aircraft displays were hosted on three servers (see Figure 4.2), one for each large monitor: *DISPLAYCPT* and *DISPLAYFO* both showed the Primary Flight Dis-

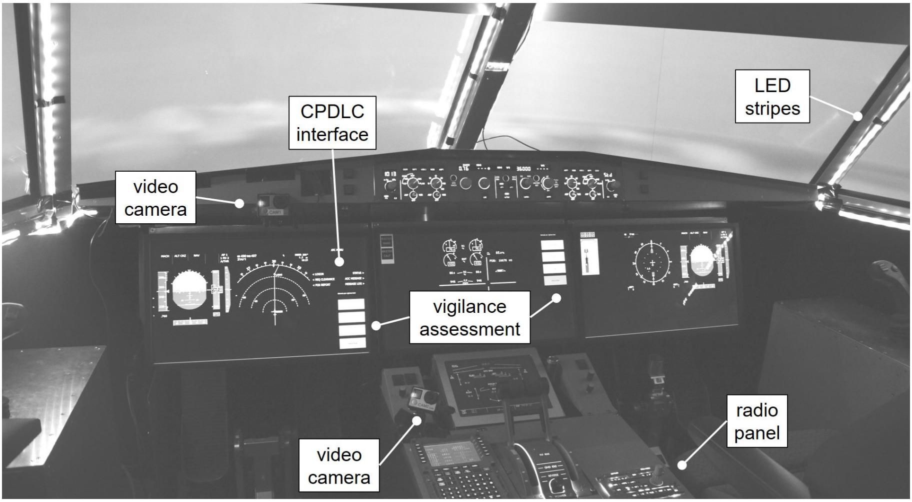

Figure 4.1.: D-AERO flight deck overview with experiment setup.

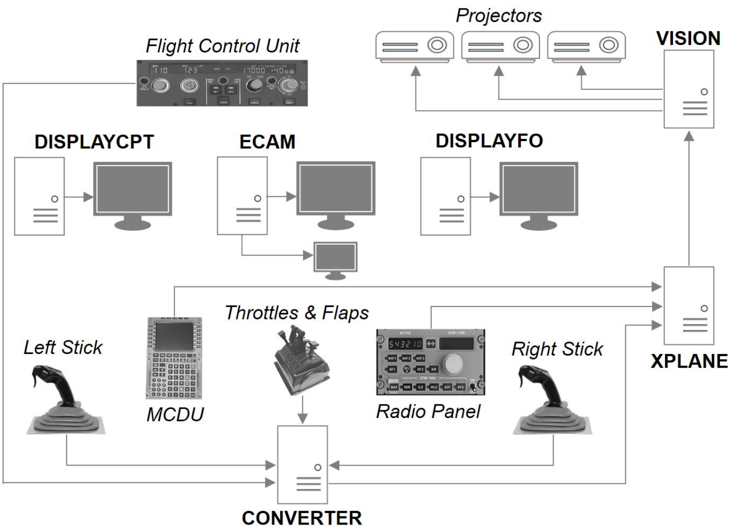

Figure 4.2.: D-AERO simulator architecture with servers, displays, and input devices. Not depicted are the datapool connections between the servers.

play (generated by open-source software XHSI) and the Navigation Display (generated with VAPS XT 4.0), *ECAM* hosted the upper and lower Electronic Centralized Aircraft Monitor (XHSI). The gear lever was simulated on *DISPLAYFO*. The Multipurpose Control and Display Unit (MCDU) with TUDA software acted as small monitor and input device to X-PLANE. Data from physical input devices (Flight Control Unit, throttle levers, side-sticks) was integrated via the datapool.

Specifically for the experiments, additional hard- and software was created and integrated through Universal Serial Bus (USB), User Datagram Protocol (UDP), or the datapool. These include a radio panel and an ATC chatter generator (STICK AND RUDDER STUDIOS' X-ATC-Chatter 1.6 plugin), vigilance assessment interfaces (see section 4.4.5), a graphical Controller-Pilot Data-Link Communication (CPDLC) interface, and PVT and interrupt task interfaces (see sections 4.1 and 4.5) on *DISPLAYCPT* and *DISPLAYFO*. Different to real aircraft, the CPDLC interface was moved to the captain's display to increase the time subjects looked into the frontal camera. Custom software running on an experiment laptop allowed the test administrator to trigger events, analyze PVT performance, and communicate with the pilots via prerecorded ATC commands (sector hand-overs) and CPDLC. Two cameras were installed on the flight deck for facial video recording and activity evaluation. To improve light conditions for these cameras (see section 4.4.3) and to simulate daylight conditions, Light-Emitting Diode (LED) stripes were attached to the cockpit fairing (see Figure 4.1). Due to limited graphics performance of VISION, a broken cloud layer was further added during the experiments to mask slow and limited loading of scenery.

The administrator station was behind the flight deck (see Figure 4.3), however separated from the flight deck by a thick curtain to create the impression of being alone on the flight deck. This station, consisting of multiple displays attached

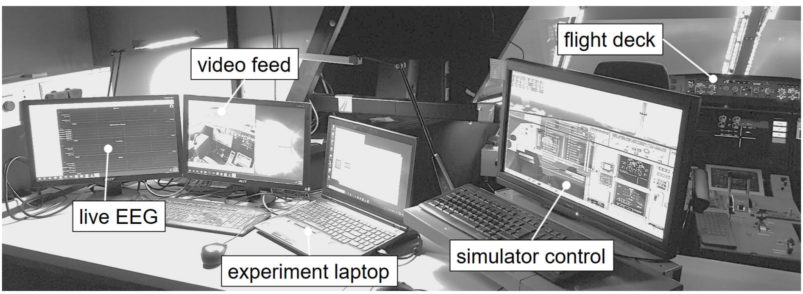

Figure 4.3.: Overview of experiment administrator station (curtain not depicted).

to two experiment laptops and the simulation software X-Plane, facilitated data recording and visualization, interaction with the pilots, event triggering (e.g. interrupt event), and condition manipulation (e.g. ATC radio chatter density). The experiment administrator was present at all times during the experiments.

4.4 Materials and Measures

Regarding the measurement methods, respectively the used hardware, a number of requirements limited the scope for potential candidate solutions. COTS devices were preferred, and they had to be mobile, non-intrusive, non-obtrusive, cheap, and easy to use. Wireless devices were preferred, their battery runtime had to satisfy experiment requirements (4 h). Raw data streams had to be accessible. Due to the explorative nature of the experiments, evidence of previous scientific value by the devices or hardware used was required. To avoid aliasing effects, sampling rates of at least twice the highest relevant frequency were required (e.g. 44 Hz for EEG gamma waves).

Regarding sensor hardware, price increases greatly with quality (see [Fun18]). With a budget of 2,000€ for this dissertation, hardware was mainly selected to enable a general proof-of-concept statement in this dissertation. If low-cost hardware can be used to draw conclusions of the vigilance state and its changes, then higher quality hardware (e.g. fNIRS sensors in the range of up to 10,000€, cf. [Sch19]) can then be used to draw differentiated conclusions. The following subsections detail the measurements methods and the used devices, objective methods to measure performance, and the interrupt task.

4.4.1 Neurological Metrics: Electroencephalogram

While professional (medical-grade) EEG caps include up to 256 electrodes (channels) to measure brain activity at different locations along the human scalp, the region of interest for this experiment is limited to the forehead (cf. section 2.5.4). Funck [Fun18] gives a high-level overview of COTS devices on the market. From this list, the InteraXon Muse™ 2016 headband was chosen (see Figure 4.4), mainly due to cost and battery run-time considerations: compared to other COTS devices, Muse™ has a long sampling time [GGP19]. This EEG headband weighs 60 g and offers a sampling rate of 256 Hz on four dry silver and conductive silicone-rubber electrode channels TP9, AF7, AF8, TP10, and one reference electrode at FpZ according to the international 10-20 system (see Figure 4.5). Data is transmitted via Bluetooth, the battery offers a runtime of up to 10 hours. [Int17]

Muse™ was originally designed as a wearable for meditation. Its software is free and allows access and visualization of both raw (potential differences) and processed data (absolute and relative band powers, artifact identification). The device was used and evaluated in various scientific studies, such as in [AKRP16, KWN+17, RWB+17]. Thus, Muse™ suits for application in the herein described experiments. It was connected to the experiment laptop via Bluetooth.

Figure 4.4.: InteraXon Muse™ 2016 device.

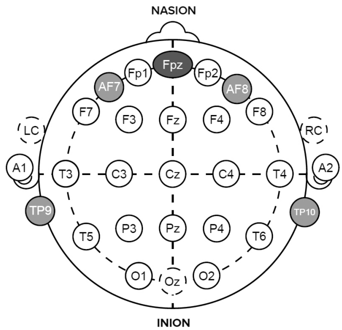

Figure 4.5.: Electrode positions, cf. [Int17].

4.4.2 Neuro-Vascular and Cardiac Metrics: Near Infrared Spectroscopy

Medical-grade fNIRS equipment as used in hospitals is expensive (> 10,000€). Due to limited budget, a low-cost alternative to record local blood oxygenation levels was found in Plux BioSignals Explorer kit in combination with an fNIRS sensor from Charles River Analytics. It was validated in a scientific study by Bracken et al. [BEFP+17]. This solution also offered continuous manufacturer support, ready-to-use sensors, data accessibility, and high signal-to-noise ratio. The sensor consists of two emitters for infrared and red light with peak emissions at 860 nm and 660 nm, and one photodiode detector with a sensitivity range between 400 nm and 1100 nm, see Figure 4.6. In the experiment setting, the fNIRS sensor was connected via cable to the BioSignals Hub, which in turn communicated via Bluetooth to the experiment computer. Data visualization and recording at 100 Hz was accomplished using Plux' free web-based software OpenSignals. The sensor was covered on its backside with black lightproof cloth to reduce light pollution. Subjects wore the sensor on the central forehead above the EEG headset.

Figure 4.6.: BioSignalsPlux fNIRS sensor consisting of two emitters and one photodiode (right), amplifier and analog digital converter (middle), and connection to BioSignals Hub (left).

As the sensor detects changes in the absorption of light, it can also be used to detect the pulse through periodic increases in Δc^{HbO_2} with each heartbeat, decreasing shortly after.

4.4.3 Ocular Metrics

To be maximally non-obtrusive to the subject, passive methods to detect eye blinks were chosen for this experiment. Available COTS eye trackers (Mirametrix S2 and Eye Tribe) were tested (see [Sch19]), and not found to be suitable for this experiment, as they were not able to uninterrubtibly record required parameters. They only allow for a limited area of (head) movement, which would have restricted subject head movement during the experiment too much. Head-mounted eye trackers had to be excluded due to interference with the chosen EEG and fNIRS headbands. Video-based methods only require a single camera mounted in front of the subjects to record the subject's face. In this experiment, a GoPro Hero4 camera was mounted directly in front of the subjects. It was recording video at 48 frames per second with a resolution of 1920x1080 pixels and a horizontal field of view of about 85° (linear field of view). These parameters were identified in own pre-testing and by [Sch19]. LED stripes were further used to increase facial lighting and contrast.

Existing facial land-mark detection algorithms were used to detect blinks and blink duration during post-processing. They capture characteristic points on a human face image, see Figure 4.7, by formulating a regression problem [Sv16]. Such algorithms are usually trained on a great number of different datasets to map from an image into landmark positions, cf. [XlT13, Sv16]. They are robust to varying illumination, video resolution, facial expressions, and moderate head rotations (landmark localization errors for state-of-the-art detectors are $<5\%$ of interocular distance) [Sv16]. For this application in particular, an algorithm developed by Soukupová and Čech [Sv16] was chosen. It was implemented by Rosebrock (see

Figure 4.7.: Facial landmarks, image after [Ros17]. Detected are face contours (points 1 - 17), eyebrows (18 - 22, 23 - 27), nose (28 - 36), eyes (37 - 42, 43 - 48), and mouth (49 - 68).

[Ros17]) in Python using open source libraries OpenCV and dlib. To detect eyelid movements (blinks), Soukupová and Čech's algorithm localizes eyes and eyelid contours, and then derives the Eye Aspect Ratio (EAR) from the facial landmarks to estimate eyelid opening state, see Figure 4.8 and Equation 4.1. When the eyelid closes, EAR is getting smaller. Additionally, a Support Vector Machine classifier is trained for blink identification and used to take into account larger temporal windows. [Sv16]

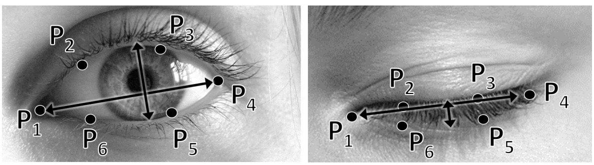

Figure 4.8.: Eye Aspect Ratio (EAR), image after [Sch19] and [Sv16]

$$EAR = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2||p_1 - p_4||} \tag{4.1}$$

Once the EAR is smaller than a defined threshold value, a blink is identified.

To account for the herein described experiment settings, the existing algorithm was expanded by Schott [Sch19] to limit EAR calculation to maximum horizontal and vertical head rotation angles to reduce false positives during head rotation (detection of eyes even if they are not visible) and to increase blink detection performance in the flight deck environment. At first, obtained EAR data was very noisy, hence comparison to a fixed EAR threshold for blink detection produced false positives. To consider small EAR changes of the eye open state over time during fatigue effects, an EAR 80% percentile moving average was calculated and used as EAR baseline (EAR_b). To detect blinks, the squared difference between actual EAR and baseline EAR was calculated (Equation 4.2). The factor k served to increase sensitivity.

$$EAR_{sq} = (EAR - EAR_b)^2 \cdot k \tag{4.2}$$

Head rotation was derived from relative distances between facial landmarks. For horizontal rotation, a line was drawn from P_1 to P_{17} (see Figure 4.9). Perpendicular to this line, a line through point P_9 was drawn. Hence, the line between P_1 and P_{17} was divided through the foot of the perpendicular P_P into two parts representing what is visible from the left ($\overline{P_1 P_P}$) and right part of the face ($\overline{P_P P_{17}}$), respectively. Since absolute values differ between persons, and also are dependent on the distance between the camera and the subject, the relation between those two distances was calculated as

$$h = \frac{\overline{P_1 P_P}}{\overline{P_P P_{17}}} \tag{4.3}$$

If the subject looked straight into the camera, both parts of the face were almost equal in size, hence $h \approx 1$. Since the head is not symmetric along the vertical axis, determining vertical head rotation was more complicated. It was approximated using the relation of chin to tip of the nose ($\overline{P_9 P_{31}}$) to tip of the nose to the foot of the perpendicular through P_9 to $\overline{P_1 P_{17}}$, P_P (as described before):

$$v = \frac{\overline{P_{31} P_P}}{\overline{P_9 P_{31}}} \tag{4.4}$$

Both relations h and v had to be within set boundaries for the algorithm to minimize false positives (detect eyes successfully). These boundaries were set for each individual and were gathered experimentally, but were generally in the ranges of $0.5 \leq h \leq 1.5$ and $0 \leq v \leq 1.5$. If h and v were outside limits, the algorithm would not calculate EAR. This extension to the original algorithm was tested and validated in desk experiments and in the simulator environments.

4.4.4 Performance Measure

To gain objective performance metrics to infer vigilance statements before and after the cruise experiment (see section 4.5), a Psychomotor Vigilance Task (PVT)

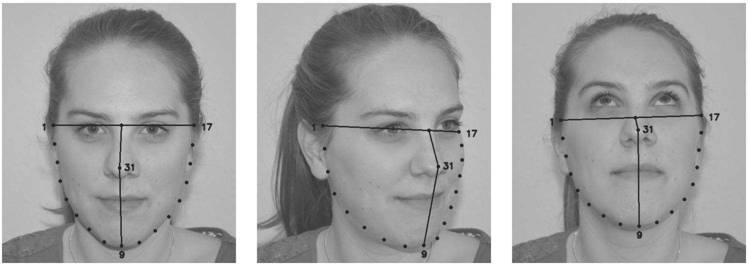

Figure 4.9.: Graphical construction of line relations between facial landmarks to constrain the blink detection algorithm to small head rotation angles. From left to right: no rotation, right head rotation, up rotation.

was used. PVTs are typically administered to track performance over time on task. A decline in performance during the PVT is then usually associated with a vigilance decrement. In this experiment the results of two 5 min PVTs were compared. Changes in performance were associated with fatiguing effects of the cruise flight task in between the PVTs.

The PVT was reimplemented in Qt 5.9.1 and administered on the left and right simulator displays for each subject (in direct line of sight for the camera). In this setup, a red square dot (approximately 2 cm × 2 cm in size) appeared at random intervals between 3 s to 8 s. Participants had to click anywhere on the screen using a COTS computer mouse connected to the respective server whenever the red dot appeared. The click made the red dot disappear. Participants received performance feedback in form of their response time in milliseconds for about one second after the dot disappeared. Response times were recorded and sent to the laptop.

4.4.5 Subjective Vigilance Assessment

To obtain subjective vigilance level assessments from the participants throughout the simulated flight, an assessment interface was placed on the front screens (one for each subject). The interface consisted of a four-point scale (buttons) with the four possible vigilance states labeled *very low*, *low*, *high*, and *very high*. As a subjective assessment of one's own vigilance state is rather difficult, a *neutral* option was omitted on purpose to prevent subject's choosing the neutral option most of the time when they were not sure. A red light above the four buttons signaled to assess the vigilance, which lit up randomly every 20 ± 4 min. Additionally, response time, measured from the time the light first illuminated to the time when the subject pressed one of the four assessment buttons, was recorded.

4.4.6 Interrupt Task: Abstracted Emergency Situation

Since participants lacked required knowledge and training for a real flight deck emergency situation, such a situation was abstracted under the paradigm of creating a situation in which the participant had to perform under time pressure. Time-constrained mental arithmetics tasks were chosen for this experiment. Triggered by passing a predefined position along the route, the left display went black, and math problems were displayed (in direct line of sight for the camera), accompanied by a sound signal to alert the subject. During this interrupt task, the aircraft proceeded along the route. Problems used basic arithmetic operations with between three and five terms in the number range up to 1000. Participants were instructed to think and calculate loud and state intermediate results. Time for each problem was limited to 12 s, and then the next problem was displayed, again accompanied by a sound signal. After 15 problems, the screen reverted back to the aircraft displays and the mission continued. Subjects did not know neither the exact time limit nor the number of problems before. Correctness of the solutions to the math problems was not assessed, as the focus of the activity was to create a time-critical activity associated with increased brain activity.

4.5 Procedure

All participants received a complete briefing package via e-mail several days before the experiment for familiarization. The experiment consisted of three main parts: experiment introduction and participant briefing, simulator session of first operational regime condition on day one, and simulator session of second operational regime condition on day two. As both conditions DPO and SPO consisted of the same tasks and were executed by the same participants, the briefing was only given once using a PowerPoint presentation. Figure 4.10 provides an overview of the complete experiment procedure with approximate durations.

The briefing consisted of an overview of the research motivation and background. The to be flown scenarios were introduced. Measures and measurement equipment were introduced in detail, and relevant instructions were given. As participants were not pilots, they received a brief introduction to the flight deck, which repeated the information they had been given before in the briefing package. Finally, participants signed the informed consent form. Following the briefing was the first condition simulator session (operating regime according to schedule in Table 4.1). Each session consisted of the application of the measurement equipment to the subject and testing. Before and immediately after the cruise flight, a 5 minute PVT was administered. After the completion, all captured data was stored;

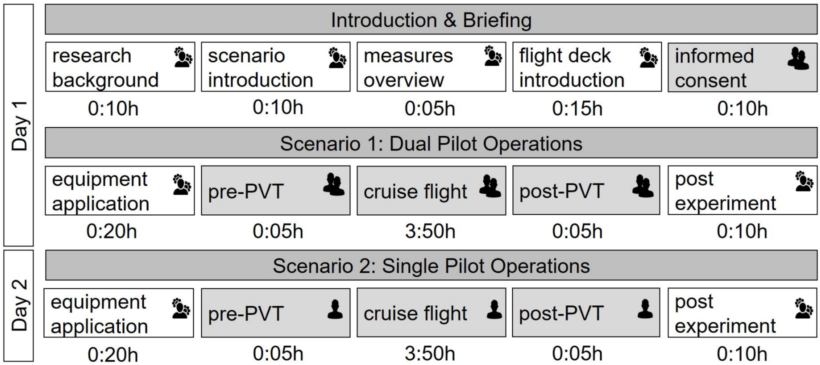

Figure 4.10.: Experiment procedure with detailed description and allocation of test administrator or subject(s). Phases marked in light grey were to be executed by the participant, those in white were led by the test administrator. Numbers below indicate approximate time allocation.

equipment was removed from the subject, cleaned, and disinfected. The second condition was administered on a separate day to minimize fatigue effects.

For the experiment, the first 4 hours of the cruise flight from Frankfurt (EDDF) to Fort McMurray (CYMM) were chosen, see Figure 4.11. Due to embedment of this research into JEPPESEN's RCO research efforts, the original mission included a flight from Frankfurt to Seattle. Due to limited range of the AIRBUS A320, the route used in the experiments ended in Fort McMurray. The simulated aircraft carried no payload and had almost full fuel tanks (18,250 kg). The flight plan was calculated with JETPLAN on 04/16/18 for a BOEING B737-800W. Relevant parameters, in particular fuel values, were adapted for the simulated A320 aircraft. The flight was optimized towards Long Range Cruise speed (Mach 0.76) at 36 000 ft, with a planned step climb to 38 000 ft at waypoint ORTAV. Conditions of the International Standard Atmosphere and no wind were applied. Participants executed pilots' tasks as presented in chapter 3. The left seat participant acted as Pilot Flying (PF), the right seat participant as Pilot Monitoring (PM) with tasks according to section 2.1.

The cruise phase itself can be split into multiple sub-phases based on events and geographical location, see Figures 4.11 and 4.12. The phases are the PVT before the flight (abbreviated as *P1* in the following), the first cruise phase (*C1*), the interrupt task (*IT*), the second cruise phase (*C2*), and the PVT after the flight (*P2*). The exact start and end times vary slightly between subjects and flights. For the first part, the pilot communicated with ATC via voice for sector hand-overs. Before entering oceanic airspace, the pilot had to switch to CPDLC. Communication was limited

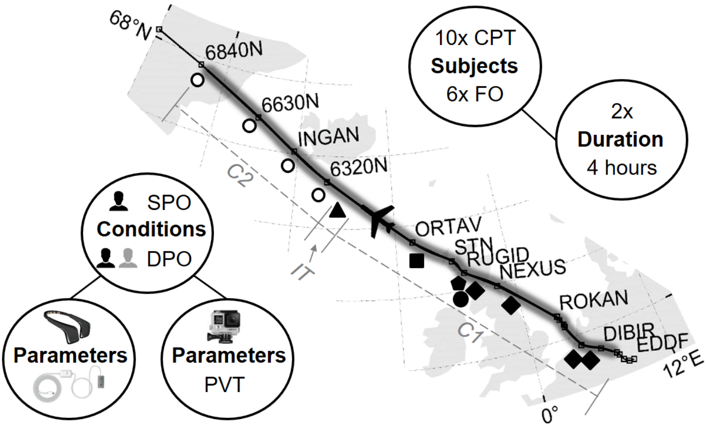

Figure 4.11.: Overview of simulated flight: highlighted is the simulated part of the route from Frankfurt (EDDF) towards Fort McMurray. Symbols represent events as shown in Figure 4.12, phases in gray text.

to the log-on process, requesting a clearance for a step climb, and half-hourly position reports. If they wanted to, the pilots could request weather information and communicate with the simulated airline's Operations Control Center (OCC) (experiment administrator) through the CPDLC interface. From waypoint STN on, the route was over water. Between waypoints 6320N and INGAN, Iceland was visible to the right. Soon after 6630N, Greenland became visible. At about 2.5 h, the interrupt (emergency) task was executed. From this point on, the only events were half-hourly CPDLC position reports.

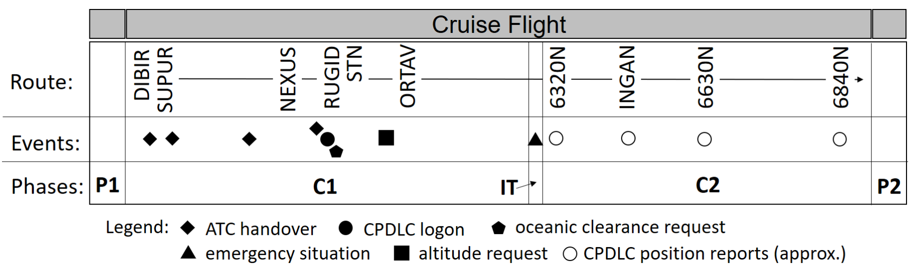

Figure 4.12.: Distance-based overview of events during cruise along the route. Events were triggered by position, except CPDLC altitude and oceanic clearance request and position reporting were at pilot's discretion.

Subjects received a briefing package and the printed flight plan for the experiment. The briefing package included a general overview of the experiment objective and procedure, a detailed description of tasks to be executed, and guides for reading the flight plan, using the radio and ATC phraseology, the CPDLC interface, the autopilot panel, and finding information on the aircraft's displays.

4.6 Data Analysis

All acquired data was recorded on two experiment laptops. EEG and fNIRS data were sent from the respective sensor devices via Bluetooth and marked with an absolute time-stamp. Bluetooth transmission times were neglected due to the analysis over very long times. Video data was later re-synchronized using audio signals sent and marked from the main experiment laptop.

A full data set from all sensors could not be obtained from all experiments. This is largely due to Bluetooth connectivity issues, disturbing sources (in particular electromagnetic interferences in the research simulator), known software bugs, and subject movements. These are detailed in the following sections.

All data analyses as well as statistical analyses were performed using MATLAB R2017b with its *Statistics and Machine Learning Toolbox*. Throughout the analyses, a significance level of 5% was assumed.

Preliminary data treatments are presented for all sensors in the following:

4.6.1 Electroencephalogram (EEG)

For the evaluation of EEG data, *Muse Direct* with *Muse*^{TM}'s built-in algorithms and signal processing pack *Muse Elements* and *Muse Lab* for live data visualization were used. Muse Elements computes the power spectral density of each frequency on each of the five channels using Fast Fourier Transform from the captured potential differences (measured in the μV range). A Hamming window of 256 samples at 220 Hz is used with a 90% overlap. The absolute band power for the given frequency band is the logarithm of the sum of the power spectral density over the respective frequency range (note that *Muse Elements* uses frequency bands differing from those stated in Table 2.1: Delta: $1-4$ Hz, Theta: $4-8$ Hz, Alpha: $7.5-13$ Hz, Beta: $13-30$ Hz, Gamma: $30-44$ Hz, cf. [Int15]). Power spectral density and absolute power band values are emitted at 10 Hz. [Int15]

Muse Elements further identifies and filters artifacts from eye blinks and jaw clenches, and considers data from its built-in accelerometers for artifact removal.

A 50 Hz notch filter was applied during the acquisition to remove environmental interferences.

Muse Elements' data was saved in a native Muse file format, and later converted into a comma-separated value file. As the artifact rejection by *Muse Elements* seemed to not be functioning correctly during certain times, additional artifact rejection was applied. This included a rejection of all values inside the Fast Fourier Transform's 90%-overlap time window when a blink or jaw clench was detected, or acceleration exceeded the following thresholds: $|a_x| \geq 0.4\,\mathrm{m/s^2}$, $|a_y| \geq 0.25\,\mathrm{m/s^2}$ and $0.95\,\mathrm{m/s^2} \leq a_z \leq 1.05\,\mathrm{m/s^2}$. These thresholds were determined through pre-testing. Values within those time windows were replaced by the values right before the onset of an artifact. As abrupt subject movements sometimes caused Muse™'s electrodes to lose skin contact for short periods of time, *Muse Elements* sometimes reported errors when calculating band powers without recovering. In case this lasted for more than 20 s, data recording was stopped manually, *Muse Direct* and *Muse Lab* were restarted manually, and data recording was resumed. Phases with no or corrupted data were later excluded from further analysis. Outliers were identified as values more than 1.5 interquartile ranges above the upper quartile or below the lower quartile, and removed from the EI data. To aid and facilitate interpretation of the large datasets of ~4 h per scenario, EI data was smoothed using a moving median over a window of 1,000 data points.

Data was re-referenced to the beginning of the experiment using absolute timestamps. $EI = \beta/(\alpha + \theta)$ were calculated for each time step from the absolute band power values, and smoothed for visualization. To make general trends visible and simplify interpretation, a linear regression fit of the form $m \cdot t + b$ was applied.

Blink artifacts were also used to calculate EBF and EBD from EEG data. These were calculated using *Muse Elements*' built-in machine learning algorithms and classifiers, and emitted at 10 Hz. As EEG data quality was generally low due to head movements and sensor quality, blink artifacts did not always reflect true eyelid movements. Whenever *Muse Elements* identified succeeding eye blinks with less than 0.75 s apart (this value was used based on video data comparisons), such blink artifacts were ignored.

4.6.2 Functional Near InfraRed Spectroscopy (fNIRS)

The photodiode converts incoming light into an electric current, which is converted to a digital value using a 16 bit resolution. Using the given transfer function for the sensor, current I_P is calculated from the digital values. Illuminance E_V is then derived with the spectral sensitivity $S = 80\,\frac{\mathrm{nA}}{\mathrm{lx}}$ of the photodiode:

$$E_V(t) = \frac{I_P(t)}{S} \tag{4.5}$$

The relation between illuminance E_V and luminous intensity I is linear and only dependent on radiant sensitive area A and solid angle Ω:

$$E_V(t) = \frac{I \cdot \Omega}{A} \tag{4.6}$$

As those two remain constant and cancel each other out in the ratio of light intensities at two times $I(t_i)$ and $I(t_{i+1})$, optical density $OD(t)$ (see subsection 2.5.4) is calculated using a linear relationship between E_V and I [Sch19]. Equation 2.4 then solved for $\Delta c^{HbR}(t)$ and $\Delta c^{HbO_2}(t)$ gives the relative concentration changes for oxygenated and deoxygenated hemoglobin for each time step. To aid interpretation, *virtual* absolute concentration values were assumed for the beginning of the experiment: $c^{HbR}(t_0) = 0$ and $c^{HbO_2}(t_0) = 0$. Virtual absolute concentrations were then calculated for $c^{HbO_2}(t)$ and $c^{HbR}(t)$, respectively:

$$c^{HbO_2}(t) = c^{HbO_2}(t_0) + \sum_{i=0}^{t} \Delta c^{HbO_2}(i) \tag{4.7}$$

In the data, periodic increases of COH (c^{HbO_2}) are visible. The time between these periodic increases reflects Heart Rate Variability (HRV). HR and HRV are not constant, but change continuously. Unexpected events (e.g. alarms, startling someone) usually increase HR. The increase to the maximum HR and thus the minimum time between R-R peaks takes a few seconds. From the raw data, local maxima in light intensities were extracted. Each local maximum represents one beat of the heart, and the time between the maxima Δt_i is HRV. This was used to calculate HR through the reciprocal of HRV:

$$HR(t_i) = \frac{60\,\mathrm{s/min}}{\Delta t_i} \tag{4.8}$$

Due to noise, not all local maxima could be successfully identified using MATLAB's algorithms. It was assumed that times between successive R-R peaks do not vary by more than 50% in young and healthy subjects, in particular not when HR is decreasing. Whenever times deviated by more, data was excluded from further analysis. Again, all results were smoothed for visualization.

4.6.3 Ocular Metrics

At first, video data was synchronized with experiment time using audio signals marking the beginning of the experiment. Then, the expanded facial landmark and blink detection algorithm was applied as described in section 4.4.3. Obtained EAR data was then imported into MATLAB for further analysis.

4.6.4 Performance

Response times, false starts, and lapses were recorded of a 5 minute PVT before and after each scenario. Mean response times and standard deviations were calculated for each PVT to be able to compare both PVTs using box plot representations. For the further evaluation, only response times were used. Evaluation only was applied for intra-subject comparisons.

4.7 Summary

Chapter 4 introduced the two-part experiment to answer the three research hypotheses. At first, null hypotheses were formulated from the global hypotheses, and sub null hypotheses for each dependent variable were derived.

The experiment was executed with 16 subjects, of which ten participants completed both conditions. Selection criteria, demographics, and their assignment during the experiments were presented. The subjects completed the same 4 hour cruise flight under two operating regime conditions: DPO and SPO. Objective performance, subjective self-assessments, and physiological parameters were recorded throughout the flights. The simulation environment at TUDA, D-AERO, was presented. The methodology and the equipment used during the experiments were presented, including the *Muse*™ EEG headband and PLUX BIOSIGNALS fNIRS sensor. The facial landmark detection algorithm used was presented, likewise the performance and subjective assessment metrics.

Finally, the experiment procedure was introduced and an overview of initial data treatment was given for each physiological parameter separately.

5 Results and Discussion

This chapter reports the experiment's results. Evaluation is performed with data from those 10 subjects who acted as left seat pilot and executed both scenarios (subjects P-01, P-03, P-04, P-06, P-07, P-09, P-10, P-12, P-13, and P-15). The evaluation is performed for each physiological parameter separately. As vigilance is highly dependent on personal factors as discussed in chapter 2, only trends are compared between subjects.

Subjective feedback indicates a significant difference in participants' ratings between the crew complement conditions Single Pilot Operations (SPO) and Dual Pilot Operations (DPO) and a desire for more engagement. Subjective vigilance assessments indicate a vigilance decrement over time during SPO, but also decreasing under DPO. The emergency event led to higher subjective vigilance assessments. Performance to Psychomotor Vigilance Tasks (PVTs) administered before and after the flights did not differ significantly. An evaluation of each of the three global hypotheses follows. Linear regression was used to determine trends. Engagement Index (EI) and Eye Blink Frequency (EBF) trends did not significantly differ from 0. Concentration of Oxygenated Hemoglobin (COH) and Heart Rate (HR) trends were significantly smaller than 0, indicating a decrease. The Eye Blink Duration (EBD) trend was found to be significantly increasing. Hypothesis 2 looks at the effects of the emergency event. EI trend was not significantly different from 0. COH and HR trends were increasing significantly. With regards to hypothesis 3, no significant differences between operating regimes in any of the five objective physiological trends are reported. Subjective vigilance assessments were significantly lower under SPO.

Objective and subjective results do not align. The findings indicate that crew complement has no significant influence on vigilance, but the task profile of today's flight deck operations during the cruise phase has.

5.1 General Observations & Subjective Feedback

Fourteen subjects completed two 4 hour simulator sessions each, two subjects completed one 4 hour session each. No one aborted. All 10 left-seat occupants completed both parts of the experiment. All left-seat subjects wore all sensors for the full experiment time. Half of the subjects indicated that the *Muse*$^{\text{TM}}$ headset was not comfortable to wear, in particular behind the ears.

While out-of-the-window graphics were of low quality, immersion was high. Although never explicitly addressed during the briefing, all but two subjects shut their smartphones off voluntarily and stowed them before the flight; the other two were asked not to use their phones.

No participant had experience in commercial cockpit operations. Two participants held gliding licenses, however, they stated that operations were quite different. During the first scenario, all participants were busy with getting familiar with the environment, the displays, Air Traffic Control (ATC) phraseology, and their exact tasks. During the second flight, task flow was more routine. When operating under DPO, subjects were chatting with each other most of the time. Over all 10 flights, about 16% of total flight time was spent in silence (with only occasional comments). Conversation topics included the experiment, cockpit operations, and aviation in general. Other topics included education and university life, personal background, interests, activities, work, and women. For all DPO conditions, conversations were extracted from the videos. Conversation topics were categorized in 5 minute intervals. These categories along with their portion of total flight time of all 10 DPO conditions are shown in Figure 5.1. Standard deviation of these values

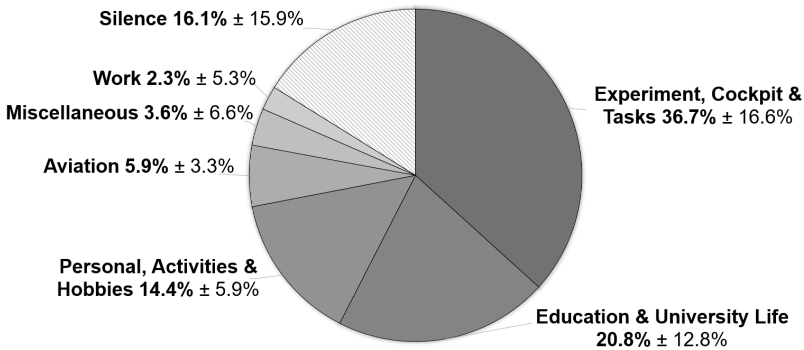

Figure 5.1.: Distribution of conversation topics over flight time for all participants during DPO. Standard deviation for all topics is given.

was high. In two DPO experiments, participants were continuously talking to each other, silence never lasted for more than two minutes, while two other participants spent more than 40% of the flight in silence.

All subjects completed a questionnaire with three five-point Likert-scale ratings and two free text questions after each flight. Answers are reported separately for the two operating regimes in Figure 5.2:

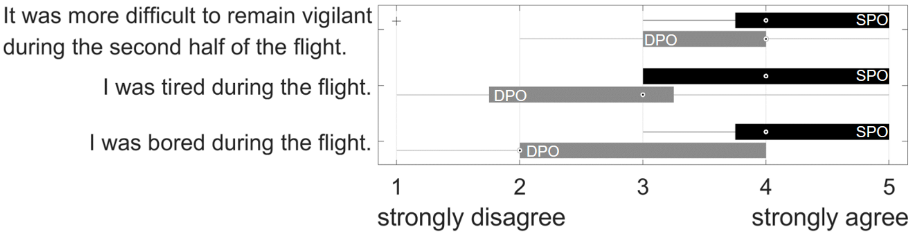

Figure 5.2.: Likert-scale box-plots of post-scenario questionnaire. Black color represents SPO condition, gray color indicates DPO condition. Shown are only answers from the 10 participants who executed both flights.

All six of the above shown data sets were non-normally distributed as assessed by the Shapiro Wilk test. Three Wilcoxon rank sum tests were performed for each question, revealing significance between SPO and DPO data for the fatigue ($Z = 2.3922, p = .0168$) and boredom question ($Z = 2.5846, p < .01$); subjects reported significantly higher boredom and fatigue during SPO. This gives a first hint towards different subjective assessments during the two operational regimes. There was no significant difference between operational regimes in the first question ($Z = 0.5751, p = .5650$).

In the questionnaire, participants were further asked what they would have needed to stay more vigilant during the flight. Figure 5.3 aggregates the top answers for each scenario with the number of occurrences per statement. The results are separated by the role participants were executing (captain vs. first officer).

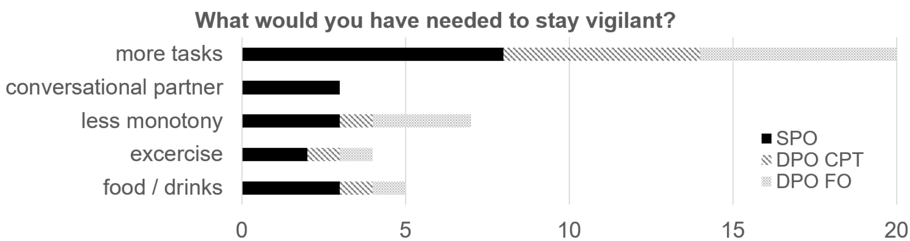

Figure 5.3.: Answers to the free text question. Solid black represents answers from the 10 participants during SPO, striped bars represent answers from the 10 participants during DPO, and checked bars represent answers from the six participants who acted as first officer during DPO.

Based on these results, the number, duration, and complexity of the tasks under both operating conditions were not enough to keep the participants vigilant. Observations further reveal that all participants created artificial tasks, such as constantly readjusting the display range of the Navigation Display to keep the next waypoint at the top.

In post-flight interviews, SPO participants reported that they had to "fight" to remain vigilant. Monotony, resulting from the task profile, the simulator graphics, and the route mostly over water, were mentioned by 8 of 10 participants during the interviews, although only 3 noted it in the questionnaires. After completion of both conditions, all participants stated in post-experiment interviews that DPO was easier, that they had higher vigilance levels, and that it felt like time progressed faster. Half of the participants acknowledged that higher vigilance levels were related to their engaging talks on unrelated topics and not related to the mission.

In the following, performance and physiological data will be analyzed to determine if objective data tells the same story.

5.2 Global Hypothesis 1: Vigilance Decrement under Reduced Crew Operations

The first global hypothesis is analyzed in the following subsection using subjective, performance, EI, COH, HR, and EBF and EBD data.

5.2.1 Results

To verify a hypothesized decline in vigilance due to the task profile of a typical North Atlantic cruise flight (number, complexity, and distribution of tasks over time, and their changes with progress en-route), and to be able to compare physiological data from multiple individuals, trends were obtained from linear regression. The statistical analyses performed in the following focus on *C1* trends only, this phase represents the task profile best: At the beginning of *C1*, subjects had to familiarize themselves with the flight deck environment and density of events (waypoint crossings, ATC communication) was high. Towards the end, fewer events occurred and time between events increased, and no major critical event occurred. Absolute physiological values are not further analyzed.

Subjective Vigilance Assessment and Objective Response Times

All subjects assessed their subjective vigilance levels 13 times during the flight (see section 4.4.5). Figure 5.4 shows the aggregated self-rated subjective vigilance for the 10 subjects who executed the SPO scenario. Appendix C.1 contains the individual assessments. The graphic shows that subjective vigilance decreases over time. Note that the times of the assessments varied slightly between participants. As expected, subjective vigilance increased after the interrupt event, which occurred between assessments 8 and 9 for all participants.

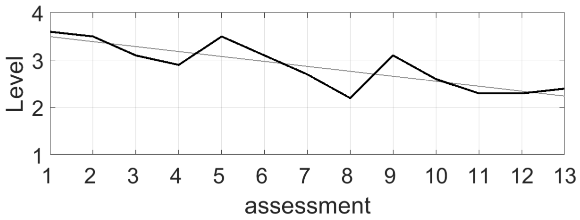

Figure 5.4.: Aggregated subjective vigilance assessments of SPO participants. The gray line represents a linear regression over all data.

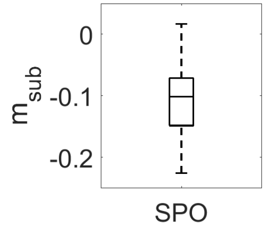

Figure 5.5.: Box-plot of linear regression coefficients m_{sub} for SPO.

A Spearman correlation analysis was performed on the time of assessment and the subjective vigilance over the complete flight, which revealed a moderate negative correlation; $r_S = -0.4414, p < .001$. This means, that subjective vigilance levels were assessed lower with increasing flight time (see also negative slope of the linear regression in Figure 5.4). A box-plot of the 10 linear regression coefficients (Figure 5.5) also shows this finding.

Response times to the assessment requests varied between participants and phases, ranging from 1.4 s to 283.2 s. Spearman's rank coefficient revealed a weak negative linear relationship between vigilance rating and response time, $r_S = -0.2346, p = .0072$.

Performance to PVT

Each subject completed two PVTs in each condition, one immediately before and one immediately after the flight. The performance-based evaluation is limited to an intra-subject comparison of PVT response times. Figure 5.6 shows box-plots of the response times of those 10 subjects who executed the SPO condition, both before and after the flight.

Figure 5.6 shows that reaction times are longer after the flight for subjects P-09, P-10 and P-12, and shorter for P-03. To statistically evaluate the effects of the flight on response times, a two-factor (*subject number* and *time*) repeated-measures Analysis of Variance (ANOVA) was conducted on the data including all outliers. Although response times were not normally distributed, as assessed by the Shapiro

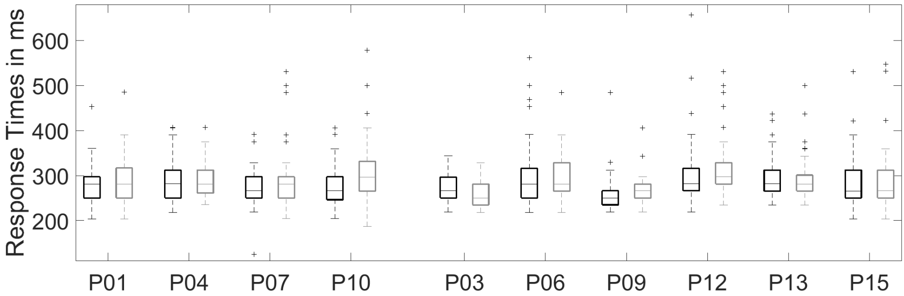

Figure 5.6.: Box-plots of response times to the PVT before (black) and after (gray) the flight in the SPO condition. Indicated are median response times, whiskers encompassing 2.7σ data, and outliers. The left four subjects started with DPO, the right six started with SPO.

Wilk test ($p < .001$) except the reaction time data obtained from P-03 before the flight ($p = .0531$), and data did not meet homoscedasticity requirements as assessed by the Levene test ($F(19{,}880) = 3.2938, p < .001$), the repeated measures ANOVA is robust. This is especially true when sample size is > 30 for each group [BAA+17, BS10].

Inter-subject comparisons are not of further interest and meaning, only the first factor, time of the PVT (before or after the flight), is relevant. The ANOVA revealed no statistically significant difference for the different conditions (before and after the flight), $F(1{,}880) = 1.44, p = .2602$. With outliers removed, the ANOVA still did not reveal statistically significant differences, $F(1{,}825) = 1.74, p = .2194$. In this case, response times were considered outliers if they were more than 1.5 the interquartile range above or below the upper or lower quartile, respectively. PVTs administered during the DPO condition report the same, non-significant, results (see Appendix C.10 for details).

Engagement Index

Data from all four electrodes were evaluated. Generally, electrode AF7 (front left) produced the most continuous data and was least impacted by artifacts. The AF8 electrode showed more data losses than AF7 due to the electrode losing skin contact. TP9 and TP10 often lost contact due to shifts in the cloth headband holding the fNIRS sensor in place. Therefore, only electrode AF7's data is reported. Similarly, and for the same reasons, only the EI of the form $\beta/(\alpha + \theta)$ is reported (see also [SBC+14]). All 10 subjects were included into the evaluation. Exemplary, P-01's data is shown in Figure 5.7 and discussed in the following. A complete set of graphs and results is given in Appendix C.2. For P-01, during *C1*, the EI had several

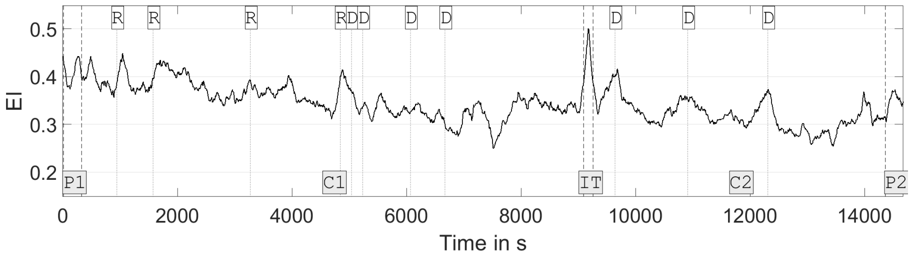

Figure 5.7.: Smoothed EI over time for P-01 during SPO with phases (bottom, dashed lines) and events (top, dotted lines; R: radio call, D: CPDLC).

peaks, but tended to decrease, which is also true for *C2*. Most local maxima can be attributed to events during the simulated flight, as marked in Figure 5.7 (time marked is the first message of each communication exchange only, answers are left out), such as radio calls, Controller-Pilot Data-Link Communication (CPDLC) messages, or cruise check activity when reaching a waypoint along the route. A linear regression, as described in section 4.6.1, was applied to the raw EI data. It verifies that EI is decreasing over time in *C1*. Overall, nine out of 10 participants show a decreasing trend of EI in *C1* (see Appendix C.2). The distribution of all *C1* regression coefficients for all participants is shown in Figure 5.8. The *C1* re-

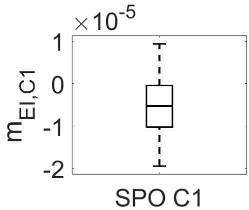

Figure 5.8.: Box-plot of EI linear regression coefficients $m_{EI,C1}$ for *C1*.

gression coefficients are normally distributed as assessed by the Shapiro Wilk-test ($p = .9039$). A one-sample t-test was performed to evaluate whether the coefficients have a mean equal to zero on a significance level of 5% (null hypothesis $H_{0,1.3}$), $t(9) = -2.1511, p = .0599$. This means that the null hypothesis cannot be rejected: the mean of the coefficient distribution is not significantly unequal to 0.

Concentration of Oxygenated Hemoglobin

Dimensionless virtual absolute concentrations $c^{HbO_2}(t)$ and $c^{HbR}(t)$, and linear regressions to COH were calculated for each participant (all graphs are aggregated in Appendix C.3). Data from P-01 is shown exemplarily in Figure 5.9:

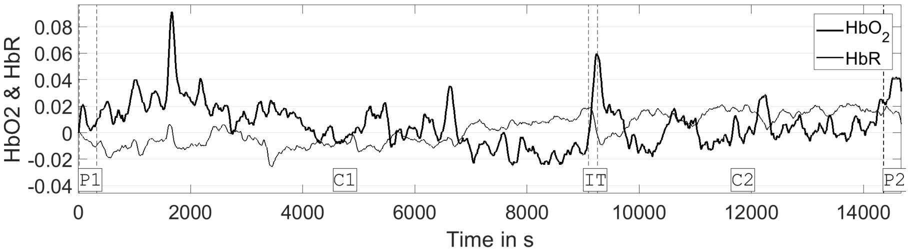

Figure 5.9.: P-01: Smoothed virtual absolute concentrations $c^{HbO_2}(t)$ and $c^{HbR}(t)$.

About half of the data from P-09 and short phases from P-12 are missing due to connection failures. P-09 was excluded from further analysis. Figure 5.9 shows opposite trends for oxygenated hemoglobin (HbO2) and de-oxygenated hemoglobin (HbR) (see section 2.5.4). These opposing trends are visible for only six of 10 participants. Figure 5.9 shows a negative trend for COH over time during *C1* for subject P-01. Nine out of 10 participants show a decreasing trend in *C1* (see Appendix C.3). Regression coefficients of *C1* were normally distributed as assessed by the Shapiro Wilk test ($p = .6664$). A one-sample t-test revealed a mean for the distribution which was significantly different from 0, $t(8) = -3.6961, p = .0061, d = -1.2320$. The effect size was large. Figure 5.10 shows a box-plot of the coefficients.

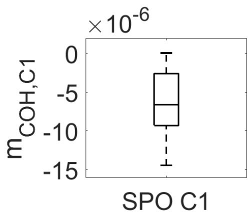

Figure 5.10.: Box-plot of COH linear regression coefficients $m_{COH,C1}$ for *C1*.

Heart Rate and Heart Rate Variability

As stated in subsection 4.4.2, Heart Rate Variability (HRV) and HR were extracted. As HR is assumed to be the simple reciprocal value of HRV (see subsection

4.6.2), only HR results will be discussed. Again, P-09 was excluded. Graphs for all participants are compiled in Appendix C.4. Absolute HR values are not of interest, as they include noise and artifacts; only trends are of interest. Linear regression was applied to the data, showing a decreasing trend for all participants during *C1*. These regression coefficients in *C1* were not normally distributed (Shapiro Wilk test: $p = .0291$). A Wilcoxon signed rank test was performed to determine if the coefficients were significantly different to 0: $z = -2.6656, p = .0077, r = 0.8885$. The mean was significantly different from 0, and the effect size was large.

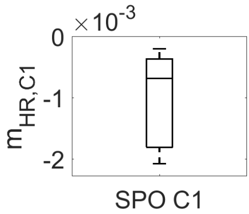

Figure 5.11.: Box-plot of HR linear regression coefficients $m_{HR,C1}$ for *C1*.

Eye Blink Frequency and Eye Blink Duration

To calculate blink rate as the number of blinks in a given time, face and eyes of the participants must be continuously detected. Due to an unexpected amount and continuity of head and body movement of the participants, and continuous partial obstruction of the face with hands (e.g. when supporting the chin with their hand, which was likely due to the uncomfortable situation in the simulator), the applied facial landmark algorithm was not able to detect eyes in about 50% of frames. This result was unexpected after the pre-tests with other subjects, which were performed with desk mounted cameras (see also [Sch19]), and in the simulator environment under experiment conditions. In particular during the DPO condition, participants were much more leaning towards each other, and thus to the very side of the camera's field of view, than those participants that took part in pre-experiment tests. Due to the long runtime (about 14 hours) of the facial landmark algorithm on the 4h long, about 60Gb videos, this only became obvious after the experiments. Two participants adjusted their seat back to an almost horizontal position during the experiment, which resulted in their eyes being outside the camera's field of view.

Instead, detected blink artifacts from the Electroencephalogram (EEG) headband were used to approximate eyelid movements. Data from all participants was used, linear regression was applied to aggregate and compare data. Graphs with EBF and EBD for all participants are compiled in Appendices C.5 and C.6.

With regards to EBF, four subjects show a positive trend (increasing EBF), while six show a decreasing trend in *C1*. *C1* EBF regression coefficients (box-plots in Figure 5.12) were normally distributed (Shapiro Wilk test: $p = .3228$), a one-sample t-test was performed. Coefficients were found to have a mean not significantly different from 0: $t(9) = -1.1258, p = .2894$. With regards to EBD, eight subjects show a positive trend in *C1* (increasing blink duration over time), two show a decreasing trend. The slopes of the two decreasing trends are of negligibly small magnitude compared to the positive coefficients. *C1* EBD regression coefficients (box-plot shown in Figure 5.13) were normally distributed as assessed by the Shapiro Wilk test ($p = .3228$), and a one-sample t-test was performed. Coefficients were found to have a mean significantly different from 0: $t(9) = 2.7163, p = .0238, d = 0.8590$. The effect size was large.

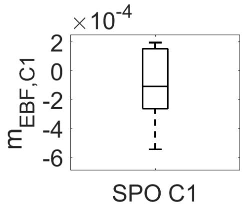

Figure 5.12.: Box-plot of EBF linear regression coefficients $m_{EBF,C1}$ for *C1*.

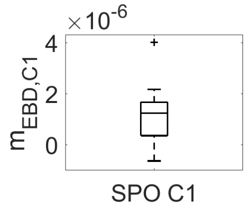

Figure 5.13.: Box-plot of EBD linear regression coefficients $m_{EBD,C1}$ for *C1*.

5.2.2 Summary of Findings and Discussion

With regards to a hypothesized vigilance decrement during the flight, one subjective parameter (vigilance self-assessment), one performance parameter (response time to a PVT), and five physiological parameters were investigated individually in this thesis: the declines in EI, COH, and HR, and the increases of EBF and EBD over time.

Due to the artificially inserted interrupt task as an abstracted emergency situation, and the focus on the task profile, only phase *C1* was evaluated. It included both engaging tasks at the beginning and long phases of no engagement towards the end. All findings are discussed in the following.

Subjective Vigilance Ratings and Response Times to Assessments

Humans react differently to monotony and boredom, and motivation to try to remain vigilant was different. The obtained results match with the expectations. From subjective data, $H_{0,1.1}$ can be rejected as the analysis revealed a moderate negative correlation between flight time and subjective vigilance levels (considered here was the whole flight, not just *C1*). As expected, flight time was proven to have a negative effect on pilot vigilance. With regards to response times, assessing vigilance was not a priority task and was only announced visually. If subjects were not looking at the displays for a few seconds, the cues went unnoticed for some time. In particular, half of the subjects took a long time to respond to the first cue as they were busy familiarizing themselves with the environment and their tasks and thus had their focus on the briefing package. The longest response time, 4:42min, was recorded from P-10 during the eighth cue. During this time, P-10 was sleeping involuntarily.

Performance to the PVT

According to the PVT response times, the flight under SPO did not have negative effects on performance. No significant differences were found, as results are interpreted as to the flight did not create an objective vigilance decrement. The null-hypothesis $H_{0,1.2}$ could therefore not be rejected. Whilst unexpected, similar results on multiple PVTs administered during long-haul flights were also reported by Thomas et al. [TGGC15] in Boeing's 2013 fatigue study, under similar conditions. The herein reported results may also be influenced by learning effects, resulting in shorter response times during the second, post-flight PVT. It might further be hypothesized that the short PVT, by its very nature, is unable to uncover vigilance decrement effects.

Engagement Index

Although nine out of 10 participants show a negative EI trend in *C1*, the mean of the distribution of regression coefficients was not found to be significantly unequal to 0, albeit close to the 5% level. This means null-hypothesis $H_{0,1.3}$ could not be rejected; flight time has no significant influence on EI. This outcome is unexpected due to the distribution of workload during the flight based on radio calls, the onset of CPDLC, and the increasing distance between waypoints (and thus cruise checks) over the Atlantic Ocean. One participant, P-10, involuntarily dozed off for a few minutes towards the end of *C1* (see gray area in Figure C.7, Appendix C.2). Evidence suggests that more subjects would have resulted in significant outcomes. Explanatory power of the obtained EI values is limited, however, due to inter-individual differences (see also Sauvet et al. [SBC+14]) and noise in the data. This noise resulted from the low-cost hardware, movement (causing dry electrodes to shift or lose skin contact), headband readjustments, and voltage fluctuations of the headset's battery during the experiments (fast battery draining and voltage fluctuations are a known issue with the *Muse*™ headband). Furthermore, the EI does not and cannot explain vigilance as a single measure, rather it is an indicator of arousal states (i.e. predominant frequency bands in overall brain activity). As elaborated in chapter 2, humans use multiple techniques to deal with boredom and monotony. Amongst them is mind-wandering, which clearly comes along with increased brain activity compared to a relaxed state. Increasing EI, while indicating increased brain activity in the beta-band, may indicate mind-wandering, and thus efforts to sustain vigilance. Due to the simple measurement methods employed in this experiment, such can only be hypothesized. If the subjects really were mind-wandering or had task-related thoughts cannot be known.

Concentration of Oxygenated Hemoglobin and Heart Rate

In theory, HbO2 and HbR should show opposing trends with a small phase difference, see subsection 2.5.4. This expected behavior can be found in the data of most subjects, but the data also contains parallel trends (see Figure 5.9: parallel trends at around 1800s and opposing trends during *IT*). The reason for this non-expected behavior is not known. Similar to EI and for the same reasons, explanatory power of COH and HR values as single measures for vigilance is limited. Increases in COH concentrations might be attributable to mind-wandering. Tables C.2 and C.3 report negative COH regression coefficients for *C1* for nine participants, and negative HR regression coefficients for *C1* for all participants. Statistical analyses revealed means significantly different from 0 with a large effect size for both dependent variables COH and HR. Null hypotheses $H_{0,1.4}$ and $H_{0,1.5}$ were rejected: Flight time has a significant influence on COH and HR, and this influence is of the expected negative nature. The influence of decreasing oxygen saturation of the environment can be assumed to be non-existent. A further evaluation of HRV through common HRV metrics, such as the Root Mean Square of Successive Differences (RMSSD), is left out due to limited data quality and questionable conclusions.

Eye Blink Frequency and Eye Blink Duration

Results regarding EBF and EBD must be treated with caution. These results rely on the correct identification of blink artifacts in the EEG, which are not completely reliable due to movements and the quality of the *Muse Elements* algorithm. A manual comparison of video data and blink artifacts revealed that blinks were not always reliably recognized. Additionally, artifacts were emitted by *Muse Elements* at approximately 10 Hz, so blink durations are not necessarily accurate. The statistical analysis failed to reject the null hypothesis $H_{0,1.6}$ (EBF is not significantly different with flight time), and rejected null hypothesis $H_{0,1.7}$ (flight time has a significant influence on EBD: increasing EBD with time). It is not known, whether these results are attributable to the correct observation of the effects, to limited data quality, or to the effects being simply not measurable in the experiment setting (or any other method or setting for this instance).

5.2.3 Conclusion

Subjective assessments indicate that there is a vigilance decrement during the flight. Table 5.1 compiles findings based on objective data only with regards to the first global hypothesis, both on a participant-by-participant and on an overall basis. The participant-related statements and their alignment or misalignment with expectations are based on the sign of the linear regression coefficients only, which do not necessarily bear universally valid information. The last line contains the statistical analyses for each parameter based on all participants.

Table 5.1.: Objective findings for Global Hypothesis 1 for phase *C1*. ✓ indicates alignment with expectation or rejection of null hypothesis, ✗ means data does not align with initial expectations (no rejection of null hypothesis).

Subject	PVT $H_{0,1.2}$	EI $H_{0,1.3}$	COH $H_{0,1.4}$	HR $H_{0,1.5}$	EBF $H_{0,1.6}$	EBD $H_{0,1.7}$
P-01	✗	✓	✓	✓	✓	✓
P-03	✗	✓	✓	✓	✓	✓
P-04	✗	✗	✓	✓	✗	✓
P-06	✗	✓	✓	✓	✗	✗
P-07	✗	✓	✓	✓	✗	✓
P-09	✗	✓	-	-	✗	✓
P-10	✗	✓	✗	✓	✓	✗
P-12	✗	✓	✓	✓	✗	✓
P-13	✗	✓	✓	✓	✓	✓
P-15	✗	✓	✓	✓	✗	✓
overall	✗	✗*	✓	✓	✗	✓

*Evidence suggests that additional subjects would have resulted in a rejection of $H_{0,1.3}$.

As discussed in subsection 2.5.3, vigilance cannot be inferred through one physiological parameter alone. Instead, multiple parameters must be combined. Yet the human physiology is sensitive to a variety of internal and external factors. Furthermore, large inter-individual characteristics make it difficult to compare results and derive universally applicable conclusions. With regards to *C1*, whilst a performance decrement could not be found in any of the 10 subjects, and the respective null hypothesis could not be rejected, data from three participants (P-01, P03, and P-13) show the expected changes in all five physiological parameters. Data from four participants (P-07, P-09, P-12, P-15) show those expected changes in four of the five parameters. Most unexpected data occurred with the two ocular parameters. These must be treated with caution as stated before.

These results are expected: subjects were engaged at the beginning of the experiment, and, in particular during the long route section between ORTAV and 6320N (see Figure 4.11), boredom was predominant towards the end of *C1*. Although the subjects being new to the flight deck environment certainly played into the results, pilots also experience a similar decline in engagement and tasks when transitioning from the take-off and climb phase to the cruise phase. It can be assumed that pilots would have been more invested in flight deck tasks.

A general statement regarding null hypothesis $H_{0,1}$ cannot be given. Based on individual data, results indicate that tasks such as communication to ATC temporarily increase vigilance, but also that vigilance decrements over time in non-engaging real environments. Statistical analyses on COH, HR, and EBD support this finding. This dissertation has shown that a vigilance decrement is measurable not only in laboratory environments with typical vigilance tasks (such as the PVT), but also in more realistic environments.

Although only three of the seven null hypotheses could be rejected on a 5% level, the results should not obscure the fact that physiological data of several subjects showed the expected behavior, and that subjective findings support the vigilance decrement hypothesis. Evidence further suggests that additional subjects would have resulted in a significance below the 5% threshold for $H_{0,1.3}$.

5.3 Global Hypothesis 2: Increased Vigilance due to an Interrupt Task

To evaluate the second global hypothesis, of interest are the short-term changes in the physiological parameters provoked by the interrupt task. The interrupt task was chosen to provoke sudden brain activity in subjects, associated with stress through time pressure. The physiological responses should consequently show a step towards higher activity (e.g. EI, COH) with the onset of the interrupt task (phase *IT*). Such can indeed be found in the EI and COH responses from some of the subjects, see for example Figures C.55 and C.71 (Appendix C). Such expected step responses are, however, not visible in the majority of the obtained data. This is for several reasons: First, physiological data is not necessarily exactly in synchronization with brain activity, which is, in particular, true for COH (as discussed in section 2.5.4: delayed Δc^{HbO_2} peak after activation onset). Second, the obtained physiological data is very noisy, which makes identification of steps difficult. Third, the exact time of the activation onset is not known. In the experiment, an audio signal announced the interrupt task, however, when subject's realized that they had to do mental arithmetics and actually started doing those is not known. And last, the magnitude of the step is dependent on individual characteristics, and might hence be too small to be identifiable.

In fact, five out of nine COH datasets (subjects) show a greatly increasing COH after the onset of *IT* for about 10 s to 15 s, which can be interpreted as a step response. With regards to EI and HR, a step-like response is only visible for two subjects. Some participants even showed decreasing EI, COH, or HR values with the beginning of *IT*.

Since step responses were difficult to identify under these conditions, and to still be able to make a general statement on hypothesized increased levels of EI, COH, and HR during a *IT*, a linear regression was applied to the data to again identify trends between the phases. For this, the physiological parameters of only the last minutes of *C1* (further identified as *C1r*) were compared to those during *IT*; duration of phase *C1r* matched the duration of phase *IT*. Linear regression coefficients were then used for statistical analyses. As subjects were quite engaged in mental arithmetics under time pressure during *IT*, it was expected that EI, COH, and HR increase significantly, while EBF and EBD drop.

5.3.1 Results

Due to insufficient data quality for EBF and EBD, and the short time frames for the evaluation of hypothesis 2, validity and reliability of results of both metrics is questionable. Both were not evaluated with regards to the second hypothesis. It must further be noted that some subjects moved their body constantly during phases of concentration.

Subjective Vigilance Assessment

Figure 5.4 indicates a subjective vigilance increase during and after *IT*, which occurred between the eighth and ninth assessment. In total, seven out of 10 participants reported a higher subjective vigilance level after *IT* than before; three reported a constant level after *IT* (see Figure C.1).

Engagement Index

Again, data from one participant, P-01, is used for illustration purposes. During *IT*, due to increased activity in the beta frequency band, P-01's EI increased and reached its global maximum, see Figure 5.7. This is also visible when looking at box-plots of EI values for each phase separately, which are reported in Figure 5.14.

Figure 5.15 shows the linear regression together with raw and smoothed EI data for *C1r* and *IT*. Appendix C.7 contains both graphs for eight other subjects as well

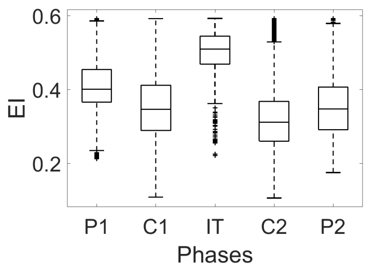

Figure 5.14.: EI box-plot for subject P-01.

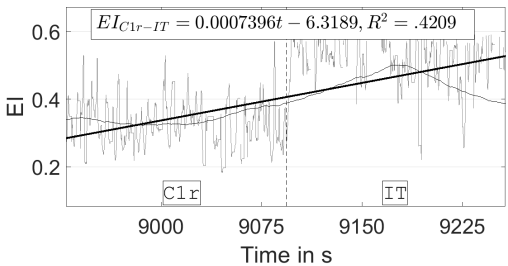

Figure 5.15.: Linear regression of EI for subject P-01 for *C1r* and *IT*.

as a table of all regression coefficients. Due to connection failures during *IT*, P-15 was not included in the analysis. Seven participants show an increasing EI with the onset of the interrupt task, while the remaining two show a decreasing EI. The coefficients as shown in Figure 5.16 were normally distributed as assessed by the Shapiro Wilk test ($p = .9581$). A one-sample t-test indicates that the sample has a mean not significantly different from 0; $t(8) = 2.2055, p = .0585$. The EI was not found to be significantly increasing.

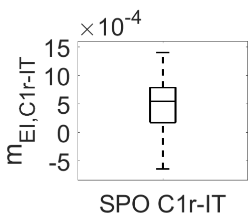

Figure 5.16.: Box-plot of EI linear regression coefficients $m_{EI,C1r-IT}$ for *C1r-IT*.

Concentration of Oxygenated Hemoglobin and Heart Rate

For the evaluation of COH and HR, data from one subject (P-09) was not considered due to a sensor failure during *IT*. For COH, Figure 5.17 shows a box-plot of regression coefficients. Appendix C.8 contains graphs for all subjects as well as a table of all regression coefficients. Figure 5.18 shows a box-plot of the linear regression coefficients for HR. Individual graphs and a table of coefficients are compiled in Appendix C.9.

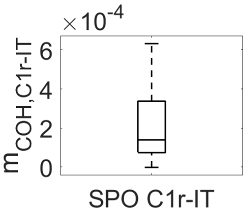

Figure 5.17.: Box-plot of COH linear regression coefficients $m_{COH,C1r-IT}$ for *C1r-IT*.

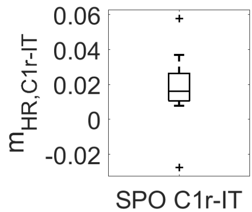

Figure 5.18.: Box-plot of HR linear regression coefficients $m_{HR,C1r-IT}$ for *C1r-IT*.

All participants showed an increasing COH with the onset of the interrupt task. A statistical analysis was performed to validate this with linear regression coefficients. They were normally distributed as assessed by the Shapiro Wilk test ($p = .0725$). A t-test revealed that the mean of the coefficients is significantly different from 0; $t(8) = 3.0023, p = .0170, d = 1.0008$. COH was increasing significantly, as all coefficients were positive. Effect size was large.

One subject shows a decreasing HR, all other eight subjects had an increasing HR with the onset of *IT*. Statistical analyses were performed on the coefficients. A Shapiro Wilk test confirmed normal distribution ($p = .2250$), hence a one-sample t-test was performed to determine if the mean was significantly different to 0: $t(8) = 2.3268, p = .0484, d = 0.7756$. With a medium effect size, this means that HR coefficients had a mean significantly different than 0. Over all subjects, HR increased over time.

5.3.2 Summary of Findings and Discussion

From Figure 5.4 it can be qualitatively derived that the interrupt task had an effect on vigilance, which manifests itself through either increased or constant subjective individual vigilance levels. Over all participants, vigilance increased (Figure

5.4, between assessments 8 and 9). Null hypothesis $H_{0,2.1}$ was rejected. Personal communication with the subjects after the experiments revealed that the interrupt task had different effects on subjects: seven felt immediately awake while three participants did not feel any different before and after the task.

With regards to EI, all subjects showed a change with the onset of *IT*, both for *C1* (box-plots in Figures 5.14, C.50 - C.64) and *C1r* (regression coefficients), although two subjects showed the change in the opposite than hypothesized direction. The statistical analysis indicates that the mean was not significantly different from 0 (meaning no change in *EI*), albeit the probability was close to the significance level of 5%. Null hypothesis $H_{0,2.2}$ could not be rejected: The EI was not found to differ significantly from *C1r* during *IT*.

As all subjects show an increasing COH, and the statistical analysis revealed that the regression coefficients were significantly different from 0, null hypothesis $H_{0,2.3}$ was rejected. COH increased significantly with the onset of the interrupt task. This result matches with expectations. Subjects had to perform mentally demanding tasks (math problems) under time pressure. Increased metabolism requires increased blood flow, which leads to a higher overall COH.

Although one subject showed a decreasing HR, which was contrary to expectations, eight subjects showed the expected behavior (increased HR during *IT*). The statistical analysis confirms this finding. Null hypothesis $H_{0,2.4}$ was rejected, meaning that HR differed significantly with the onset of *IT*. Table 5.2 summarizes all findings.

Table 5.2.: Overview of findings on Global Hypothesis 2. ✓ indicates rejection of respective null hypothesis, ✗ means failure to reject null hypothesis.

	EI $H_{0,2.2}$	COH $H_{0,2.3}$	HR $H_{0,2.4}$	EBF $H_{0,2.5}$	EBD $H_{0,2.6}$
overall	✗*	✓	✓	-	-

*Evidence suggests that additional subjects would have resulted in a rejection of $H_{0,2.2}$.

5.3.3 Conclusion

Based on the subjective statements, the failure to reject the EI-based null hypothesis, and the successful rejection of the COH and HR null hypotheses, a general statement towards $H_{0,2}$ cannot be made. When looking at trends and individual subjects, obtained data matches with the expectations. The interrupt task had effects on pilot vigilance, although not all effects were shown to be significant.

The employed method to detect changes in the physiological parameters between *C1r* and *IT* returns a statement if there was a change in those parameters (via the slope of the regression curve). What it cannot determine, is, if this change occurred at the onset of *IT*, and thereby confirming a hypothesized causal relationship between *IT* and changed the respective physiological parameter. An analysis of raw and smoothed data, however, shows that those changes occur at or close to the onset of *IT*, indicating that the linear regression method is valid.

5.4 Global Hypothesis 3: Comparison of Operating Regimes

Answering the third hypothesis is, due the complexity and variability of human physiology, not as straightforward. As physiological parameters were collected on different days, many influencing variables, such as duration and quality of sleep, play into the data. Absolute parameters are not compared as this might lead to unfounded statements. The evaluation focuses on relative changes within the two conditions per subject.

5.4.1 Results

Subjective Vigilance Assessment

Figure 5.19 reports the aggregated self-rated subjective vigilance for the 10 participants who executed both scenarios, over time. Individual assessments are compiled in Appendix C.1. For comparison, both operating regime conditions are displayed. As the graphic shows, while there still is a subjective vigilance decrement

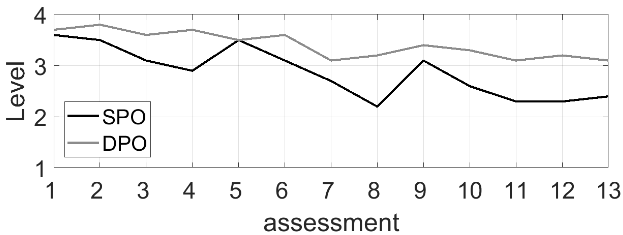

Figure 5.19.: Comparison of aggregated subjective vigilance assessments for all 10 participants who completed both scenarios. Black lines represent the SPO condition, gray lines the DPO condition.

over time in the DPO condition, it is of smaller magnitude than in the SPO condition. Figure 5.20 shows a box-plot with all absolute values. Neither the absolute

SPO nor the absolute DPO vigilance level assessments were normally distributed as assessed by the Shapiro Wilk test. The Wilcoxon signed rank test revealed a significant difference in absolute subjective vigilance assessments between the two operating conditions: $Z = -5.9825, p < .01$. Subjective vigilance assessments were significantly higher in the DPO condition.

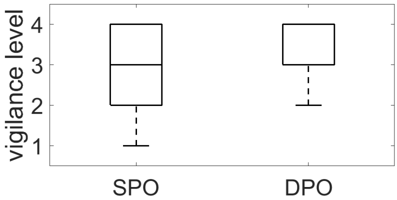

Figure 5.20.: Box-plots of absolute vigilance assessments for both conditions.

Response times in the DPO condition vary between 1.3 s and 66.7 s. A correlation analysis between vigilance assessment and response time during DPO again revealed a weak negative correlation; $r_S = -0.2359, p = .0069$.

Engagement Index

For illustration purposes, Figure 5.21 shows the EI over time for P-01 in both conditions SPO and DPO; graphs for all other participants are compiled in Appendix C.2. Trends are similar, in particular during *IT* and *C2*. Each peak in EI during *C2* corresponds to CPDLC position reports. While the EI during SPO is almost always lower than P-01's EI during DPO, the absolute values do not bear universally valid information, as the two conditions were recorded on different days, and many influencing factors were not controlled (such as level of fatigue before the flight, amount of caffeine consumed, etc.). For the majority of participants, absolute EI

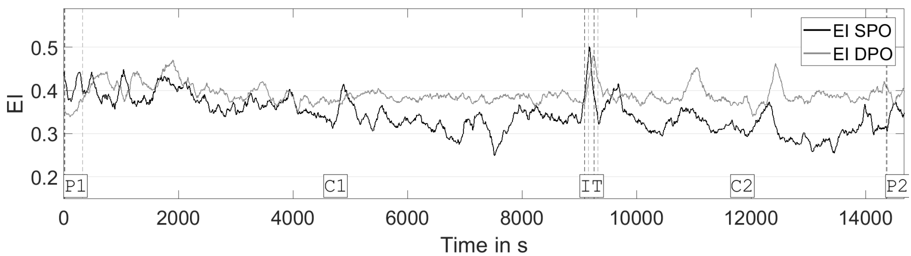

Figure 5.21.: Comparison of EI between SPO and DPO conditions for P-01.

values were greater in the SPO scenario than in the DPO scenario, which seems to contradict the hypothesis of lower EI values during SPO due to less engagement. Instead of absolute values, trends were compared between conditions. From the linear regression analysis performed on EI data (see subsection 5.2.1), regression coefficients, m, for *C1* have been calculated for each condition and subject.

Coefficients in the SPO scenario were normally distributed as assessed by the Shapiro Wilk test ($p = .9039$), the same applied to the DPO coefficients ($p = .4951$). A box-plot of coefficients (Figure 5.22) shows that coefficients in the SPO scenario are of higher negative value (albeit the differences are very small), indicating that the negative trend of EI is greater under SPO when compared to DPO. A paired sample t-test was conducted to compare *C1* correlation

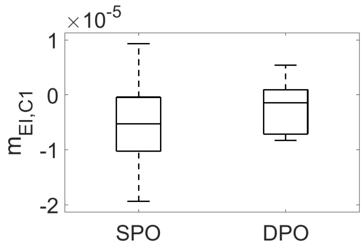

Figure 5.22.: Box-plot of EI linear regression coefficients for C1 in SPO and DPO conditions for all 10 participants.

coefficients in SPO and DPO conditions. No significant differences were found; $t(9) = -1.2253, p = .2516$.

Concentration of Oxygenated Hemoglobin

Exemplarily, Figure 5.23 shows the COH over time for P-03 in both conditions SPO and DPO; graphs for all other participants are compiled in Appendix C.3. It

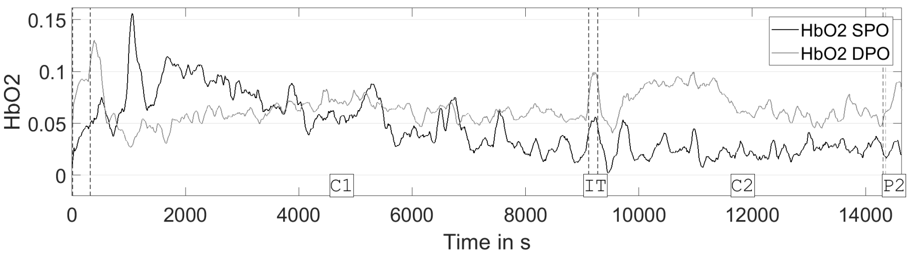

Figure 5.23.: Comparison of COH between SPO and DPO conditions for P-03.

can be seen that during *C1*, COH remains more or less constant during DPO (gray color). It decreases in the SPO condition.

Again, only trends (expressed as coefficients, m, of a linear regression performed on COH data in phase *C1*) were compared. These coefficients were normally distributed in both conditions as assessed by the Shapiro Wilk test (SPO: $p = .7493$, DPO: $p = .6187$). Figure 5.24 shows a box-plot of linear regression coefficients for eight of the 10 participants during *C1* under both conditions; note that for P-01 and P-09 no or very limited COH data existed for the DPO scenario and the SPO scenario, respectively. Both subjects were therefore excluded for this analysis. Again, the coefficients in the SPO condition were of higher negative value. A

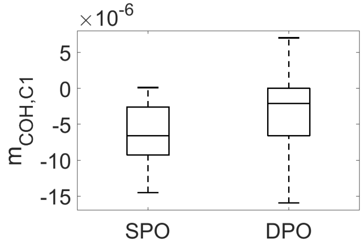

Figure 5.24.: Box-plot of COH linear regression coefficients for C1 in SPO and DPO conditions for nine participants.

paired sample t-test revealed no significant differences between conditions, however; $t(7) = -1.2477, p = .2474$.

Heart Rate

The same P-01 and P-15 were excluded as HR was derived from the same sensor. Coefficients both in the SPO and DPO condition were normally distributed as assessed by the Shapiro Wilk test ($p = .0791$ and $p = .1164$). Figure 5.25 shows a box-plot of linear regression coefficients for eight of the 10 participants during *C1* under both conditions. A paired t-test was performed, again revealing no significant differences between conditions, $t(7) = -0.9821, p = .3588$.

Eye Blink Frequency and Eye Blink Duration

Linear regression coefficients on EBF were normally distributed in both conditions as assessed by the Shapiro Wilk test (SPO: $p = .3228$, DPO: $p = .5495$). EBD coefficients were also normally distributed (SPO: $p = .4128$, DPO: $p = .4469$). Figures 5.26 and 5.27 show box-plots of EBF and EBD linear regression coefficients for all participants during *C1* under both conditions. EBF coefficients in the SPO

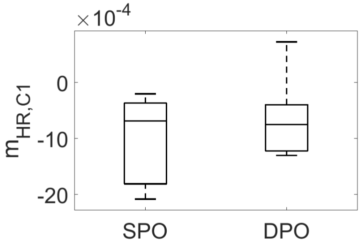

Figure 5.25.: Box-plot of HR linear regression coefficients for C1 in SPO and DPO conditions for eight participants.

condition were greater than under the DPO condition, indicating increased blink frequencies under SPO compared to under DPO. A paired sample t-test revealed no significant differences between conditions; $t(9) = 0.6514, p = .5311$. With regards to EBD coefficients, a paired sample t-test revealed no significant differences between conditions; $t(9) = 1.0629, p = .3155$.

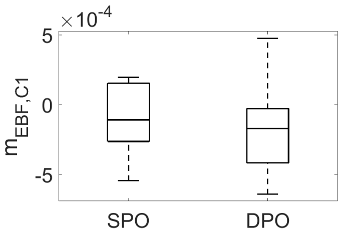

Figure 5.26.: Box-plot of EBF regression coefficients for C1.

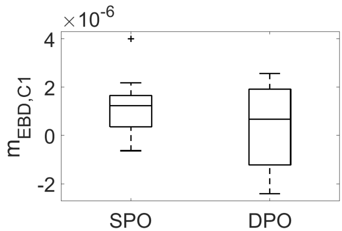

Figure 5.27.: Box-plot of EBD regression coefficients for C1.

5.4.2 Summary of Findings and Discussion

Section 5.1 already detailed subjective feedback from participants with regards to global hypothesis 3. In particular, participants rated boredom and fatigue to be higher during the SPO condition; both boredom and fatigue are closely linked with vigilance (see subsection 2.5.1), hence vigilance should be lower during SPO. Subjective vigilance assessments during the experiments confirm this: a significant difference between absolute subjective vigilance assessments was shown. Null hypothesis $H_{0,3.1}$ was rejected. However, this result should be treated with caution because of potential multifold interpretations of *vigilance* by the participants.

While subjects reported high vigilance levels mostly due to less boredom through the presence of a second person, their vigilance levels towards the flying and monitoring task were not necessarily higher than in the SPO condition. After the experiment, half of subjects raised the concern of being so deeply engaged with the other person, that their (subjective) Situation Awareness (SA) was sometimes even lower than during SPO. The distraction created through the presence of the second subjects was positive for engagement, but negative for SA, as outlined in section 2.1. Observations of the participants and their conversation topics during the experiments validate this finding. In particular when discussing controversial topics (such as the cost-benefit relation of comprehensive coverages for one's car), participant's mission awareness was observed as low.

With regards to response times, lower maximum response times in comparison to the SPO condition were expected. Vigilance assessment requests were first noticed by the second subject five times.

Regarding EI, linear regression coefficients were smaller in the SPO conditions, but the difference of EI trends was not significant. The analysis revealed no significant differences of COH and HR trends between conditions. The experiment failed to reject null hypotheses $H_{0,3.2}$, $H_{0,3.3}$, and $H_{0,3.4}$; the condition had no significant influence on the magnitude of the EI, COH, and HR trends. This is not expected; the reasons, however, may include data quality, different environmental conditions, and subject preconditions. It should be noted that HR is unspecific and vigilance effects are impossible to isolate. Thus, HR is potentially not suitable for this analysis. The experiment also failed to reject null hypotheses $H_{0,3.5}$ and $H_{0,3.6}$: Both EBF and EBD trends were not significantly different between operating regimes.

Table 5.3 summarizes the findings towards hypothesis 3. All hypotheses could not be rejected; this thesis did not reveal significant effects of the operational regime on pilot vigilance based on objective parameters. Only subjective parameters and self-rated vigilance levels show significant differences.

Table 5.3.: Overview of findings Global Hypothesis 3. A ✓ indicates rejection of respective null hypothesis, a ✗ means failure to reject null hypothesis based on the statistical tests against significance.

EI $H_{0,3.2}$	COH $H_{0,3.3}$	HR $H_{0,3.4}$	EBF $H_{0,3.5}$	EBD $H_{0,3.6}$
✗	✗	✗	✗	✗

According to these results, null-hypothesis $H_{0,3}$ could not be rejected. The operational regime had no significant effect on the objective change of pilot vigilance; objective vigilance under SPO is not significantly lower than under DPO.

5.5 Conclusion

These findings reveal the underlying, real problem to vigilance on the flight deck (regardless of the operating regime) and are useful towards the design of future commercial SPO: It is not removing the second human from the flight deck itself that causes a vigilance decrement. It was already shown in this thesis that SPO is common practice today in general, business, and military aviation (see section 1.1), hence the difference between SPO and DPO must be within acceptable limits with regards to overall safety, or the regulation authorities would not allow SPO. The real problem, it seems, is the vigilance decrement caused by the unsatisfactory human-automation teaming on today's flight decks and little human engagement only at distinct times in the progress of the flight over time. The current task profile over time does not satisfy vigilance requirements. This is supported by the fact that subjective responses and objective findings contradict each other. While the tasks were objectively manageable (all participants correctly executed all tasks), and carry the potential to increase vigilance levels temporarily, they do not serve to engage the human continuously throughout the flight. This causes the human to mentally disconnect in between required tasks. This happens also under DPO as the results indicate. The analysis found a significant difference between absolute subjective vigilance ratings between the conditions, with generally lower vigilance ratings under SPO. The analysis also revealed a decline in vigilance under DPO. So while two pilots can talk to each other and create opportunities for engagement, even today's flying-related tasks do not engage pilots enough during cruise to prevent boredom. Although it can be assumed that real pilots would have been more invested in flight deck operations than the subjects, this study's subjects might be closer to the future, less experienced SPO/RCO pilot.

To investigate interrelations and parameter trends under the different conditions, absolute physiological data should further be taken into account and compared between operating regimes in future experiments. For such experiments to be valid, influencing factors must be eliminated or controlled. Pilots should be used to validate the findings with regards to the participants themselves. Women, specifically, should be included to investigate potential gender differences of physiological parameters, in particular towards reduced vigilance. Finally, it is advisable to repeat the whole experiment under real flight conditions. Real conditions add stress to the operator, so physiological parameters might be more distinct. More robust hardware might be necessary for the real flight deck environment.

5.6 Summary

Chapter 5 presented and discussed the results from the simulator experiment. General observations and subjective feedback were reported, indicating a significant difference in participants' ratings between the two crew complement conditions SPO and DPO. Participants desired more engagement during the cruise phase. In the DPO condition, they spent most of the time chatting with each other on experiment-related and also on unrelated topics.

Regular subjective vigilance assessments indicate a vigilance decrement over time during SPO. Absolute values were significantly higher under DPO, but also declining. An abstracted emergency event led to higher subjective vigilance assessments. Performance to PVTs administered before and after the flights did not show significant differences.

An evaluation of each of the three global hypotheses followed. For the first hypothesis, only phase *C1* was analyzed. To determine and compare trends, linear regression was used. EI and EBF trends were not found to be significantly different from 0, while COH and HR trends were significantly different from 0, correlation coefficients were negative indicating a decrease. EBD trend was found to be significantly increasing.

Hypothesis 2 looked at the effects of the emergency event. Again, the trend of EI was not found to be significantly different from 0, indicating no significant changes. COH and HR trends were found to be significantly increasing. EBF and EBD were left out for data quality. Hypothesis 3 found no significant differences between operating regimes in any of the five objective physiological trends.

While objective and subjective results contradict each other, the findings were interpreted in the direction that the crew complement had no significant influence on pilot vigilance, but the very nature of today's human engagement only at designated times during the cruise phase had. Generating meaningful engagement throughout the mission is key for future SPO / RCO.

6 Concept of Operations for Single Pilot Vigilance

The conclusions drawn from the experiments indicate that the lack of continuous human engagement on the flight deck, and, even more, the current unsatisfactory integration of the human operator into flight deck automation are a barrier towards the implementation of commercial Single Pilot Operations (SPO) and the current efforts to further reduce the number of flight crew during long-haul flights. Sections 5.2 and 5.4 showed that there was a vigilance decrement on the flight deck when typical current operations are replicated, both under SPO and Dual Pilot Operations (DPO). Even higher levels of automation on the flight deck, as likely required for commercial SPO due to the risk of human incapacitation, will make matters worse.

This *irony of automation* (see section 2.4) must be addressed in a human-centric Concept of Operations (ConOps) for commercial SPO, which provides continuous opportunities for the human operator to actively and meaningfully engage in the operation of the complex system *aircraft* (see subsections 2.4 and 2.5.5). Such a ConOps for the cruise phase of commercial SPO was developed as part of this thesis; it is presented and elaborated on in this chapter. The goal was to achieve progress on those human factors challenges existing concepts deal with (see section 2.2), and to specifically address the vigilance-related findings from the study described in the previous chapters.

Assumptions and requirements will be derived for the expected operating environment. The ConOps is developed using an approach adapted from the approach to automation design proposed by Harrison, Johnson and Wright [HJW03]: Three agents, a mission manager, autonomous systems, and a ground operator are defined. The focus lies on the human operator's task and engagement profile during cruise, which chapter 5 found to be at the root of the vigilance decrement problem. Based on operational modes, functions are allocated among agents. Individual tasks are derived to keep the mission manager engaged in mission-related tasks to keep vigilance levels high, both subjectively and objectively. Besides, efficiencies towards the organization may be achieved through the empowerment of the human operator. Mission manager tasks include core tasks, Total Mission Management (TMM), Operations Control Center (OCC) support functions, and non-job-related tasks.

6.1 Assumptions towards Future Single Pilot Operations

A set of high-level assumptions towards future flight deck operations will help defining a new ConOps, cf. [NSK18]. This section introduces the most important.

First, it is assumed that one human operates on the flight deck. This is motivated by human's unique characteristics: Humans are creative, adaptable and bring common sense, and they can better deal with the unexpected, unanticipated, or complex situations than machines. Managing risks is the main reason why humans are still on the flight deck [DRPD17], cf. [LBB+17, MKL+15].

It is further assumed that airline operations and their goals remain the same as today. The air transportation system is assumed to have implemented changes as laid out in the Single European Sky ATM Research Programme (SESAR) and Next Generation Air Transportation System (NextGen), in particular 4D-trajectories, self-separation, and System Wide Information Management technologies.

The existence, economic availability, security, and usage of high bandwidth communication between air and ground and air to air are assumed. This includes the use of Controller-Pilot Data-Link Communication (CPDLC). Next, autonomous flight capabilities are assumed. The ability to monitor (reliably detect and diagnose in real-time) the operator's health to react to pilot incapacitation, cf. [LBB+17], is assumed.

6.2 Justification and Requirements Derivation

Future SPO must meet several requirements. A selection is detailed in the following.

Safety is the most important requirement that an SPO ConOps must facilitate. Safety could be measured e.g. in terms of number of fatal accidents per vehicle-miles or similar metrics (see [OK16]) and in terms of required probabilities of occurrence of events per operational hour ($10^{-9}/h$ for flight control functions, cf. [SAE96, FAA00]). Federal Aviation Administration (FAA) AC 25.1309 may guide the design and evaluation process to achieve the necessary level of safety.

Requirement 1. *The ConOps shall enable operations at least as safe as current 14 CFR Part 121 commercial aviation.*

Next to safety is efficiency. The ConOps shall enable efficient operations in three dimensions: airline cost, trip reliability, and human operator workload derived from the previously described experiment.

Requirement 2. *Airline direct operating cost shall be less than today.*

Requirement 3. *Trip reliability, measured in terms of percentage of canceled and delayed flights, shall be comparable to current Part 121 operations.*

Requirement 4. *Operator workload shall always be within acceptable limits at all times.*

The latter requirement relates to the Yerkes-Dodson-law (section 2.3) and a derived minimum level of human engagement on the future flight deck.

Previous studies on SPO, e.g. by Etherington et al. [EKB+16], and challenges reported in section 3.5 and Appendix A.2 regarding safety, loss of redundancy, and social aspects during cruise strongly indicate that removing the second pilot "will require significant redesign of automation and increased levels of automation support" [EKB+16]. Today's operations and flight deck design cannot provide the required levels of safety with only one pilot. Instead, an SPO concept must fundamentally change the role of the human operator, the nature and schedule of work undertaken on the flight deck, and the flight deck itself, cf. [HSS15, WG15, LDRDD12, DP05]. In fact, even today there is a mismatch in pilot roles, requirements, actual needs, and tasks on the flight deck [NSK18]. A new ConOps must address this mismatch by redefining the operator's roles and tasks.

Requirement 5. *Operator role, tasks, human-human and human-automation interaction shall be designed to ensure high levels of vigilance, operator engagement, mode awareness, and recognize human cognitive capabilities and limitations, cf. [BKK+17].*

This requirement relates to both the tasks themselves and to the schedule in which they are executed. The human must be allowed times of mental disconnection from the mission and to relax, but the ConOps must not create boredom (see subsection 2.5.5).

Furthermore, the operating environment (flight deck) must be adapted to new roles, tasks and needs:

Requirement 6. *The future flight deck shall be designed for one operator according to their role, tasks, and needs.*

6.3 Definition of Agents, their Roles, and their Goals in the future System

This section details rationale for and the concept itself. The development process consisted of four steps: context familiarization, vision development, incubation, and synthesis (for a detailed description, see [NSK18]).

The ConOps includes three agents (an entity, human or machine, which may act independently, cf. [CB09, Min06]) that are directly involved in mission execution: a human operator (*mission manager* as single pilot), autonomous systems, and a ground operator. The proposed architecture, depicted in Figure 6.1, is based on concept category G (cf. section 2.2). In the herein presented concept, operator roles are newly defined, and broken away from traditional pilot roles. A description of each agent follows. While autonomous systems and ground support are only touched on, the focus lies on the human.

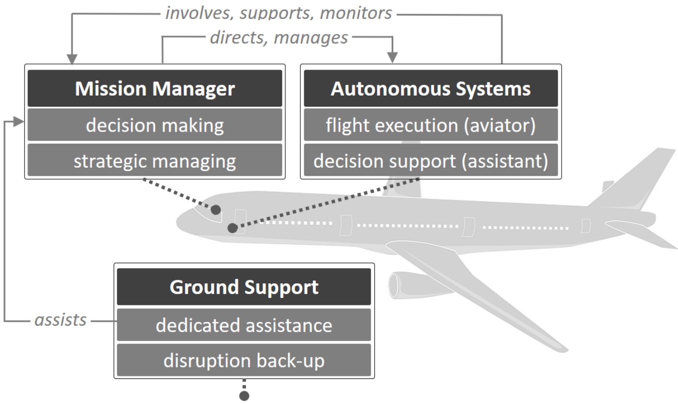

Figure 6.1.: Agents involved in future SPO and their relations.

6.3.1 Mission Manager

Traditionally, pilots' tasks are categorized into aviating, navigating, communicating, and managing systems (ANC+S). With the introduction of sophisticated automation, systems monitoring and systems management play a growing role today. It is unfitting to call pilots "aviators" or "pilots" anymore. On the contrary, they are much more automation managers [Bha10], exception handlers [Sch15], or system monitors, as the experiment conducted as part of this thesis has shown (see chapters 4 and 5). In particular, the role of system monitor is critical, as humans are not made for monitoring tasks [CS15]. The experiment proved that vigilance decrements during typical monitoring tasks.

Against this background, it makes sense to redefine the human operator's role. Due to their unique characteristics, they are put in the center of the concept, with automation built around the human. Human operators will no longer operate under the traditional ANC+S scheme, but instead focus on *Total Mission Management*. In fact, they are largely decoupled from direct flight control functions. To better match the new role, pilots will be called Mission Managers (MMs) in the future. This terminology has been used before, e.g. by Lucas et al. to describe "agent-based planning & control" [LRRK04] for intelligent autonomous unmanned aerial vehicle architectures.

The role of the human operator shifts the focus towards strategical, macro-managemental functions. This is the first step towards a meaningful opportunity for human engagement. Following future developments in air traffic management, literature recommendations, legal considerations, and social and ethical considerations, the MM remains to be the responsible instance for the mission and thus an essential part of the normal and non-normal operation, cf. [SGW17, MKL+15, Don01]. As Pritchett, Kim and Feigh [PKF14] and Woods [Woo88] argue, the human will also have the authority and means to take over flight control functions, acting as safety net. This allocation of responsibility and authority is made at the design level of the SPO concept and thus fixed for all possible situations, cf. [DRPD17]. Summarizing, the definition of the future operator reads as follows:

Definition 1. *The MM is the authoritative decision-making and responsible instance of a mission. A MM defines mission objectives and manages mission planning and execution on a strategic level safely and efficiently. The MM acts as the link between operating context and automation, cf. [Har07]. The MM is a safety net to the mission and manages corner cases.*[1] *[SNS18]*

The definition of the operator role incorporates sub-roles. In relation to those defined by Cummings and Bruni [CB09], the MM can be seen as a moderator, initiating operations and decision-making processes. To fulfill their role successfully, MMs must build and maintain *mission awareness*. Mission awareness can be seen as a broader term for Situation Awareness (SA) regarding the current mission, includes mode awareness, and requires a certain degree of operator vigilance. The MM also has the sub-role team member, which highlights the need for communica-

[1] With regards to the given definition, a *mission* is defined as "to transport both passengers and cargo from a departure airport A to a destination airport B", cf. [PRP+95]. *To manage* is defined as "to handle or direct with a degree of skill" and "to make and keep compliant" (Merriam Webster Dictionary: "to manage", accessed May 24, 2017, https://www.merriam-webster.com/dictionary/manage). The scope of *safely* and *efficiently* has been given in section 6.2. *Strategic mission execution* refers to mostly knowledge-based cognitive planning and deciding functions that serve to achieve overall mission success, cf. [LRRK04]. Hence, MMs achieve their goals by directing and observing ("managing") other agents in the system.

tion, coordination, and shared resources. They are also human individuals, hence anthropometrics must be integrated carefully. [PRP+95]

Likely, future MM have not much in common with today's pilots, as the job evolves from aviators to strategic mission and disruption managers. Similarly, job requirements, candidate traits, and training curricula will have to be adjusted.

6.3.2 Autonomous Systems

Autonomous Systems (AS) are key to enabling the role transformation from pilot to MM, and take repetitive tasks away from the human operator to allow for more meaningful engagement. With technological capabilities assumed, AS are the executing instance regarding the mission. Summarized under this umbrella term are all systems operating that take over traditional ANC+S and micro-management tasks. AS are (to a certain degree) self-governed and self-directed systems that have intelligence-based capabilities [CKO+14]. AS perform the two core functions *aviator* and *assistant* and are defined as:

Definition 2. *AS execute the mission and operate the aircraft safely and efficiently, in accordance with mission objectives ("aviator"), and based on MM decisions. AS actively involve the MM into the operation, support decision making ("assistant"), and facilitate human engagement to prevent vigilance decrement and boredom.*

AS support MM decision making through adopting the role of *generator* according to Cummings and Bruni [CB09]. The main function is to implement MM, joint, or own decisions (decision implementation usually affects the flight execution function). AS execute micro-management tasks such as diagnosing and resolving malfunctions. At all times, AS observe time, cost, and comfort constraints set by the MM. Besides, they monitor the MM throughout the mission. If required, AS can execute a mission abortion autonomously.

6.3.3 Ground Operator(s)

While the focus of this thesis is on the aircraft side only, Ground Operators (GOs) are defined as follows: GOs are a single point of service entity, similar to current airline OCC, however with an enhanced service portfolio. GO collaborate closely with the MM: being humans in the system, they are a go-to-point to the MM. Usually, each mission is accompanied by a GO who monitors MM and AS; a GO serves multiple missions at the same time. The role definition reads as follows:

Definition 3. *GOs provide dedicated assistance to the MM to support them in obtaining their goals. GOs absorb workload peaks and temporarily substitute the MM during MM indisposition.*

6.4 Modes of Operation

Although AS take over flight execution functions, the MM always remains the responsible instance. With only one human operator remaining on the flight deck, and natural times of operator indisposition occurring (e.g. biological breaks, phases of reduced concentration, fatigue), three different Modes of Operation (MoOs) are introduced for future flight deck operations. The MoOs describe a framework for function allocation among the three agents, but are always centered on the MM and define their involvement in the execution of the mission. As required, the MoO is adaptable to current MM needs during the mission. Such a mode change schedule is negotiated before and continuously during the mission. Mode changes may be implemented manually by the MM whenever required, according to airline Standard Operating Procedures (SOPs), or based on AS determining MM incapacitation. Being the last authority on board, the MM may override AS-initiated mode changes. The three MoOs are described in the following:

Full Mission Awareness and Involvement (FMAI) (mission focus) is required during phases of strategic decision-making, critical and non-normal operations, and to build and maintain awareness. Critical phases include, in particular, taxiing, take-off, and landing. Mission focus implies that the MM devotes all cognitive resources to the current mission and can make decisions with no lead time. It does not mean that the MM executes all operations.

Reduced Mission Awareness and Involvement (RMAI) (support focus), in particular during the cruise phase, allows the MM to direct attention and resources to other job-related and non-job-related tasks (detailed in the following section). This mode serves to even out workload during the mission, and to stimulate operator vigilance through human engagement in meaningful tasks. Phases of reduced involvement are interrupted by phases of FMAI to facilitate human strategic decision-making and mission awareness. Whenever non-normal events occur, the MM must change their focus to the mission and transition to the FMAI mode. RMAI tasks must be easily interruptible and always have a lower priority than FMAI tasks.

Minimal Mission Awareness and Involvement (MMAI) (rest focus) occurs during involuntary indisposition (health issues, incapacitation) or preplanned voluntary indisposition (sleep, controlled rest, biological break). GOs are informed.

Mode changes, independent of manual, schedule- or need-based change, will always be communicated in real-time to the MM and the assigned GO. If the MM transitions to the FMAI mode, the change will be communicated with appropriate lead time to allow for situation familiarization if possible.

A key point is to uphold *mode awareness* among all agents, a challenge that has historically not been mastered to a satisfying degree in aviation as accidents such as Asiana flight 214 prove, see also [HD95]. According to Sarter and Woods [SW95], mode awareness describes the user's knowledge of a system that contains information on the system's states and behaviors, and on parameters that characterize them. Generally speaking, mode awareness refers to the correct linking of all required information into the mental model depending on the task [BDG+14]. As the MM will shift away cognitive capability from the mission towards other tasks during RMAI, it is of importance that the system communicates transparently what it is doing and why, and which interactions or decisions from the human are required at each point in time.

6.5 The Mission Manager's Tasks

Based on the requirements for SPO derived from the vigilance experiment (in short: continuous meaningful human engagement), the new role definition, and the introduction of MoO, the Mission Manager (MM)'s tasks are derived in the following. They are made up of four functional categories the MM will engage in. With even higher automation levels than today, the main objective of these tasks is to meaningfully engage the operator more with the operation. The immediate benefits are twofold: objective vigilance decrements over time are reduced due to continuous and meaningful human engagement, and subjective and objective vigilance estimations are aligned through the upvalue of role and tasks on the flight deck. The new task profile helps to minimize the gap between subjective and objective results found in the previously described experiment.

The four categories likely increase the human operator's workload above today's workload profile and above suggested future automation-approaches during the cruise phase in literature. Still, this approach was chosen following Schutte et al. [SGC+07] and Hancock [Han13] to even out workload across the whole mission, to prevent boredom, and to increase subjective job satisfaction, increasing performance.

Core functions are functions pilots are executing on today's flight decks, such as lowering the landing gear before landing. Such functions will remain on future flight decks, they might, however, be reallocated to other agents. Some will remain with the human operator to facilitate mode and situation awareness. The next section elaborates on the reallocation method and its results.

Total Mission Management (TMM) includes all functions directly related to and derived from the new role. Three functions, planning & scheduling, implementation, and monitoring & evaluation, cf. [CFG+17], are introduced in section 6.5.2. Specifically, they serve to update the human operator's mental model (mode awareness) through active involvement (decisions and actions) in tasks that have significant consequences on the overall mission, its safety, and efficiency [SGC+07].

Operations Control Center (OCC) support and other job-related functions during phases of RMAI facilitate human engagement in meaningful tasks during long cruise periods, cf. [Sch15]. Airline operations support functions relate to functions currently executed in OCCs. Section 6.5.3 details the method used to investigate which of those functions can be transferred to the MM; results and additional job-related tasks are also presented in section 6.5.3.

Non-job-related functions executed during RMAI or MMAI serve to stimulate motivation and balance adequate levels of vigilance and fatigue. Examples of such tasks are provided in section 6.5.4.

6.5.1 Dual Pilot Operations Task Reallocation

The function reallocation for core functions on the flight deck must aim at generating the highest benefit in terms of reducing operator Mental Workload (MWL) during phases of high MWL, and facilitating meaningful operator vigilance and involvement during times of low MWL. Therefore, the design decision has been made to largely decouple the MM from direct flight control functions during normal operations. This also enables the MM to engage in strategic management functions. The function reallocation must further be aimed at reducing those automation-induced problems described in section 2.4.

Reallocation Method

Parasuraman, Sheridan and Wickens [PSW00] outline an automation design framework (Figure 6.2), which is derived from Endsley and Kaber's Level of Automation (LoA) taxonomy, cf. [EK99]. The framework provides an objective basis for the decision of which system functions (information acquisition, information analysis, decision and action selection, and action implementation) should be automated and to what extent, and its evaluation. Within each function class, various levels of automation may be applied from fully manual (level 1) to fully automated (level 10), cf. [SV78, PSW00] and Appendix A.1. Evaluation criteria focus on human performance consequences through MWL or SA metrics of a specific design decision [PSW00]. Deciding on and evaluation of the LoA is an iterative process, which usually results in a range of possible automation levels. Using this framework, each pilot task during cruise, identified in section 3.3, was analyzed and reallocated.

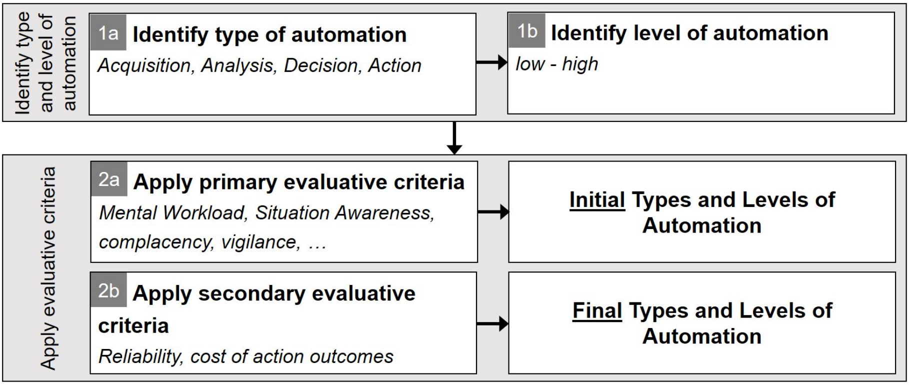

Figure 6.2.: Automation design framework after PARASURAMAN, SHERIDAN AND WICKENS [PSW00], [own illustration].

Reallocation Results

Exemplary, a possible reorganization of operations in the "periodic checks" plan (see Appendix B), which suits the above stated function allocation goals, is presented in Table 6.1. For each operation, the LoA range recommendation is given for each of the four function classes (information acquisition, analysis, decision, implementation). Results and consequences from the human performance evaluation are detailed. All reorganization results are incorporated into the scenarios presented in section 6.7. Similar to today's operations, information acquisition usually is automated. Since AS must execute a (mission abort) trajectory autonomously in case of MM indisposition, analysis can be performed by AS, however the MM shall be involved to facilitate mission awareness. In general, decisions are split between AS with a maximum level of 7 (includes communication of results to human) and the MM, again to facilitate mission awareness. Concurrent with the MM role definition, strategic decisions are made by the MM with AS support ("assistant" component). Tasks to be carried out by the MM will become part of future SOPs. Implementation is part of the AS function domain.

Table 6.1.: Proposed function allocation in terms of levels of automation (cf. Appendix A.1, Table A.1) for the operations in the "periodic checks" plan. Levels refer to *acquisition | analysis | decision | implementation*.

Operation	Levels	Human Performance Evaluation
check displays	10|4-10|7|5-7	AS self-test, communicate non-normal states. MM shall check regularly.
update log	10|10|10|7-9	AS log parameters automatically, communicate selected parameters.
compare against predictions	10|7-9|7|4-7	AS perform task but include MM in decision-making process. Full automation during MM indisposition.
check engines	10|10|3-7|7	AS self-test. MM shall check regularly. AS involve MM in case of non-normal status.
plan escape route	2-10|2-7|2-7|7	MM shall plan escape route. AS may require MM to plan route. In case of indisposition, AS execute autonomously.
check weather	10|10|4|4-7	AS acquire and display weather automatically, joint decision-making of suitability.

Reallocation Consequences

The reallocation results lead to three main findings: first, the reallocation of core functions guarantee that the MM continues to be the responsible and authoritative instance for the mission. Although several function categories are mostly automated, AS transparently communicate information, not data, on the current state, methods, and results to enable the MM to assume any function if required, and to facilitate operator mode awareness. Second, the reallocation results in less physical and psychomotoric tasks such as manually dialing-in values on the autopilot control panel, but in an increased number of cognitive tasks. AS must be designed to optimally support this increased amount and complexity of cognitive work. The acceptance of this work profile with the new mission managers must be validated, and other forms of work involving less cognitive work must be introduced to ensure a balance of work forms on the flight deck and to counter operator fatigue (see section 6.5.4). And third, through the increased use and autonomy of AS, the MM has less core tasks to accomplish during cruise than today. The presented reallocation allows the MM to perform additional, different, and meaningful functions on the flight deck while still being in control of the mission. These functions include TMM (see section 6.5.2) and airline operations support (discussed in section 6.5.3).

6.5.2 Total Mission Management

Concurrent with the new role, the MM performs new functions: management of mission planning and execution. TMM consists of three phases as described below. All three phases are iterated continuously during the mission.

Planning & Scheduling: The MM is actively involved in and responsible for mission planning, an enhanced function today partly executed by dispatchers. In this phase, the MM is supported by the "assistant" component of AS. Mission management includes flight planning and monitoring, and the management of crew, maintenance events, weather events, passenger and cargo, turnaround and disruptions, and resources through negotiating the MoO schedule. Flight planning includes route and fuel planning, flight plan generation and filing with Air Traffic Control (ATC), alternate airports evaluation, and weather monitoring. The responsibilities of today's pilots are expanded beyond the flight itself.

Plan & Schedule Implementation: Plans need to be implemented. Strategic targets and constraints are derived for the safe and efficient operation and mission execution, which the MM delegates to AS for implementation and execution ("aviator" component). Based on the MoO schedule, functions are allocated among agents, cf. [LBB+17]. As Schutte et al. [SGC+07] suggest, the MM will deliver position reports and initiate significant speed and trajectory changes, which serves to update the operator's mental model [SGC+07] and replenishes vigilance levels.

Monitoring & Evaluation: The third phase is the continuous monitoring and evaluation of the performance of other agents and external grand-scale factors relating to the mission. This monitoring task must not be mistaken with today's typical flight monitoring tasks.

6.5.3 Airline Operations Support and Other Job-Related Functions

OCC functions are the primary means to create continuous opportunities for human engagement on the future flight deck. Concurrent with their new role, MMs share greater responsibility. Besides the need for continuous opportunities for meaningful human engagement to maintain vigilance, operations support functions and other job-related tasks may deliver additional benefits through the empowerment of MMs to enable the execution of strategic planning functions, cf. [Joi11]. They may raise MM mission awareness through mission and flight planning, and help the organization focus on passenger service. To be efficient, high importance lays in the

MM viewing such additional functions as meaningful and mission-related [Bou06], which is a requirement for the new job profile (see subsection 6.3.1).

This section gives a high-level introduction to OCC goals, functions, and tasks. Furthermore, this section details the analysis of which of the current OCC functions could be transferred to a MM and lists related functions.

OCC Purpose, Goals, and Functions

Commercial airlines try to develop optimal flight schedules to maximize revenue [KLL+07, CRO14]. Unpredictable events often disrupt this plan and associated revenue objectives on the day of operation [CLLR10]. To deal with this problem, airlines run a specialized department called the Operations Control Center (OCC). This airline entity monitors the execution of the operational plan, anticipates and minimizes the impact of irregular operations, and solves problems to ensure economical, operational, and commercial efficiency [Bru16, Baz10, Int14]. The main focus lies on short term (tactical) operational planning and the exercise of operational control on the day of operations. [CRO14, Baz10]

An OCC is composed of specialized human teams contributing to disruption management. The most common functions include flight dispatch, aircraft control, crew control, passenger services and cargo control, and operations management, cf. [CRO14, Jep07, Cla98]. A description of each function is included in Appendix D.1. From these functions, a total of 16 different tasks were identified. These are, in no particular order: flight planning, crew briefing, movement & flight control, maintenance control, weather analysis, crew planning, crew tracking, ATC coordination, load planning, performance analysis, passenger coordination, cargo and catering coordination, station control, emergency handling, IT support, and sales and marketing (see Appendix D.2 for details).

Function Transfer Analysis Method

Each of the 16 tasks identified was evaluated against three categories of evaluation criteria: task complexity, task interruptibility, and operator and task autonomy (see also [Grä17]). As these criteria are multidimensional, tasks were rated against each dimension:

Task Interruptibility is essential as the MM is always responsible for the current mission. The MM must be able to address non-normal events immediately, hence, any task performed by the MM during RMAI must be interruptible. Tasks must not impede with taking over control of the aircraft and its systems at any time. [Ina03] Tasks should be dividable into sub-tasks; intermediate results must be obtainable.

Task complexity has disparate dimensions including overall complexity, time criticality of results, task duration, required knowledge and skills to perform a given task, and the number of input variables. It is hypothesized, that higher task complexity requires longer time to disengage from the support task and transition to FMAI. Higher task complexity, however, might lead to higher acceptance of new tasks as adequate engagement.

Operator and Task Autonomy describe the degree as to how task execution depends on continuous data-link communication to the OCC and human input. In light of potential communication interruptions, higher operator and task autonomy leads to higher suitability for a function transfer.

Individual dimensions were weighted. The higher the weight, the higher is its influence in the transferability evaluation result. If a dimension is formulated inversely, it is weighted negatively. The higher the overall score, the more suitable a function is to be transferred to MMs.

Function Transfer Analysis Results and Interpretation

Table 6.2 lists the resulting scores for all tasks in descending order. Based on the scores, tasks may be categorized into three categories: low scores (< 5, "non-transferable"), medium-range scores ($6-10$, "transferable with reservations") and high scores (> 10, "transferable"). The full analysis is detailed in Appendix D.3.

Table 6.2.: Results of the OCC task transferability analysis. (T): transferable, (TR): transferable with reservations, (N): non-transferable.

Task	Score	Task	Score
Maintenance Control	13 (T)	Movement & Flight Control	10 (T)
Weather Analysis	13 (T)	Crew Briefing	8 (TR)
Performance Analysis	12 (T)	Crew Planning	8 (TR)
Sales & Marketing	12 (T)	ATC Coordination	7 (TR)
Crew Tracking	11 (T)	Passenger Coordination	7 (TR)
IT Support & Databases	11 (T)	Station Control	1 (N)
Cargo & Catering Coordination	10 (T)	Flight Planning	0 (N)
Load Planning	10 (T)	Emergency Handling	-2 (N)

Discussion of Results

A few observations can be made from the transferability analysis: Non-transferable OCC functions are either complex (flight planning: many input variables, iterative processing), non-interruptible (emergency handling), or require physical presence (station control). Functions transferable with reservations may become workload-heavy and time-critical during non-normal events (crew planning, passenger coordination). Functions indirectly related to flight operations, such as weather analysis, performance analysis, and sales and marketing, and those that primarily consist of assessing the current situation (crew tracking) are generally transferable. Flight planning functions, despite the low rating, still make sense on the future flight deck. Their potential lies in increasing mission awareness and empowering the human: The MM gains an overall picture of why a trajectory was selected, and of environmental conditions. Increased use of pre-calculated trajectory options will help to reduce operational complexity and make flight planning suitable for execution on the flight deck.

Two functions are further chosen to be transferred: load planning and weather analysis, due to their potential to increase mission awareness. Finally, maintenance planning serves human engagement and again extends the scope of MM responsibility. While being a classic support function for airline operations, it gives the MM more responsibility on the flight deck.

The proposed new functions on the flight deck expand the role of the human from managing the current mission to being a valuable resource in the whole airline process. Characteristics of current OCC personnel become relevant for the new job. Again, the acceptance of such airline support tasks on the flight deck by future MMs must be validated.

Additional non-mission-related, but job-related functions

Additional functions are mostly executed already by today's pilots. Such functions include fleet management and technical pilot functions. MMs engage in continuing education on skills, systems, or mission-related aspects, and information gathering. The latter includes reading and implementing updates communicated through Notices to Airmen (NOTAMs), as well as aircraft and company operating manuals. Further, MMs engage in layover planning tasks.

6.5.4 Non-Job-Related Functions

Functions described in the previous sections add efficiency from the organizational point of view. In the interest of airline efficiency, MMs are required to work a given minimum time on operations support and job-related functions. Human performance will degrade when the human is forced to work over extended periods of time due to fatigue and vigilance decrement effects. The MM must be given opportunities to relax, entertain themselves, or even sleep; and to do so self-selectively, cf. [Han13]. Such engagement serves to stimulate motivation and to balance required levels of vigilance and forms of work. Activities could include watching movies and playing games, eating, looking out of the window, and napping. The human is purposefully given times of disengagement.

6.6 The Future Flight Deck

Current commercial flight decks incorporate the capability to be operated by a single pilot (cf. section 2.1). They do not support the envisioned roles and tasks of a MM. As these differ from today, a new flight deck must be designed. This section provides insight into a possible flight deck layout solution developed as part of JEPPESEN's Reduced Crew Operations (RCO) research, which was implemented at TECHNISCHE UNIVERSITÄT DARMSTADT (TUDA). Both hardware and software requirements, and practical and human factors considerations drive the design process for a future flight deck for SPO. The work environment can be downsized, and the single operator can be placed in the center of the flight deck. The flight deck must be large enough to not increase the feeling of loneliness and boredom. To be able to execute TMM and OCC support functions, the MM needs access to a large, configurable screen supporting multiple interaction paradigms. Although modern cockpits including the BOEING B787 and B777X already provide large (touch) screens, their installation angle and distance does not allow for continuous ergonomic interaction. To enable the MM, a new display concept was implemented, supporting both traditional information displaying and also typical desk work. Figure 6.3 shows a render of such an envisioned flight deck. A large,

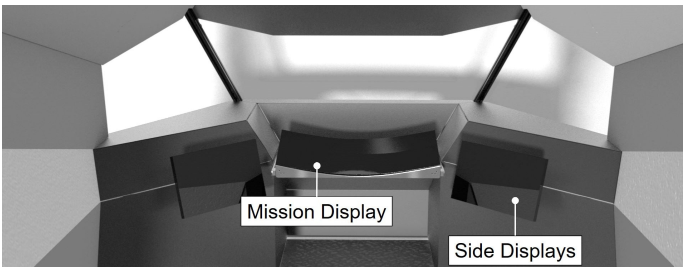

Figure 6.3.: Render of proposed flight deck layout solution as implemented at TUDA. Image after: [KKL+18])

central, curved *mission display* may be slid through different angles to allow for interaction as a *display desk* or traditional information presentation as *monitor*.

6.7 Operational Scenario

To demonstrate the ConOps, a comprehensive scenario was developed, which bases on Jeppesen's Digital Aviation portfolio demonstration. It is meant to showcase individual agents, their roles, functions, and collaboration. It includes several normal and non-normal events with various impacts on the mission to specifically highlight how the MM manages the mission. In this scenario, the MM must oversee and manage a scheduled trans-Atlantic mission from Frankfurt (ICAO code: EDDF) to Seattle (KSEA). The trajectory was calculated using JetPlan with Boeing B787-8 performance data. Weather and North Atlantic tracks were taken for August 25th, 2017. The total flight time during cruise is 9:10 hours. All phases are described briefly.

First, the MM reviews the mission plan on their mobile device. Once on the flight deck, MM identify themselves through biometrics. They proceed with flight deck initialization, GO introduction, and mission planning and review. The outcomes are goals and constraints for the mission for AS. The MM establishes a first MoO schedule, allocating time for mission support, OCC support, job-related, and non-job-related functions. Figure 6.4 shows a schematic example.

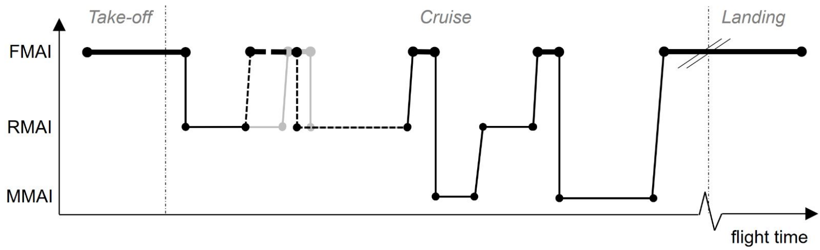

Figure 6.4.: Schematic MoO schedule including regular TMM periods. OCC support accounts for most of the cruise time. Black dotted lines represent an updated schedule due to a non-normal event from the original schedule (represented in gray lines).

With the relevant clearances received, the MM actively initiates taxi-out, take-off, and climb-out, monitoring execution by AS subsequently. While AS continue to execute the agreed trajectory, the MM actively transitions through the operational modes introduced in section 6.4 during cruise. Under normal operations, the MM spends most of the time in RMAI mode to uphold vigilance and generate efficiencies for the organization. This satisfies the subjects' desire for more continuous engagement as described in the previous chapter. As per SOPs, an airline may require MMs to engage in support functions for a given minimum time, however exact task scheduling is self-directed. Regularly under routine operations, or based on external factors, AS requests the MM to transition to FMAI to rebuild mission awareness and make mission-relevant decisions. This includes updating trajectory and constraints as required by ATC, and actively initiating trajectory changes. The MoO schedule is adapted regularly to account for external events and MM (vigilance) needs. When MoO schedule and overall conditions permit, the MM may transition to MMAI to rest or engage in job-unrelated activities.

In the described scenario, several non-normal events need to be managed. These include, among others, a maintenance event to take care of, a weather event requiring rerouting, and a passenger medical event requiring mission abortion and deviation to Calgary. The MM consults with the GO to help with decision-making and resolving issues. Upon reaching the last en-route waypoint, AS perform descent, arrival, approach, landing, and taxi-in procedures, again actively initiated and monitored by the MM. At the end of the mission, the MM signs final mission documents and checks out.

6.8 Summary

Based on the vigilance-related findings from the previously described experiment, this chapter gave an introduction into a new ConOps for SPO. This new ConOps satisfies the subject's desire for more continuous and meaningful engagement while enabling the safe operation by a single operator.

Assumptions and requirements were derived. Three agents, a mission manager, autonomous systems, and a ground operator were defined. The focus was on the human operator's task and engagement profile during cruise, which was found to be at the root of the vigilance decrement problem (see chapter 5). Tasks were derived to keep the mission manager engaged in mission-related tasks to keep vigilance levels high, both subjectively and objectively. Besides, efficiencies towards the organization may be achieved through the empowerment of the human operator. Mission manager tasks include core tasks, Total Mission Management (TMM), OCC support functions, and non-job-related tasks.

In a next step, the described ConOps should be implemented into the existing RCO demonstration environment at TUDA. To evaluate the ConOps towards vigilance, the same subjective, performance-based, and physiological vigilance-measures should be taken during a long-haul flight simulation.

7 Conclusion and Outlook

This chapter summarizes the research performed in this dissertation, and concludes its impact. An outlook is given towards further research with regards to operator vigilance under Single Pilot Operations (SPO).

7.1 Summary

The goals of the herein described research were to provide additional insight and knowledge to the discussion on the feasibility of commercial SPO. This was accomplished by evaluating pilot vigilance during the uneventful cruise phase in a non-laboratory, realistic flight deck environment, and by developing a new, human-centric Concept of Operations (ConOps). The ConOps takes into account the lessons learnt from literature reviews on the chosen topic and the obtained results from the evaluation.

Firstly, the current state of Dual Pilot Operations (DPO) and SPO in commercial aviation were presented, along with the regulatory view and a critical evaluation of commercial DPO. The results of a comprehensive literature review on existing SPO concepts were presented in form of a classification of concepts. A review of these categories followed; research gaps related to the topic at hand were identified.

Situation Awareness (SA), Mental Workload (MWL), and automation concepts were introduced briefly, as they laid the foundation for vigilance theory, which was elaborated in detail. The focus laid on vigilance estimation methods using physiological measures.

To better understand today's flight deck operations and evaluate MWL during cruise, a Hierarchical Task Analysis (HTA) was performed. MWL of today's pilots during the cruise phase was assessed as being minimal. Bringing the findings from the concept analysis, in particular the research gaps, together with those from the task analysis, the research gap for the thesis at hand along with research hypotheses was formulated. To confirm or reject these hypotheses, an experiment was designed, implemented, and executed. The goal was to evaluate operator vigilance during the cruise flight in a realistic flight deck environment, and compare vigilance levels, respectively the decrement in vigilance, between DPO and SPO.

With pilots not available, engineering students acted as pilots and executed a simulated 4 hour commercial cruise flight under both operating regimes (independent variable) in an AIRBUS A320 research simulator at TECHNISCHE UNIVERSITÄT DARMSTADT (TUDA). Performance, subjective assessments, Engagement Index (EI), Concentration of Oxygenated Hemoglobin (COH), Heart Rate (HR), Eye Blink Frequency (EBF), and Eye Blink Duration (EBD) were used as dependent variables. Physiological parameters were measured using low-cost Commercial-Off-The-Shelf (COTS) hardware. The experiment revealed that self-rated vigilance levels decreased with increasing flight time and no critical events; there was, however, no significant differences in performance to 5-minute Psychomotor Vigilance Tasks (PVTs) administered before and after the flight.

From the start of the simulation until the beginning of the abstracted emergency event (first cruise phase *C1*), linear regressions of the physiological parameters were calculated. Slope coefficients revealed trends. While the mean of coefficients for EI and EBF were not found to be significantly different from 0, the corresponding null hypotheses were rejected for COH, HR, and EBD trends. Three subjects showed the hypothesized trends in all five physiological parameters. Vigilance was, therefore, found to significantly decrement over time when inferred through COH, HR, and EBD; EI trends did not indicate a significant vigilance decrement. Evidence suggests that more subjects would have resulted in statistical significance for EI trends. Vigilance increased temporarily when executing tasks such as communication and checks.

The effect of an abstracted emergency event was determined next. Again linear regression was used to identify changes in three parameters. Again, the mean of EI regression coefficients was not found to be significantly different from 0, but those for COH and HR were. Both parameters increased significantly with the onset of the abstracted emergency event, hence vigilance increased.

Last, means of the slope coefficients were compared between the two operating regimes. Although trends were visible towards the expected outcomes, none of the five parameters showed significant differences between DPO and SPO. Vigilance trends were not significantly different between operating conditions. Absolute subjective vigilance ratings showed a significant difference between operating regimes.

Based on the experiment findings, a new ConOps for commercial SPO was developed to create continuous engagement opportunities for the human operator throughout the mission. Assumptions were stated, requirements were derived and justified. All agents, in particular the Mission Manager (MM), Autonomous Systems (AS), and Ground Operators (GOs) were defined. Three Modes of Operation (MoOs) govern the interaction between those agents. Then, the main part of the concept, the MMs' tasks, was presented in detail. The task derivation process was explained including the definition of the problems, the description of applied methods, and the detailing and discussion of results. Four categories of tasks were identified: core tasks (today's tasks), Total Mission Management (TMM), airline Operations Control Center (OCC) support, and non-job-related functions. To put the developed concept into context, the associated future flight deck was described based on a functional demonstrator, which was built at the TUDA as part of this dissertation. An operational scenario was introduced to demonstrate the developed ConOps.

7.2 Conclusion

This thesis contributes twofold to the discussion on the feasibility of commercial SPO: firstly, it provides insight into the broad topic of vigilance during particularly low-workload phases during SPO and Reduced Crew Operations (RCO). This human factors problem will become of greater importance with higher and increasing levels of automation on the flight deck.

Contrary to prevalent methods, which use only short datasets $< 1\,\mathrm{s}$ for the evaluation of event-related potentials, this dissertation evaluated long-time continuous physiological data to detect changes in those parameters over time. It was shown on a 95% confidence level, that the change in vigilance levels in two crew complement conditions were not significantly different based on objective physiological data. The effect of removing the second pilot is negligible. As subjective vigilance assessments still showed differences between operating regimes, and, in particular, that subjective vigilance was shown to decrease during SPO, the nature of the cruise phase with respect to human engagement and the human's role must be changed. Such a change must facilitate meaningful, self-directed opportunities for human engagement, which is key for continued vigilance.

Second, this thesis provides the first attempt to designing an SPO concept towards the until now underestimated human factors, including the operator's role (as MM), their vigilance, and measures to keep vigilance levels high through meaningful human engagement as learned from the simulator experiment.

7.3 Outlook

The work conducted in this thesis represents a starting point to the exploration of operator vigilance on the flight deck with regards to commercial SPO. Several challenges and questions in this field, however, remain to be open and should be addressed in future work. Towards the discussion of the general feasibility of RCO and SPO, this thesis provided a first look at quantifying vigilance levels during SPO in a realistic flight deck environment. Furthermore, a new SPO concept which addresses the relevant human factors challenges was introduced.

These two provisions form the basis for further research. With regards to understanding and quantifying vigilance levels, the herein conducted experiment should be expanded. To obtain significant statements, the number of participants must be increased. To increase applicability and to uncover further challenges, the experiments should be conducted with pilots during real flights. Initial tests have already begun: Between October and December 2019, Australian airline QANTAS in partnership with BOEING conducted three ultra-long-haul test flights from New York and London to Sydney. The knowledge gained through this dissertation, and selected sensor equipment (the MUSE headband) was used in these test flights to determine pilot fatigue and vigilance.

While the measurement equipment was chosen for low cost, easiness to use, and non-intrusiveness (see section 4.4), medical-grade equipment would greatly improve the signal (in particular resolution and signal-to-noise ratio), and therefore better represent the actual physiological states.

Each sensor and method were evaluated separately. For future applications, multiple (physiological and performance) parameters should be combined and fused into one model to better determine vigilance. Interactions between and effects of the individual sensors must be determined. Real-time capability should be added to the model.

To quantify the benefits of the proposed new ConOps, it should first be implemented into a simulation environment. The same vigilance experiment should be conducted under this new, third, condition $\mathrm{SPO}_{\mathrm{new}}$ to validate lower vigilance decrements and higher overall vigilance levels. In particular, the MoO paradigm and its proposed benefits should be validated.

Although the focus of this dissertation clearly was on a manned flight deck, the obtained results are also applicable to possible future Zero Pilot Operations along with a ground operator monitoring aircraft under this operating regime. The proposed SPO-ConOps already relies on the aircraft having autonomous flight execution capability, hence the herein proposed ConOps can be seen as an intermediate step towards Zero Pilot Operations. As those operations are likely supervised by a ground station, ground operators face very similar, if not the same, challenges with regards to vigilance. Further research with a focus on Zero Pilot Operations and ground station crew should be conducted.

References

[A4A16] A4A Airlines 4 America. Airlines 4 America Quarterly Passenger Airline Cost Index: United States Passenger Airlines. http://airlines.org/data/a4a-quarterly-passenger-airline-cost-index-u-s-passenger-airlines/, 2016. Visited on July 7th, 2016.

[AABM17] M. Abtahi, A.M. Amiri, D. Byrd, and K. Mankodiya. Hand Motion Detection in fNIRS Neuroimaging Data. *Healthcare*, 5(2), 2017.

[AAI15] Final Report AAIS Case No. AIFN/0022/2012 Pilot Incapacitation. Technical report, United Arab Emirates General Civil Aviation Authority Air Accident Investigation Sector, September 2015.

[AAI17] Final Report AAIS Case No. AIFN/0011/2017 Pilot Incapacitation. Technical report, United Arab Emirates General Civil Aviation Authority Air Accident Investigation Sector, 2017.

[AC11] Paul Atchley and Mark Chan. Potential Benefits and Costs of Concurrent Task Engagement to Maintain Vigilance - A Driving Simulator Investigation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 53(1):3–12, February 2011.

[Ada56] J.A. Adams. Vigilance in the detection of low intensity visual stimuli. *Journal of Experimental Psychology*, 52:204–208, 1956.

[ADSG71] J. Annett, K.D. Duncan, R.B. Stammers, and M.J. Gray. Task Analysis. Training Information No. 6, 1971.

[AFB06] Ulf Ahlstrom and Ferne J. Friedman-Berg. Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics*, 36(7):623 – 636, 2006.

[AFS16] AFS Aurora Flight Sciences. ALIAS Program Results. http://www.aurora.aero/wp-content/uploads/2016/10/ALIAS-Brochure_X8.pdf, October 2016. Visited on November 23rd, 2018.

[AFS17] AFS Aurora Flight Sciences. Robotic Co-Pilot Autonomously Flies and Lands a Simulated Boeing 737. http://www.aurora.aero/wp-content/uploads/2015/09/APR-343_ALIAS-737-Simulator.pdf, May 2017. Visited on November 23rd, 2018.

[AIB+07] H. Ayaz, M. Izzetoglu, S. Bunce, T. Heiman-Patterson, and B. Onaral. Detecting cognitive activity related hemodynamic signal for brain computer interface using functional near infrared spectroscopy. In *3rd International IEEE/EMBS Conference on Neural Engineering*, pages 342–345, May 2007.

[Air19] Air Line Pilots Association. ALPA Whitepaper: The Dangers of Single-Pilot Operations, July 2019.

[AKRP16] N. Armanfard, M. Komeili, J.P. Reilly, and L. Pino. Vigilance lapse identification using sparse EEG electrode arrays. In *2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–4, May 2016.

[AMM87] J. Aasman, G. Mulder, and L.J.M. Mulder. Operator Effort and the Measurement of Heart-Rate Variability. *Human Factors*, 29(2):161–170, 1987.

[ASB+12] H. Ayaz, P.A. Shewokis, S. Bunce, K. Izzetoglu, B. Willems, and B. Onaral. Optical brain monitoring for operator training and mental workload assessment. *NeuroImage*, 59(1):36 – 47, 2012. Neuroergonomics: The human brain in action and at work.

[Aus13] Australian Transport Safety Bureau. In-flight uncontained engine failure Airbus A380-842, VH-OQA. Aviation Occurrence Investigation AO -2010 - 089, Australian Transport Safety Bureau, June 2013.

[BAA+17] M.J. Blanca, R. Alarcön, J. Arnau, R. Bono, and R. Bendayan. Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29(4):552–557, Nov 2017.

[Bai83] Lisanne Bainbridge. Ironies of automation. *Automatica*, 19(6):775–779, 1983.

[BAV+14] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*, 44:58 – 75, July 2014. Applied Neuroscience: Models, methods, theories, reviews. A Society of Applied Neuroscience (SAN) special issue.

[Baz10] Massoud Bazargan. *Airline Operations and Scheduling*. Ashgate Publishing, Ltd., 2nd edition, 2010.

[BB09] Carsten Borchers and Martin Borchers. Akteure im Cockpit der Zukunft. In Gerhard Faber, editor, *Die 4. Jetgeneration der Verkehrsflugzeuge - zunehmende Automatisierung*, Darmstadt, 2009.

[BCF+13] B.K. Burian, B. Christopher, D. Fry, S. Pruchnicki, and E. Silverman. Jet Single-Pilot Simulation Study Scenario Overviews, Task Analyses, and Concurrent Task Timelines. Technical Report NASA/TM-2013-216608, NASA Ames, Hanover, MD, October 2013.

[BD01] K.A. Brookhuis and D. De Waard. *Stress, Workload and Fatigue*, chapter Assessment of drivers' workload: performance, subjective and physiological indices, pages 321 – 333. Lawrence Erlbaum Associates, 2001.

[BDG+14] E. Brandenburg, L. Doria, A. Gross, T. Günzler, and H. Smieszek. *Grundlagen und Anwendungen der Mensch-Maschine-Interaktion: 10. Berliner Werkstatt Mensch-Maschine-Systeme 10.-12. Oktober 2013*. Universitätsverlag der TU Berlin, 2014.

[BEFP+17] B.K. Bracken, S. Elkin-Frankston, N. Palmon, M. Farry, and B. de B Frederick. A system to monitor cognitive workload in naturalistic high-motion environments. In *H-Workload 2017: The first international symposium on human mental workload*, Dublin, Ireland, June 28-30 2017. Dublin Institute of Technology.

[BEH+09] C. Bruder, H. Eißfeldt, J. Hörmann, L. Jörn, C. Stern, U. Teegen, and O. Zierke. Aviator 2030 - Fähigkeitsrelevante Aspekte zukünftiger ATM-Systeme aus Sicht beruflicher Experten. Teil 1 Konzeptentwicklung. Technical Report DLR FB 2009-02, DLR, Hamburg, January 2009.

[Ber12] Palash Bera. Does cognitive overload matter in understanding BPMN models? *Journal of Computer Information Systems*, 52(4):59–69, 2012.

[BGS08] S. Banbury, M. Gauthier, and A. Scipione. Intelligent Adaptive Systems. Literature-Research of Design Guidance for Intelligent Adaptive Automation and Interfaces. Contract Report DRDC Toronto CR 2007-075 Contract Number: W7711-067983/001/TOR, Defence R & D Canada - Toronto, Kanata, ON, April 2008.

[Bha10] Hemant Bhana. Trust but verify. *AeroSafety World*, 5(5):13–14, 2010.

[Bil96] Charles E. Billings. *Aviation automation: the search for a human-centered approach*. Human factors in transportation. Lawrence Erlbaum Associates, 1996.

[BJS14] K.D. Bilimoria, W.W. Johnson, and P.C. Schutte. Conceptual Framework for Single Pilot Operations. In *International Conference on Human-Computer Interaction in Aerospace*, 2014.

[BKK+17] R.E. Bailey, L.J. Kramer, K.D. Kennedy, C.L. Stephens, and T.J. Etherington. An Assessment of Reduced Crew and Single Pilot Operations in Commercial Transport Aircraft Operations. In *36th Digital Avionics Systems Conference 2017*, 2017.

[BL91] Sheila B. Banks and Carl S. Lizza. Pilot's Associate: A Cooperative, Knowledge-Based System Application. *IEEE Expert: Intelligent Systems and Their Applications*, 6(3):18–29, June 1991.

[BLBJ15] S.L. Brandt, J. Lachter, V. Battiste, and W. Johnson. 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015 Pilot Situation Awareness and its Implications for Single Pilot Operations: Analysis of a Human-in-the-Loop Study. *Procedia Manufacturing*, 3:3017 – 3024, 2015.

[BLF+76] C.E. Billings, J.K. Lauber, H. Funkhouser, G. Lyman, and E.M. Huff. NASA Aviation Safety Reporting System. Tech. Rep. TM-X-3445, NASA Ames Research Center, Moffett Field, CA, 1976.

[BLRS18] S.L. Brandt, J. Lachter, R. Russell, and R.J. Shively. *Advances in Neuroergonomics and Cognitive Engineering*, chapter A Human-Autonomy Teaming Approach for a Flight-Following Task, pages 12–22. Advances in Intelligent Systems and Computing 586. Springer, 2018.

[BM99] Gregory Bedny and David Meister. Theory of Activity and Situation Awareness. *International Journal of Cognitive Ergonomics*, 3(1):63–72, 1999.

[BMRW75] K.S. Bibby, F. Margulies, J.E. Rijnsdorp, and R.M.J. Withers. Man's role in control systems. In *Proc. 6th IFAC Congress*, Boston, MA, 1975.

[BMSH14] C. Bogler, J. Mehnert, J. Steinbrink, and J. Haynes. Decoding Vigilance with NIRS. *PLOS ONE*, 9(7):1–12, 07 2014.

[BNP18] Lukas Bier, Simon Negash, and Timo Pommeranz. Preventing monotony caused fatigue with Driver-Vehicle-Interaction-A design concept. In *64. Kongress der Gesellschaft für Arbeitswissenschaft*, Frankfurt, February 2018.

[BNS+06] L.M. Bergasa, J. Nuevo, M.A. Sotelo, R. Barea, and M.E. Lopez. Real-time system for monitoring driver vigilance. *IEEE Transactions on Intelligent Transportation Systems*, 7(1):63–77, March 2006.

[BOB09] P. Belobaba, A. Odoni, and C. Barnhart, editors. *The global airline industry*. John Wiley & Sons, Chichester, UK, April 2009.

[Boe16] Boeing Commercial Airplanes Flight Services Marketing. 2016 Pilot & Technician Outlook. http://www.boeing.com/resources/boeingdotcom/commercial/services/assets/brochure/pilottechnicianoutlook.pdf, July 2016. Visited on January 14th, 2019.

[Boe19] Boeing. Pilot and Technician Outlook 2019. https://www.boeing.com/resources/boeingdotcom/commercial/market/pilot-technician-services/assets/downloads/2019_pto_infographic.pdf, 2019. Visited on July 29th, 2019.

[Bou06] W. Boucsein. *Enzyklopädie der Psychologie. Themenbereich D: Praxisgebiete, Serie III: Wirtschafts-, Organisations- und Arbeitspsychologie, Band 2: Ingenieurpsychologie*, chapter Psychophysiologische Methoden in der Ingenieurpsychologie, pages 317–358. Hogrefe, Göttingen, 2006.

[Boy14] Guy André Boy. Requirements for Single Pilot Operations in Commercial Aviation: A First High-Level Cognitive Function Analysis. In *Proceedings of the poster workshop at CSD&M 2014*, pages 227–234, 2014.

[BPR+13] B.K. Burian, S. Pruchnicki, J. Rogers, B. Christopher, K. Williams, E. Silverman, G. Dreschler, A. Mead, C. Hackworth, and B. Runnels. Single-Pilot Workload Management in Entry-Level Jets. Technical Report NASA/TM-2013-216557, NASA, August 2013.

[Bru16] Peter J. Bruce. *Understanding Decision-making Processes in Airline Operations Control*. Taylor & Francis, 2016.

[BS92] C.T. Bennett and M. Schwirzke. Analysis of accidents during instrument approaches. *Aviat Space Environ Med.*, 63(4):253–61, April 1992.

[BS10] Jürgen Bortz and Christof Schuster. *Statistik für Human- und Sozialwissenschaftler*. Springer, Berlin-Heidelberg, 7th edition, 2010.

[BTFM00] M. Bonner, R. Taylor, K. Fletcher, and C. Miller. *Human Performance, Situation Awareness and Automation: User-centred Design for the New Millenium*, chapter Adaptive automation and decision aiding in the military fast jet domain, pages 154–159. Omnipress, Madison, WI, 2000.

[Buc66] Leslie Buck. Reaction time as a measure of perceptual vigilance. *Psychological Bulletin*, 65(5):291–304, 1966.

[BWHA18] Lukas Bier, Philipp Wolf, Hanna Hilsenbek, and Bettina Abendroth. How to measure monotony-related fatigue? A systematic review of fatigue measurement methods for use on driving tests. *Theoretical Issues in Ergonomics Science*, 0(0):1–38, 2018.

[CB09] Mary L. Cummings and Sylvain Bruni. *Collaborative Human-Automation Decision Making*, pages 437–447. Springer, Berlin, Heidelberg, 2009.

[CBL+13] D. Comerford, S.L. Brandt, J. Lachter, S.-C. Wu, R. Mogford, V. Battiste, and W.W. Johnson. NASA's Single-Pilot Operations Technical Interchange Meeting Proceedings and Findings. Technical Report Report no. NASA-CP-2013-216513, NASA Ames Research Center, Moffett Field, CA, April 2013.

[CCMF93] P. Cabon, A. Coblentz, R. Mollard, and J.P. Fouillot. Human vigilance in railway and long-haul flight operation. *Ergonomics*, 36(9):1019–1033, 1993.

[Cen09] Alessandro Cento. *The Airline Industry: Challenges in the 21st Century*. Physica-Verlag Heidelberg, 1st edition, 2009.

[CEU03] P.P. Caffier, U. Erdmann, and P. Ullsperger. Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology*, 89(3):319–325, May 2003.

[CFG+17] J. Castle, C. Fornaro, D. Genovesi, E. Lin, D.E. Strauss, T. Wadewitz, and D. Edridge. Flying solo - how far are we down the path towards pilotless planes? Research report, UBS, August 2017.

[CGT16] M.L. Cummings, F. Gao, and K.M. Thornburg. Boredom in the Workplace: A New Look at an Old Problem. *Human Factors*, 58(2):279–300, March 2016.

[Civ13] Civil Aviation Authority. Global Fatal Accident Review 2002 - 2011. Technical Report CAP 1036, Civil Aviation Authority, London, June 2013.

[CKO+14] M. Clark, K. Kearns, J. Overholt, K. Gross, B. Barthelemy, and C. Reed. Air Force Research Laboratory Test and Evaluation, Verification and Validation of Autonomous Systems Challenge Exploration Final Report. Technical report, Air Force Research Laboratory, Wright-Patterson AFB, OH, November 2014.

[Cla98] Michael D.D. Clarke. Irregular airline operations: a review of the state-of-the-practice in airline operations control centers. *Journal of Air Transport*

Management, 4(2):67 – 76, 1998.

[CLLR10] J. Clausen, A. Larsen, J. Larsen, and N.J. Rezanova. Disruption management in the airline industry - Concepts, models and methods. *Computers & Operations Research*, 37(5):809 – 821, 2010.

[Com12] Comisión de Investigación de Accidentes e Incidentes de Aviación Civil. Incident Report IN-021/2011. Technical report, CIAIAC, June 2012.

[Com14] Committee on Autonomy Research for Civil Aviation; Aeronautics and Space Engineering Board; Division on Engineering and Physical Sciences; National Research Council. *Autonomy Research for Civil Aviation: Toward a New Era of Flight*. Washington, D.C., 2014.

[CRO14] A.J.M. Castro, A.P. Rocha, and E. Oliveira. *A New Approach for Disruption Management in Airline Operations Control*, volume 562. Springer, Berlin Heidelberg, 2014.

[CS02] Maurizio Corbetta and Gordon L. Shulman. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3):201–215, 2002.

[CS15] Stephen M. Casner and Jonathan W. Schooler. Vigilance impossible: Diligence, distraction, and daydreaming all lead to failures in a practical monitoring task. *Consciousness and Cognition*, 2015(35):33–41, September 2015.

[CSB+15] S. Coelli, R. Sclocco, R. Barbieri, G. Reni, C. Zucca, and A.M. Bianchi. EEG-based index for engagement level monitoring during sustained attention. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1512–1515, Aug 2015.

[CSC15] M.L. Cummings, A. Stimpson, and M. Clamann. Functional Requirements for Onboard Intelligent Automation in Single Pilot Operations. In *AIAA Infotech @ Aerospace, AIAA SciTech Forum*, number AIAA 2016-1652, San Diego, CA, USA, 2015.

[Csi90] M. Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*. Harper Perennial, New York, 1990.

[CT12] Michele Chinosi and Alberto Trombetta. BPMN: An introduction to the standard. *Computer Standards & Interfaces*, (34):124–134, June 2012.

[DAR14] DARPA. ALIAS Seeks to Provide Portable, Flexible Advanced Autopilot Capabilities. http://www.darpa.mil/news-events/2014-04-18, April 2014. Visited on June 14th, 2018.

[DDP+14] G. Derosière, S. Dalhoumi, S. Perrey, G. Dray, and T. Ward. Towards a Near Infrared Spectroscopy-Based Estimation of Operator Attentional State. *PLOS ONE*, 9(3):1–9, 03 2014.

[Dek02] Sidney Dekker. Frequently asked questions about new technology and human error. In C. Weikert, editor, *Human factors and safety in aviation: Procedings of a conference September 26-27*, Lund, 2002. Swedish Centre for Aviation Research and Development.

[Dek13] Sidney Dekker. *The Field Guide to Understanding Human Error*. Ashgate Publishing, Ltd., 2013.

[deW96] Dick. deWaard. *The Measurement of Drivers' Mental Workload*. PhD thesis, Traffic Research Centre, University of Groningen, Glimmen (Haren), June 1996.

[DG98] D. Dinges and R. Grace. PERCLOS, a valid psychophysiological measure of alertness as assessed by psychomotor vigilance. Technical Report FHWAMCRT-98-006, United States NHTSA, 1998.

[DH99] Sidney Dekker and Eric Hollnagel, editors. *Coping with computers in the cockpit*. Ashgate, Aldershot, UK, 1999.

[Dia04] Dan Diaper. *The Handbook of Task Analysis for Human-Computer Interaction*. Lawrence Erlbaum, Mahwah, NJ, 2004.

[DKC+15] A.V. Dao, K. Koltai, S.D. Cals, S.L. Brandt, J. Lachter, M. Matessa, D.E. Smith, V. Battiste, and W.W. Johnson. Evaluation of a Recommender System for Single Pilot Operations. *Procedia Manufacturing*, 3:3070 – 3077, 2015.

[DLR08] DLR. PILAS (PILoten-ASsistenz-System für Luftfahrzeuge). http://www.dlr.de/fl/desktopdefault.aspx/tabid-1150/1741_read-5015/, 2008. Visited on May 26th, 2017.

[Don01] Craig Donald. *People in Control: Human factors in control room design*, volume 60 of *IET Control Engineering Series*, chapter Vigilance, pages 35–49. IET, Herts, UK, 2001.

[Dou11] Michael J. Doughty. Consideration of Three Types of Spontaneous Eyeblink Activity in Normal Humans: during Reading and Video Display Terminal Use, in Primary Gaze, and while in Conversation. *Optometry and Vision Science*, 78(10):712–725, 2011.

[DP82] David Roy Davies and Raja Parasuraman. *The psychology of vigilance*. Academic Press, 1982.

[DP05] Stephen Deutsch and Richard W. Pew. Single Pilot Commercial Aircraft Operation. Technical Report No. 8436, BBN Technologies, Cambridge, MA, November 2005.

[DP17] S. Dixit and R. Prasad. *Human Bond Communication: The Holy Grail of Holistic Communication and Immersive Experience*. Wiley, Hoboken, NJ, 2017.

[DRPD17] K.R. Driscoll, A. Roy, D.S. Ponchak, and A.N. Downey. Cyber safety and security for reduced crew operations (RCO). In *Aerospace Conference, 2017 IEEE*, Big Sky, MT, March 2017.

[DRST11] U. Dulleck, A. Ristl, M. Schaffner, and B. Torgler. Heart Rate Variability, the Autonomic Nervous System, and Neuroeconomic Experiments. *Journal of Neuroscience, Psychology, and Economics*, 4:117–124, 05 2011.

[EBH+16] S.L. Estes, K.J. Burns, J.R. Helleberg, K.M. Long, M.E. Pollack, and J.L. Stein. Digital Copilot: Cognitive Assistance for Pilots. In *The 2016 AAAI Fall Symposium Series: Cognitive Assistance in Government and Public Sector Applications*, number FS-16-02, pages 141–144, 2016.

[Edw77] Elwyn Edwards. Automation in civil transport aircraft. *Applied Ergonomics*, 8(4):194–198, December 1977.

[EJ12] Mica R. Endsley and Debra G. Jones. *Designing for Situation Awareness: An Approach to User-Centered Design*. CRC, 2nd edition, 2012.

[EK95] M.R. Endsley and E.O. Kiris. Situation awareness global assessment technique (SAGAT) TRACON air traffic control version user guide. Technical report, Texas Tech University, Lubbock, TX, 1995.

[EK97] Mica R. Endsley and David B. Kaber. The use of level of automation as a means of alleviating out-of-the-loop-performance problems: A taxonomy and empirical analysis. In P. Seppala, T. Luopajarvi, C.H. Nygard, and M. Matilla, editors, *13th Triennial Congress of the International Ergonomics Association*, pages 168–170, Helsinki, 1997. Finnish Institute of Occuptational Health.

[EK99] Mica R. Endsley and David B. Kaber. Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, 42(3):462–492, 1999.

[EKB+16] T.J. Etherington, L.J. Kramer, R.E. Bailey, K.D. Kennedy, and C.L. Stephens. Quantifying pilot contribution to flight safety for normal and non-normal airline operations. In *Proceedings of the 35th Digital Aviation Systems Conference, September 25-29, 2016, Sacramento, CA*, 2016.

[ENB+16] J. Eisert, F. Di Nocera, C. Baldwin, J. Lee, J.S. Higgins, W.S. Helton, and P. Hancock. Vigilance and Fatigue: A Double Sided Coin? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1):1563–1568, 2016.

[End88] Mica R. Endsley. Design and Evaluation for Situation Awareness Enhancement. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting October 1988*, volume 32, pages 97–101, 1988.

[End96] Mica R. Endsley. *Automation and human performance: Theory and applications*, chapter Automation and situation awareness, pages 163–181. Lawrence Erlbaum, 1996.

[End17] Mica R. Endsley. From here to autonomy: Lessons learned from human-automation research. *Human Factors*, 59(1):5–27, 2017. PMID: 28146676.

[Eng01] Kai Engels. *Realisierung und Untersuchung der Kommunikationsstruktur einer Simulationsarchitektur für einen verteilten Forschungssimulator*. Phd thesis, TU Darmstadt, Darmstadt, 2001.

[ES01] D. Ellinghaus and B. Schlag. Beifahrer. Eine Untersuchung über die psychologischen und soziologischen Aspekte des Zusammenspiels von Fahrer und Beifahrer. Technical report, UNIROYAL-Verkehrsuntersuchung, November 2001.

[FAA00] *System Safety Handbook*. Federal Aviation Administration, December 2000.

[FAA05] Statistical Data for the Boeing-747-400 Aircraft in Commercial Operations. Technical Report DOT/FAA/AR-04/44, Federal Aviation Administration, Washington, D.C., January 2005.

[FAA15] Flight Standards Information Management System (CHG380). Technical Report Document 8900.1 Volume 3: General Technical Administration (Chapter 2: Exemptions, Deviations, Waivers, and Authorizations), FAA, Washington, D.C., March 2015.

[FAA16] Fly the aircraft first. https://www.faa.gov/news/safety_briefing/2016/media/SE_Topic_16-08.pdf, 2016. Visited on January 16th, 2018.

[FAC+03] F.O. Flemisch, C.A. Adams, S.R. Conway, K.H. Goodrich, M.T. Palmer, and P.C. Schutte. The H-Metaphor as a Guideline for Vehicle Automation and Interaction. Technical Memorandum NASA/TM-2003-212672, NASA Langley Research Center, Hampton, VA, 2003.

[Fat15] Younes Rakhshani Fatmehsari. Designing and Implementing a Portable Near-Infrared Imaging System for Monitoring of Human's Functional Brain Activity. Master's thesis, University of Manitoba, 2015.

[FMPS99] F.G. Freeman, P.J. Mikulka, L.J. Prinzel, and M.W. Scerbo. Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology*, 50(1):61 – 76, 1999.

[FMS+00] F.G. Freeman, P.J. Mikulka, M.W. Scerbo, L.J. Prinzel, and K. Clouatre. Evaluation of a Psychophysiologically Controlled Adaptive Automation System, Using Performance on a Tracking Task. *Applied Psychophysiology and Biofeedback*, 25(2):103–115, 2000.

[FSWL98] K. Funk, C. Suroteguh, J. Wilson, and B. Lyall. Flight deck automation and task management. In *SMC'98 Conference Proceedings. 1998 IEEE International*

Conference on Systems, Man, and Cybernetics (Cat. No.98CH36218), volume 1, pages 863–868, October 1998.

[Fun91] Ken Funk. Cockpit Task Management: Preliminary Definitions, Normative Theory, Error Taxonomy, and Design Recommendations. *The International Journal of Aviation Psychology*, 1(4):271–285, 1991.

[GGP19] A. Garcia, J.M. Gonzalez, and A. Palomino. Data Acquisition System for the Monitoring of Attention in People and Development of Interfaces for Commercial Devices. In Vanessa Agredo-Delgado and Pablo H. Ruiz, editors, *Human-Computer Interaction 4th Iberoamerican Workshop, HCI-Collab 2018, Popayan, Colombia, April 23-27, 2018*, pages 83–97, 2019.

[GHLT14] J. Graham, C. Hopkins, A. Loeber, and S. Trivedi. Design of a single pilot cockpit for airline operations. In *Systems and Information Engineering Design Symposium (SIEDS), 25-25 April 2014*, 2014.

[GOW+14] O. Grauers, M. Offredi, J. Williams, S. Mueller-Diveky, T. Dautermann, D. Mosquera, W. Rouwhorst, O. Luecke, and W. Huson. D10.1-1 Candidate Solutions for the Electronic Standby Pilot. Report ACROSS/WP10/BAE/TECH/DEL/0003, January 2014.

[Haa07] Andrea Haarmann. *Der Einsatz psychophysiologischer Variablen bei der adaptiven Automatisierung am Beispiel einer Flugsimulationsaufgabe*. Dissertation, Bergische Universität Wuppertal, Wuppertal, 2007.

[Han07] Pat Hanlon. *Global Airlines*. Butterworth-Heinemann, Oxford, 3rd edition, 2007.

[Han13] P.A. Hancock. In Search of Vigilance: The Problem of Iatrogenically Created Psychological Phenomena. *American Psychologist*, 68(2):97–109, 2013.

[Han16] Theo Hankers. *Approach To Advanced Pilot Workflow Support Using Integrated Mobile Devices*. Dissertation, TU Braunschweig, Braunschweig, Germany, 2016.

[Har07] Don Harris. A human-centred design agenda for the development of single crew operated commercial aircraft. *Aircraft Engineering and Aerospace Technology*, 79(5):518–526, 2007.

[Har11] Don Harris. *Human Performance on the Flight Deck*. Ashgate Publishing, Ltd., 2011.

[Har13] Don Harris. Distributed cognition in flight operations. In *EPCE'13 Proceedings of the 10th international conference on Engineering Psychology and Cognitive Ergonomics: applications and services - Volume Part II*, pages 125–133, Las Vegas, NV, 2013.

[HD95] Dan Hughes and Michael Dornheim. Automated Cockpits Special Report, Parts I & II. *Aviation Week & Space Technology*, January 30-February 6 1995.

[HDWM99] W.S. Helton, W.N. Dember, J.S. Warm, and G. Matthews. Optimism, Pessimism, and False Failure Feedback: Effects on Vigilance Performance. *Current Psychology: Developmental, Learning, Personality, Social*, 18(4):311–325, Winter 1999/2000 1999.

[Hen95] K.C. Hendy. Situation Awareness and Workload: Birds of a Feather? In *AGARD AMP Symposium on Situational Awareness: Limitations and Enhancements in the Aviation Environment, Brussels, 24-28 April 1995*, 21-1, April 1995. AGARD.

[HHK+16] K.M. Haus, C. Held, A. Kowalski, A. Krombholz, M. Nowak, E. Schneider, G. Strauß, and M. Wiedemann. *Praxisbuch Biofeedback und Neurofeedback*. Springer, Berlin, 2nd edition, 2016.

[HJW03] M.D. Harrison, P.D. Johnson, and P.C. Wright. *Handbook of Cognitive Task Design*, chapter Relating the automation of functions in multi-agent control systems to a systems engineering representation. Lawrence Erlbaum, Mahwah, NJ, USA, 2003.

[HM88] P.A. Hancock and N. Meshkati, editors. *Volume 52. Human Mental Workload*, volume 52 of *Advances in Psychology*. North Holland, 1st edition, January 1988.

[Hoc00] Jean-Michel Hoc. From human - machine interaction to human - machine cooperation. *Ergonomics*, 43(7):833–843, 2000.

[HR15] William S. Helton and Paul N. Russell. Rest is best: The role of rest and task interruptions on vigilance. *Cognition*, 134:165–173, 2015.

[HSS15] D. Harris, N.A. Stanton, and A. Starr. Spot the difference: Operational event sequence diagrams as a formal method for work allocation in the development of single-pilot operations for commercial aircraft. *Ergonomics*, 58(11):1773–91, June 2015.

[HTG+11] S. Haufe, M.S. Treder, M.F. Gugler, M. Sagebaum, G. Curio, and B. Blankertz. EEG potentials predict upcoming emergency brakings during simulated driving. *Journal of Neural Engineering*, 8(5):056001, 2011.

[IAT15] Airline Cost Management Group (ACMG) Enhanced Report Fiscal Year 2014. Technical report, International Air Transport Association, September 2015.

[IBI+07] M. Izzetoglu, S.C. Bunce, K. Izzetoglu, B. Onaral, and K. Pourrezaei. Functional brain imaging using near-infrared technology. *IEEE Engineering in Medicine and Biology Magazine*, 26(4):38–46, July 2007.

[ICA05] Annex 2 to the Convention on International Civil Aviation: Rules of the Air. Technical Report 10th Edition, International Civil Aviation Organization, Montreal, QC, July 2005.

[ICA12] Doc 9966, Fatigue Risk Management Systems - Manual for Regulators, Appendix B. Manual for regulators, International Civil Aviation Organization, Montreal, QC, 2012.

[Ina03] Toshiyuki Inagaki. *Handbook of cognitive task design*, chapter Adaptive automation: Sharing and trading of control, pages 147–169. Erlbaum, Mahwah, NJ, 2003.

[Int14] International Air Transport Association. Airline Cost Conference Geneva: Operations Control (OCC). http://www.iata.org/whatwedo/workgroups/documents/acc-2014-gva/occ-3-occ.pdf, August 2014. Visited on July 5th, 2017.

[Int15] MuseIO Available Data. http://developer.choosemuse.com/tools/available-data, 2015. Visited on November 26th, 2018.

[Int17] InteraXon. Muse - Technical Specifications, Validation and Research Use. Technical report, InteraXon, 2017.

[Jep07] The Integrated Operations Control Center (IOCC). Technical report, Jeppesen, Neu-Isenburg, Germany, April 2007.

[Jep08] PILoten-Assistenz-System für Luftfahrzeuge (PILAS): Abschlußbericht. Fördervorhaben des BMWA im Rahmen des Luftfahrtforschungsprogramms 2003-2008 Förderkennzeichen: 20H0302C. Final report, Jeppesen GmbH, Neu-Isenburg, Germany, June 2008.

[JFM15] D. Javaux, F. Fortmann, and C. Moehlenbrink. Adaptive Human-Automation Cooperation: A General Architecture for the Cockpit and its Application in the A-PiMod Project. Technical report, 2015.

[JLF+12] W. Johnson, J. Lachter, M. Feary, D. Comerford, V. Battiste, and R. Mogford. Task Allocation for Single Pilot Operations: A Role for the Ground. In *Int. Conf. Human-Computer Interact. Aeronaut.*, pages 1–4, Brussels, 2012.

[JM94] Tzyy-Ping Jung and S. Makeig. Estimating level of alertness from EEG. In *Proceedings of 16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1103–1104 Vol. 2, 1994.

[Joi11] Joint Planning and Development Office. Concept of Operations for the Next Generation Air Transportation System. Version 3.2. Technical report, Joint Planning and Development Office, Next Generation Air Transportation System (NextGen), Washington, D.C., 2011.

[JSC+95] T. Joubert, S.E. Salle, G. Champigneux, J.Y. Grau, P. Sassus, and H. Le Doeuff. The Copilote Electronique project: First lessons as explanatory development starts. In *Proceedings of the 3rd international workshop on human-computer teamwork (Human-Electronic Crew: Can we trust the team?). Cambridge, UK, 27-30 September 1994*, Cambridge, UK, 1995.

[KA92] B. Kirwan and L.K. Ainsworth. *A Guide To Task Analysis*. Taylor & Francis Ltd, London, UK, September 1992.

[Kab96] David B. Kaber. *The effect of level of automation and adaptive automation on performance in dynamic control environments*. Doctoral dissertation, Texas Tech University, Lubbock, TX, August 1996.

[KDV09] C. Keinrath, M. Dorneich, and J. Vašek. Designing for the future: A Cognitive-Adaptive-Man-Machine-Interface . In *Annual Meeting of Human Factors and Ergonomics Society Europe Chapter, Linkoeping, Sweden, October 14-16*, Linkoeping, 2009.

[KJP14] B.D. Kidwell, W.D. Miller Jr., and R. Parasuraman. Automation Complacency: Using Non-Invasive Brain Stimulation to Change Attention Allocation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 58, pages 240–44, September 2014.

[KK04] Niklas Kohl and Stefan E. Karisch. Airline Crew Rostering: Problem Types, Modeling, and Optimization. *Annals of Operations Research*, 127(1):223–257, 2004.

[KKI17] J.-H. Kim, D.-W. Kim, and C.-W. Im. Brain Areas Responsible for Vigilance: An EEG Source Imaging Study. *Brain Topography*, 30(3):343–351, May 2017.

[KLL+07] N. Kohl, A. Larsen, J. Larsen, A. Ross, and S. Tiourine. Airline disruption management - Perspectives, experiences and outlook. *Journal of Air Transport Management*, 13(3):149 – 162, 2007.

[KLR+00] M. Kohl, U. Lindauer, G. Royl, M. Kuhl, L. Gold, A. Villringer, and U. Dirnagl. Physical model for the spectroscopic analysis of cortical intrinsic optical signals. *Phys Med Biol*, 45(12):3749–3764, December 2000.

[KMJ18] M.A. Kamran, M.M.N. Mannann, and M.Y. Jeong. Differential Path-Length Factor's Effect on the Characterization of Brain's Hemodynamic Response Function: A Functional Near-Infrared Study. *Frontiers in Neuroinformatics*, 12:37, 2018.

[Koe99] Harry S. Koelega. *Enzyklopädie der Psychologie*, chapter Aufmerksamkeit, pages 403–478. Hogrefe Verlag, 1999.

[KrK+07] K. Kaida, T. Åkerstedt, G. Kecklund, J.P. Nilsson, and J. Axelsson. Use of Subjective and Physiological Indicators of Sleepiness to Predict Performance during a Vigilance Task. *Industrial Health*, 45(4):520–526, 2007.

[KrS+06] G. Kecklund, T. Åkerstedt, D. Sandberg, M. Wahde, T. Dukic, A. Anund, and M. Hjälmdahl. State of the art review of driver sleepiness. DROWSI project report, deliverable 1.1, 2006.

[KRS+15] M.T. Koltz, Z.S. Roberts, J. Sweet, H. Battiste, J. Cunningham, V. Battiste, K.-P.L. Vu, and T.Z. Strybel. An Investigation of the Harbor Pilot Concept for Single Pilot Operations. *Procedia Manufacturing*, 3:2937 – 2944, 2015.

[KSZ15] M. Körber, W. Schneider, and M. Zimmermann. Vigilance, boredom proneness and detection time of a malfunction in partially automated driving. In *2015 International Conference on Collaboration Technologies and Systems (CTS)*, pages 70–76, June 2015.

[KW14] Bryan Kolb and Ian Q. Whishaw. *An Introduction to Brain and Behavior*. Worth Publishers, New York, NY, 4th edition, 2014.

[KWN+17] O.E. Krigolson, C.C. Williams, A. Norton, C.D. Hassall, and F.L. Colino. Choosing MUSE: Validation of a Low-Cost, Portable EEG System for ERP Research. *Frontiers in Neuroscience*, 11:109, 2017.

[LBB+14] J. Lachter, S.L. Brandt, V. Battiste, S.V. Ligda, M. Matessa, and W.W. Johnson. Toward Single Pilot Operations: Developing a Ground Station. In *HCI-Aero '14: Proceedings of the International Conference on Human-Computer Interaction in Aerospace*, number 19 in Proceedings of the HCI-AERO 2014 conference, Silicon Valley, California, July 30-August 1, 2014., Moffett Field, CA, USA, July 2014. NASA Ames Flight Deck Display Research Group, ACM New York.

[LBB+17] J. Lachter, S.L. Brandt, V. Battiste, M. Matessa, and W. Johnson. Enhanced Ground Support: Lessons from Work on Reduced Crew Operations. *Cognition, Technology & Work*, July 2017.

[LBM+14] J. Lachter, V. Battiste, M. Matessa, Q.V. Dao, R. Koteskey, and W.W. Johnson. Toward Single Pilot Operations: The Impact of the Loss of Non-verbal Communication on the Flight Deck. In *Proceedings of the International Conference on Human-Computer Interaction in Aerospace*, HCI-Aero '14, pages 29:1–29:8, New York, NY, USA, 2014. ACM.

[LC05] Saroj K.L. Lal and Ashley Craig. Reproducibility of the spectral components of the electroencephalogram during driver fatigue. *International Journal of Psychophysiology*, 55(2):137 – 143, 2005.

[LCH+14] C.T. Lin, C.H. Chuang, C.S. Huang, S.F. Tsai, S.W. Lu, Y.H. Chen, and L.W. Ko. Wireless and Wearable EEG System for Evaluating Driver Vigilance. *IEEE Transactions on Biomedical Circuits and Systems*, 8(2):165–176, April 2014.

[LCLD12] José León-Carrión and Umberto León-Domínguez. Functional Near-Infrared Spectroscopy (fNIRS): Principles and Neuroscientific Applications. In Peter Bright, editor, *Neuroimaging*, chapter 3. IntechOpen, Rijeka, 2012.

[LDB09] L.D. Loukopoulos, R.K. Dismukes, and I. Barshi. *The Multitasking Myth: Handling Complexity in Real-World Operations*. Ashgate Studies in Human Factors for Flight Operations. Routledge, 1st edition, 2009.

[LDRDD12] E. Letsu-Dake, W. Rogers, M.C. Dorneich, and R. DeMeyers. *Advances in Human Aspects of Aviation*, chapter Innovative Flight Deck Function Allocation Concepts for NextGen, pages 301–310. CRC Press, Boca Raton, FL, 2012.

[Lee06] John D. Lee. *Handbook of Human Factors and Ergonomics*, chapter Human Factors and Ergonomics in automation design, pages 1570–1590. John Wiley & Sons, Iowa, 2006.

[LFM+15] S.V. Ligda, U. Fischer, K. Mosier, M. Matessa, V. Battiste, and W.W. Johnson. *Engineering Psychology and Cognitive Ergonomics Volume 9174 of the series Lecture Notes in Computer Science*, chapter Effectiveness of Advanced Collaboration Tools on Crew Communication in Reduced Crew Operations, pages 416–427. Springer International Publishing Switzerland, 2015.

[LGR+16] J. Liu, A. Gardi, S. Ramasamy, Y. Lim, and R. Sabatini. Cognitive pilot-aircraft interface for single-pilot operations. *Knowledge-Based Systems*, 112:37–53, August 2016.

[LL07] Sy Levine and Leslie Jae Lenell Levine. An onboard pilot and remote copilot for aviation safety, security and cost. In *Proceedings of the 26th Digital Avionics Systems Conference*, Culver City, CA, October 2007. SAFELANDER.

[Lou13] Nicholas A. Lounsberry. Pilot Supply and Demand: A Quantitative Forecast Examining Changing Industry Dynamics. Master thesis, University of North Dakota, Grand Forks, ND, August 2013.

[LRH+03] N.S. Lawrence, T.J. Ross, R. Hoffmann, H. Garavan, and E.A. Stein. Multiple neuronal networks mediate sustained attention. *Journal of Cognitive Neuroscience*, 15(7):1028–1038, 2003.

[LRRK04] A. Lucas, R. Ronnquist, P. Rechter, and S. Karim. Development & Flight Testing of an Intelligent, Autonomous UAV Capability. In *AIAA 3rd "Unmanned Unlimited" Technical Conference, Workshop and Exhibit*. AIAA, September 2004.

[LS04] John D. Lee and Katrina A. See. Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, 46(1):50–80, 2004.

[LS11] Joseph B. Lyons and Charlene K. Stokes. Human-Human Reliance in the Context of Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 54(1):112–121, 2011.

[LS19] Alan Levin and Harry Suhartono. Pilot Who Hitched a Ride Saved Lion Air 737 Day Before Deadly Crash. https://www.bloomberg.com/news/articles/2019-03-19/how-an-extra-man-in-cockpit-saved-a-737-max-that-later-crashed, March 2019.

[LSSO95] J.F. Lubar, M.O. Swartwood, J.N. Swartwood, and P.H. O'Donnell. Evaluation of the effectiveness of EEG neurofeedback training for ADHD in a clinical setting as measured by changes in T.O.V.A. scores, behavioral ratings, and WISC-R performance. *Biofeedback Self Regul.*, 20(1):83–99, March 1995.

[Luc14] Steven J. Luck. *An Introduction to the Event-Related Potential Technique*. MIT Press, 2nd edition, 2014.

[Mac57] Norman H. Mackworth. Some factors affecting vigilance. *Adv. Sci.*, 53:389–393, 1957.

[Mac64] J.F. Mackworth. Performance decrement in vigilance, threshold, and high-speed perceptual motor tasks. *Can J Psychol*, 18:209–223, 1964.

[MBB+07] E. Molteni, A.M. Bianchi, M. Butti, G. Reni, and C. Zucca. Analysis of the dynamical behaviour of the EEG rhythms during a test of sustained attention. In *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1298–1301, August 2007.

[McA12] Robert McArthur, editor. *Translational Neuroimaging*. Academic Press, London, UK, 1st edition, 2012.

[McC11] Laurie Kelly McCorry. Physiology of the autonomic nervous system. *Am J Pharm Educ*, 71(4), 2011.

[McC16] Marlis McCollum. The Solo Pilot Gets a Second Set of Eyes. https://www.mitre.org/publications/project-stories/the-solo-pilot-gets-a-second-set-of-eyes, August 2016. Visited on September 15th, 2016.

[MDL81] J.L. McLucas, F.J. Drinkwater, and H.W. Leaf. Report of the President's Task Force on Aircraft Crew Complement. Technical report, US Government Printing Office, Washington, D.C., 1981.

[MDWS00] G. Matthews, D. Davies, S. Westermann, and R. Stammers. *Human performance, cognition, stress and individual differences*. Psychology Press, 2000.

[MFBK17] K.L. Mosier, U. Fischer, B.K. Burian, and J.A. Kochan. Autonomous, context-sensitive, task management systems and decisions support tools I: Human-autonomy teaming fundamentals and state of the art. NASA Technical Memorandum, NASA, 2017. in press.

[MG16] Amina Malik and Volker Gollnick. Impact of Reduced Crew Operations on Airlines - Operational Challenges and Cost Benefits. In *16th AIAA Aviation Technology, Integration, and Operations Conference, AIAA Aviation, (AIAA 2016-3303)*. AIAA, 2016.

[MGB+11] L.K. McIntire, C. Goodyear, N. Bridges, R.A. McKinley, M. Merritt, K. Griffin, and J. McIntire. Eye-Tracking: An Alternative Vigilance Detector. Technical report, 2011.

[MHR04] A.H. Midkiff, R.J. Hansman, and T.G. Reynolds. Air Carrier Flight Operations. Technical Report ICAT-2004-3, International Center for Air Transportation, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, July 2004.

[MI93] Scott Makeig and Mark Inlow. Lapse in alertness: coherence of fluctuations in performance and EEG spectrum. *Electroencephalography and Clinical Neurophysiology*, 86(1):23 – 35, 1993.

[Min06] Marvin Minsky. *The Emotion Machine - Commonsense Thinking, Artificial Intelligence, and the Future of the Human*. Simon & Schuster, New York, 2006.

[MK05] Philip Marsden and Mark Kirby. *Handbook of Human Factors and Ergonomics Methods*, chapter Allocation of functions, pages 34–1 – 34–7. Taylor & Francis, London, 2005.

[MKL+15] N. McDonald, A. Kay, P. Liston, R. Morrison, and M. Ryan. *An Integrated Framework for Crew - Centric Flight Operations*, pages 436–447. Springer International Publishing, Cham, 2015.

[MMGM14] L.K. McIntire, R.A. McKinley, C. Goodyear, and J.P. McIntire. Detection of vigilance performance using eye blinks. *Applied Ergonomics*, 45(2, Part B):354 – 362, 2014.

[MMM+13] L.K. McIntire, R.A. McKinley, J. McIntire, C. Goodyear, and J. Nelson. Eye Metrics: An Alternative Vigilance Detector for Military Operators. *Military Psychology*, 25(5):502–513, 2013.

[MPG99] C.A. Miller, M. Pelican, and R. Goldman. Tasking interfaces for flexible interaction with automation: Keeping the operator in control. In *Proceedings of the International Conference on Intelligent User Interfaces, Redondo Beach, CA, January 5th-8th 1999*, Redondo Beach, CA, January 1999.

[MRJBI15] G. Matthews, L.E. Reinerman-Jones, D.J. Barber, and J. Abich IV. The Psychometrics of Mental Workload: Multiple Measures Are Sensitive but Divergent. *Human Factors*, 57(1):125–143, 2015.

[MSV+17] M. Matessa, T. Strybel, K. Vu, V. Battiste, and T. Schnell. Concept of Operations for RCO/SPO. Technical report, NASA, Moffett Field, CA, 2017.

[NCT07] F. Di Nocera, M. Camilli, and M. Terenzi. A Random Glance at the Flight Deck Pilots' Scanning Strategies and the Real-Time Assessment of Mental Workload. *Journal of Cognitive Engineering and Decision Making*, 1(3):271–285, 2007.

[NK06] Christa Neuper and Wolfgang Klimesch, editors. *Event-Related Dynamics of Brain Oscillations*, volume 159. Elsevier, 1st edition, 2006.

[NK14] Stefan M. Neis and David M. Klaus. Considerations Toward Defining Medical "Levels of Care" for Commercial Spaceflight. *New Space*, 2(4):165–177, December 2014.

[Nof09] Shimon Y. Nof, editor. *Springer Handbook of Automation*. Springer Berlin Heidelberg, 2009.

[Nor07] R.M. Norman. Economic Opportunities and Technological Challenges for Reduced Crew Operations. Technical Report under contract NNL06AA04B Task NNL06AA63T, NASA Langley Research Center, 2007.

[NSK18] S.M. Neis, J. Schiefele, and U. Klingauf. Classification and Review of Conceptual Frameworks for Commercial Single Pilot Operations. In *Proceedings of the 37th Digital Avionics Systems Conference*, London, UK, September 2018. IEEE.

[NTS14] Descent Below Visual Glidepath and Impact With Seawall, Asiana Airlines Flight 214, Boeing 777-200ER, HL7742, San Francisco, California, July 6, 2013. Accident Report NTSB/AAR-14/01 PB2014-105984, National Transportation Safety Board, Washington, D.C., June 2014.

[OBM07] Marlene Oscar-Berman and Ksenija Marinković. Alcohol: Effects on Neurobehavioral Functions and the Brain. *Neuropsychology Review*, 17(3):239–257, Sep 2007.

[OK77] James F. O'Hanlon and Gene R. Kelley. *Comparison of Performance and Physiological Changes Between Drivers who Perform Well and Poorly During Prolonged Vehicular Operation*, pages 87–109. Springer US, Boston, MA, 1977.

[OK16] Robert P. Ocampo and David M. Klaus. Comparing the Relative Risk of Spaceflight to Terrestrial Modes of Transportation and Adventure Sport Activities. *New Space*, 4(3):190–197, September 2016.

[OMG11] Business Process Model and Notation (BPMN). http://www.omg.org/spec/BPMN/2.0, 2011. Visited on March 4th, 2018.

[ONE07] ONERA. Innovative future air transport system. Final Activity Report IFATS/ONERA/07/D700.9-PSR, European Union, July 2007.

[Onk94] Reiner Onken. Human-centred cockpit design through the knowledge-based cockpit assistant system CASSY. In *NATO DRG Panel Workshop on Improving Function Allocation for Integrated System Design*, Soesterberg, NL, 1994. TNO Institute for Perception.

[OS10] Reiner Onken and Axel Schulte. *System-Ergonomic Design of Cognitive Automation*. Springer, 2010.

[OSE06] B.S. Oken, M.C. Salinsky, and S.M. Elsas. Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology*, 117(9):1885–1901, August 2006.

[Par87] Raja Parasuraman. Human-computer monitoring. *Human Factors*, 29(6):695–706, 1987.

[Par98] Raja Parasuraman. *The attentive brain*. MIT Press, Cambridge, 1998.

[PBB95] A.T. Pope, E.H. Bogart, and D.S. Bartolome. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology*, 40(1):187 – 195, 1995. EEG in Basic and Applied Settings.

[PC72] K.A. Provins and P. Cunliffe. The Relationship Between E.E.G. Activity and Handedness. *Cortex*, 8(2):136 – 146, 1972.

[PCC09] A. Picot, S. Charbonnier, and A. Caplier. *Recent Advances in Biomedical Engineering*, chapter Monitoring drowsiness on-line using a single encephalographic channel, pages 145–164. In-Tech, 2009.

[PKF14] A.R. Pritchett, S.Y. Kim, and K.M. Feigh. Measuring Human-Automation Function Allocation. *Journal of Cognitive Engineering and Decision Making*, 8(1):52–77, 2014.

[PMS93] R. Parasuraman, R. Molloay, and I.L. Singh. Performance Consequences of Automation Induced Complacency. *International Journal of Aviation Psychology*, 3(1):1–23, February 1993.

[PNHS08] N. Pattyn, X. Neyt, D. Henderickx, and E. Soetens. Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiology & Behavior*, 93(1):369 – 378, 2008.

[PPG+12] R.E. Propper, J. Pierce, M.W Geisler, S.D. Christman, and N. Bellorado. Asymmetry in Resting Alpha Activity: Effects of Handedness. *Open Journal of Medical Psychology*, 1:86–90, 2012.

[PR97] Raja Parasuraman and Victor Riley. Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, 39(2):230–253, 1997.

[Pra99] Scott Prahl. Optical absorption of hemoglobin. https://omlc.org/spectra/hemoglobin/, 1999. Visited on December 2nd, 2018.

[PRP+95] M.T. Palmer, W.H. Rogers, H.N. Press, K.A. Latorella, and T.S. Abbott. A Crew-Centered Flight Deck Design Philosophy for High-Speed Civil Transport (HSCT) Aircraft. Technical Memorandum 109171, NASA Langley Research Center, Hampton, VA, January 1995.

[PSKD12] V.L. Pop, E.J. Stearman, S. Kazi, and F.T. Durso. Using Engagement to Negate Vigilance Decrements in the NextGen Environment. *International Journal of Human Computer Interaction*, 28, 2012.

[PSW00] R. Parasuraman, T.B. Sheridan, and C.D. Wickens. A Model for Types and Levels of Human Interaction with Automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30(3):286–297, May 2000.

[PSW08] R. Parasuraman, T.B. Sheridan, and C.D. Wickens. Situation Awareness, Mental Workload, and Trust in Automation: Viable, Empirically Supported Cognitive Engineering Constructs. *Journal of Cognitive Engineering and Decision Making*, 2(2):140–160, June 2008.

[Rec10] Jan Recker. Opportunities and constraints: the current struggle with BPMN. *Business Process Management Journal*, 16(1), 2010.

[RM01] M.A. Regan and E. Mitsopoulos. Understanding Passenger Influences on Driver Behaviour: Implications for Road Safety and Recommendations for Countermeasure Development. Technical Report 180, Monash University Accident Research Centre, Clayton, 2001.

[Ros89] Alan H. Roscoe. *Vigilance and Performance in Automatized Systems*, volume 49, chapter Flight Deck Automation and Pilot Workload, pages 111–122. Springer Netherlands, Dordrecht, 1989.

[Ros17] Adrian Rosebrock. Eye blink detection with OpenCV, Python, and dlib. https://www.pyimagesearch.com/2017/04/24/eye-blink-detection-opencv-python-dlib/, April 2017. Visited on October 27th, 2018.

[Rou88] William B. Rouse. Adaptive Aiding for Human/Computer Control. *Human Factors*, 30(4):431–443, August 1988.

[RSV05] J. Rong, T. Spaeth, and J. Valasek. Small Aircraft Pilot Assistant: Onboard Decision Support System for SATS Aircraft. In *AIAA 5th ATIO and16th Lighter-Than-Air Sys Tech. and Balloon Systems Conferences*, Arlington, VA, September 2005. American Institute of Aeronautics and Astronautics.

[RWB+17] E. Ratti, S. Waninger, C. Berka, G. Ruffini, and A. Verma. Comparison of Medical and Consumer Wireless EEG Systems for Use in Clinical Trials. *Front Hum Neurosci.*, 11(398), August 2017.

[SAE96] SAE. ARP4761, Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment, December 1996.

[Sat93] Paul M. Satchell. *Cockpit Monitoring and Alerting Systems*. Ashgate, 1993.

[SBC+14] F. Sauvet, C. Bougard, M. Coroenne, L. Lely, P. Van Beers, M. Elbaz, M. Guillard, D. Léger, and M. Chennaoui. In-Flight Automatic Detection of Vigilance States Using a Single EEG Channel. *IEEE Transactions on Biomedical Engineering*, 61(12):2840–2847, Dec 2014.

[SBD+03] A.M. Strijkstra, D.G.M. Beersma, B. Drayer, N. Halbesma, and S. Daan. Subjective sleepiness correlates negatively with global alpha (8-12 hz) and positively with central frontal theta (4-8 hz) frequencies in the human resting awake electroencephalogram. *Neuroscience Letters*, 340(1):17 – 20, 2003.

[SBH+16] G. Sadler, H. Battiste, N. Ho, L. Hoffmann, J. Lyons, W. Johnson, R. Shively, and D. Smith. Effects of Transparency on Pilot Trust and Agreement in the Autonomous Constrained Flight Planner. In *35th Digital Aviation Systems Conference, Sacramento, 25 - 29th September 2016*, September 2016.

[SBL+16a] R.J. Shively, S.L. Brandt, J. Lachter, M. Matessa, G. Sadler, and H. Battiste. Application of Human-Autonomy Teaming (HAT) Patterns to Reduced Crew Operations (RCO). In Don Harris, editor, *13th International Conference, EPCE 2016, Held as Part of HCI International 2016, Toronto, ON, Canada, July 17-22, 2016, Proceedings*, Lecture Notes in Computer Science, pages 244–55. Springer International Publishing, 2016.

[SBL+16b] R.J. Shively, S.L. Brandt, J. Lachter, M. Matessa, G. Sadler, and H. Battiste. *Engineering Psychology and Cognitive Ergonomics*, chapter Application of Human-Autonomy Teaming (HAT) Patterns to Reduced Crew Operations (RCO), pages 244–255. Lecture Notes in Computer Science. Springer, 2016.

[SBS94] J.A. Stern, D. Boyer, and D. Schroeder. Blink rate: A possible measure of fatigue. *Human Factors*, 36(2):285–297, 1994. PMID: 8070793.

[Sce01] M.W. Scerbo. *Stress, workload and fatigue*, chapter Stress, workload and boredome in vigilance: A problem and an answer, pages 267–278. Erlbaum, Mahwah, NJ, 2001.

[Sch99] Paul C. Schutte. Complemation: An Alternative to Automation. *Journal of Information Technology Impact*, 1(3):113–118, 1999.

[Sch00] Paul C. Schutte. Distance: A Metric For Flight Deck Design and Function Allocation. In *Proceedings of the IEA 2000/HFES 2000 Congress*, 2000.

[Sch14] André Scholz. *EEG-basierte Erfassung von Aufmerksamkeitsprofilen*. Grundlagenforschung zur Reduzierung der psychischen Belastung am Arbeitsplatz.

Springer Gabler, Wiesbaden, 2014.

[Sch15] Paul C. Schutte. *Engineering Psychology and Cognitive Ergonomics*, volume 9174 of *Lecture Notes in Computer Science*, chapter How to Make the Most of Your Human: Design Considerations for Single Pilot Operations, pages 480–491. Springer International Publishing, July 2015.

[SDL16] A. Schulte, D. Donath, and D.S. Lange. *Engineering Psychology and Cognitive Ergonomics*, volume 9736 of *Lecture Notes in Computer Science*, chapter Design Patterns for Human-Cognitive Agent Teaming, pages 231–243. Springer International Publishing, 2016.

[SdS11] Donald L. Schomer and Fernando H. Lopes da Silva. *Niedermayer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Wolters Kluwer, Lippincott Williams & Wilkins, Philadelphia, PA, 6th edition, 2011.

[Sel08] Selex Galileo. CAMMI Cognitive Adaptive Man-Machine Interface, December 2008.

[SGC+07] P.C. Schutte, K.H. Goodrich, D.E. Cox, E.B. Jackson, M.T. Palmer, A.T. Pope, R.W. Schlecht, K.K. Tedjojuwono, A.C. Trujillo, R.A. Williams, J.B. Kinney, and J.S. Barry. The Naturalistic Flight Deck System: An Integrated System Concept for Improved Single-Pilot Operations. NASA Tech Report NASA/TM-2007-215090, NASA Langley Research Center, Hampton, VA, December 2007.

[SGW17] P.C. Schutte, K. Goodrich, and R. Williams. *Advances in Human Aspects of Transportation*, chapter Synergistic Allocation of Flight Expertise on the Flight Deck (SAFEdeck): A Design Concept to Combat Mode Confusion, Complacency, and Skill Loss in the Flight Deck, pages 899–911. Springer International Publishing, Cham, 2017.

[SH92] Herman H. Samson and R.Adron Harris. Neurobiology of alcohol abuse. *Trends in Pharmacological Sciences*, 13:206 – 211, 1992.

[SH95] K. Smith and P.A. Hancock. Situation awareness is adaptive, externally directed consciousness. *Human Factors*, 37(1):137–148, 1995.

[SHS14] N.A. Stanton, D. Harris, and A. Starr. Modelling and Analysis of Single Pilot Operations in Commercial Aviation. In *Proceedings of the International Conference on Human-Computer Interaction in Aerospace*, HCI-Aero '14, pages 18:1–18:8, New York, NY, USA, 2014. ACM.

[SHS16] N.A. Stanton, D. Harris, and A. Starr. The future flight deck: Modelling dual, single and distributed crewing options. *Applied Ergonomics*, 53, Part B:331 – 342, 2016. Transport in the 21st Century: The Application of Human Factors to Future User Needs.

[SI15] D. Surangsrirat and A. Intarapanich. Analysis of the meditation brainwave from consumer EEG device. In *SoutheastCon 2015*, pages 1–6, April 2015.

[SK17] Daniela Schmid and Bernd Korn. A Tripartite Concept of a Remote-Copilot Center for Commercial Single Pilot Operations. In *AIAA SciTech Forum*, Grapevine, TX, January 2017. AIAA.

[SLFJ91] N. D. Sundberg, C. A. Latkin, R. F. Farmer, and S. Jihad. Boredom in young adults: Gender and Cultural Comparisons. *Journal of Cross-Cultural Psychology*, 22(2):209–223, 1991.

[SLH11] R.F. Schmidt, F. Lang, and M. Heckmann. *Physiologie des Menschen: Mit Pathophysiologie*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2011.

[SMBR00] L.J. Skitka, K. Mosier, M. Burdick, and B. Rosenblatt. Automation Bias and Errors: Are Crews Better Than Individuals? *International Journal of Aviation Psychology*, 10(1):85–97, 2000.

[SNS18] S.M. Sprengart, S.M. Neis, and J. Schiefele. Role of the human operator in future commercial Reduced Crew Operations. In *Proceedings of the 37th Digital Avionics Systems Conference*, London, UK, September 2018.

[SO98] Michael Strohal and Reiner Onken. Intent and error recognition as part of a knowledge-based cockpit assistant. In *Proc. SPIE 3390, Applications and Science of Computational Intelligence*, volume 3390, pages 287–299, 1998.

[SRM+15] R. Stocker, N. Rungta, E. Mercer, F. Raimondi, J. Holbrook, C. Cardoza, and M.A. Goodrich. An Approach to Quantify Workload in a System of Agents. In R.H. Bordini, E. Elkind, G. Weiss, and P. Yolum, editors, *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), May 4-8, 2015, Istanbul, Turkey*, number ARC-E-DAA-TN21367, pages 1041–1050, Richland, SC, May 2015. NASA Ames Research Center.

[SS01] A. Schulte and P. Stütz. *Engineering Psychology and Cognitive Ergonomics*, volume 5, chapter Cognitive Concepts in Mission Management for Air-To-Ground Attack Aircraft, pages 23–30. Ashgate, Aldershot, 2001.

[SS06] J. Smallwood and J.W. Schooler. The restless mind. *Psychological Bulletin*, 132(6):946–958, 2006.

[SSJ+07] P.M. Salmon, N.A. Stanton, D.P. Jenkins, G.H. Walker, M.S. Young, and A. Aujla. *Engineering Psychology and Cognitive Ergonomics Volume 4562 of the series Lecture Notes in Computer Science*, chapter What Really Is Going on? Review, Critique and Extension of Situation Awareness Theory, pages 407–416. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[SSP02] A. Schlögl, M. Slater, and G. Pfurtscheller. Presence research and EEG. 2002.

[SSS+09] E.A. Schmidt, M. Schrauf, M. Simon, M. Fritzsche, A. Buchner, and W.E. Kincses. Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accident Analysis & Prevention*, 41(5):1087 – 1093, 2009.

[SSS+11] E.A. Schmidt, M. Schrauf, M. Simon, A. Buchner, and W.E. Kincses. The short-term effect of verbally assessing drivers' state on vigilance indices during monotonous daytime driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14(3):251 – 260, 2011.

[SSS17] S. Samima, M. Samima, and D. Samanta. Detecting Vigilance in People Performing Continual Monitoring Task. In P. Horain, C. Achard, and M. Mallem, editors, *Intelligent Human Computer Interaction*, pages 202–214, Cham, 2017. Springer International Publishing.

[SSWG06] P. Salmon, N. Stanton, G. Walker, and D. Green. Situation awareness measurement: A review of applicability for {C4i} environments . *Applied Ergonomics*, 37(2):225 – 238, 2006.

[Sta06] Neville A. Stanton. Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics*, 37(1):55 – 79, 2006. Special Issue: Fundamental Reviews.

[STR18] STRESS. Stress final project results report. Technical Report H2020-SESAR-2015-1, SESAR, 2018.

[SV78] Thomas B. Sheridan and William L. Verplank. Human and computer control of undersea teleoperators. Technical report, MIT Man-Machine Systems Laboratory, Cambridge, MA, August 1978.

[Sv16] Tereza Soukupová and Jan Čech. Real-Time Eye Blink Detection using Facial Landmarks. In Luka Čehovin, Rok Mandeljc, and Vitomir Štruc, editors, *21st Computer Vision Winter Workshop*, Rimske Toplice, Slovenia, February 2016.

[SW95] Nadine B. Sarter and David D. Woods. How in the World Did We Ever Get into That Mode? Mode Error and Awareness in Supervisory Control. *Human Factors*, 37(1):5–19, March 1995.

[SW13] Felix Scholkmann and Martin Wolf. General equation for the differential pathlength factor of the frontal human head depending on wavelength and age. *Journal of Biomedical Optics*, 18(10), October 2013.

[SWB97] N.B. Sarter, D.D. Woods, and C.E. Billings. *Handbook of Human Factors & Ergonomics*, chapter Automation surprises, pages 1926–1943. Wiley, New York, NY, 2nd edition, 1997.

[Tat14] William O. Tatum. Ellen R. Grass Lecture: extraordinary EEG. *The Neurodiagnostic Journal*, 54 1:3–21, 2014.

[TBS15] D.R. Thomson, D. Besner, and D. Smilek. A Resource-Control Account of Sustained Attention: Evidence From Mind-Wandering and Vigilance Paradigms. *Perspectives on Psychological Science*, 10(1):82–96, 2015. PMID: 25910383.

[TGGC15] L.C. Thomas, C. Gast, R. Grube, and K. Craig. Fatigue Detection in Commercial Flight Operations: Results Using Physiological Measures. *Procedia Manufacturing*, 3:2357 – 2364, 2015. 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015.

[TH01] Heiko Tietze and Volker Hargut. Zweidimensionale Analyse zur Beurteilung des Verlaufs von Ermüdung. Technical report, Interdisziplinäres Zentrum für Verkehrswissenschaften an der Universität Würzburg, Würzburg, 2001.

[THBK07] W.O. Tatum, A.M. Husain, S.R. Benbadis, and P.W. Kaplan. *Handbook of EEG interpretation*. Demos Medical Publishing, 2007.

[WBM11] A. Wittmer, T. Bieger, and R. Müller, editors. *Aviation Systems: Management of the Integrated Aviation Value Chain*. Springer Texts in Business and Economics. Springer Berlin Heidelberg, 2011.

[WC80] Earl L. Wiener and Renwick E. Curry. Flight-deck automation - promises and problems. *Ergonomics*, 23(10):995–1011, November 1980.

[WCF84] E.L. Wiener, R.E. Curry, and M.L. Faustina. Vigilance and Task Load: In Search of the Inverted U. *Human Factors*, 26(2):215–222, 1984.

[WD95] L. Wioland and P. Doireau. Detection of human error by an outside observer: A case study in aviation. In L. Norros, editor, *5th European conference on cognitive science approaches to process control, Espoo, Finland*. VTT Automation Finnish Psychological Society, 1995.

[WDH96] J.S. Warm, W.N. Dember, and P.A. Hancock. *Automation and human performance. Theory and applications. Human factors in transportation.*, chapter Vigilance and workload in automated systems, pages 183–200. Erlbaum, Hillsdale, NJ, 1996.

[Wei13] Sam Weigel. A Day in the Life of an RJ Pilot: Flying for an industry in flux, one regional jet pilot tells all. https://www.flyingmag.com/pilots-places/pilots-adventures-more/day-life-rj-pilot/, June 2013. Visited on July 30th, 2019.

[WG15] Cynthia A. Wolter and Brian F. Gore. A Validated Task Analysis of the Single Pilot Operations Concept. Technical Report NASA/TM-2015-218480, NASA Ames Research Center, 2015.

[Wil02] Glenn F. Wilson. An Analysis of Mental Workload in Pilots During Flight Using Multiple Psychophysiological Measures. *The International Journal of Aviation Psychology*, 12(1):3–18, 2002.

[WM01] Nicola Wright and Amanda McGown. Vigilance on the civil flight deck: incidence of sleepiness and sleep during long-haul flights and associated changes in physiological parameters. *Ergonomics*, 44(1):82–106, 2001.

[Woo88] David D. Woods. *Tasks, Errors, and Mental Models*, chapter Coping with Complexity: The Psychology of Human Behaviour in Complex Systems, pages 128–148. Taylor & Francis, Inc., Bristol, PA, USA, 1988.

[WPM08] J.S. Warm, R. Parasuraman, and G. Matthews. Vigilance Requires Hard Mental Work and Is Stressful. *Human Factors*, 50(3):433–441, June 2008.

[XlT13] X. Xiong and F. De la Torre. Supervised Descent Method and Its Applications to Face Alignment. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 532–539, June 2013.

[YBWH15] M.S. Young, K.A. Brockhuis, C.D. Wickens, and P.A. Hancock. State of science: mental workload in ergonomics. *Ergonomics*, 58(1):1–17, December 2015.

[YD08] R.M. Yerkes and J.D. Dodson. The Relation of Strength of Stimulus to Rapidity of Habit Formation. *Journal of Comparative Neurological Psychology*, 18:459–482, 1908.

[YHYH09] J.-R. Yen, C.-C. Hsu, H. Yang, and H. Ho. An investigation of fatigue issues on different flight operations. *Journal of Air Transport Management*, 15(5):236 – 240, 2009.

[YS05] M.S. Young and N.A. Stanton. *Handbook of Human Factors and Ergonomics Methods*, chapter Mental workload, pages 39–1. CRC Press, 2005.

Related Student Theses

All theses listed were supervised and mentored by the author of this dissertation.

[Eid17] Matthias Eiden. Identification of the second pilot's contribution to safe and efficient flight performance in commercial aviation. Master thesis, TU Darmstadt, Darmstadt, Germany, May 2017.

[Fun18] Maximilian Funck. Ermittlung von Vigilanz im zukünftigen Airline Cockpit. Master thesis, TU Darmstadt, Darmstadt, Germany, July 2018.

[Grä17] Lucas Gräff. Transferability analysis of tasks of the airline operation center to future pilots. Bachelor thesis, TU Darmstadt, Darmstadt, Germany, October 2017.

[KKL+18] M. Keller, R. Krieg, X. Ludewig, P. Menner, and J. Schmelz. Development and Construction of the Interior and Infrastructure of a cockpit simulator. Advanced Design Project, TU Darmstadt, Darmstadt, Germany, April 2018.

[Sch19] Caroline Schott. Analyse von Methoden zur Ermittlung von Vigilanz im zukünftigen Cockpit. Master thesis, TU Darmstadt, Darmstadt, Germany, February 2019.

A Supplementary Material to the State of the Art

This Appendix contains Parasuraman, Sheridan and Wickens' [PSW08] levels of automation taxonomy, a summary of Single Pilot Operations (SPO)-related challenges and a descriptive list of past and current research projects on SPO.

A.1 Levels of Automation

Table A.1 lists the Level of Automation (LoA) description used in section 6.5.1. It is derived from Endsley and Kaber's Levels of Automation taxonomy, cf. [EK99].

Table A.1.: Parasuraman, Sheridan and Wickens' levels of automation taxonomoy.

LoA	Description
10	Computer decides everything, acts autonomously, ignores the human.
9	informs the human only, if the computer decides to
8	informs the human only if asked.
7	executes automatically, then necessarily informs the human
6	allows the human a restricted time to veto before automatic execution
5	executes that suggestion if the human approves
4	suggests one alternative
3	narrows the selection down to a few
2	Computer offers a complete set of decision / action alternatives
1	No computer assistance, human must take all decisions and actions.

Kaber [Kab96] showed that the levels of automation and the level of Situation Awareness (SA) correlate and that intermediate levels of automation result in high levels of SA, see also [EK95].

A.2 Challenges of Single Pilot Operation

Removing the second pilot from the flight deck poses significant challenges towards the design of a future flight deck. In short, SPO introduce a single point of failure in an aircraft [DP05, SGC+07]. Hence, a major challenge is to provide a system guaranteeing not less than the current level of reliability and safety. Among the most prominent challenges are unacceptable high workload in critical phases for the single pilot and a perceived reduction of safety and security. All of the in subsection 2.2.2 presented concepts, except concept A (simple removal of the second pilot), are an attempt to solve some of these challenges, e.g. through adding automation to reduce pilot workload. In doing so, new challenges are introduced, such as additional complexity, new failure modes, and human-autonomy collaboration issues as discussed in chapter 2. This section highlights critical challenges that SPO / Reduced Crew Operations (RCO) and the presented concept categories come with. All of these challenges must later be addressed in developing a Concept of Operations (ConOps).

The (potential) challenges are categorized into general, on-board, and ground-based challenges. It should be noted that challenges may overlap with each other, cf. [Com14]. If no specific citation is given, the challenges are discussed in [CBL+13, Com14] and in the ACROSS project.

A.2.1 General Challenges

First and foremost, SPO for commercial aviation is not certifiable under the current Federal Aviation Administration (FAA) regulations. In particular, this relates to the minimum number of flight crew and pilot incapacitation (14 CFR §121.385c, AC 25.1523: considering the pilot as a system under 14 CFR 25.1309, probability of pilot incapacitation must not exceed $10^{-9}/h$), airworthiness standards (e.g. hazard categories and probabilities, redundancy, system reliability and availability, aircraft and flight deck design, cf. AC 25.1309-1A), maximum flight duty requirements (14 CFR §121.500), and oxygen mask requirements (14 CFR §121.333). Also, this includes methods of verification and validation of increasingly complex systems and algorithms (e.g. adaptive / non-deterministic systems, cf. [Com14], where system-level intelligence is divided between humans and systems). Current FAA guidance assumes that modern avionics add complexity and thus increase workload, and that there are differences in operating complexity between different sets of regulations (e.g. airworthiness standards detailed in Federal Aviation Regulations Parts 23 vs. 25). Both assumptions need not necessarily be true, however. [CBL+13]

Further challenges and barriers to the implementation of SPO and RCO are briefly discussed in the following (in no particular order):

Acceptance and Perception of a RCO concept by society, by media, by pilot unions, and by insurance companies is questionable and might hinder the introduction of this concept, cf. [CBL+13, CFG+17]. Trust in (intelligent) autonomous systems must be engendered. LACHTER ET AL. [LBB+17] see stakeholder acceptance as one of the two most difficult roadblocks to the introduction of SPO.

Legal issues such as accountability questions arise, impeding degree and speed of adoption of the new concept. COMERFORD ET AL. [CBL+13] propose developing an automation policy to guide design, operation, and management of highly automated systems. Furthermore, if a single pilot is to be monitored for pilot incapacitation or similar, the legal and ethical treatment of performance data must be addressed.

Authority delegation levels between air and ground stations must be determined. Delegating authority away from the aircraft raises several safety and security issues (discussed in detail in [DRPD17]).

Uncertainty of cost savings: Although cost savings are forecast under SPO / RCO (cf. [MG16, CFG+17]), it is not sure whether SPO yields savings. Research, development, certification (including validation and verification of autonomous systems), training, and operation have financial cost associated [CBL+13, CFG+17, DRPD17]. Furthermore, the reaction of insurance companies to SPO is not clear. [CFG+17] suggest that insurance premiums will decrease. MALIK AND GOLLNICK [MG16] indicate a 4-7% decrease in direct operating cost for an airline with SPO, CASTLE ET AL. [CFG+17] forecast an 11-28% profit increase for European airlines (up to 56% for American airlines).

Secure and enhanced communication capabilities (e.g. in terms of bandwidth) are presupposed for most concepts discussed in section 2.2.2. However actually providing available, persistent, reliable, secure, data-intensive, and cost-effective datalink communication remains a challenge, cf. [DRPD17]. It is unlikely that the amount of currently allocated bandwidth for civil aviation purposes by the Federal Communications Commission is increased, thus available bandwidth is limited [Com14, DRPD17]. Again, LACHTER ET AL. [LBB+17] see the development and certification of secure communication methods as one of the two most difficult roadblocks to the introduction of SPO.

Cyberphysical threats to highly automated aircraft must be addressed adequately [DRPD17]. Increased reliance on communication aggravates the challenge, cf. [CFG+17].

Pilot training might become a challenge as it must be developed. Additionally, the concept of apprentice training gets lost [CSC15, CBL+13]. Skill degradation of traditional flying tasks might become an issue.

Complexity and resilience management for Autonomous Systems (AS) remains a challenge.

Vehicle diversity, interoperability (mixed-equipage), and transitioning to SPO within the existing airspace system might be an issue. As airframes have life cycles spanning decades, new technologies might have to be backward compatible with legacy systems.

A.2.2 Flight-Deck Related Challenges

Several challenges occur on the flight deck, mostly relating to the loss of (human) redundancy, cf. [LBB+17]. They include (in no particular order):

Pilot incapacitation includes detecting any form of incapacitation (both mentally and physically, e.g. pilot asleep, deceased, intoxicated, unconscious, under the influence of drugs, or experiencing mental breakdown, cf. [CBL+13]) through reliable, minimally-invasive, real-time health monitoring (physiological and cognitive), handling of pilot incapacitation (e.g. accountability issues when handing control to automation), and potentially ending the incapacitated state and handing back control to the single pilot. Similarly, single pilot availability at the work station due to biological needs and fatigue remains a challenge.

Pilot workload, social loneliness, and boredom are issues that need to be addressed. Both over- and underload must be avoided. Pilot vigilance and performance might be compromised through reduced social interaction and missing social and peer pressure on the flight deck [LBM+14], as hypothesized in this thesis. Pilot boredom could become a major challenge. Detecting and countering malicious intents is an additional challenge.

Human error & error management under SPO / RCO is challenging. The aviation industry has early acknowledged that humans err and that human abilities (cognitive and physical) are limited, cf. [NK14]. In SPO, a direct and real-time human redundancy in the form of error checking is missing. Recognizing and dealing with human errors must therefore be addressed.

Complex automation introduces further challenges. Besides those discussed in section 2.4, these are in particular function allocation, the allocation of roles and responsibility, and Single Pilot Resource Management.

New technology integration, maturity, and reliability e.g. of voice recognition and gesture control remain an issue. Creating hardware redundancy (e.g. in the form of multiple displays) must be addressed adequately to fulfill certification standards.

A.2.3 Ground Station Related Challenges

Likely, a manned ground station will exist to help overcome some of the above mentioned challenges (such as excessive workload). The ground station itself introduces new challenges, such as:

Limited support of international flights e.g. of polar routes with limited satellite coverage is a challenge.

Air-ground authority questions need to be solved.

Timely ground operator SA acquisition and maintenance might be a challenge, as well as emotional disengagement from other aircraft in case of a dedicated assistance request by a single pilot. Ground Operator (GO) performance might be lower as the GO is not directly involved.

A.3 Research Projects on Single Pilot Operations

This section lists past and current research projects dealing with automating the functions of the second pilot (concept category E as used in section 2.2). Most are listed in [BGS08] and [OS10].

Pilot's Associate and Rotorcraft Pilot's Associate cf. [MPG99, BL91]. This first cockpit assistance system project was started in 1986 by the U.S. Air Force in cooperation with the Defense Advanced Research Projects Agency (DARPA). Pilot's Associate is a complex system including hard- and software components and provides a variety of decision support services for the pilot [OS10].

Copilote Electronique [JSC+95]. Also in the military domain and parallel to Pilot's Associate, this French project was designed to provide an in-flight mission re-planning decision aid for an advanced combat aircraft [OS10].

COGPIT (Cognitive Cockpit) [BTFM00] was a project in the United Kingdom. It may be considered an assistant system as long as the work objective is known by the system, cf. [OS10]. It provides associative and substituting assistance.

ASPIO (Assistant for Single Pilot IFR Operation) [Onk94] is a civil project started in the late eighties developed at the University of German Armed Forces. It was designed mainly to support the approach and landing phases of Instrument Flight Rules (IFR) flights [OS10].

CASSY (Cockpit ASsistant System) [Onk94] is the further development of ASPIO and was the first cockpit assistant system worldwide, which was successfully demonstrated in-flight. It provides alerting and associative assistance, being exclusively advisory. CASSY acts like an additional crew member, staying passive as long as everything is working normally. [OS10] The structure of CASSY, and extensive results on the simulation and flight trials are compiled in [OS10].

CAMA (Crew Assistant for Military Aircraft) [SS01, SO98] is the corresponding application for military transport flights, including modules for military applications.

PILAS (PIloten-Assistenz-System) [Jep08] was a German project for a pilot assistance system started in 2003. Its aim was to allow the application of new flight profiles in the context of future air traffic management, enabling the crew to perform more demanding operations than today without reducing safety. [DLR08]

CAMMI (Cognitive Adaptive Man Machine Interface) [KDV09, Sel08] is part of E.U.'s ARTEMIS project and started in 2008. CAMMI is a joint-cognitive system that optimizes pilot workload by helping to perform required tasks.

ALIAS (Aircrew Labor In-Cockpit Automation System) [DAR14] is funded by the Defense Advanced Research Projects Agency (DARPA). ALIAS is a tailorable, drop-in, removable kit for existing aircraft, capable of operating an aircraft from takeoff to landing. The level of automation is adjustable to suit pilot preference. The system is successfully tested on various civil and military aircraft, including helicopters, by Aurora Flight Sciences. [AFS16, AFS17]

ESP (Electronic Standby Pilot) is part of the European ACROSS project. It was envisioned to take over control in case of pilot incapacitation and land the aircraft at the nearest suitable airport (through the help of a ground station).

SAPA (Small Aircraft Pilot Assistant) [RSV05] is an on-board pilot decision aid system intended to assist pilots in both information-processing and decision-making using artificial intelligence techniques.

Digital Copilot developed by the MITRE Corporation [McC16], is a cognitive assistant that determines when information is required based on flight context and automatically provides it to the pilot at the appropriate time [EBH+16]. It provides a single pilot with benefits of a crewed operation through a tablet device. [McC16]

IFATS (Innovative Future Air Transport System) sponsored by the European Union, had the goal of achieving the greatest possible automation in the future air transport system by largely replacing humans (pilots, air traffic controllers) and their previous tasks with technical systems [ONE07, BEH+09].

B Supplementary Material to the Flight Crew Task Analysis

This section shows the graphical representations of Dual Pilot Operations (DPO) plans and operations. Note that not all sub-plans are depicted here.

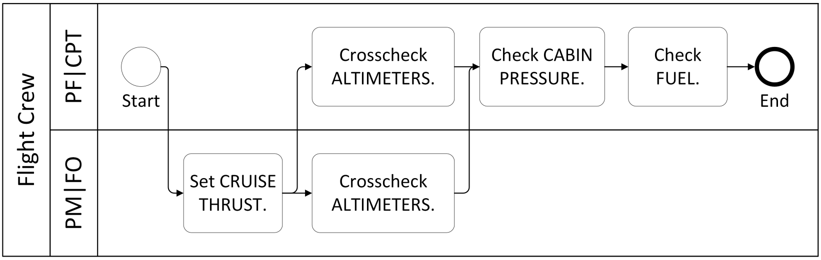

Figure B.1.: Cruise check: plan and associated operations.

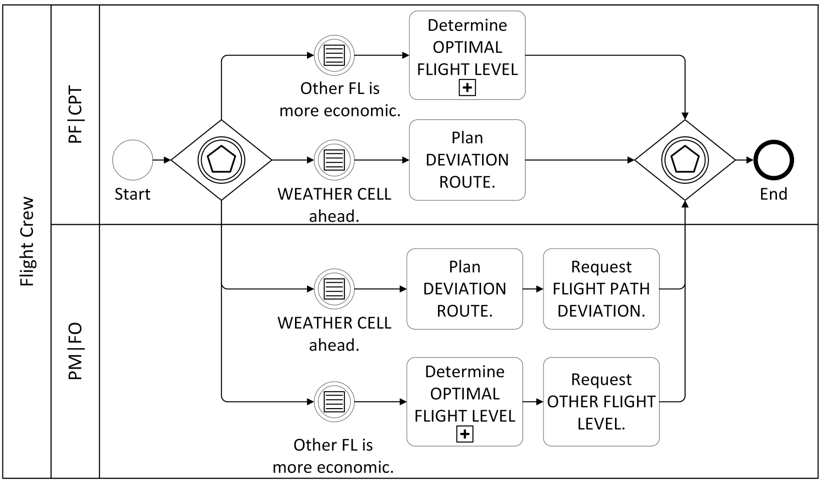

Figure B.2.: En-route flight optimization plan and associated operations.

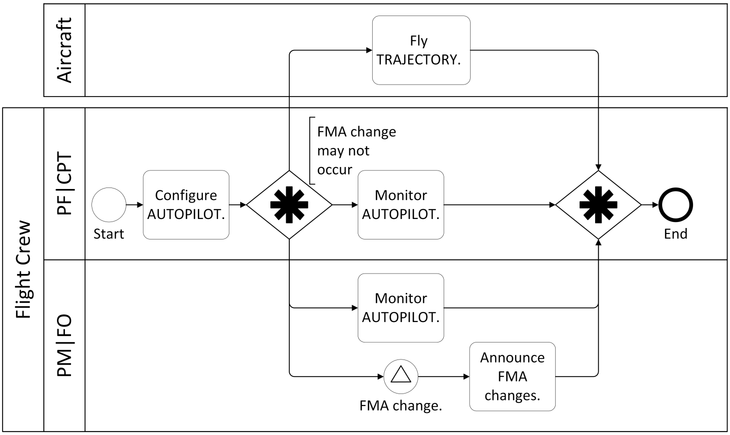

Figure B.3.: Autopilot use: plan and associated operations.

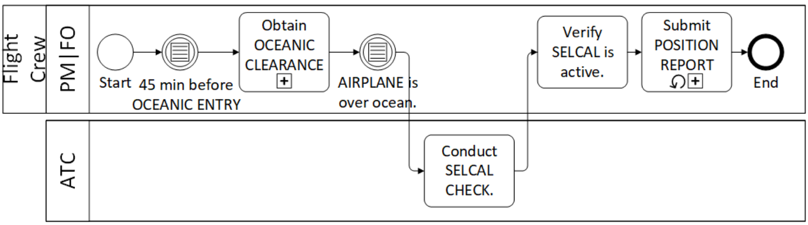

Figure B.4.: Transoceanic procedures execution plan and associated operations.

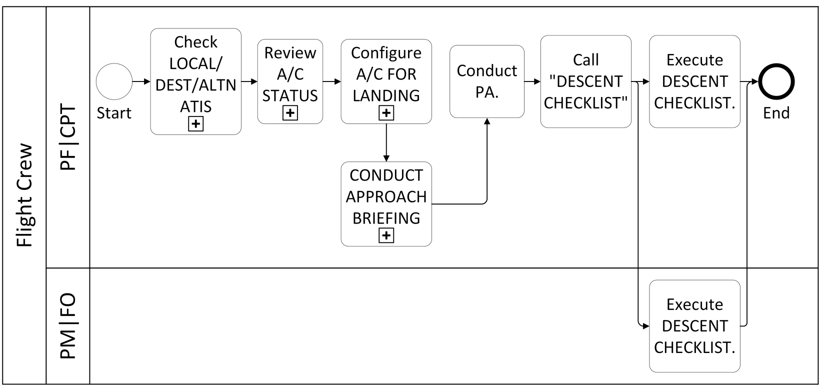

Figure B.5.: Before descent procedure execution plan and associated operations.

C Supplementary Material to the Experiment Results

Chapter 5 reports and discusses all obtained experiment results, however, only data for one single subject was shown exemplarily. This Appendix contains graphs and results for those subjects not explicitly shown in chapter 5.

C.1 Individual Vigilance Assessments

Figure C.1 reports the self-rated subjective vigilance for the 10 participants who executed both scenarios, over time. For comparison, answers from both operating regime conditions are displayed. Additionally, (normalized) response times to the vigilance assessment requests are depicted with error bars.

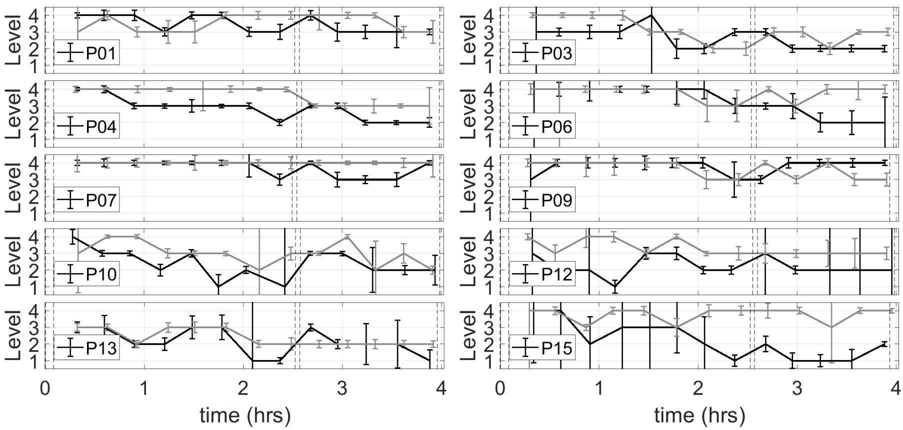

Figure C.1.: Comparison of subjective vigilance assessments. Vertical dotted lines define the five phases, error bars denote normalized response times to the assessment request (the distance between level grid lines equals 7.5 s). Black lines represent the SPO condition, gray the DPO condition. Gray color indicates time in which the subject was asleep.

C.2 Engagement Indices (AF7) over time

The following graphs show the Engagement Index (EI) over time for the subjects P-03 to P-15 for both conditions (operating regimes) SPO (black color) and DPO (gray). Data for P-01 is shown in subsections 5.2.1 and 5.4.1, respectively.

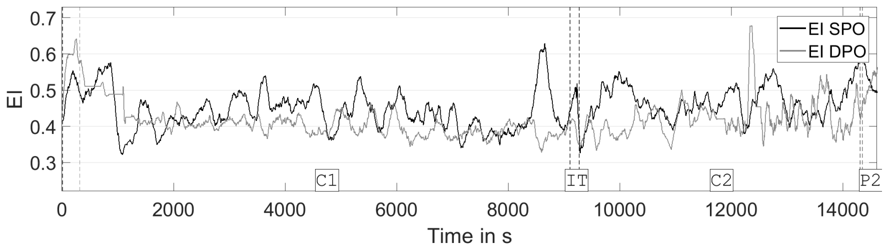

Figure C.2.: EI over time for subject P-03 with phases.

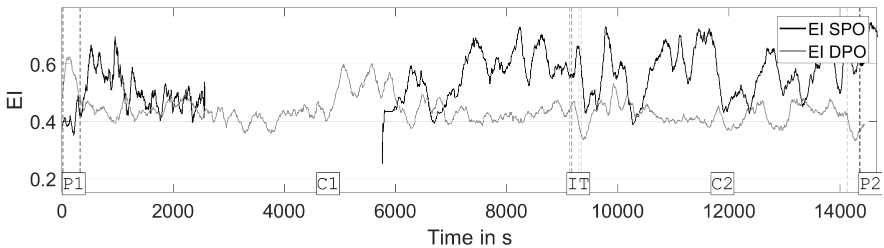

Figure C.3.: EI over time for subject P-04 with phases.

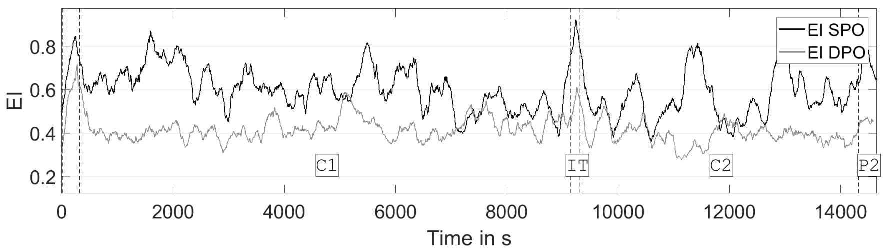

Figure C.4.: EI over time for subject P-06 with phases.

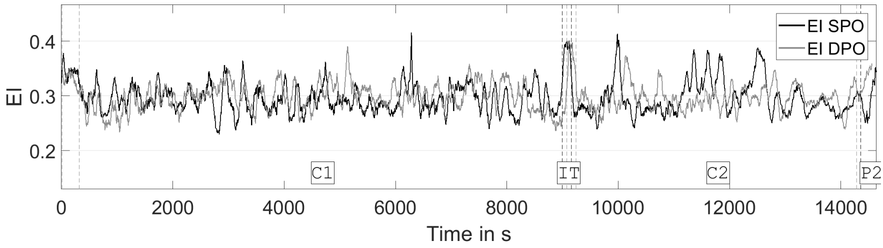

Figure C.5.: EI over time for subject P-07 with phases.

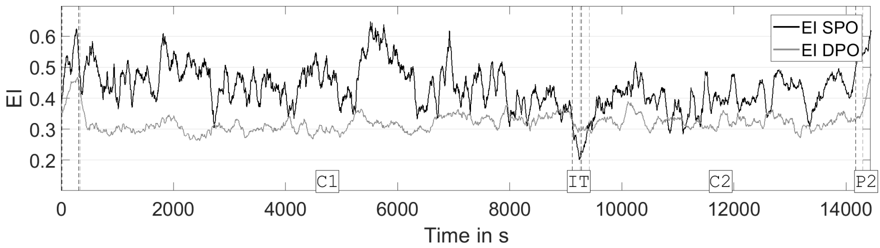

Figure C.6.: EI over time for subject P-09 with phases.

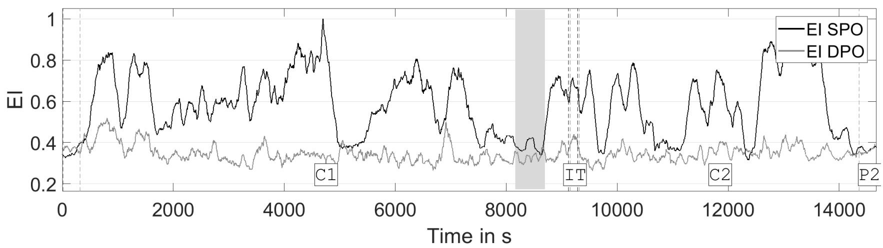

Figure C.7.: EI over time for subject P-10 with phases.

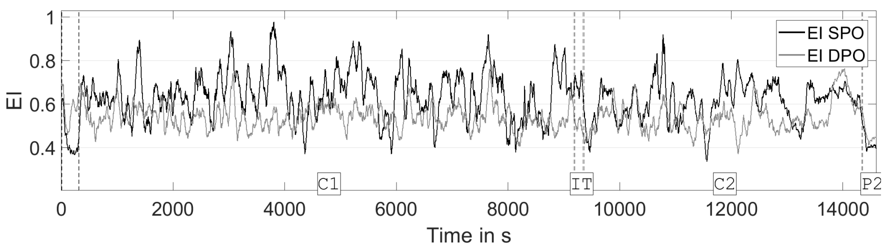

Figure C.8.: EI over time for subject P-12 with phases.

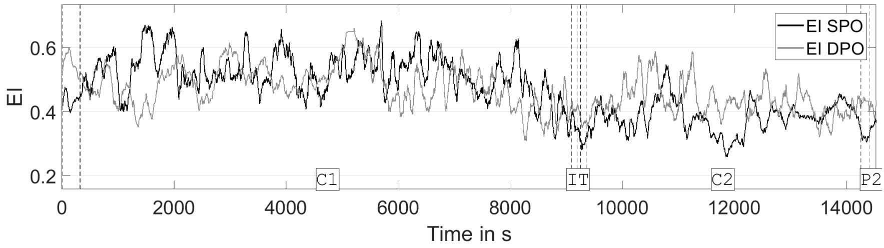

Figure C.9.: EI over time for subject P-13 with phases.

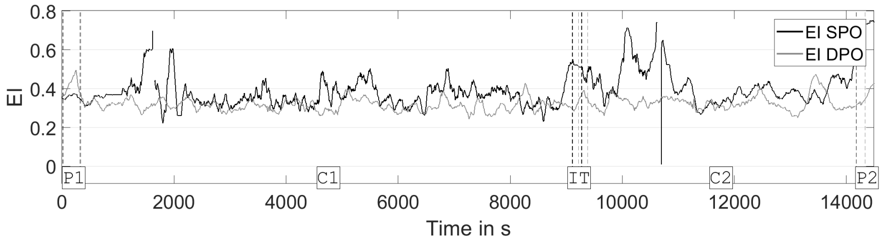

Figure C.10.: EI over time for subject P-15 with phases.

Table C.1 lists the linear regression coefficients for all 10 participants:

Table C.1.: Dimensionless linear regression coefficients, m (slope), for EI for all 10 SPO scenarios, phase *C1*. All values multiplied by $1 \cdot 10^{-5}$.

ID	C1 coefficient	ID	C1 coefficient
P-01	−1.02686	P-03	−0.32315
P-04	0.93158	P-06	−1.93803
P-07	−0.04247	P-09	−0.85439
P-10	−1.57955	P-12	−0.02618
P-13	−0.74696	P-15	−0.16628

C.3 Concentration of Oxygenated Hemoglobin over Time

The following graphs show dimensionless virtual absolute values of oxygenated hemoglobin (HbO2) over time for the remaining participants in both operating regimes SPO (black color) and DPO (gray).

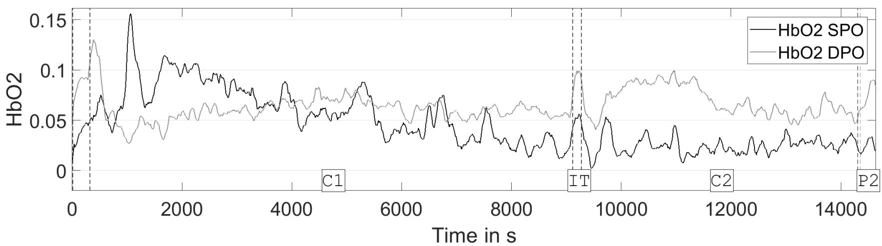

Figure C.11.: HbO2 over time for subject P-03.

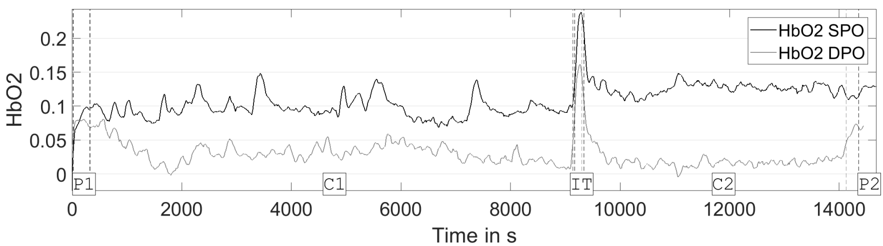

Figure C.12.: HbO2 over time for subject P-04.

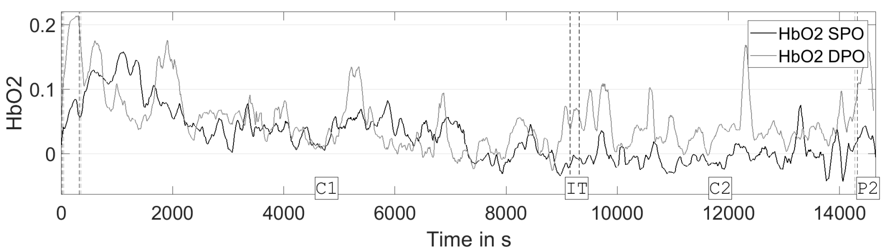

Figure C.13.: HbO2 over time for subject P-06.

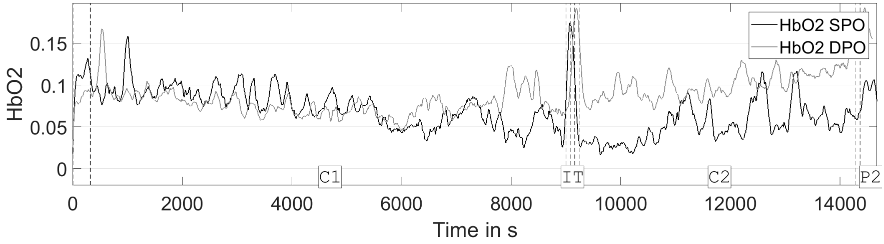

Figure C.14.: HbO2 over time for subject P-07.

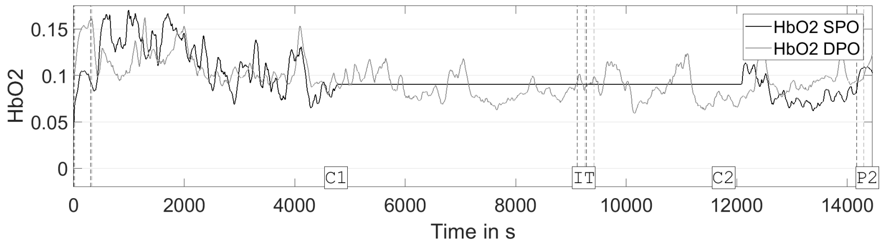

Figure C.15.: HbO2 over time for subject P-09.

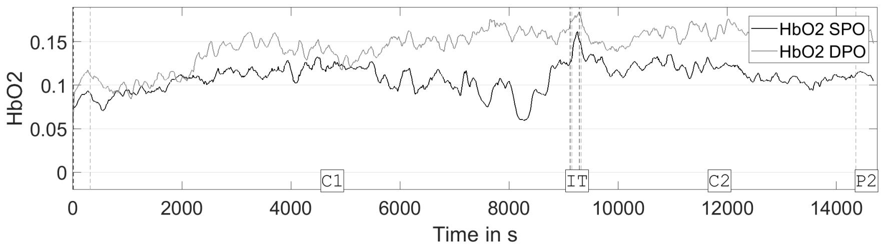

Figure C.16.: HbO2 over time for subject P-10.

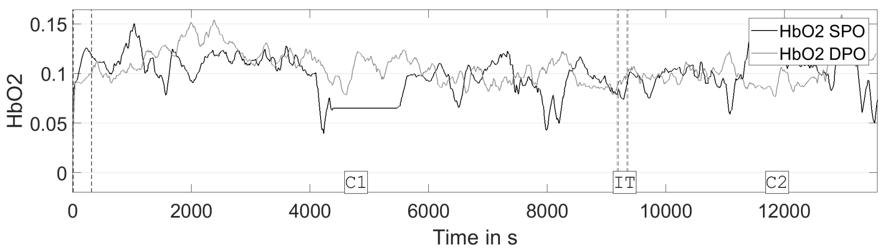

Figure C.17.: HbO2 over time for subject P-12.

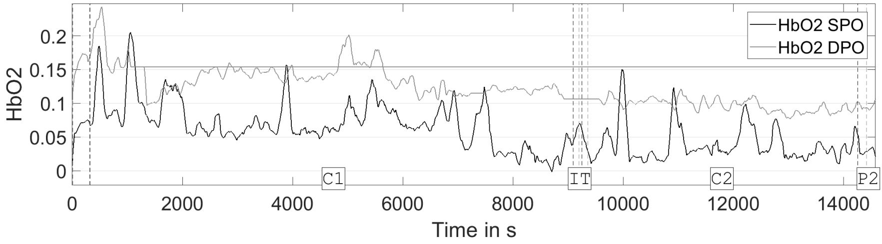

Figure C.18.: HbO2 over time for subject P-13.

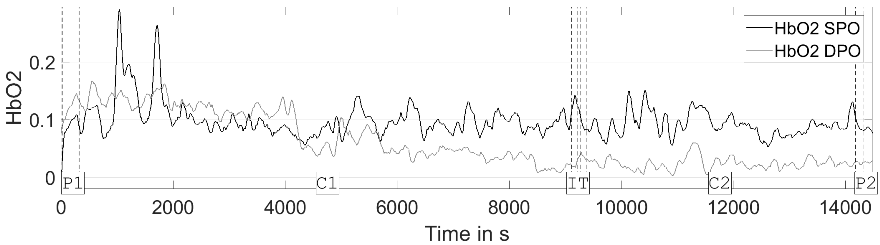

Figure C.19.: HbO2 over time for subject P-15.

Table C.2 lists the regression coefficients (m, slope of the curve) of all 10 subjects for both phases *C1* and *C2*.

Table C.2.: Dimensionless linear regression coefficients, m (slope), for COH for all 10 SPO scenarios, phase *C1*. All values multiplied by $1 \cdot 10^{-6}$.

ID	C1 coefficient	ID	C1 coefficient
P-01	−5.39787	P-03	−8.26972
P-04	−0.36810	P-06	−12.3812
P-07	−6.60801	P-09	no data
P-10	0.09566	P-12	−3.33020
P-13	−7.58596	P-15	−4.31609

C.4 Heart Rate over Time

The following graphs show the virtual absolute values of Heart Rate (HR) over time for the remaining participants in both operating regimes SPO (black color) and DPO (gray). Gaps in the graphs result from the application of the maximum 50% difference on succeeding RR-peaks cutoff limit as described in subsection 4.6.2.

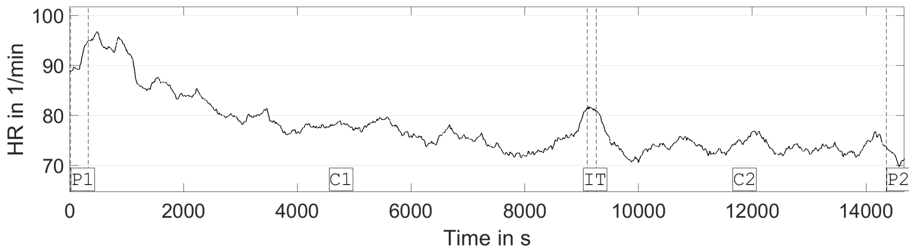

Figure C.20.: Heart Rate over time for P-01.

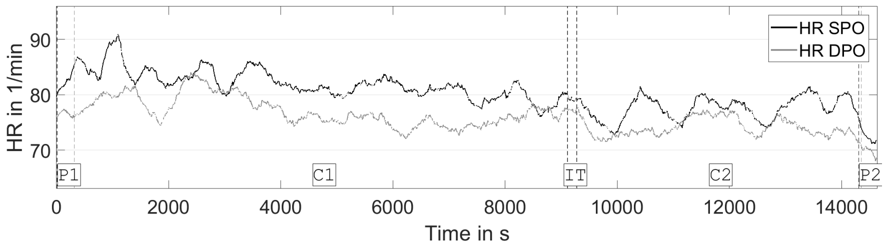

Figure C.21.: HR over time for subject P-03.

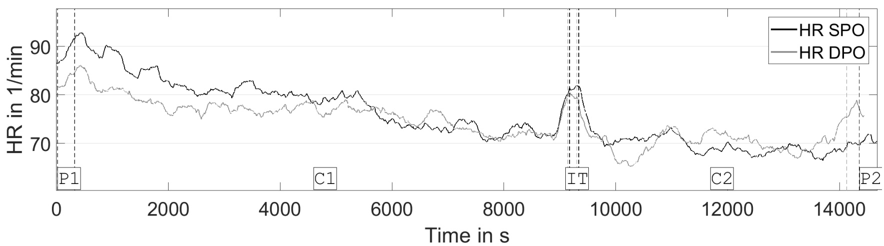

Figure C.22.: HR over time for subject P-04.

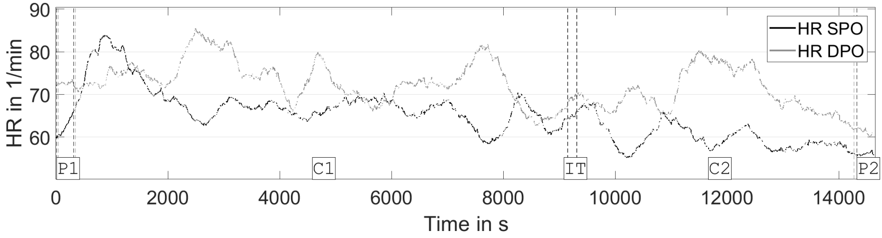

Figure C.23.: HR over time for subject P-06.

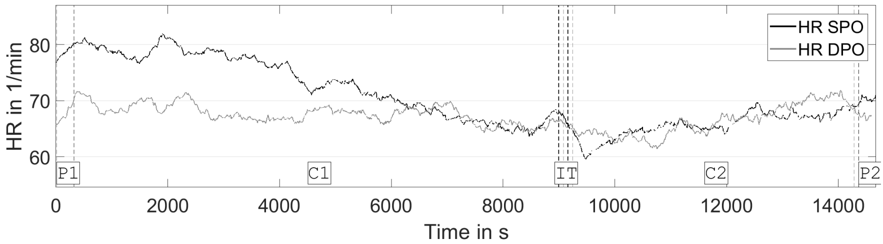

Figure C.24.: HR over time for subject P-07.

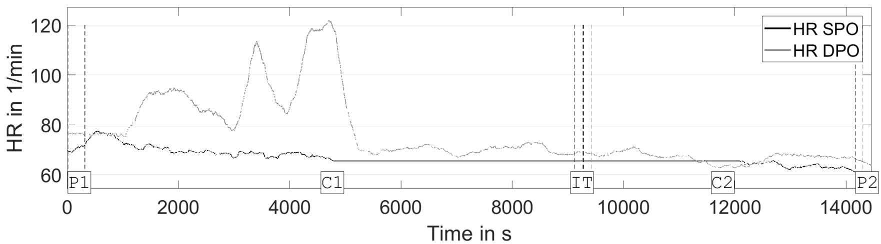

Figure C.25.: HR over time for subject P-09.

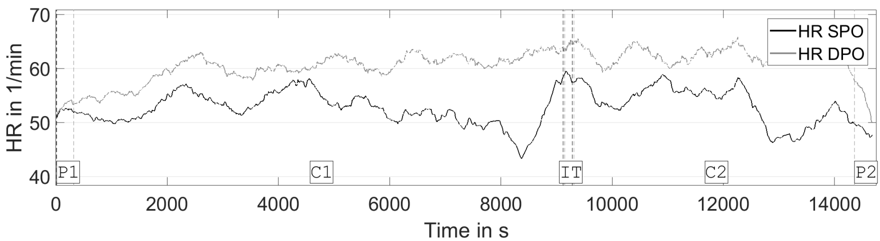

Figure C.26.: HR over time for subject P-10.

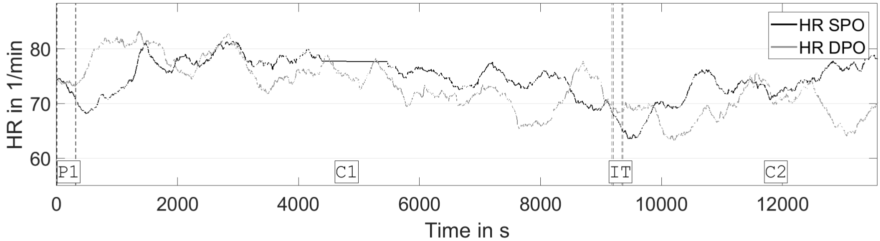

Figure C.27.: HR over time for subject P-12.

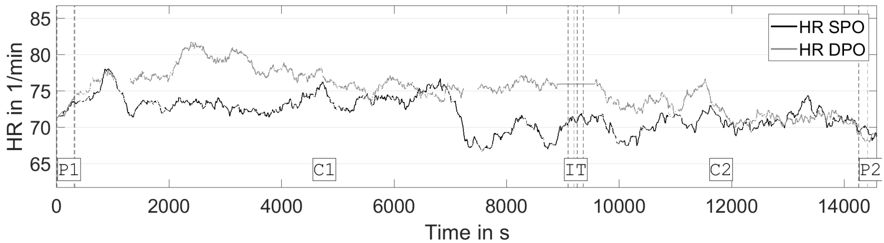

Figure C.28.: HR over time for subject P-13.

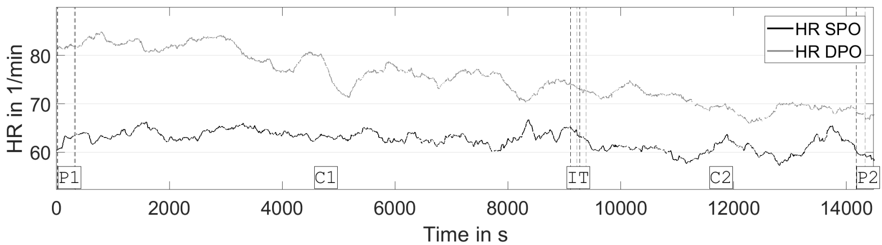

Figure C.29.: HR over time for subject P-15.

Table C.3.: Dimensionless linear regression coefficients, m (slope), for HR for all 10 SPO scenarios, phase *C1*. All values multiplied by $1 \cdot 10^{-3}$.

ID	C1 coefficient	ID	C1 coefficient
P-01	−2.14843	P-03	−0.80819
P-04	−2.05521	P-06	−1.56364
P-07	−2.08159	P-09	no data
P-10	−0.36318	P-12	−0.37125
P-13	−0.56395	P-15	−0.20130

C.5 Eye Blink Frequency over Time

The following graphs show the Eye Blink Frequency (EBF) over time for the subjects P-03 to P-15 for both conditions (operating regimes) SPO (black color) and DPO (gray).

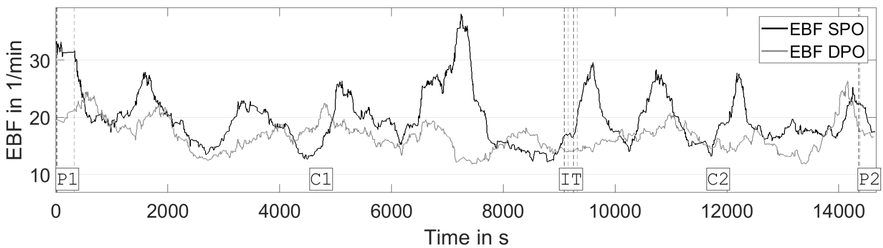

Figure C.30.: EBF over time for subject P-01 with phases.

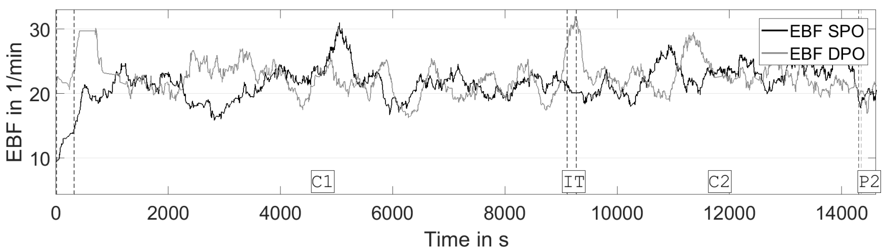

Figure C.31.: EBF over time for subject P-03 with phases.

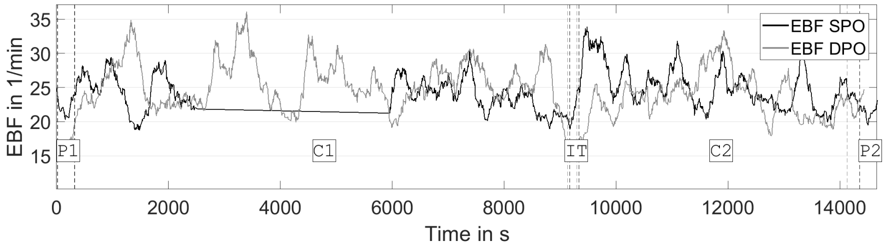

Figure C.32.: EBF over time for subject P-04 with phases.

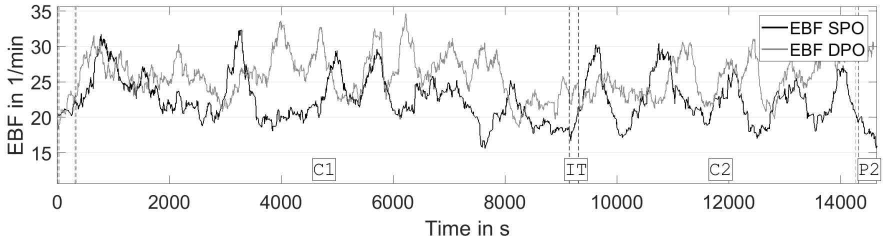

Figure C.33.: EBF over time for subject P-06 with phases.

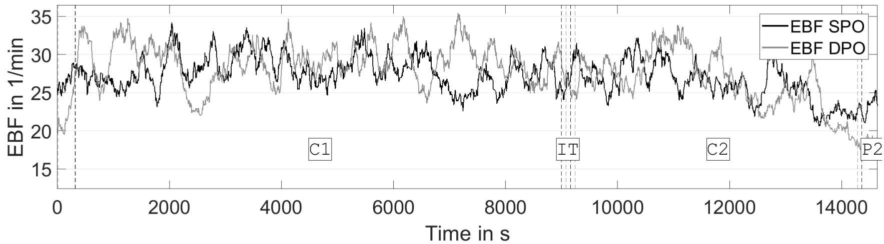

Figure C.34.: EBF over time for subject P-07 with phases.

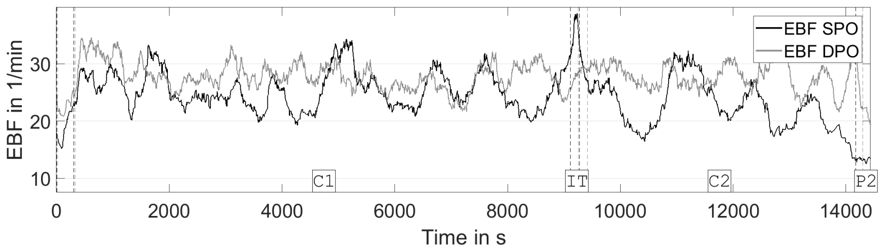

Figure C.35.: EBF over time for subject P-09 with phases.

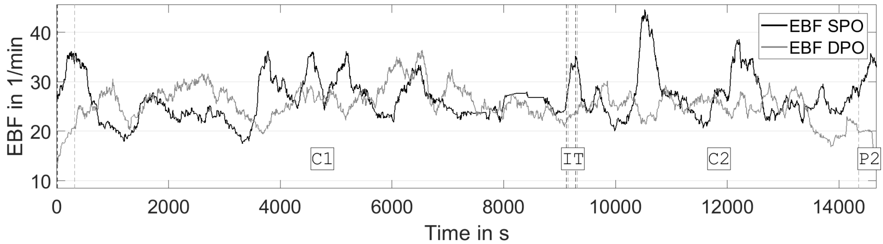

Figure C.36.: EBF over time for subject P-10 with phases.

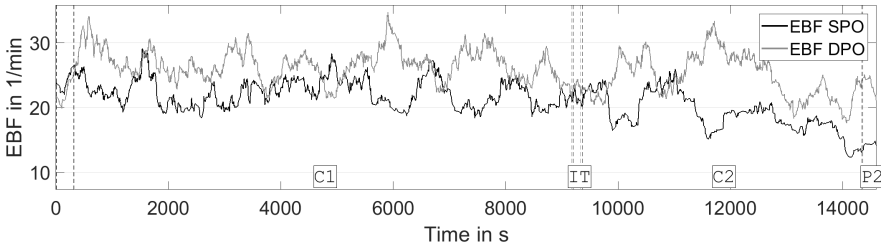

Figure C.37.: EBF over time for subject P-12 with phases.

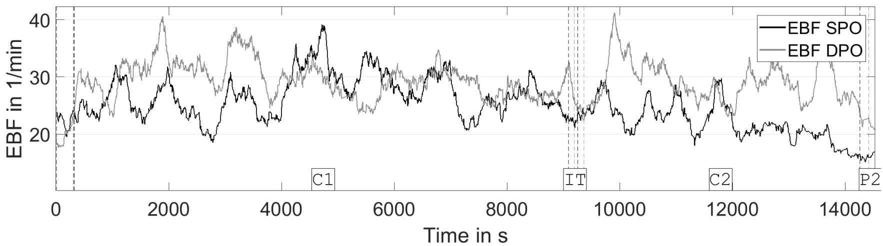

Figure C.38.: EBF over time for subject P-13 with phases.

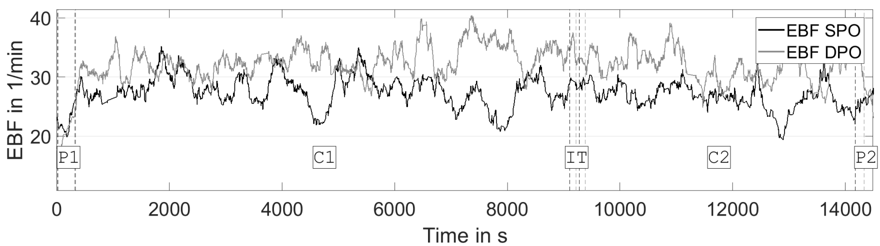

Figure C.39.: EBF over time for subject P-15 with phases.

Table C.4.: Dimensionless linear regression coefficients m (slope) for EBF for all 10 SPO subjects, phase *C1*. All values multiplied by $1 \cdot 10^{-4}$.

ID	C1 coefficient	ID	C1 coefficient
P-01	1.14967	P-03	1.52230
P-04	2.63429	P-06	−5.44782
P-07	−1.14931	P-09	−2.18064
P-10	1.83954	P-12	−1.02651
P-13	1.95738	P-15	−2.81628

C.6 Eye Blink Duration over Time

The following graphs show the Eye Blink Duration (EBD) over time for the subjects P-01 to P-15 for both conditions (operating regimes) SPO (black color) and DPO (gray).

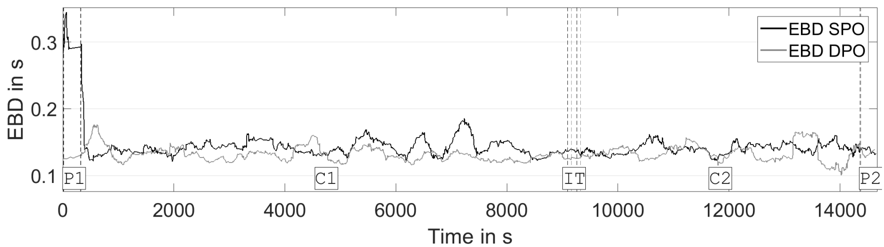

Figure C.40.: EBD over time for subject P-01 with phases.

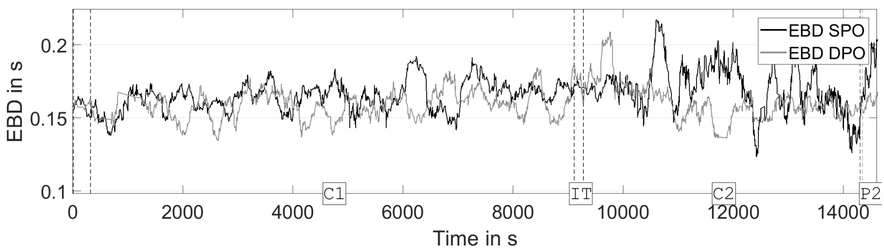

Figure C.41.: EBD over time for subject P-03 with phases.

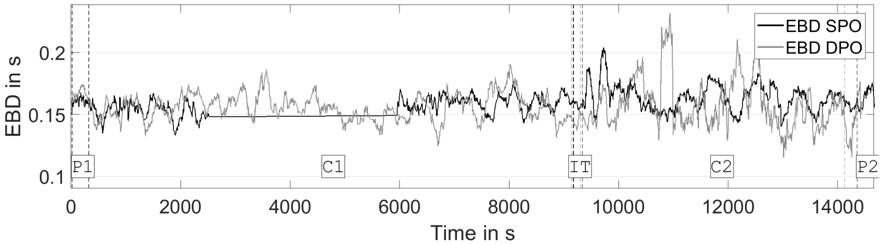

Figure C.42.: EBD over time for subject P-04 with phases.

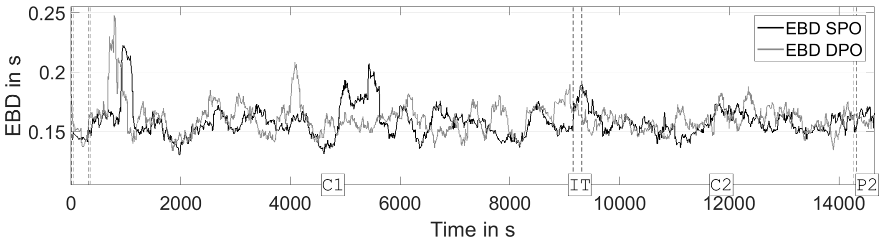

Figure C.43.: EBD over time for subject P-06 with phases.

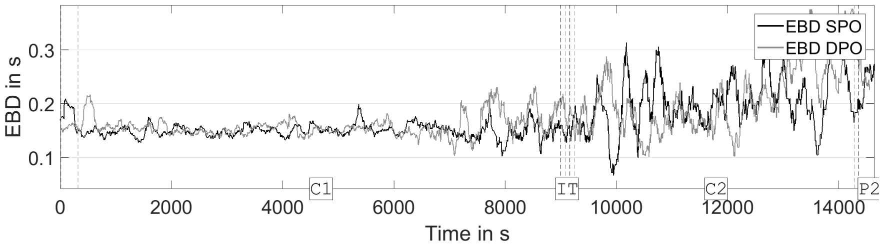

Figure C.44.: EBD over time for subject P-07 with phases.

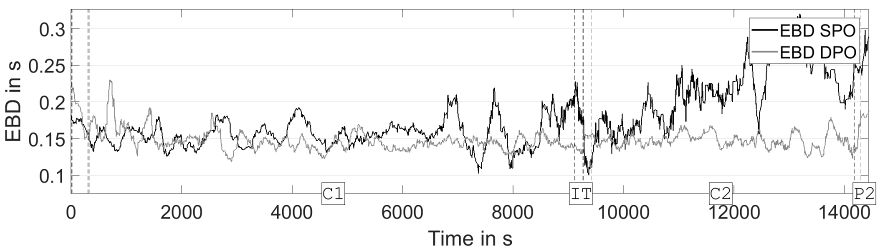

Figure C.45.: EBD over time for subject P-09 with phases.

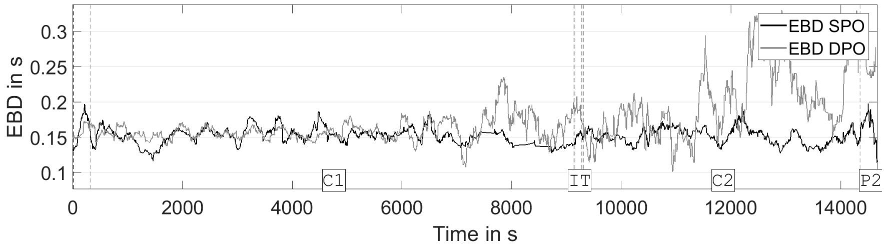

Figure C.46.: EBD over time for subject P-10 with phases.

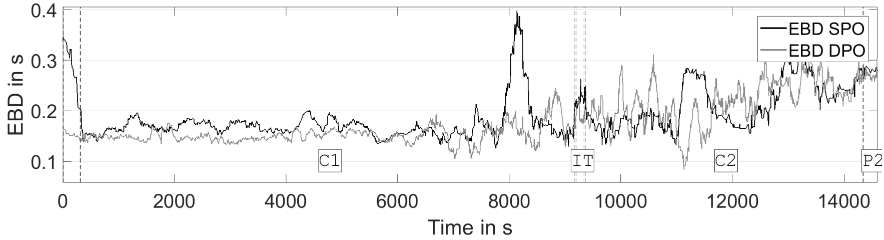

Figure C.47.: EBD over time for subject P-12 with phases.

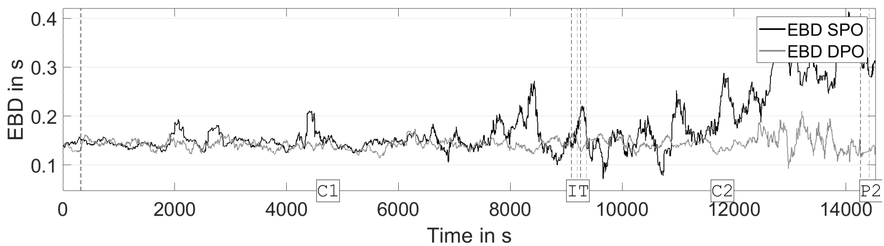

Figure C.48.: EBD over time for subject P-13 with phases.

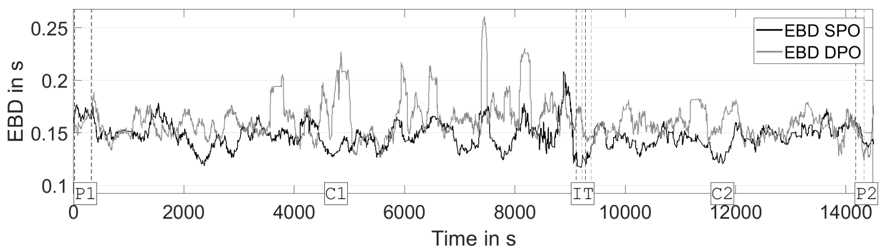

Figure C.49.: EBD over time for subject P-15 with phases.

Table C.5.: Dimensionless linear regression coefficients m (slope) for EBD for all ten SPO subjects, phase *C1*. All values to be multiplied with $1 \cdot 10^{-6}$.

ID	C1 coefficient	ID	C1 coefficient
P-01	1.37474	P-03	1.38735
P-04	1.10432	P-06	−0.63370
P-07	0.35476	P-09	1.65793
P-10	−0.34116	P-12	4.00836
P-13	2.17585	P-15	0.42102

C.7 Engagement Indices (AF7): Comparison Between Phases

This section contains all boxplots of EI values grouped into the five phases *P1*, *C1*, *IT*, *C2*, and *P2*, as well as the linear regression plots of *C1r* and *IT* phases.

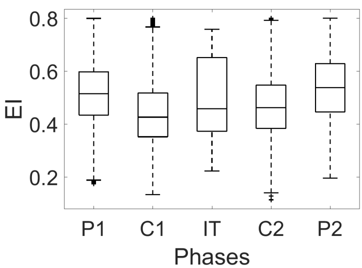

Figure C.50.: Box-plot of EIs of P-03 during SPO.

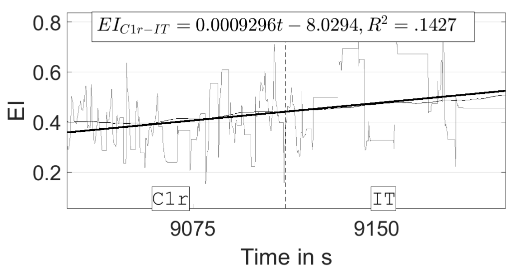

Figure C.51.: Linear regression of EI for subject P-03 for *C1r* and *IT*.

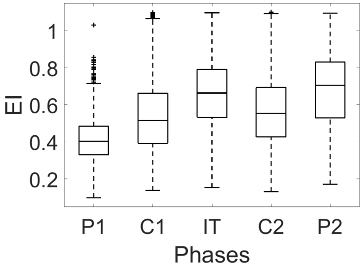

Figure C.52.: Box-plot of EIs of P-04 during SPO.

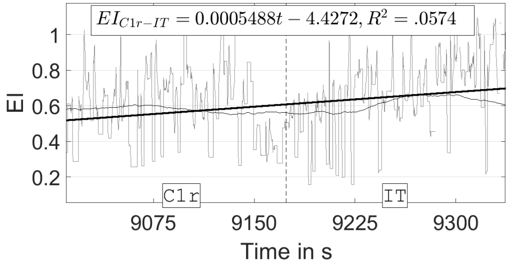

Figure C.53.: Linear regression of EI for subject P-04 for *C1r* and *IT*.

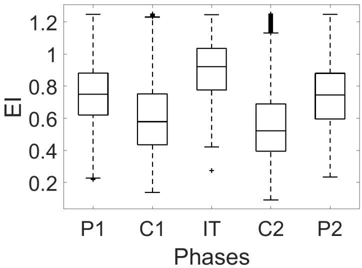

Figure C.54.: Box-plot of EIs of P-06 during SPO.

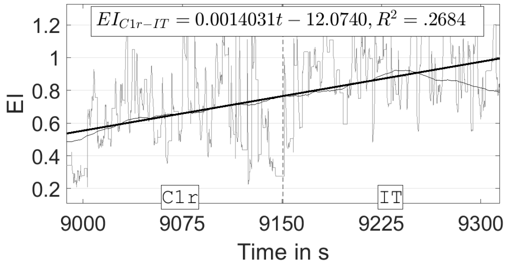

Figure C.55.: Linear regression of EI for subject P-06 for *C1r* and *IT*.

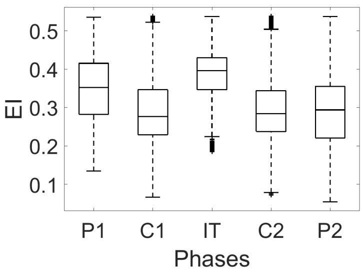

Figure C.56.: Box-plot of EIs of P-07 during SPO.

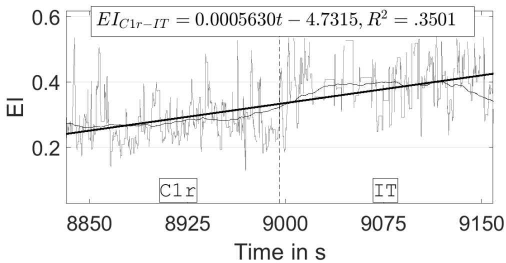

Figure C.57.: Linear regression of EI for subject P-07 for *C1r* and *IT*.

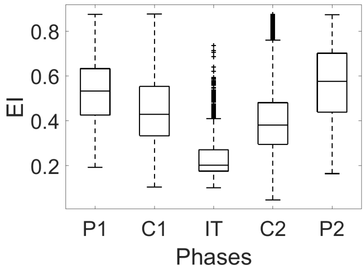

Figure C.58.: Box-plot of EIs of P-09 during SPO.

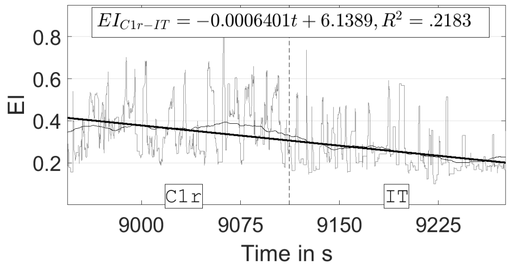

Figure C.59.: Linear regression of EI for subject P-09 for *C1r* and *IT*.

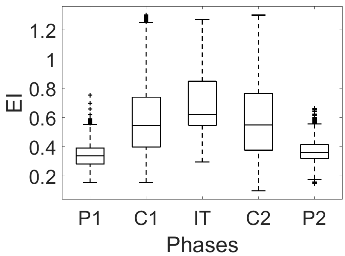

Figure C.60.: Box-plot of EIs of P-10 during SPO.

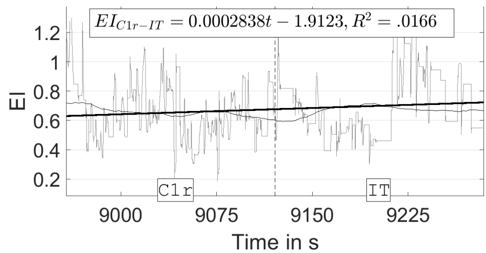

Figure C.61.: Linear regression of EI for subject P-10 for *C1r* and *IT*.

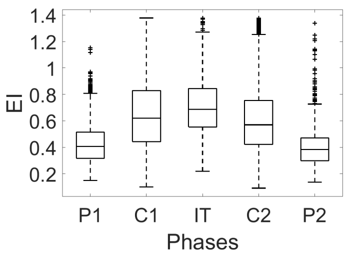

Figure C.62.: Box-plot of EIs of P-12 during SPO.

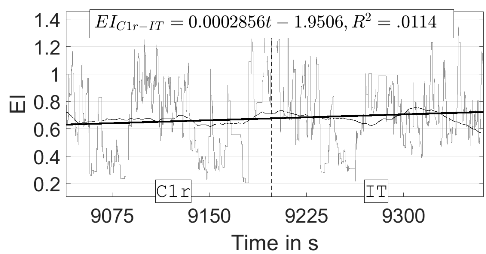

Figure C.63.: Linear regression of EI for subject P-12 for *C1r* and *IT*.

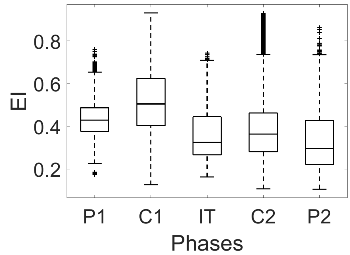

Figure C.64.: Box-plot of EIs of P-13 during SPO.

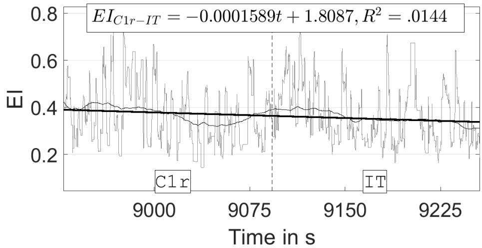

Figure C.65.: Linear regression of EI for subject P-13 for *C1r* and *IT*.

Table C.6 lists the regression coefficients for nine participants (excluding P-15).

Table C.6.: Linear regression coefficients m (slope) for EI for nine participants over phases *C1r* and *IT*. All values multiplied by $1 \cdot 10^{-4}$.

ID	coefficient	ID	coefficient	ID	coefficient
P-01	7.3957	P-03	9.2956	P-04	5.4881
P-06	14.0306	P-07	5.6300	P-09	−6.4012
P-10	2.8378	P-12	2.8559	P-13	−1.5889

C.8 Concentration of Oxygenated Hemoglobin: Phase Comparison

This section contains all boxplots of HbO2 values grouped into the five phases *P1*, *C1*, *IT*, *C2*, and *P2*, as well as the linear regression plots of *C1r* and *IT* phases.

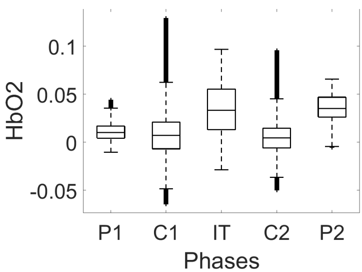

Figure C.66.: HbO2 box-plot subject for P-01.

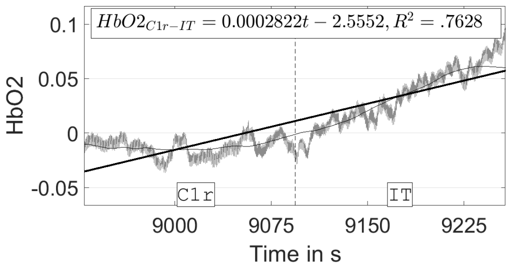

Figure C.67.: Linear regression of HbO2 for subject P-01 for *C1r* and *IT*.

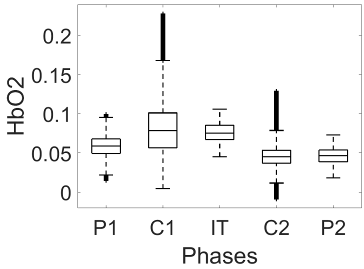

Figure C.68.: Box-plot of HbO2 of P-03 during SPO.

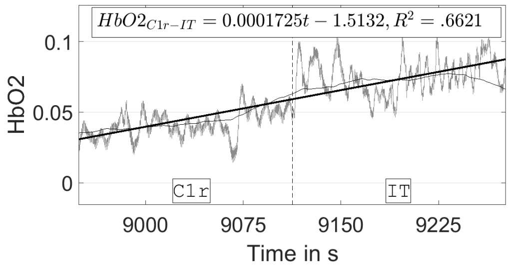

Figure C.69.: Linear regression of HbO2 for subject P-03 for *C1r* and *IT*.

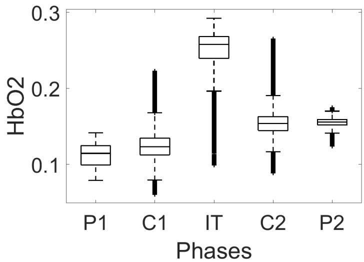

Figure C.70.: Box-plot of HbO2 of P-04 during SPO.

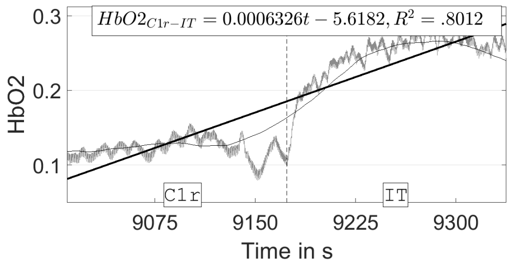

Figure C.71.: Linear regression of HbO2 for subject P-04 for *C1r* and *IT*.

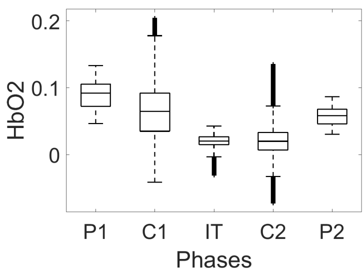

Figure C.72.: Box-plot of HbO2 of P-06 during SPO.

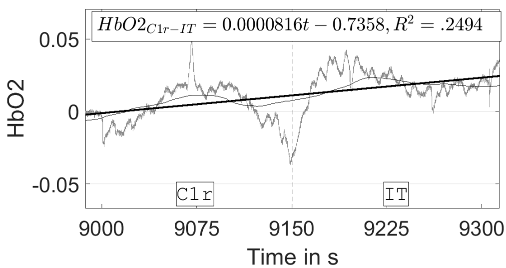

Figure C.73.: Linear regression of HbO2 for subject P-06 for *C1r* and *IT*.

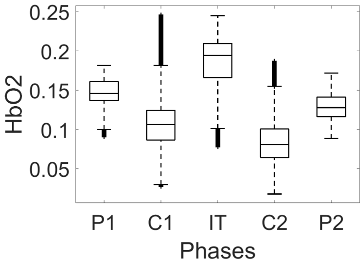

Figure C.74.: Box-plot of HbO2 of P-07 during SPO.

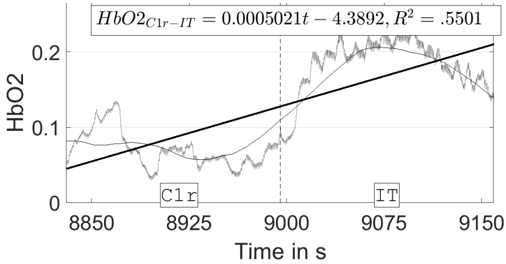

Figure C.75.: Linear regression of HbO2 for subject P-07 for *C1r* and *IT*.

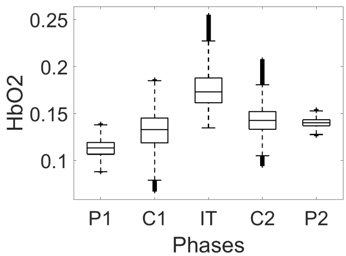

Figure C.76.: Box-plot of HbO2 of P-10 during SPO.

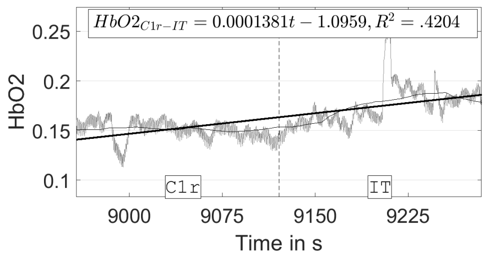

Figure C.77.: Linear regression of HbO2 for subject P-10 for *C1r* and *IT*.

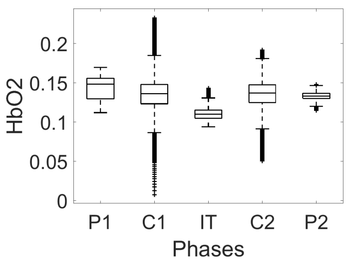

Figure C.78.: Box-plot of HbO2 of P-12 during SPO.

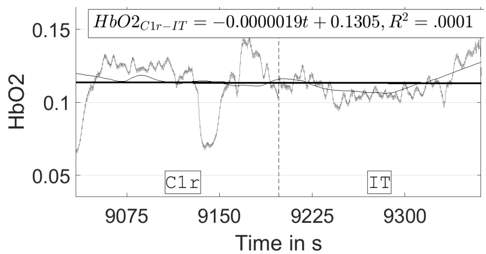

Figure C.79.: Linear regression of HbO2 for subject P-12 for *C1r* and *IT*.

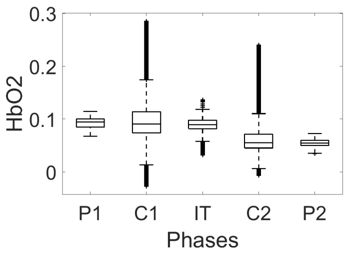

Figure C.80.: Box-plot of HbO2 of P-13 during SPO.

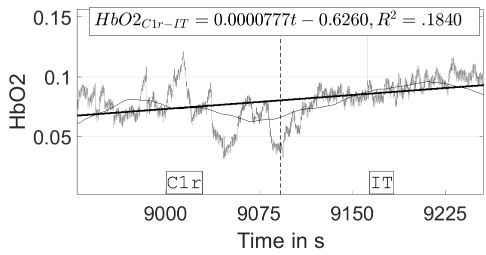

Figure C.81.: Linear regression of HbO2 for subject P-13 for *C1r* and *IT*.

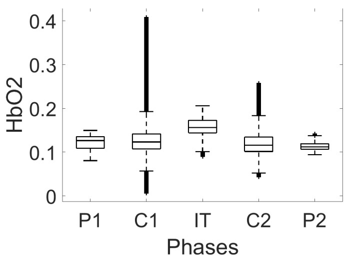

Figure C.82.: Box-plot of HbO2 of P-15 during SPO.

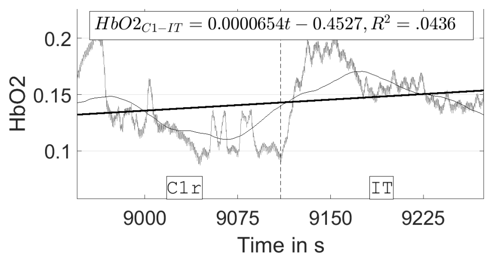

Figure C.83.: Linear regression of HbO2 for subject P-15 for *C1r* and *IT*.

Table C.7 lists the regression coefficients m for the nine participants (excluding P-09).

Table C.7.: Linear regression coefficients m (slope) for COH for nine participants over phases *C1r* and *IT*. All values multiplied by $1 \cdot 10^{-4}$.

ID	coefficient	ID	coefficient	ID	coefficient
P-01	2.8221	P-03	1.7255	P-04	6.3264
P-06	0.8163	P-07	5.0213	P-10	1.3807
P-12	−0.0187	P-13	0.7768	P-15	0.6538

C.9 Heart Rate: Comparison Between Phases

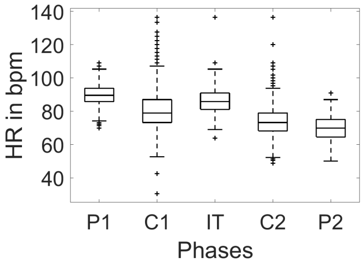

Figure C.84.: HR box-plot for subject P-01.

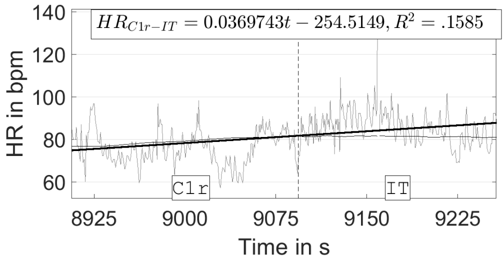

Figure C.85.: Linear regression of subject P-01 HR for *C1r* and *IT*.

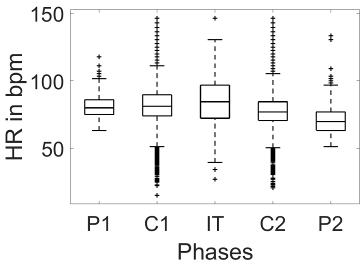

Figure C.86.: Box-plot of HR of P-03 during SPO.

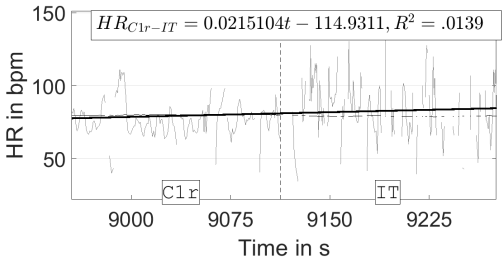

Figure C.87.: Linear regression of HR for subject P-03 for *C1r* and *IT*.

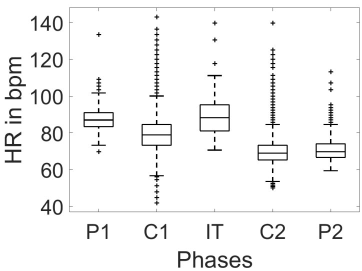

Figure C.88.: Box-plot of HR of P-04 during SPO.

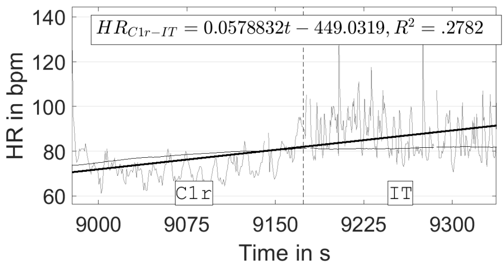

Figure C.89.: Linear regression of HR for subject P-04 for *C1r* and *IT*.

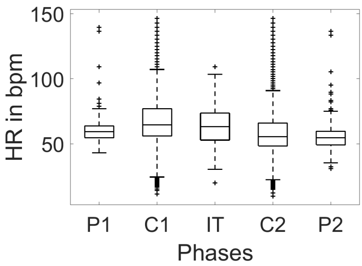

Figure C.90.: Box-plot of HR of P-06 during SPO.

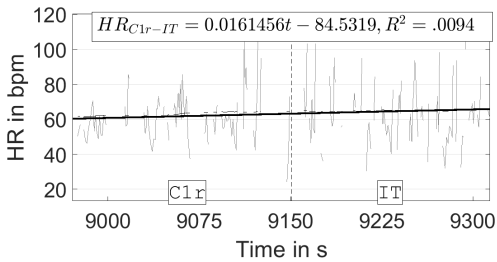

Figure C.91.: Linear regression of HR for subject P-06 for *C1r* and *IT*.

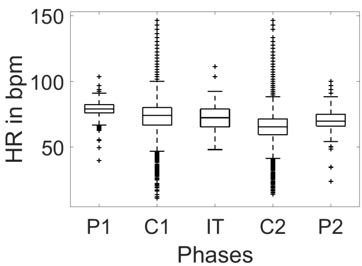

Figure C.92.: Box-plot of HR of P-07 during SPO.

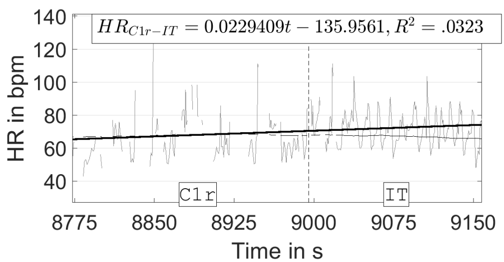

Figure C.93.: Linear regression of HR for subject P-07 for *C1r* and *IT*.

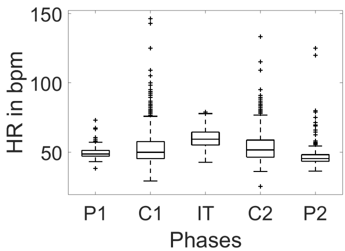

Figure C.94.: Box-plot of HR of P-10 during SPO.

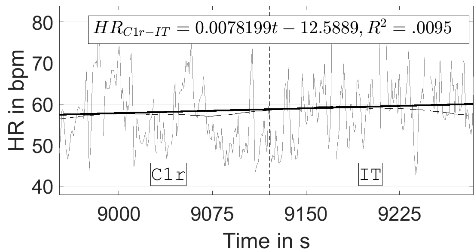

Figure C.95.: Linear regression of HR for subject P-10 for *C1r* and *IT*.

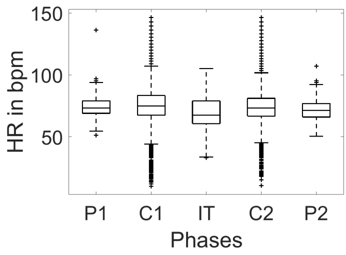

Figure C.96.: Box-plot of HR of P-12 during SPO.

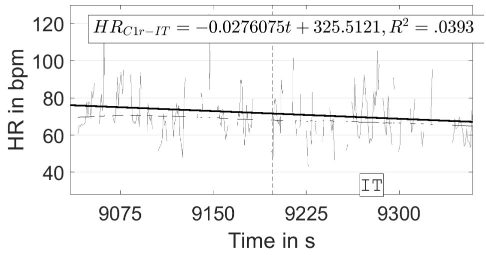

Figure C.97.: Linear regression of HR for subject P-12 for *C1r* and *IT*.

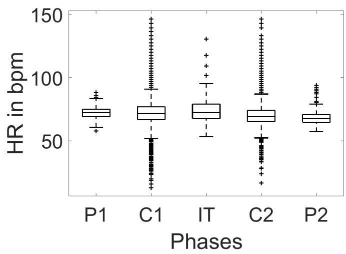

Figure C.98.: Box-plot of HR of P-13 during SPO.

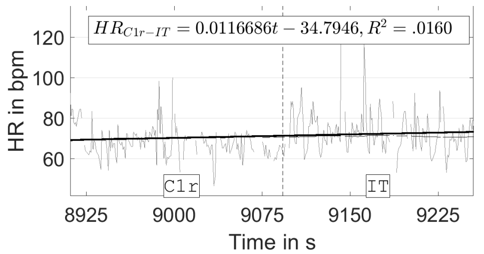

Figure C.99.: Linear regression of HR for subject P-13 for *C1r* and *IT*.

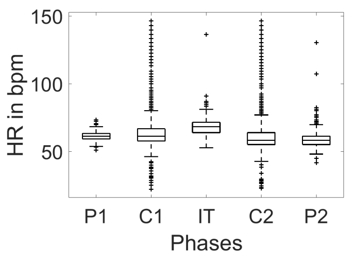

Figure C.100.: Box-plot of HR of P-15 during SPO.

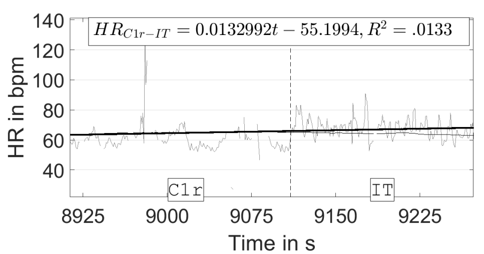

Figure C.101.: Linear regression of HR for subject P-15 for *C1r* and *IT*.

Table C.8 lists the regression coefficients m for the nine participants (excluding P-09).

Table C.8.: Linear regression coefficients m (slope) for COH for nine participants over phases *C1r* and *IT*. All values multiplied by $1 \cdot 10^{-4}$.

ID	coefficient	ID	coefficient ID	coefficient	
P-01	0.0370	P-03	0.0215	P-04	0.0579
P-06	0.0161	P-07	0.0229	P-10	0.0078
P-12	−0.0276	P-13	0.0117	P-15	0.0133

C.10 Performance under Dual Pilot Operations (DPO)

A direct comparison of response times to the Psychomotor Vigilance Task (PVT) from both crew complement conditions is not meaningful, and would lead to wrong conclusions. As the two conditions were administered on different days, too many uncontrollable variables play into the response times. Still, an analysis of response times before and after the DPO condition is executed towards the first hypothesis and to determine if the flight had an effect on response times after the flight.

Figure C.102 shows box-plots of the response times in the DPO scenario of those 10 subjects who executed both scenarios, before and after the flight.

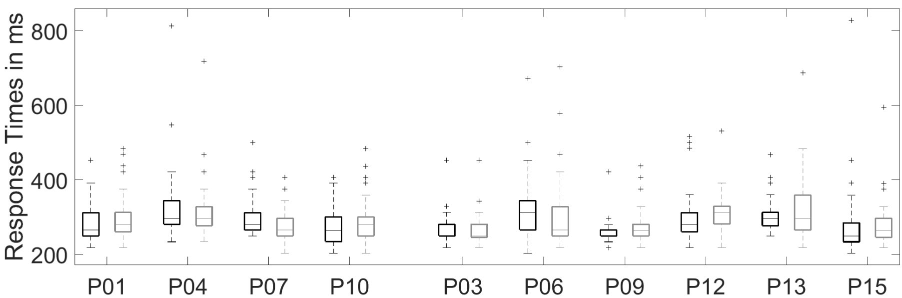

Figure C.102.: Box-plots of response times to the PVT before (black) and after (gray) the flight in the DPO condition. Indicated are median response times with the whiskers encompassing 2.7σ data. The left four subjects started with DPO, the right six started with SPO.

Again, the effects of the 4 hour flight on response times are evaluated. Data was not normally distributed as assessed by the Shapiro-Wilk test ($p < .001$), homoscedasticity requirements were also not met as assessed by the Levene test ($F(19{,}880) = 2.7411, p = .0001$). Still, due to its robustness (see section 5.2.1), a two-factor (*subject number* and *time*) repeated-measures Analysis of Variance (ANOVA) was conducted. For the intra-subject factor, time of the PVT, the ANOVA reports again no statistically significant difference: $F(1{,}880) = 0.11, p = .7466$. The same result was obtained when performing the ANOVA on data with outliers removed: $F(1{,}825) = 0.01, p = .9285$.

Again, the DPO condition does not have effects on the performance after the flight.

D Supplementary Material to the Developed Concept of Operations

D.1 Operations Control Center Functions

The most common functions of a typical Operations Control Center (OCC) may be summarized as follows, cf. [CRO14, Jep07, Cla98], see also FAA AC 120-101:

Flight Dispatch includes flight planning, Air Traffic Control (ATC) coordination, weight and balance planning, operations support (e.g. navigation specialists, operations analysts), and weather services. They prepare flight plans and request new slots to ATC entities. They ensure a serviceable aircraft and qualified crew, and monitor flight operations from origin to destination. They communicate with flight crews, and coordinate and implement plans to resolve off-schedule operations. In the U.S., the flight dispatcher shares legal responsibility with the pilot in command.

Aircraft Control manage the resource *aircraft*. It is a central coordination role. In disruptive situations, they try to minimize delays through changing aircraft or rerouting flights.

Crew Control manage the resource *crew*. They schedule and track flight and cabin crews. In case of disruptions, crew control will use reserve crews and update the crew roster.

Maintenance Control is responsible for any unplanned maintenance services and short-term maintenance scheduling. They provide technical advice and arrange the movement of parts.

Passenger Services and Cargo Control consider and minimize the impact of any OCC decision on passengers. Passenger Services act as liaison between Airport Passenger Services and the OCC, they coordinate passenger re-accommodation and over-sales. Cargo Control maximize freight uplift capability and coordinate timely delivery.

Operations Management oversee the workflow and decisions of an OCC. They report operations status to upper management, represent upper management in operational decisions, and define the daily operational plan. Operations Coordination implement agreed plans to handle disruptions and communicate these plans to all OCC functional groups.

Communication is a central function of an OCC, both internally between functional groups to minimize overall disruption effects and externally with other air transport system stakeholders such as ATC, flight crew, regulatory bodies, airline airport offices, or maintenance centers.

D.2 Operations Control Center Tasks

Based on subsection 6.5.3, more details are given for the identified airline OCC tasks:

Flight Planning generating and filing flight plans, fuel calculations.

Crew Briefing communicating and discussing the flight plan with flight crew.

Movement & Flight Control supervising fleet and individual flights and aircraft.

Maintenance Control planning and coordinating of maintenance events, spare parts management, providing technical expertise.

Weather Analysis

Crew Planning crew assignment, crew management.

Crew Tracking supervising crew check-ins.

ATC Coordination communicating with ATC.

Load Planning generating load and trim manifest, communicating with ground personnel.

Performance Analysis supervising operational airline performance, developing new processes.

Passenger Coordination communicating with passengers, minimizing impact of disruptions on passengers, and handling disruptions.

Cargo & Catering Coordination maximizing cargo capacity, coordinating loading process, coordinating catering.

Station Control planning and coordinating ground personnel, coordinating passenger handling, coordinating ramp control.

Emergency Handling managing emergencies.

IT Support & Databases providing maintenance and support.

Sales & Marketing conducting analyses and market studies.

D.3 Transferability Analysis of Operations Control Center Tasks

Each OCC task (see previous section) was rated in each of the three categories task interruptibility, task complexity, and operator and task autonomy (see section 6.5.3 and [Grä17]) against each dimension on a three-point scale: no or low (-, value: 1), medium (o, value: 2), and high (+, value: 3). Additionally, individual dimensions were weighted, the higher the weight, the higher its influence in the transferability evaluation result. Dimensions are weighted negatively in case a higher rating has a negative influence.

The higher the overall score, the more suitable a task is to be transferred from today's OCCs to Mission Managers (MMs).

Table D.1.: Results of the OCC task transferability analysis.

	Complexity					Interruptibility			Autonomy		
Task	Complexity	Time criticality	Duration	Required knowledge	number of input variables	Interruptibility	Divisibility	Intermediate results obtainable	Continuous data-link dependence	number of people involved	**Score**
Weight	2	-2	-1	-1	-1	3	2	2	-1	-1	
Flight Planning	+	o	+	+	+	o	o	-	+	o	0
Crew Briefing	-	o	-	-	o	o	+	+	o	o	8
Movement & Flight Control	o	+	o	-	-	+	+	+	+	o	10
Maintenance Control	o	-	o	o	o	+	+	+	o	o	13
Weather Analysis	o	-	o	+	o	+	+	+	o	-	13
Crew Planning	o	+	o	-	+	+	+	+	+	o	8
Crew Tracking	-	o	o	-	-	+	+	+	o	o	11
ATC Coordination	o	o	o	o	+	o	+	+	o	o	7
Load Planning	o	+	o	o	o	+	+	+	o	-	10
Performance Analysis	o	-	o	+	+	+	+	+	o	-	12
Passenger Coordination	o	+	o	o	+	+	+	+	o	+	7
Cargo & Catering Coordination	o	+	o	-	o	+	+	+	o	o	10
Station Control	+	+	o	o	+	o	o	o	+	+	1
Emergency Handling	+	+	o	o	+	-	o	o	+	+	-2
IT Support & Databases	o	-	o	o	o	o	+	+	o	-	11
Sales & Marketing	o	-	+	+	+	+	+	+	-	-	12