An approach to goal directed information management on the flight deck

Ein Ansatz für zielorientiertes Informationsmanagement im Flugzeugcockpit
Zur Erlangung des akademischen Grades Doktor-Ingenieur (Dr.-Ing.)
Genehmigte Dissertation von Sebastian Michael Sprengart aus Seeheim-Jugenheim
Tag der Einreichung: 29.06.2021, Tag der Prüfung: 27.10.2021

1. Gutachten: Prof. Dr.-Ing. Uwe Klingauf
2. Gutachten: Prof. Dr.-Ing. Ralph Bruder
Darmstadt – D 17

TECHNISCHE UNIVERSITÄT DARMSTADT

Mechanical Engineering Department

Institut für Flugsysteme und Regelungstechnik

An approach to goal directed information management on the flight deck
Ein Ansatz für zielorientiertes Informationsmanagement im Flugzeugcockpit

Accepted doctoral thesis by Sebastian Michael Sprengart

1. Review: Prof. Dr.-Ing. Uwe Klingauf
2. Review: Prof. Dr.-Ing. Ralph Bruder

Date of submission: 29.06.2021
Date of thesis defense: 27.10.2021

Darmstadt – D 17

Für meine Partnerin, meine Familie, meine Freunde und alle Kollegen, die mich während dieser Arbeit unterstützt oder ermutigt haben. Danke für diese tolle und ereignisreiche Zeit, in der ihr meine Begeisterung geteilt habt - egal ob freiwillig oder auch manchmal unfreiwillig.

Sebastian Sprengart, Januar 2022

Erklärungen laut Promotionsordnung

§8 Abs. 1 lit. c PromO

Ich versichere hiermit, dass die elektronische Version meiner Dissertation mit der schriftlichen Version übereinstimmt.

§8 Abs. 1 lit. d PromO

Ich versichere hiermit, dass zu einem vorherigen Zeitpunkt noch keine Promotion versucht wurde. In diesem Fall sind nähere Angaben über Zeitpunkt, Hochschule, Dissertationsthema und Ergebnis dieses Versuchs mitzuteilen.

§9 Abs. 1 PromO

Ich versichere hiermit, dass die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Quellen verfasst wurde.

§9 Abs. 2 PromO

Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, 29.06.2021

S. Sprengart

Abstract

Today's flight decks are the result of an evolutionary development process. With every design step the level of automation and amount of available information increased. Considerable challenges in relation to information management are documented. Looking into the future, the introduction of Reduced Crew Operation (RCO) may aggravate these challenges. Removing the second pilot will remove cognitive capacity, which is a central factor for today's information management.

The increasingly relevant question of how to properly manage information on the flight deck in the future is evaluated in this thesis.

A development process based on the "Ergonomics of Human System Interaction" standard is pursued to answer this question. Challenges related to information management and existing efforts in this domain are identified. Based on a context analysis of future operations a new concept of operations centered on the human operator on the flight deck is developed. *Mission manager* is established as a new job title for pilots.

Requirements towards information management on the future flight deck are derived. An information management concept for a human operator interacting with a highly automated aircraft is proposed. Goal oriented Information Management (GoIM) is developed to describe this interaction. The implementation of the concept of operations and GoIM in hard- and software is described. Hypotheses towards the parameters of effectivity, efficiency, and satisfaction of usability are formulated. An overall beneficial rating of usability is hypothesized.

Evaluation of the concept is performed using a purpose-built RCO research simulator. Users are tasked to perform a two-part evaluation study, consisting of an application usability test and a scenario-based evaluation, utilizing the software implementation. It is shown with a 95 % confidence level, that the GoIM concept offers acceptable usability. Backed by positive user feedback GoIM is shown to be a potential solution for future information management on the flight deck.

Recommendations towards the further development of the herein proposed concept are given. A closer user interaction and more focused design work is advised. Further broadening the concept and performance-based evaluation and validation is recommended.

Zusammenfassung

Heutige Flugzeugcockpits sind das Resultat eines evolutionären Entwicklungsprozesses. Mit jedem Designschritt hat sich der Automationsgrad und der Umfang der verfügbaren Informationen erhöht. Probleme in Bezug auf das Informationsmanagement sind bekannt. Mit Blick in die Zukunft wird die Einführung von Reduced Crew Operation (RCO) dies weiter verschärfen. Das Entfernen des zweiten Piloten wird eine Reduktion der kognitiven Kapazität zur Folge haben.

Die zunehmend relevante Fragestellung, wie Informationen im Cockpit der Zukunft verwaltet werden, ist Gegenstand dieser Thesis.

Ein auf der Norm "Ergonomie der Mensch-System-Interaktion" basierter Entwicklungsprozess wird zur Beantwortung der Frage herangezogen. Herausforderungen im Bereich des Informationsmanagements werden identifiziert. Basierend auf einer Kontextanalyse zukünftiger Operationen wird ein neues Operationskonzept für den menschlichen Akteur im Cockpit beschrieben. *Mission manager* wird als neuer Tätigkeitstitel für Piloten eingeführt.

Anforderungen an die Gestaltung des Informationsmanagements im zukünftigen Cockpit werden abgeleitet. Ein Konzept zur Beschreibung der Interaktionen zwischen Mensch und hoch automatisierten Luftfahrzeugsystemen wird beschrieben. Goal oriented Information Management (GoIM) wird zur Beschreibung dieser Interaktionen entwickelt. Eine Implementierung von GoIM in Hard- und Software findet statt. Hypothesen hinsichtlich Effektivität, Effizienz und Nutzerzufriedenheit werden formuliert. Eine insgesamt positive Bewertung der Gebrauchstauglichkeit wird hypothesiert.

Die Evaluierung des Konzeptes wird in einem speziell angefertigtem RCO Flugsimulator durchgeführt. Die Nutzer werden durch eine zweistufige Evaluierung geführt, bestehend aus Gebrauchstauglichkeits- und einer Szenario basierten Evaluierung. Die entwickelte Software wird hierbei eingesetzt. Mit einem 95 % Konfidenzintervall wird GoIM eine akzeptable Gebrauchstauglichkeit attestiert. Gestützt durch positives Feedback wird GoIM als eine mögliche Lösung für Informationsmanagement im zukünftigen Cockpit gesehen.

Empfehlungen für nachfolgende Arbeiten werden gegeben. Engere Einbindung der Nutzer in den Entwicklungsprozess, sowie die Weiterentwicklung des Konzeptes und weitere Validierungen werden angeraten.

Contents

Nomenclature

Symbols and subscripts

Symbols

Notation	Description	Unit
λ	Effectiveness score	-
a	Action script/routine	-
b	Time correction factor	-
G	Goal state	-
H	Hypothesis	-
I	Initial state	-
O	Operator	-
P	Set of prepositional state variables	-
q	Factor for time limit	-
S	World state	-
t	Time	s

Mathematical notations

Notation	Description	Notation	Description
$\wedge$	Logical "and"	IQR	Interquartile Range
μ	Expectation value	Mdn	Median
$\subset$	Subset	p	p-value
$\subseteq$	Proper subset	r	Effect size
τ	Kendall's τ	W	Wilcoxon test statistic

Subscripts

Notation	Description
$0...9$	Numbered indices
a	Alternative
e	Effect
exp	Experienced user
$I, II, ...$	Numbered indices
i, n	Indices
$initial$	Initial value
m	Mission
max	Maximum
MCH	Related to MCH rating
pc	Preconditions
$react$	Reaction time
ref	Reference value
$SLOW$	Related to slow-down event in scenario
SUS	Related to SUS rating
$task$	Related to a task
VHF	Related to CPDLC event in scenario

Acronyms

A*	A* search algorithm
AC	Advisory Circular
ACROSS	Advanced Cockpit for Reduction of StreSs and workload
ACU	Artificial Cognitive Unit
ADL	Action Description Language
ADS-B	Automatic Dependent Surveillance - Broadcast
AIME	Automation and Information Management Experiment
AIMS	Airplane Information Management System
AMgr	Agenda Manager
ANCS	Aviate Navigate Communicate Manage Systems
ANOVA	Analysis of Variance
ANSP	Air Navigation Service Provider
AOC	Airline Operation Center
API	Application Programming Interface
ARINC	Aeronautical Radio Incorporated
ARP	Aerospace Recommended Practice
ASRS	Aviation Safety Reporting System
ATC	Air Traffic Control
ATM	Air Traffic Management
ATSU	Air Traffic Service Unit
BGS	Boeing Global Services
BPMN	Business Process Model and Notation
CA	Captain
CAMA	Crew Assistant Military Aircraft
CAST	Commercial Aviation Safety Team
CASSY	Cockpit ASsistant SYstem
CFR	Code of Federal Regulations
COTS	Commercial off-the-shelf
CPDLC	Controller-Pilot Data Link Communications
CRAN	Comprehensive R Archive Network
CTM	Crew Task Management
DIKW	Data-Information-Knowledge-Wisdom
DLR	Deutsches Zentrum für Luft- und Raumfahrt
DPO	Dual Pilot Operations
DOT	U.S. Department of Transportation
EASA	European Aviation Safety Agency
ECAM	Electronic Centralized Aircraft Monitoring

EDDF	Frankfurt Airport
EFB	Electronic Flight Bag
EICAS	Engine Indication and Crew Alerting System
ESP	Electronic Standby Pilot
EU	European Union
EUROCONTROL	European Organization for the Safety of Air Navigation
FAA	Federal Aviation Administration
FCOM	Flight Crew Operating Manual
FCU	Flight Control Unit
FMS	Flight Management System
FO	First Officer
FSR	Institute of Flight Systems and Automatic Control
GATS	Global Air Transportation System
GOAP	Goal Oriented Action Planning
GoIM	Goal oriented Information Management
HAT	Human-Autonomy Teaming
HMI	Human-Machine Interface
HTA	Hierarchical Task Analysis
IATA	International Air Transport Association
ICAO	International Civil Aviation Organization
IID	Information and Interaction Design
ILS	Instrument Landing System
IMP	Information Management Panel
JCAB	Japan Civil Aviation Bureau
JSON	JavaScript Object Notation
KBOS	Boston Logan Airport
KLAX	Los Angeles International Airport
LED	Light-Emitting Diode
MVC	Model View Controller
NASA	National Aeronautics and Space Administration
NextGen	Next Generation Air Transportation System
NTSB	National Transportation Safety Board
MCH	Modified Cooper-Harper
MCP	Mode Control Panel
PA	Pilot's Associate
PF	Pilot Flying
PIC	Pilot in Command
PILAS	PILoten-ASsistenz-System für Luftfahrzeuge

PM	Pilot Monitoring
PNF	Pilot Not Flying
QML	Qt Modeling Language
RCO	Reduced Crew Operation
RPA	Rotorcraft Pilot's Associate
RTCA	Radio Technical Commission for Aeronautics
SA	Situation Awareness
SAE	Society of Automotive Engineers
SERA	Standardised European Rules of the Air
SESAR	Single European Sky ATM Research Programme
SID	Standard Instrument Departure
SMM	Safety Management Manual
STAR	Standard Terminal Arrival Route
STRIPS	Stanford Research Institute Problem Solver
SUS	System Usability Scale
SWIM	System Wide Information Management
TBO	Trajectory Based Operation
TCP	Transmission Control Protocol
TUDA	Technische Universität Darmstadt
UAV	Unmanned Aerial Vehicle
UI	User Interface
US	United States
USB	Universal Serial Bus
UTC	Coordinated Universal Time
UX	User Experience
VHF	Very High Frequency

1. Introduction

The motivation outlines how technological development on the flight deck has left the human operator behind and expands on the consequences thereof. Derived from the motivation is the subsequently described objective of this thesis. Eventually, the structured approach based on the DIN EN ISO 9241-210 “Human-centered design for interactive systems” standard [DIN11] is described. The chapters of this thesis are mapped to the individual process steps.

1.1. Motivation

Modern flight decks are the result of a continuous evolutionary development process throughout the last century of aviation. During this development process the advancements of technology allowed a steady decrease of crew size on the flight deck [Lac+14a]. A reduction in crew size was achieved by reallocating human functions to (partially) automated systems onboard the aircraft [SHS16], increasing overall system complexity [You+16a]. Eventually, a reduction of crew size from five to two was made possible and is still the prevailing flight deck configuration for more than three decades in the commercial aviation domain [Boy14]. The increasing presence of automated systems influenced the role and responsibilities of the crew [Bil91; Boy98], the role of pilots has changed from being aviators to managers of systems [Har07; BBR07; Boy14; Sat93; Sch15]. Considering the proposed changes under Next Generation Air Transportation System (NextGen) [Joi11] and Single European Sky ATM Research Programme (SESAR) [Sin15a], the degree of automation is likely to further increase. In conjunction with higher responsibilities for the individual flight crew members, the task of actually flying the aircraft is increasingly loosing relevance [Joi11; Dua+15].

The initial motivation behind the development of advanced automation technology was to improve precision and economy of operations, as well as lowering training requirements [SWB97]. While it achieved to alleviate some of the routine pilot tasks, it also created “unexpected” [SWB97] problems: Automation did not reduce workload across the board, as it initially appeared [Wie88; FSW10]. Previously physical workload turned into increased cognitive, or mental, workload for the pilot [BBR07; SW92], turning into a problem in high workload situations [Dek04]. Summarized under the term of *clumsy* automation [Wie88] is the circumstance that workload is further reduced in low-workload phases such as cruise while it is increased in high workload phases because the pilot must assess the situation while simultaneously instructing and monitoring automated systems [FSW10; SWB97; SW91]. As a result pilots controversially tend to deactivate automation functions in high workload situations [CW02; Bur+13].

Connections to the management of information between human operator and automated systems can be drawn: Successful cooperation between a human operator and automated systems requires the operator to maintain an accurate mental model and understanding of the current system state [SW97; JR95; UJ02]. Mismatches between the mental model and the state of automated systems can lead to *mode errors* [SMW07] which are followed by *automation surprises* [SWB97] when actions performed by the automation do not match expectations [SWB97; SW92]. An example of a fatal breakdown in information management is the tail first landing in Nagoya in 1994 [Air96; BBR07]. The required data to assess the situation is usually available on the flight deck [WPR02; DBS13], in fact the number of sources and amount is growing [You+16a; UJ02]. Bridging the gap between data and information [Ack89] requires undesirable cognitive effort which is ultimately needed to update the mental model of the system [End95; Sch+07]. An effective assessment of safety critical information is often not feasible [Sch+12; Eth+19], Air-Transat flight 236 [Avi04] can be seen as an example [DBS13].

In the light of a series of accidents [Air96; CJP96; Hal+96] attention was drawn to the interfaces and interactions between operators and aircraft systems [ASS96]. A report compiled by the Federal Aviation Administration (FAA) [ASS96] in 1996 summarizes the recognized shortfalls of Human-Machine Interfaces (HMIs) on modern flight decks, published alongside a set of recommendations to mitigate those. In 2013 a second report under the guidance of the FAA was published, focusing on the changes made since 1996 [AMR13]. One of the conclusions drawn from this process is that although there have been improvements, the previously discussed issues, like automation surprises and mode errors, are still (partly-) persistent [AMR13]. A reluctance to facilitate large changes to the human machine system, which would be required to overcome those issues, exists on both, the manufacturer and operator side, not lastly due to long development cycles and high development cost in aviation [SG16; BBR07; ASS96].

At the same time, driven by economical and operational needs [Dao+15; SHS16; Boe16a], the the long resting trend of further reducing the crew on commercial flight decks became subject of research and has been actively researched, for example by National Aeronautics and Space Administration (NASA) [CSC16; Com+13]. The Swiss investment bank UBS suggests total savings of over $15 billion for the world-wide aviation market

[CF17] through the introduction of Reduced Crew Operation (RCO) to the commercial aviation domain. References towards a potential introduction of RCO have been made by Airbus and Boeing [Tré19; Bat18]. Experience with operations performed by a single pilot do exist in military and general aviation flying [SHS14; UJ02]. In contrast to that it has also been recognized that the transition to a single cockpit member is more complex than the previous reductions [DP05; Bai+17]. While flight decks are *technically* certified for (temporary) operations with one pilot under FAR/CS 25.1523-1 [Fed05], research has shown that the current flight deck design is insufficient for RCO [Eth+16; Eth+18; Sch15; Har07]. In this regard the barrier to larger changes is not the technology, but its application when developing automation and user interfaces [Bai+17] – the consideration of the human factors side [Har07; SHS14; SHS16].

Under the assumption that today's cockpit design will not suffice for RCO, we are at a pivotal point where a chance can be taken to learn from the past experiences and to rethink the cockpit architecture for the better [BBR07; Com+13; CF17]. Especially the role of the human operator under RCO should be considered [Han13; Boy14]: It can be argued that the change of the human operator's role was dictated by technology [Abb14]. Over time they became observers of automated systems, a task for which a human is less suitable than an machine would be [Lov95; Dek04; Har07; CS15], which is ultimately a result of past allocation strategies [Sch+07]. The human operator "has been automated out-of-the-loop" [UJ02]. Once aviators, the human operators became managers of systems on the flight deck [Har07; Boy14], in this process the designers took advantage of the human ability to adapt [UJ02]. However, humans should not be forced to adapt to technology, it should be the other way around [Sch+07; SWB97]. Following the argumentation of Norman, who rewrote the motto of the 1933 Chicago World's Fair from "Science finds, industry applies, man conforms" to "people propose, science studies, technology conforms" [Nor93], an approach that considers the human operator first should be chosen [SWB97].

A summary of the previous paragraphs can be drawn: Challenges in the interaction between human operators and automated aircraft systems exist and arguably there is a connection to the way the operator's role was changed by technology. Based on this the aim of this thesis is derived: Conceptualize and validate a concept for how a human operator under RCO will cooperate with automated aircraft systems to successfully reach a common mission goal. In this endeavour the emphasis is first placed on the *role of the human operator* on a future flight deck to avoid a limbo state between being an aviator and a system manager. With this role definition in place the development of an *information management* concept to support the human operators in their role is focused on.

1.2. Objective of this work

The objective of this work is to answer the question of how a human-centric information management system could be designed to allow a high degree of *usability* for the human operator when interacting with highly automated aircraft systems in future commercial aviation operations. This question is asked under the assumption of a different operational environment, mainly characterized by the assumption of RCO and a time horizon beyond 2050 and the associated advances in technology. In this context the role of the human operator is redefined based on the previous considerations.

The intended outcome of answering the previously formulated question is a first proof of concept of how *a* potential solution for information management for use in future operations could be conceptualized. It is not the intent to provide a definite answer on how *the* future information management on the flight deck will work. This work is intended as a contribution to the academic body of flight deck research without constraining itself to just incremental or evolutionary changes. The knowledge and learnings built up in existing research are put into action. Much more, this work is intended to look beyond what is practically possible today, sparking new ideas, encouraging new ways of thinking, and understanding the challenges to improve upon in future work. As part of this effort the following research questions are sought to be answered:

- What will be the role and tasks of human operators on the future flight deck?
- How can human operators and automated aircraft systems cooperate in reaching mission goals?
- How acceptable is a conceptualized solution for the human operators?

Answering those questions is done by empirical investigation, expert consultations, and eventually a simulator study. A novel information management concept is proposed. An implementation in hard- and software is performed. During the analysis and development of the concept the previously stated questions are sought to be answered. In a first evaluation study using a purpose-built research simulator the developed concept is tested. First feedback about the usability of the concept is gathered through this evaluation. Future areas of research and work are identified.

1.3. Methodology and structural approach

For the framing development process conducted in this thesis, a user-centered approach for designing interaction systems is chosen. The followed process is oriented at the DIN EN ISO 9241-210 “Human-centered design for interactive systems” standard [DIN11]. Depicted in Figure 1.1 is the design process, as described by DIN EN ISO 9241-210.

First the context in which the developed interface is going to be used is analyzed (1). Derived from this initial design step are requirements (2), which flow into the following design stage: the creation of design solutions (3). Previously created design solutions are evaluated against the formulated requirements (4). When necessary, another iteration of the design process is initiated, starting at the appropriate design step (5). Eventually, when the requirements are met, the process is stopped (6).

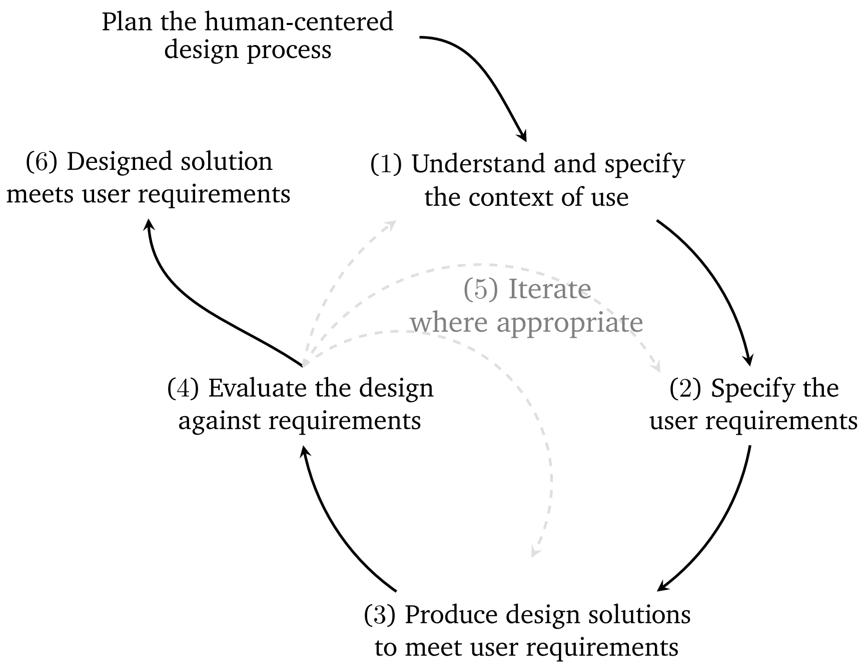


Figure 1.1.: Design process as described in DIN EN ISO 9241-210 "Human-centered design for interactive systems" [DIN11], depiction modified by author.

Additionally to the guidance provided in DIN EN ISO 9241-210 [DIN11], Cooper et al.'s *About Face* [Coo+14], in which the process of *goal-directed interaction design* is followed, is used as practical guidance. Shown in Figure 1.2 is a graphical representation of the chapters in the thesis at hand, matched with the design steps of DIN EN ISO 9241-210.

In Chapter 1 the motivation to this thesis is given. The research objective is stated and a research question is formulated, the overarching method is described. In Chapter 2, the state of the art, an introduction to flight deck information management is provided. Challenges connected to information management on the flight deck are identified. Related work is summarized. A classification of this thesis relative to the identified work is per-

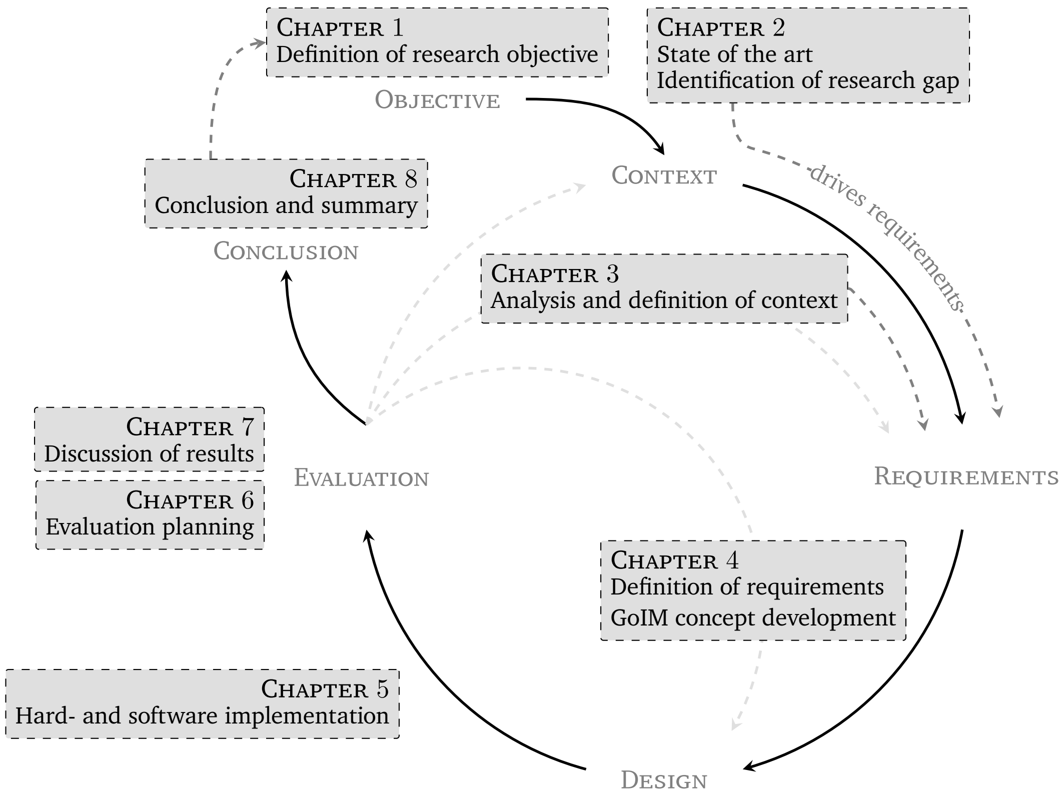


Figure 1.2.: Chapters of the thesis at hand, mapped to the design steps described in DIN EN ISO 9241-210 “Human-centered design for interactive systems” [DIN11].

formed. Utilized psychological models are introduced. Today’s Global Air Transportation System (GATS) is outlined.

In Chapter 3 the expected usage context of the developed information management concept is defined. An overview of the efforts made to understand today’s operating context is given. Expected changes to the GATS as well as background and challenges of RCO are discussed. Assumptions for the subsequent development process are described. The chapter is concluded with a role definition of the future operator, the *mission manager*. Chapter 4 describes the development of the information management concept. As indicated in Figure 1.2, requirements are derived from the existing body of research and related work in Chapter 2. Additional requirements are derived from the context definition in Chapter 3. Based on the requirements, Goal oriented Information Management (GoIM) is developed as an information management concept for the future flight deck in Chapter 4.

The implementation of the concept in hard- and software is described in Chapter 5. In the following Chapter 6 the evaluation of the developed concept is planned. Hypotheses are derived from the objective of this thesis. The evaluation setup and utilized methods are described. A two-fold simulator evaluation is performed. The results of the evaluation are discussed in Chapter 7. Chapter 8 completes the development cycle performed in this thesis. An answer to the research question stated in Chapter 1 is given. A conclusion is drawn, recommendations for future research and development efforts are provided.

2. State of the art

In this chapter the background required for the subsequent chapters of this thesis is provided. The Global Air Transportation System (GATS) and its actors are outlined. Aircraft operating cost and the responsibilities of the cockpit crew are emphasized. Psychological models utilized in this work are presented. A method for automated action planning is introduced. The second half of this chapter focusses on information management in aviation. A general terminology for information management is established to allow a common understanding of the terms used throughout this thesis. Challenges related to flight deck information management are summarized. An overview of existing work in this domain leads to the identification of the targeted research gap.

2.1. Global Air Transportation System

The different elements of the GATS are introduced in the following to facilitate a better understanding of the content presented in chapter 3. An emphasis is placed on the pilots operating an aircraft and the operating cost of an aircraft. Shown in Figure 2.1 is graphical representation of the described actors and their relationships.

Centrally shown in Figure 2.1 are the demand creators *passengers and cargo*, which are both creating a demand for transportation in exchange for money [SG16]. Those demands are fulfilled by *airlines*, whose main contribution is the service of transporting goods and people, with the traditional target market of long-distance service [SG16]. In 2017 airlines transported over 4.1 billion passengers, and 53.9 million tonnes of cargo, representing 35 % of the global trade by value [Int18a]. Airlines are mostly economic driven organizations, seeking to maintain profitability, although out of national interest certain airline types may be subsidized by a government [Hir11]. Cost reductions are sought, especially given the *hyper-competitive* [Wri13] market they are operating in [SG16]. The central element operated by airlines is a fleet of *aircraft*, type and manufacturer depending on airline type and operated network [SG16]. Shown between airline and aircraft in Figure 2.1 are *pilots*, employed by airlines they operate the aircraft and are in contact with an Air Navigation Service Provider (ANSP) during operations. *Airports* serve as the

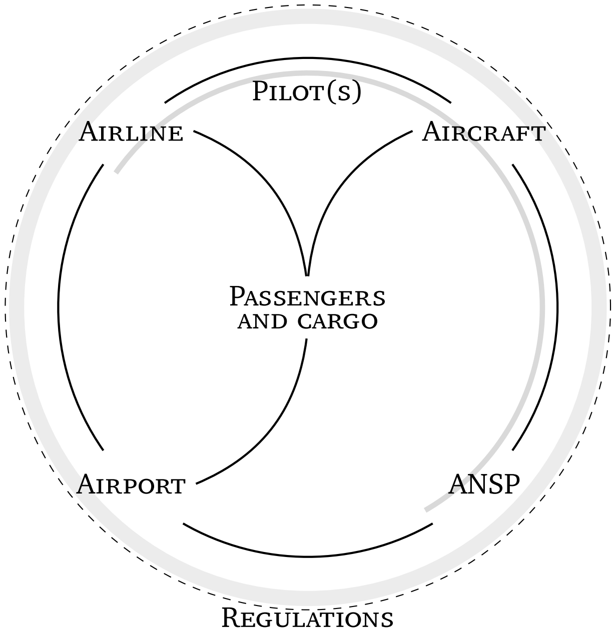


Figure 2.1.: Actors in the Global Air Transportation System and their relations, adapted from the lecture “Systemic Evaluation of the Air Transportation System” at Technische Universität Darmstadt (TUDA), with permission of the lecturer.

connection between airplanes, airlines, and passengers [Hir11]. The relationship between airports and airlines is on a commercial basis, in which airlines negotiate and pay for time slots and services, which are provided by the airport from which they are operating [SG16]. Safe and efficient operations in the GATS are facilitated through ANSPs [SG16]: Supporting a safe and efficient traffic flow in the GATS is supported through *Air Traffic Management (ATM)* [Eur04]. An underlying activity is the management of information, maintenance of communication systems, and surveillance of traffic flow [SG16], for which guidance is provided in International Civil Aviation Organization (ICAO) Doc. 4444 [Int16]. The executive activities are summarized under *Air Traffic Control (ATC)*, which ensures collision prevention, separation between aircraft, and adherence to airspace restrictions [Hir11; SG16]. Rules regarding the operation of ATC are formalized by the ICAO under Annex 11 [Int18b]. Services provided through the ANSPs are paid for by airspace users [Hir11].

The GATS is bound by a set of national and international *regulations*. On a national level regulations are instated by national aviation authorities (e.g., the Federal Aviation

Administration (FAA) in the United States (US)), internationally treaties between individual countries or organizations with constitutive roles, like the International Air Transport Association (IATA) or ICAO exist [SG16]. One aspect of the regulatory framework is the airworthiness of aircraft. With the current level of safety, the target probability for a catastrophic event is estimated at less than 1×10^{-9} per flight hour [Fed88]. Showing compliance with this level of safety is a mandatory step in the certification of an aircraft. Important applicable regulatory frameworks are the CS-25 [Eur18] issued by the European Aviation Safety Agency (EASA), the Aerospace Recommended Practice (ARP) series [S-1a; S-1b] issued by the Society of Automotive Engineers (SAE), DO-178C and DO-254 published by the Radio Technical Commission for Aeronautics (RTCA) [Rad12; Rad00], or the ICAO Annex 8 [Int18c]. Certification of airworthiness is issued by publicly controlled authorities, such as FAA, EASA, or Japan Civil Aviation Bureau (JCAB) [Men13; SG16]. Other regulations are covering the operation of aircraft and the operator's responsibilities, for example the Code of Federal Regulations (CFR) Part 91 [Fed20] or ICAO Doc. 4444 [Int16].

2.1.1. Operating cost of an aircraft

Aircraft are commonly the most expensive single item ($85.8 million to $425.8 million [Boe18]) of an airline, with an expected usage time of 20 to 25 years [Hir11]. The general operating cost of an aircraft is divided into direct and indirect cost [Hir11]. According to [Fed16a], the percentage of direct operating to total cost is about 52 % and 43 % for large passenger and cargo carriers[1], respectively. Absolute numbers reported in 2013 for all passenger aircraft range from an average of $4709 to $11 138 on direct operating cost per block-hour[2] [Fed16a]. Included in direct operating cost are crew salaries, crew training, fuel, maintenance, airport and ANSP service charges, and ground service [Hir11]. Indirect cost on the other hand are associated with aircraft ownership, which can be a result of leasing fees, insurance payments, an overhead airline cost [Hir11; Dog10]. Fuel, maintenance, and crew cost represent the largest cost fractions [SHS16; SUT16]. They offer the biggest leverage for an overall operating cost reduction for airlines by establishing more efficient crew utilization approaches or by replacing fuel-inefficient aircraft [Hir11; Fed16a]. Over the last years a significant fluctuation in fuel prices (up to 40 %) and a trend of increasing crew cost (between 6 % and 16 % for 2014/2015) [SUT16], limit airline planning security.

[1] Yearly revenue greater than $1 billion (Group III) [Fed16a].
[2] Time between closing of doors at departure gate, until opening the doors at the arrival gate [Mas18].

2.1.2. Role and responsibilities of a commercial airline pilot

A commercial passenger airline flight is usually conducted by a crew of two pilots at the same time [Men13]. This mode of operation will be referred to as Dual Pilot Operations (DPO). They can be supported by possible stand by pilots onboard the aircraft for long-haul flights that would exceed 8 h of duty time [Fed05]. The roles and the division of responsibilities between the Pilot in Command (PIC) and the Pilot Monitoring (PM) [Fli14] are examined in the following. The legal responsibilities of the PIC are outlined.

Roles in the cockpit In a commercial two pilot cockpit environment a clear distinction between pilot roles and ranks is made. While both crew members in the cockpit are qualified as pilots, only one of them is flying the aircraft at a time. The pilot currently controlling the aircraft is denoted as Pilot Flying (PF), while the other pilot is denoted as Pilot Not Flying (PNF) or the more recent term PM [Fli14]. During a flight the role of the PF may be transferred between crew members at any time through positive exchange of roles [Fed15; WG15]. While the PF is concerned with flying the aircraft at all times, even when the autopilot is currently engaged, the PM is monitoring the PF and is responsible for tasks such as communication or supporting the PF where needed [Fed15; Hut95]. Independent of the currently performed role, pilots are traditionally ranked as either Captain (CA) or First Officer (FO) [WG15].

Legal responsibilities of the Pilot in Command The PIC is the final authority onboard the aircraft and responsible for safety and security [Int05]. Independent of who is currently the acting PF, the CA as the PIC retains the final authority [WG15; Jen+99]. Declared by ICAO under Annex 2, “Rules of the Air”, par. “2.3.1 Responsibility of pilot-in-command” the PIC’s responsibility is defined as [Int05]:

> The pilot-in-command of an aircraft shall, whether manipulating the controls or not, be responsible for the operation of the aircraft in accordance with the rules of the air, except that the pilot-in-command may depart from these rules in circumstances that render such departure absolutely necessary in the interests of safety.

Extended by Annex 2, “Rules of the Air”, par. “2.4 Authority of pilot-in-command of an aircraft” [Int05]:

> The pilot-in-command of an aircraft shall have final authority as to the disposition of the aircraft while in command.

In accordance with ICAO regulations, national and international regulations exist: For flights operating in, out or within the European Union (EU) Standardised European Rules of the Air (SERA) [Eur12], for US flights CFR 91.3 [Fed]. The PIC is considered to take all actions necessary to guarantee a safe flight conduction, e.g. familiarizing with *all* available information concerning the planned flight during pre-flight [Int05].

2.2. Psychological models

In this section the psychological models and theories utilized in this thesis are introduced. The concept of workload and the multiple resource theory by Wickens are presented. The situation awareness model by Endsley is described. Eventually, the human behavior model after [Ras83] is outlined.

2.2.1. Workload

Workload describes the amount of work attributed to an individual [SJ98], in an aviation context this commonly relates rather to mental than to physical workload. What is described as mental workload is the relation of cognitive resources demanded by a specific task and available cognitive resources of the human operator [PSW08; Wic02]. Characterization of mental workload is challenging [CR07], direct detection of mental workload is considered impossible [XS00]. One inherent problem of measuring an individuals mental workload, is the strong influence of the individual's information processing, motivation, or personality traits [Mes88; ML88].

Multiple resource model

Although the multiple resource model, as proposed by Wickens [Wic81], is *not* a workload theory [Wic02], it provides a way to determine the impact of multiple time shared tasks on task performance in high workload situations [TV06]. Opposed to the theory of having a single resource available for all processing demands [Mor67], the model assumes the availability of multiple resources. Resource attribution, and performance of multiple parallel tasks, is not just dependent on the quantitative demands (difficulty/complexity) of a task, but qualitative demands [Wic81]. Qualitative demands are grouped into three, and later four distinct dichotomous dimensions [Wic08]: processing stages, perceptual modalities, visual channels and processing codes [Wic02]. A visual representation of the model is shown in Figure 2.2.

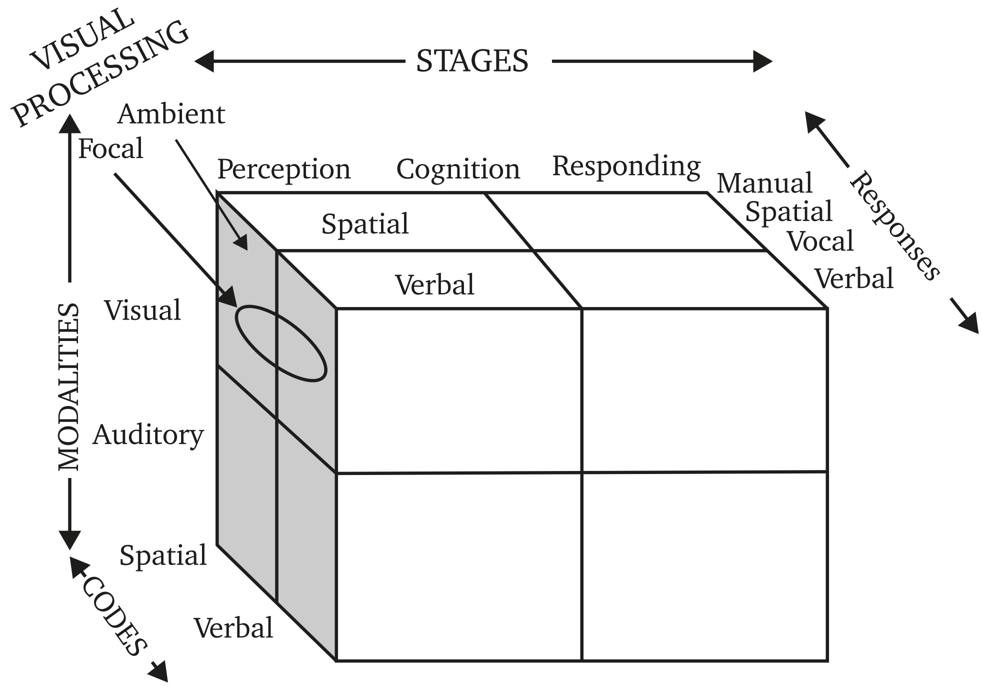


Figure 2.2.: Three dimensional representation of the multiple resource model, the fourth dimension of visual processing is shown within visual resources [Wic02]. Depiction taken from [Wes15] with permission of the author.

The first dimension, processing stages, is divided into the levels of perception and cognition, and responding [Wic02]. Wickens describes that an air traffic controller would be able to acknowledge an aircraft's state change (responding to communication) while maintaining an accurate airspace picture (perception and cognition). The dimension of perceptual modalities is used to describe how information is perceived, in this model either visually or auditory [Wic02]. Related to that is the dimension of visual channels, distinguished between focal and ambient vision [Wic02]. Processing codes distinguish between spatial responding and processing (moving a joystick) and verbal processing and responding (having a conversation) [Wic02].

According to this model, two tasks demanding similar resources (same level in any given dimension) will interfere with each other, potentially resulting in higher workload [TV06]. Aside from being used to predict performance levels, the model can be used as a decision aid when designing interaction modalities for complex environments [Wic02].

Influence on and of workload

Investigated by [Spe71] is the effect of workload on operator task strategies when performing safety-critical tasks. Observed in this study were ATC controllers, which were given the task to define and execute a landing sequence for a number of aircraft. As observed by [Spe71] controllers tend to choose less economical control strategies for low-workload phases, but change to more economical ones (for example shortening voice messages) when workload increases. Through this feedback the workload is influenced and partially compensated [Spe71]. Another influence on the individually available cognitive resources, and the experienced workload levels, is described in [TV06]. Skill level and abilities influence the supply of processing resources [TV06]. This observation is supported by [JCM03], where neurophysiological experiments showed lower brain activity for more proficient subjects, when compared to less proficient ones, indicating a greater availability of cognitive resources [JCM03]. Opposed to the previously described influences on workload, the experienced workload itself also impacts task performance. Described by Yerkes and Dodson in 1908, the Yerkes-Dodson Law is a relation between a stress level and task performance, an observation based on experiments performed on mice [Dia+07]. Observed is a relationship between arousal level as shown in Figure 2.3, where task performance increases with increasing arousal level, but starts to decline after a certain point for complex tasks [YD08]. The performance for a simple task on the other hand increases until it reaches a boundary value and remains constant [YD08]. Hence, to avoid the so called *hours of boredom and moments of terror*, a moderate level of arousal should be kept to improve human proficiency [Sch+07].

Experimental workload determination

In the following two methods to be utilized during the evaluation as described in chapter 6 are outlined. The secondary task method is introduced as an in-experiment method of determining workload. The subjective Modified Cooper-Harper (MCH) rating scale is introduced as a post-experiment measurement.

Secondary task method Workload can only be measured through indirect measurement of highly correlated variables [XS00]. The secondary task method is such an indirect measurement technique, for which the concept of mental workload as outlined in subsection 2.2.1 is employed. Based on the theory that mental workload is the relation between the available and required mental resources [PSW08], the workload imposed on a human operator from a primary task can be indirectly determined by measuring the performance of a secondary task. When users are able to perform well on a secondary task, which is

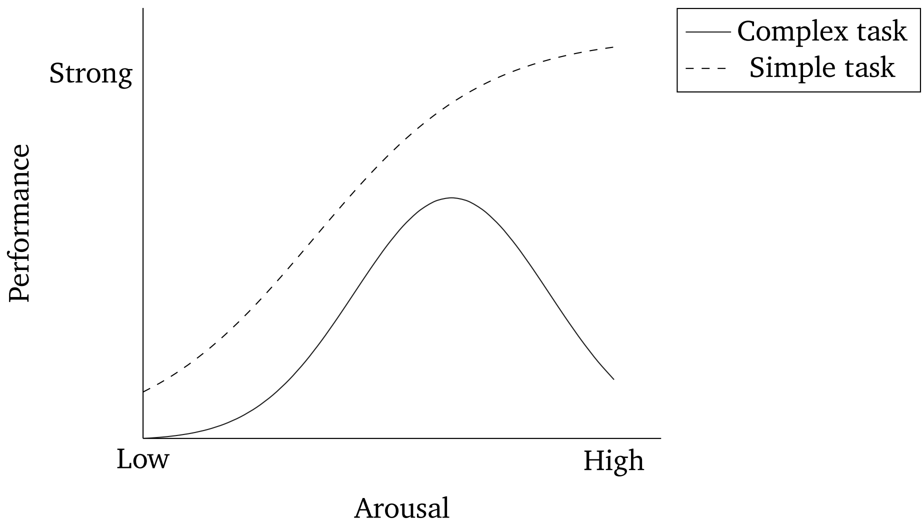


Figure 2.3.: Yerkes-Dodson [YD08] Law, showing relation between arousal and performance for complex (solid line) and simple tasks (dashed line), own representation adapted from [Dia+07].

competing for the same mental resources, it can be assumed that they have spare mental capacity left [TV06; OE86]. Secondary task methods are suitable for environments where direct performance measurements are unavailable or hard to get [TV06]. When applied in an experimental context, the user should be instructed that the primary task has higher priority. During experiment design it should also be ensured that both tasks are competing for the same resources (e.g., both are competing for the visual channel) [TV06; OE86; Gaw08].

Modified Cooper-Harper The Cooper-Harper scale was initially developed to rate aircraft handling qualities during flight tests [CH69; Gaw08]. Over time different variations of the scale were published, including a modified version, the MCH as proposed in [WC83]. Instead of asking pilots about controllability of the system, the experienced mental workload is polled [Gaw08]. Sensitivity to different workload types is reported [CW83; CW84; SRW86; Wol78]. This altered version is suitable to be used in qualitative rating of experienced mental workload.

When using this scale, the user is asked a series of up to three binary Yes/No questions. Depending on the answer of each question, the user will end up with a general statement about the experienced mental workload, which they have to further specify in a more granular three point scale [Gaw08]. The scale used during the evaluation is presented in Appendix D, section D.1. Recommendations for using the scale are provided in [WC83]: The scale should be used for overall workload assessments rather than individual subsystem evaluations. It should be submitted shortly after the task was completed, training should be provided to the participants prior to using the scale. Clear definitions of the terminology are advisable. Nonparametric analysis techniques are recommended in [Hil+92].

2.2.2. Situation awareness

Following the definition of Endsley, the term of situation awareness describes a state of knowledge that a person acting in a certain environment achieves. More specifically situation awareness is defined by Endsley as [End88; End95; End06]:

> Situation awareness is 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.

The state of knowledge is differentiated from the process of attaining situation awareness, which is referred to as *situation assessment* [End95]. Building and maintaining situation awareness requires an operator to not only perform the following steps once: “The development of SA [situation awareness] is a dynamic and ongoing process [...].” [End06]. In the context of dynamic decision making [End95], the process is described in three different levels [End88]: perception, comprehension, and projection.

Perception is the first step when acquiring situation awareness. During this step the status, attributes and dynamics of elements in the environment are perceived [End95]. Depending on the operator’s function and goals the required set of information changes [End06]. A pilot will require more information about the system status of the own aircraft, such as performance or pending warnings, whereas an ATC controller will require a different set of information to build situation awareness [End95]. With each element attributes are associated to describe the perceived elements (e.g., color, size, speed, location) [End95]. Achieving the first level of situation awareness is mainly limited by the human ability to perceive multiple items accurately in parallel, imposed by attention and working-memory capacity [End06].

Comprehension, as the second step in building situation awareness, builds upon the elements perceived in the first step. Whereas elements in the first step where perceived incoherently, their combination into a holistic picture is part of the second level [End95]. From this combined view of the situation, the significance of objects and events can be derived [End06]. In comparison to the first level of situation awareness, the experience level of the individual operator has a greater impact when integrating the various elements [End95; End06]: A novice operator is able to achieve the same situation awareness on the first level than a more experienced operator, but might fall comparably short on the second level.

During the *projection* step the previously performed steps are incorporated to predict the events in the near term future for the given system [End95]. When predicting future system states, the operator relies on his own mental model of the controlled system. During this process the mental model is *run* with previously perceived inputs to predict the future outcome [WGM00]. Through this knowledge, and based on the operator's individual goals, future actions can be planned to achieve the most favorable outcome [End95; End06]. Feedback about the outcome of a performed action is then perceived in the first level through a changed environmental state [End95].

Across all three levels the process is influenced by individual factors (personal experience, goals, or abilities) [EB94; OHa97; GT96; SJE03] and environmental factors (system design, stressors, or workload) [End06] which influence the ability to acquire situation awareness [End95]. Design recommendations to support situation awareness exist in the body of research [End95; End16; Wic08; PSW08; EJ12; End06].

2.2.3. Human behavior model

Understanding the cognitive behavior of human operators supports the creation of a human-centered design solution as pursued in this thesis. The human behavior model by Rasmussen is chosen. Rasmussen distinguishes three levels of human behavior: skill, rule, and knowledge based behavior [Ras83; Ras86]. On the lowest level the *skill-based* behavior describes behavior that typically does not underlay any conscious control. Eventually, human activities are the result of a sequence of these actions [Ras86]. On the next level, *rule-based* behavior responses to known situations are generated, based on received signs from the environment [SLB10]. Although an explicit knowledge of what rule is used to determine a sequence of actions is not necessarily present, a higher level of consciousness exists on this level [Ras86].

If the human operator is faced with a situation for which no rules or skill-based behaviors exist, the highest level of *knowledge-based* behavior is activated. In this context knowledge refers to the possession of a mental model of the system, alternatively the term *model-based* is proposed for this level [Ras86]. A depiction of this level is provided in Figure 2.4.

On this level information is interpreted as symbols, abstract information about the environment. Goals are explicitly formulated and used to derive alternative solutions to reach them. After having decided on an alternative a plan is formulated, which is then put into action through the lower levels. [Ras83; Ras86]

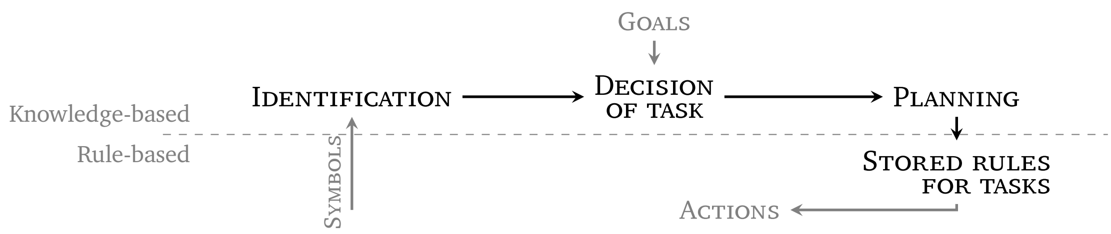


Figure 2.4.: Knowledge-based level of human behavior as depicted in [Ras83], including parts of the rule-based level and the activating information perceived as symbols.

2.3. Automated action planning

A method for automated planning of actions is used as part of the implementation described in chapter 5. The chosen method is Goal Oriented Action Planning (GOAP) as described by [Ork03]. GOAP was originally designed for controlling agent behavior in video games [Ork06]. Through their high complexity and present dynamics video games offer a different environment, than the ones for which planning techniques are developed in an academic context [Neu+19]. GOAP is based on the Stanford Research Institute Problem Solver (STRIPS) described by [FN71], which is referred to as the first major planning system [Rus+16]. Extensions to the formal language of STRIPS exist, for example the Action Description Language (ADL) in which constraints present in STRIPS are relaxed [Ped89; Rus+16]. In the following a description of the underlying principles of GOAP is given based on the nomenclature of STRIPS in [FN71].

STRIPS distinguishes between *world states* and *operators* [FN71; Neu+19]: The world state is described through a set of propositional variables (P). Operators (O) represent (parameterized [FN71]) action routines (a) which can be performed to change variables in the world state. They contain information about *preconditions* (P_{pc}) and *effects* (P_e) as shown in Equation 2.1:

$$O = [P_{pc}, P_e] \tag{2.1}$$

Both, preconditions and effects, are formulated as propositional variables. The action associated with an operator can be executed when the current world state ($S_0 = [P_0]$) meets the preconditions. The effects of the application cause a transition to another world state (S_1). In STRIPS the world state is maintained as a list of *facts* which are added or removed through operators [FN71]. In GOAP a fixed set of modifiable state variables is maintained [Ork05; Neu+19]. Equation 2.2 depicts this relation:

$$\begin{aligned} \text{if} \quad & P_{pc} \subseteq S_0 \\ \text{then} \quad & S_0 \times O \rightarrow S_1 \quad \text{with} P_e \subseteq S_1 \end{aligned} \tag{2.2}$$

For the planning process an initial world state (I) and a goal world state (G) are defined. A search algorithm is used to find a sequence of actions ($[a_0, \ldots, a_n]$) which need to be performed to successively transfer the initial state I into the goal state G [FN71]. GOAP uses the A* search algorithm (A*) planning algorithm [HNR68] under consideration of action cost [Ork03]. During execution of the generated *plan* (the sequence of actions) [FHN72] a constant monitoring of the environment allows reacting to changes. This functionality is part of GOAP [Ork05], whereas STRIPS gained this functionality through the STRIPS execution monitor “PLANEX” [FHN72; FN94].

2.4. Information management in aviation

In a computerized, information-driven economical world, the management of information itself becomes increasingly important [Nor14]. The growing amount of information on the flight deck and inherently connected challenges towards information management for the human operator are recognized [JR95; SMC10; JGW10; You+16a] and will be discussed in the remainder of this chapter. Challenges related to information management are not exclusive to aviation. A broader view is provided by [Nor14]:

> We are in the midst of what some people call “the information explosion”, but there is too much information for anyone to assimilate, the information is of doubtful quality, and perhaps most important, the things we collect statistics about are primarily those things that are easiest to identify and count or measure - which may have little or no connection with those factors of greatest importance.

In the following section an introduction to information management in aviation is given. A definition of information management as used in this thesis and the related terminology based on the Data-Information-Knowledge-Wisdom (DIKW) hierarchy are provided. Challenges related to information management in aviation are summarized, examples of work in this area are identified. A classification of this thesis in relation to existing work is performed in subsection 2.4.5.

2.4.1. Definition of information management

With the term *information management* being a "[…] broad conceptual term […]" [Det10], the usage in this thesis will mainly be limited to information management in the sense of human-machine interaction. Information management on a computer-system integration level (e.g., the Boeing B-777 Airplane Information Management System (AIMS) based on the ARINC-629 standard [Aer94; Mor14], or advanced Electronic Flight Bag (EFB) systems falling under the definitions of [Int18d; Fed17]) is explicitly not part of this thesis. Organizations tend to rely on their own definitions of what information management is meaning for their business. Exemplary, the FAA defines information management in the FAA Order 1375.1E as [Fed11]:

> The leading, planning, organizing, structuring, describing, and controlling of the collection of information (developed from one or more data sources) and monitoring that information throughout its life-cycle; including the distribution of information to one or more audiences, and reviewing users needs to incorporate future best practices.

Information are described as enterprise assets, which need to be adequately managed to ensure accuracy and quality [Fed11]. A more generally applicable definition of information management and its objective is given by [Det10]:

> […] the management of the processes and systems that create, acquire, organize, store, distribute, and use information. […] to help people and organizations access, process and use information efficiently and effectively.

Based on the observation of flight deck activities [Han16] proposes a differentiation of information management into: access, processing, transmission, storage, and disposal. From this description the depiction shown in Figure 2.5 is derived. Highlighted through the gray box are the steps considered most relevant to this thesis. A relation to the way pilots manage cockpit tasks and Crew Task Management (CTM) [Fun91] is pointed out by [Han16]. The information management process starts by *accessing* an information source, inputting information. After that, the retrieved information is *processed*, e.g. as part of a decision making process. If required a *transmission* of information takes place, during which the agent might elect to share the processed information with other agents (humans or machines equally considered) in the system. Information can be *stored* for later use if potential future relevancy for the just processed information remains. If no further relevancy for the information exist (e.g., the frame of relevance has passed), the step of *disposal* concludes the information management process. Highlighted in Figure 2.5 is the focus of this work as described in section 1.2, which consists of the steps of access, process, and transmission.

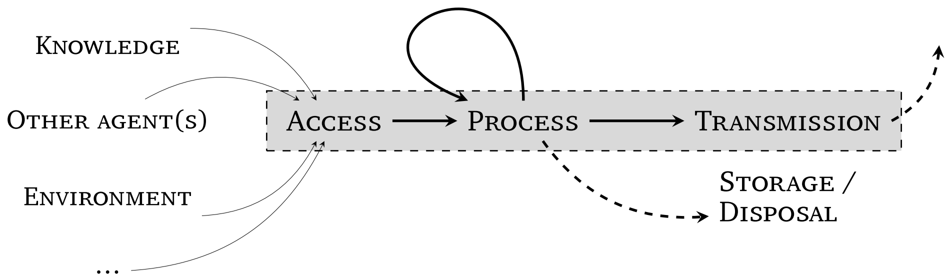


Figure 2.5.: Graphical depiction of the information management process as described by [Han16]. Shown is the process for one agent from accessing information until eventually transmitting, storing, or disposing it. Depiction by author, based on [Han16].

Examples of information management applications in civil aviation On a global scale, information management is recognized as *essential* by the ICAO, which devoted the Information Management Panel (IMP) to support the development of a global information management framework for aviation [Int15]. As part of the European research program Single European Sky ATM Research Programme (SESAR), System Wide Information Management (SWIM) is developed, allowing increased cooperation between airspace users [Sin15a].

Just as information management exists on a macroscopic scale, information management systems are a part of the flight deck itself. The Flight Management System (FMS) system for example is a key-element regarding flight deck information management, which is utilized to manage the aircraft's flight path during a mission [Abb14]. In this critical role, the FMS is also a critically viewed component [BM04], regarding the interaction between human operator and automated systems on the flight deck, to which the remainder of this chapter is attributed. The Engine Indication and Crew Alerting System (EICAS)[3], is a primary example for the integration of different information sources for the crew. Displayed on this system are engine and system parameters with supporting synoptic displays to enhance situation awareness and the understanding of system parameters for the crew [JGW10]. Electronic checklists help guiding the crew through required steps [Boo01]. A more recent addition to the flight deck environment are EFBs [Int18d; Fed17],

[3]EICAS is the terminology utilized by Boeing, whereas Airbus uses the term Electronic Centralized Aircraft Monitoring (ECAM) to describe a comparable system.

on which software supporting operations can be executed. Either designed as part of aircraft avionics (and hence requiring certification) or as portable tools (such as laptops or tablets), they are able to provide a range of functions available to the flight crew [Han16; Beg12].

2.4.2. Data-Information-Knowledge-Wisdom hierarchy & terminology

Data, information, knowledge, and wisdom – terms being used in conjunction with information management require precise distinction or differentiation. To facilitate a common terminology and understanding of these terms throughout this work a definition is provided. A commonly used representation, and description, of the relation between those four terms is the so called DIKW hierarchy, a central model of information management [Row07].

The origins of the DIKW hierarchy are attributed to the work of [Ack89], [Adl86] and [Zel87] [Fri08]. Since then there is an ongoing discussion in the realm of Information Sciences about the applicability of the DIKW hierarchy and the underlying definitions and assumptions [Fri08; Ma12]. With the exception of the more philosophical level *wisdom* [Row07; Jas04] the different layers and their relations are described in the following.

Data, the lowest level in the DIKW hierarchy, describes properties or events as discrete objective facts with symbols, often in an unorganized, quantified format with low value per se [Row07; PS13]. A distinct property of data is that it has to be true to be considered as data [Fri08]. More generally, the term data comprises every value observable by a sensor, independent of whether it is a human sensory channel or an instrument [Ack89]. An example for this category would be the current engine fuel flow rate [Har13].

Information is the second layer in the DIKW hierarchy. Contrary to just being an observation, information is able to answer questions beginning with *who*, *what*, *when* and *how many*. In order to do so, information is either generated from data by organizing and transforming it into an appropriate form, or by inferring it from the underlying data set [Ack89]. Information is a subset of the refined underlying data layer [Fri08], making raw data valuable, useful, and relevant for a specific purpose or context [Row07]. The value is defined by the context and purpose it is received and used in [PS13]. The term “information” itself is not self-evident [Ma12]. Referring to the previous example, information would be the remaining amount of fuel or the expected endurance, inferred from available aircraft sensory data [Har13].

Knowledge is built on information, also referred to as *actionable-information*, allowing more effective acting when compared to data or information [Jas04]. Knowledge can be described as *know-how* [Ack89]. Knowledge involves factors outside the previously introduced layers of the DIKW hierarchy, such as experience and aggregated contextual information [PS13]. Knowledge can be distinguished into tacit and explicit knowledge [Row07]: Tacit knowledge is described as *know-how* (know how to do a landing flare), it resides within the individual itself and is complex to record and share. Explicit knowledge is described by *know-that* expressions (know that a wing is used to generate lift), and is the kind of knowledge usually found in books [Jas04].

2.4.3. Information management challenges on the flight deck

In academia challenges in relation to information management on the flight deck are well known, and just as well documented and researched. Mainly in the wake of introducing advanced FMSs to the commercial aviation domain, challenges regarding the human-machine interaction and the prevailing information management in the cockpit became apparent and sparked interest in the scientific community. For example, studies like the ones conducted by Sarter and Woods ([SW92; SW95; SW97; SWB97]) pointed out shortcomings of the introduced systems. A general debate about challenges regarding human-machine interaction in highly automated environments started.

Under the wide and open interpretation of information management many studies contributed directly, or indirectly, to this body of research. Additional traction was gained when a study was initiated by the FAA in 1994 as a reaction to the tail first landing of China Airlines 140 at Nagoya, Japan[4]. The study was focused on the interfaces between flight crews and aircraft systems, and was released in 1996. In the final report the discovered issues were summarized and recommendations given [ASS96]. In 2013 a follow-up report by the FAA was released, in which the findings and recommendations given in the report of 1996 are revisited with an emphasis on flight path management systems [AMR13].

In the following paragraphs different aspects of information management challenges and related research efforts are summarized. They will be used in chapter 4 to derive and drive the requirements for the developed information management concept.

Role of the human operator A contributing factor to the challenges in information management is the role of the human operator on today's flight decks. Throughout the last century of aviation a continuous development of the flight deck lead to a decrease of crew size to two pilot [Lac+14a; Boy14]. Human functions were allocated to aircraft systems, resulting in an increase of overall system complexity [You+16a]. At the same time role and responsibilities, described through the Aviate Navigate Communicate Manage Systems (ANCS) schema [JR95; Abb93; ST96], of the crew changed [Bil91; Boy98].

[4] In 1994 a China Airlines flight crashed in the proximity of Nagoya Airport during approach, before reaching the runway. Accident report: [Air96]

Previously aviators (A) they primarily became managers of systems (S) [Har07; BBR07; Boy14; Sat93; Sch15]. In this role they "fly" the aircraft 3.5 min to 7 min per flight [CSC16]. For the remainder of the flight they monitor the automated systems, a task less suited for humans [Lov95; Dek04; Har07; CS15]. This is a result of the pursued allocation strategies [Sch+07] eventually dictating a change of the operators role through technology [Abb14].

Automation and system design The introduction of advanced automation, which lead to the previously described role change of the human operator, was accompanied by "unexpected" [SWB97] effects. One of these effects is that the introduced automation did not only reduce workload for the human operator [Wie88; FSW10]. Instead, previously physical workload became cognitive, or mental, workload [BBR07; SW92]. Titled *clumsy automation* [Wie88] is the circumstance in which workload is lowered during low workload phases and further increased during high workload phases through an additional overhead caused by monitoring and instruction systems [FSW10; SW91]. A tendency to deactivate automation in high workload situations has been observed [CW02; Bur+13], in contrast to that cases in which *complacency* led to an overreliance on automation exist [Bil91]. Another factor contributing to information management challenges is the growing number of information sources and amount [You+16a; UJ02]. This leads to a situation where the required information is technically available but the human operator is required to invest additional cognitive effort to identify and use it [Sch+07]. Safety critical information can be overlooked [Sch+12; Eth+19].

Teaming between human operator and automated systems Another aspect of information management challenges is the teaming of the human operator and automated systems. A successful cooperation requires a shared set of system and mission state information [SW97; JR95; UJ02]. A mismatch in this area can lead to *mode errors*, false assumptions about the current mode of the automated system [SMW07]. In this situation actions performed by the operator can lead to unexpected results or *automation surprises* [SWB97]. The operator is surprised by actions performed by automated systems [SWB97; SW92]. The previously outlined accident in Nagoya is attributed to this phenomenon [BBR07]. More recent examples of this are still being reported by pilots [CSC16]. Consequently, the proper *teaming between human operators and automated systems* is of special interest, which is summarized under the term Human-Autonomy Teaming (HAT) [Bra+18b; CW02; UJ02; Eth+16]. The growing need for proper HAT is recognized, especially with a potential introduction of Reduced Crew Operation (RCO) on the horizon. It is stated by Bailey et al. that "[...] present-day automation design paradigm is not sufficient for

SPO/RCO. [...] Autonomy must effectively team with the human to keep them in the loop and situation aware." [Bai+17]. Robust *autonomous* systems are proposed that allow delegation of tasks from the pilot but require an effective teaming with the human operator [Bai+17].

Task management Connected to this is the area of *task management* on the flight deck, which eventually is connected to information management [Han16]. In this domain, research is concentrated on how flight deck crews manage different and sometimes concurring tasks [LDB09]. Preceding work was performed by [Fun91] where the term CTM was coined. A taxonomy of task management errors is presented in [Fun91; Fun97; ST96; CMF96]. Further preceded is this work by the work on information categorization performed by [JR95] and the definition of flight deck functional categories by [Abb93].

2.4.4. Related work

An overview of existing information management challenges is provided in subsection 2.4.3. Based on this foundational work a variety of approaches to overcome or mitigate the impact of the identified challenges emerged. As stated in section 1.2, this thesis seeks to contribute to this body of research. A cross-section of existing related work field is summarized in the following.

A common approach to the reported challenges are technological solutions to support the pilot(s) during mission execution. A recent example for this category is the Automation and Information Management Experiment (AIME) study is performed by National Aeronautics and Space Administration (NASA) in response to a set of recommendations issued by the Commercial Aviation Safety Team (CAST) [Joi14]. In simulator trials with 11 airline crews, in a DPO setting, flight deck systems for aircraft state display and prediction are evaluated [Eva+16; You+16a; You+16b]. Evaluated are maneuver envelope displays [Lom+15], trajectory prediction [Shi+17], energy state alerting, and system interaction synoptics [UD15; You+16a]. Aside from that other examples of supporting technological can be identified: A new display concept for monitoring engine parameters in response to Air Transat 236 [Avi04] is proposed in [DBS13]. [Dao+15] describes a diversion airport recommendation system. A mobile device-based information management support system, covering the pilot's information needs before arriving at the airport until preflight, is described and evaluated in [Han16]. Recommendations for an improved Mode Control Panel (MCP) are suggested in [BM04]. These examples highlight how new technology is proposed for the flight deck to overcome isolated issues. No changes to the underlying concept of operations and little changes to the flight deck are made, allowing for a lower entry barrier (cf. subsection 2.1.1). A recent example of such a system in service is described in [AS18; AWG20]. Differentiated to this thesis are those solutions by being more near-term focused. In contrast to what is described in section 1.2, those solutions fall into the category of incremental or evolutionary changes.

A second category are assistant systems with varying levels of authority. In contrast to the previous category, those solutions are geared towards a broader support of the pilot(s) rather than providing isolated support. Similar ot the previous category a wide range of examples is identified (cf. [OS10; Ban+08]) across all aviation domains. Examples from the military and commercial domain are summarized in the following.

As part of the research on CTM [Fun97] an agent-based program titled Agenda Manager (AMgr) was developed. The program was designed to monitor pilots in respect to their stated goals through task management support. Evaluations in a part-task simulator showed potential to reduced undetected flight crew errors [Fun+97]. The Cockpit ASsistant SYstem (CASSY) [Onk96] is an example for an advisory system in the commercial aviation domain. The system acts exclusively in an advisory mode as a “third crew member” [OS10]. It maintains a set of alternative flight plans and presents them on request or conflict detection, goals are not explicitly communicated to the system but are derived from regulations and the definition of mission success [OS10]. CASSY was evaluated in simulator and flight trials with the Deutsches Zentrum für Luft- und Raumfahrt (DLR) [OP94; Ger+95]. The continuation of CASSY is Crew Assistant Military Aircraft (CAMA) for military air transport usage scenarios [OS10]. CAMA extends CASSY with military-domain specific capabilities such as terrain following. Simulator and flight trials with the German Air Force were performed [SS99; OW01]. Additional examples of cockpit assistant systems are found in the military aviation domain: The Pilot’s Associate (PA) program [BL91] is one of the earlier assistant systems [OS10]. PA is able to execute plans when authorized by the pilot [BL91]. Goals are not explicitly communicated to the system which turned out to be a weakness [OS10]. Rotorcraft Pilot’s Associate (RPA) is based on the earlier PA program [OS10] and is designed as a cockpit assistant for Apache helicopters [MH99]. RPA manages information presentation in the cockpit and allowed for dynamic task allocation between the pilots and automation [MH99]. Concepts of cockpit-based assistant system for guiding multiple Unmanned Aerial Vehicles (UAVs) in a military context [Sch13] are evaluated in [BS18; Bra+18a]. So called Artificial Cognitive Units (ACUs) [OS10; Sch13] are used to allow task-based guidance of UAVs [US11; Cla+12; CS14]. Additional examples from the military domain are found in [BTF00] and [UJ02].

Examples of research projects in the commercial aviation domain are observed over the last two decades. PILoten-ASsistenz-System für Luftfahrzeuge (PILAS) is a cockpit assistant system designed for helicopters, developed to allow the execution of advanced trajectories with a focus on future ATM operations [Deu08]. The PILAS planner provides a two pilot crew with an updated flight plan if a conflict situation is detected [Deu08].

As part of the EU funded research initiative Advanced Cockpit for Reduction of StreSs and workload (ACROSS) [Eur13b] concepts for RCO were investigated. The investigated concepts range from unintentional RCO through the incapacitation of one/all crew member(s) to intentional RCO operations [LM14]. A suite of support systems for the pilot(s) are identified as supporting means during mission conduction. The identified solutions are described as the *six ACROSS pillars*, a publicly available summary of the identified solutions is provided in [MdL18]. An emergency landing system called Electronic Standby Pilot (ESP) is designed to take over control of the aircraft in case of a full crew incapacitation, supported by a ground element [Gra+14b]. The ANCS schema is explicitly referenced in the six pillars, no essential changes to the human operator's role are assumed [MdL18].

The majority of work on assistance systems has been performed in the military domain [Ban+08]. Fewer systems were developed for the operationally different commercial aviation domain, to which this thesis seeks to contribute. Compared to the first category of targeted support systems, the assistant systems introduce larger changes to the flight deck environment. Changes to the role of the human operator are incidental, arguably introduced by technology comparable to what is described in section 1.1.

Eventually, related work focusing on the role of the human operator is summarized. In line with the objective and assumptions stated in section 1.2 work attributed to the domain of RCO is emphasized. An overview of existing concepts for RCO for commercial aviation is provided in [NKS18; Nei20] and [SS20]. A categorization of the concepts is performed by [NKS18], where two main categories are described: *replacement* design approach, and *displacement* design approach. The replacement approach focuses on replacing the second human operator with advanced automation either onboard the aircraft [Har07; Com+13; DP05; Gra+14a; MG16] or with a combination of onboard and ground-based automation solutions [SHS16; SK17; SHS14]. The displacement approach seeks to move the second operator from the airborne segment of operations to the ground where they can support one or multiple flights at a time. The current concept of operations for RCO as proposed by NASA [Mat+17] falls into this category. Based on how the division of labor is performed, a number of subcategories within the displacement approach are identified in [NKS18]: [SK17; SHS16; Bra+15; WG15; Lig+15; BJS14; Lac+14a].

The majority of the existing work is at a conceptual or modeling stage (cf. [SS20]). A limited number of practical implementations of concepts for evaluation purposes exist. Empirical evaluation for the NASA concept of operations focussing on the ground segment exist [Bra+15]. Aside from introducing an additional display, the flight deck and the assumed role of the human operator remains largely unchanged [Bat+18; Mat+18]. Concepts exploring the possibility of fundamental changes to the role of the human operator (e.g., [Sch+07]) have not been implemented nor tested in a respective environment.

2.4.5. Research gap

Based on the observations made during the review of related work a differentiation of this thesis to the existing body of research is performed. The first contributing differentiator is the approach to the human-centric development process. Comparable work often starts by defining system requirements based on the *current* role of the human operator. Resulting changes are typically of incremental or evolutionary nature, a comprehensible approach given the high certification cost and entry barriers for extensive changes [BBR07; ASS96]. Changes to the role of the human operator are incidental, potentially again through the introduction of new technology to the existing flight deck. In contrast to that a step back is taken in this thesis as outlined in section 1.2: A definition of the human operator's role is performed first. This definition is based on the understanding of the future context of commercial aviation operations, shaped through RCO and changes to the future GATS. Arguably, the chosen approach alone does not sufficiently differentiate this thesis from existing work, for example [Sch+07].

The second contributing factor is the implementation depth of the developed concept. As outlined in subsection 2.4.4, the majority of work beyond evolutionary change is found to be in concept or model stages, existing on paper. The effort described in here intends to push beyond that by completing a first development cycle (cf. section 1.3): With the definition of the context and the role of the operator a normal development process is followed. Requirements are being derived from the existing body of research, an information management concept is developed. Realization of the concept is performed in hard- and software through a purpose built RCO simulator and the accompanying software implementation of an information management system. In this environment a first evaluation of the developed information management concept is performed. Concepts with a comparable, or deeper, implementation and evaluation depth are identified in subsection 2.4.4 An example is [Lac+14a; Bat+18; Mat+18] where evaluations with the ground station are performed, or [Onk96] where flight tests took place. However, those examples fall short regarding the first contributing differentiator: The role definition of the human operator remains unchanged, the physical operating environment experiences minor changes. Highlighted by those examples is the main differentiator of this thesis: the combination of the chosen approach of the human-centric development process and the implementation depth. Through the completion of a first development and evaluation cycle a foundation for future research to build upon is formed.

3. Analysis of current and future context

An analysis of the expected future operational context based on the current context is performed. This step represents the first step of the DIN EN ISO 9241-210 standard [DIN11] used in this work. The context in which today's pilots are operating is initially reviewed. Next, the expected changes to the larger Global Air Transportation System (GATS) are summarized. Known, or at least proposed, influences through research programmes such as Single European Sky ATM Research Programme (SESAR) are taken into consideration. The concept of Reduced Crew Operation (RCO) is also discussed. After that, the future operating context is further defined. A new concept of operations for the human operator on the flight deck is described, the term mission manager is introduced.

3.1. Today's operational context

Although the herein developed concept for information management is targeted at a future scenario, the understanding of today's operational context is required. Understanding the current operational context allows for a better extrapolation of the relevant future context. Consequently, in preparation of this work different efforts were made: A task analysis of a commercial aviation flight deck crew under normal operating conditions was performed. Following that, a survey-based analysis of pilot personas, following the persona concept of Cooper et al. [Coo+14], was conducted. Lastly, high workload situations under normal operating conditions were identified.

3.1.1. Pilot's task analysis

The task analysis was performed using the Hierarchical Task Analysis (HTA) method [AD67; Ann03; Sta06]. Under the top level goal of "conduct safe and efficient flight" pilot tasks were broken down. The analysis was performed for nominal Dual Pilot Operations (DPO), assuming a normal trans-oceanic flight. Flight Crew Operating Manuals (FCOMs) [Air13; Boe14], video recordings, existing task analyses [Lac+12; WG15; CML14; ST96], and subject matter expert interviews were used to create the task analysis. The resulting task analysis is largely airframe or manufacturer agnostic and applicable for a wide range of commercial aviation aircraft. Identified tasks and processes performed by pilots during a flight mission were translated into the Business Process Model and Notation (BPMN) 2.0 standard [Obj11]. The results are maintained in a Microsoft Access database which is connected to a Microsoft Visio visualization using the BPMN standard.

3.1.2. Pilot's persona identification

While the task analysis provides information about *what* the pilots are doing during today's operations, it provides no information about *who* the pilots are. To gain a better understanding of the users themselves, a survey among pilots was performed by [Wer17] under the supervision of the author. In an online survey consisting of 82 close-ended, 13 open-ended, and 9 free-text questions 43 pilots were interviewed, yielding 33 complete data sets. The participants were asked about their attitude towards modern technology, interface design, and the pilot's role on the flight deck. Demographics and personal desires as well as goals were assessed in order to identify user groups and create personas [Coo+14]. Two personas were identified: a more experienced, older, pilot and a younger, less experienced, pilot [Wer17]. In contrast to the older pilot the younger pilot is found to be more technological affine and shows a higher openness towards advanced automation [Wer17]. For the design in this thesis the younger pilot was more strongly considered, as they are believed to be the target audience for proposed time frame (cf. section 1.2).

3.1.3. Identification of high workload situations

The task analysis provides a concise overview of the tasks a flight deck crew is *supposed* to do over the course of a idealized mission. What it fails to represent are challenging situations, in the sense of high cognitive workload, that are likely to occur during normal operations. Following the reasoning given in section 1.1, the effects of these situations on the flight crew are likely to aggravate under RCO. A general understanding of such situations is required.

With support of student teams under the supervision of the author [Has+17; Boe+19] representative situations were identified. Situations were identified from existing literature and academia, mainly based on the literature summarized in subsection 2.4.3. Accident and incident databases were searched for situations in which information management or high cockpit workload was identified as a contributing factor. The Aviation Safety Reporting System (ASRS)[1] provided by National Aeronautics and Space Administration

[1] https://asrs.arc.nasa.gov/ - last accessed on Sunday February 10, 2019

(NASA) and the National Transportation Safety Board (NTSB) accident database[2] are utilized. Subject matter experts were interviewed by the author. In semi-structured interviews a total of four pilots (three active, one retired) were asked about situations they experienced as stressful and what actions were taken during these situations. The interview guideline utilized by the author during the interviews is attached in section A.1. Non structured interviews were performed with one additional pilot and an industry aviation safety expert. The situations identified during the research phase were documented and pooled into use cases. During the creation of the use cases attention was paid that they remain as *universal* as possible and are not constrained to specific geographic or temporal constraints. Use cases should not represent rare or abnormal conditions but instead represent a cross section of potential everyday high stress situations. They should be *challenging*. Lastly, the use cases should be *relevant* for future commercial aviation operations, given the underlying assumptions of this work. Only *pilot-centered* use cases are considered, meaning that the effects of a situation can be controlled, or mitigated, by the pilots on the flight deck. A rating of the expected workload during each use case was performed in [Boe+19], the use cases were ranked by expected workload. The rating was performed using the cohesive model of workload [Mes88]. The identified use cases are summarized in section A.2. The description of the influencing factors for the rating of use cases and the results are provided in section A.3. The identified use cases were used to broaden the understanding of the context. Learnings and recommendations from literature associated with the identified use cases are reflected in subsection 4.1.2. The evaluation scenario described in section 6.2 incorporates elements of the identified use cases.

3.2. Future Global Air Transportation System

An underlying assumption of this thesis, as mentioned in section 1.2, is a changed operational environment. The transition from todays GATS, as described in section 2.1, into the future is described. A brief introduction to the Air Traffic Management (ATM) research initiatives of SESAR (Europe) and Next Generation Air Transportation System (NextGen) (United States (US)) is given. The concept of RCO is outlined.

[2]https://ntsb.gov/_layouts/ntsb.aviation/index.aspx - last accessed on Sunday February 10, 2019

3.2.1. Future changes to Air Traffic Management[3]

In the past decades a number of research initiatives in the ATM domain were initiated. The premier motivation for setting up these (inter)national research initiatives is to increase efficiency of the air transportation system, and preparing it for future developments. One of the bigger challenges the world wide air transportation is about to face is a growing traffic demand: A study initiated in 2010 by European Organization for the Safety of Air Navigation (EUROCONTROL) predicted a growth of air traffic above Europe of 80 % from 9.4 million to 16.9 million flights in 2030 [Eur10]. As a result of high volatility of air traffic demand since 2008, economic regression, and a lower than expected expansion rate of Middle East hubs, the following report from 2013 indicated lower than initially projected numbers [Sin15a]. The projected number of air traffic across Europe is now at 14.4 million flights in 2035 according to the most likely scenario [Eur13a]. Either way the growing traffic demands will impact airspace and airport capacity. This in turn will increase the likelihood of delayed or non-accommodated flights [Eur13a]. The U.S. Department of Transportation (DOT) estimates that air-traffic delay cost for the economy accumulates to $20 billion each year in the US [US 17].

To counteract the negative effects induced by increasing air-traffic, new technologies and management principles are required. Research of these technologies and development of appropriate procedures is driven by different initiatives. The two most prevalent research initiatives are the European SESAR [Sin15a] initiative and NextGen [Fed16b; Joi11; US 17] in the US. Connected to an increase of operational efficiency, both research initiatives strive to maintain and increase the current safety level. Furthermore, the environmental impact of aviation is sought to be reduced [Fed16b; Sin15a]. Key changes proposed by those research initiatives reach across all sectors of ATM:

- Under SESAR a concept named System Wide Information Management (SWIM) is being proposed [CR13; Eur17], allowing the sharing of operationally-relevant information between all ATM participants.
- In the US, as part of NextGen, a higher connectivity through an increased utilization of Controller-Pilot Data Link Communications (CPDLC) and Automatic Dependent Surveillance - Broadcast (ADS-B) is intended [Fed16b].
- Operationally, a higher predictability of aircraft trajectories is desired. As a consequence both SESAR and NextGen consider the concept of Trajectory Based Operation (TBO) a key feature [Sin15a; Joi11].
- More responsibility is delegated to the aircraft cockpit, for example through self-separation of traffic [Joi11].

[3] The background information given here reflects the situation prior to the global COVID-19 pandemic and its impact on the aviation market. The long-term effects remain to be assessed at the time of writing.

- An increased use of automated systems and autonomy to support ground-based and airborne ATM systems is envisioned [Joi11; Sin15b].

3.2.2. Reduced Crew Operation

In section 1.2 RCO is introduced as an assumption for this thesis. In section 1.1 a rationale is provided why the introduction of RCO is a pivotal point for larger changes to the flight deck environment and the role of the human operator. In this thesis the term RCO is used to described the operation of a single operator on the flight deck of a commercial airliner. This should not be confused with the total number of pilots onboard the aircraft for example during a long-haul flight (cf. subsection 2.1.2). The potential presence of more than one pilot onboard the aircraft, for example in crew rest, is not ruled out.

Background

From an airline perspective the motivation for RCO is economically driven: Labor cost is one of the biggest operational cost drivers [SHS16; SUT16], a reduction of cockpit personnel will have an impact on the overall cost structure of an airline as outlined in subsection 2.1.1. In 2017 the Swiss investment bank UBS estimated savings for commercial airlines to be about $15 billion over the next two decades, assuming a reduction of cockpit crew size to a single person [CF17]. This is further amplified by the currently developing situation on the job market, regarding the availability of qualified cockpit personnel. As predicted by the Boeing Company in 2016 a total of 617 000 new pilots will be required between 2016-2035, partially due to an increased demand of emerging markets in the Asia Pacific region [Boe16a]. A pilot shortage is expected. Already today the consequences of such a shortage are already visible and are becoming an emerging threat, especially for smaller carriers [Ost17].

Challenges

The introduction of RCO in commercial aviation is faced by a number of challenges [Com+13; ADN14]. Maintaining a comparable level of safety as found in commercial aviation today is challenging [DP05; Bai+17]. Effectively, a single point of failure is introduced through the remaining single pilot [Sch+07]. Issues when removing a pilot from today's flight deck are described in [Eth+16] and corresponding publications. Especially the questions of how to handle pilot incapacitation, the impact on workload, and newly found automation errors, for the remaining pilot are raised [SHS16; Sch15; Har07; Dek02]. The development of the required technology is an additional challenge [SHS16].

An example of this is the development of required communication technology: Existing concepts, as well as the concept described in section 3.3, often require high connectivity (cf. [NKS18; SS20]). In contrast to that, an increase of available bandwidth for aviation purposes is questioned in [ADN14]. The certification of secure communication methods [Lac+17] or cost-effectivity [Dri+17] are discussed in this context. Other technological challenges are related to the development and certification of advanced automation for the flight deck: While technical feasibility of such systems has been demonstrated, see subsection 2.4.4, additional development is likely required [ADN14]. The certification of nondeterministic and/or adaptive systems poses an additional barrier. Current software certification standards for aviation (e.g., DO178B/C [Rad12]) do not support the verification and validation of such systems [ADN14]. Challenges in the design of human-machine interaction, as summarized in subsection 2.4.3, will have to be addressed [SHS16; MG16]. Reliable, non-encumbering, technology for surveiling the human operator's physiological and cognitive health is required [Bai+17; Com+13]. Related data privacy issues must be resolved [Bai+17].

Concerns regarding the social isolation faced by pilots under RCO are discussed in [Com+13]. Potential performance decrements through boredom, lack of social cues, and missing peer pressure are described [Lac+14b]. Social concerns on a macroscopic scale affect the perception of RCO by the flying public [ADN14], insurance companies [Sch+07], and pilot unions [MG16; Car18]. Building up acceptance and trust with all stakeholders (pilots, public, official authorities) is a key challenge towards the introduction of RCO [Lac+17; SS20].

Current efforts towards RCO

As of today, a number of concepts and theoretical considerations towards the implementation of RCO in a commercial aviation context are published. An overview of concepts is provided in subsection 2.4.4. In contrast to military and general aviation, where RCO is part of the routine operations [Har07; CF17; UJ02; SHS14; Lac+14a], and the exception of some applications in the business aviation domain [Emp10], RCO is not being performed in the context of commercial aviation. The emerge of RCO concept in the aviation industry itself becomes apparent through announcements made by Airbus [Tré19; Kin19], Boeing [Bat18], and the China-Russia Commercial Aircraft International Company [Eis19].

3.3. Future concept of operations for the human operator on the flight deck

From the larger future context described in the previous section the focus is now shifted towards the environment for which an information management concept is designed, the flight deck. Here the role of the human operator under RCO is reviewed more closely. As it was previously layed out in section 1.1 there is a need to redefine the human operator's role on the flight deck. As part of a human-centered approach this is considered to be the first step towards a new information management solution. Based on the herein described concept of operations for the human operator a suitable solution can be developed.

The already existing concept of operations for future operations, as summarized in subsection 2.4.4, are considered non-satisfactory for the approach of the work at hand. It was decided to create a new role definition for the human operator to be used and eventually carried forward beyond just the *concept* stage. The herein described concept of operations was first published by the author in [SNS18].

3.3.1. Objectives

A set of three high-level objectives are defined which the future concept of operations should satisfy. In descending order of importance the objectives are safety, economic efficiency and improved utilization of the human operator as a resource for flight operations.

Safety Safety is and needs to remain the highest priority in the aviation domain. As of today air travel is one of the safest travel modalities with an average of one fatal accident per 6.7 million flights [Int18a]. Following the guidelines of the Federal Aviation Administration (FAA), the chance for a fatal flight outcome must be as low as 1×10^{-9} (or less) per flight hour [Fed88]. Not meeting the already established standards and expectations towards aviation safety of today is not an option for a future concept of operations and will render it unacceptable from a political and social point of view. The primary goal *must* be to provide a higher, or at least the same, level of safety as in today's operations. Yet it also has to be noted that a full and comprehensive evaluation of the achieved level of safety is not possible within the scope of this thesis. Decisions potentially impacting safety are carefully weighted instead. Later, an alignment with published safety standards such as the International Civil Aviation Organization (ICAO) Safety Management Manual (SMM) [Int13b] should be performed.

Economic efficiency Economic efficiency is a key-enabler for a change in airline flight deck operations. The transition from two cockpit crew members to RCO will incur a substantial amount of cost, associated with the expensive process of flight deck certification [BBR07; Pal+95; SG16]. A distribution of this entry cost between airlines, airframe manufacturers, and suppliers is likely to occur. As many of these companies are operating benefit oriented, a return on their investment within a reasonable amount of time is naturally expected. Without the provision of (significant) cost benefits associated with the herein discussed concept an entry into service is highly unlikely as there is no economic incentive. Similar to the objective of safety it is therefore imperative to provide an economic viable and efficient solution.

Improved utilization of the human operator Closely related to the previous objective is the improved utilization of the human operator as a resource, especially when considering that crew cost is one of the biggest cost drivers for airlines as described in section 2.1. Derived from that is the goal to create a better match between (unique) human capabilities and the job profile under which they will be operating. As opposed to further following the path of designing technological systems and forcing the pilot to adapt to them, the goal is to design technology in a way that best caters human capabilities and compensates human weaknesses. This objective is also reflected in the overarching objective of this thesis as stated in section 1.2. By first defining the operational context of the operator before designing any technical systems, the first step towards achieving this goal and putting the human into the center of the development process is already taken.

3.3.2. Assumptions and limitations

Assumptions and limitations are formulated to delimit the following role definition for the human operator and the subsequent development in chapter 4. The assumptions are made in response to the challenges summarized in subsubsection 3.2.2.

Operational In general, the operational environment in which the described concept of operations will be applied does not significantly differ from today's operations. One major assumption however is that RCO is accepted by airlines, passengers (society), and pilots. No major changes to the operational modes and routines of airlines are assumed.

On the other side, changes in ATM with a greater deviation from today's mode of operations are expected. Most prevalently, the operational changes as proposed under SESAR and NextGen, as summarized in section 3.2, are assumed to be in place. In particular, four dimensional trajectories, self-separation, and revised information management and sharing principles are assumed to be available [US 17; Sin15a].

Technological A generally higher technological level than in today's civil commercial aviation operations is assumed to be available. First and foremost, a higher level of connectivity in the aviation domain is expected to be available. This includes, but is not limited to, the existence of a *reliable, global, high bandwidth, and secure* data link coverage throughout all phases of flight. Related to that is that an assumed shift towards data link based communication methods has taken place. The majority of airspace users and Air Traffic Service Units (ATSUs) is using data link based communication over the "traditional" ways of communication such as voice communication.

Another area in which technological advances are assumed are automatic aircraft operations. Extrapolating the trends observed today, the assumption is made that *fully automatic* aircraft operations are technically possible and are being successfully conducted for cargo operations. Yet, there is no concept nor any intent to conduct unmanned passenger flights on which a human operator is still required due to social concerns. Lastly, the assumption is made that sophisticated equipment exists, which can reliably assess the physiological and psychological (e.g., vigilance) state of the human operator. A limitation is made regarding the implementation of the required technological solutions. In the scope of the thesis at hand no implementation details of technological systems are considered, the conceptual character is superficial.

Legal Challenges regarding the certification process or the legal framework are disregarded. Reasoning behind this decision is that the legal landscape for an envisioned entry into service in 2050 is unpredictable and likely to change based on lobbying efforts or changing international regulations.

Economical No considerations regarding the economical implications are made. The cost associated with certification, training requirements, or manufacturing is disregarded.

3.3.3. Associated agents

Shown in Figure 3.1 are the three primary agents and their relations as viewed in the future concept of operations. The overall system is divided into two domains: airborne and ground-based. On the airborne side the human operator, called mission manager, and aircraft systems are acting. On the ground an additional agent generically depicted as a ground station is located. In the following each of the three agents and their relations to other agents are outlined.

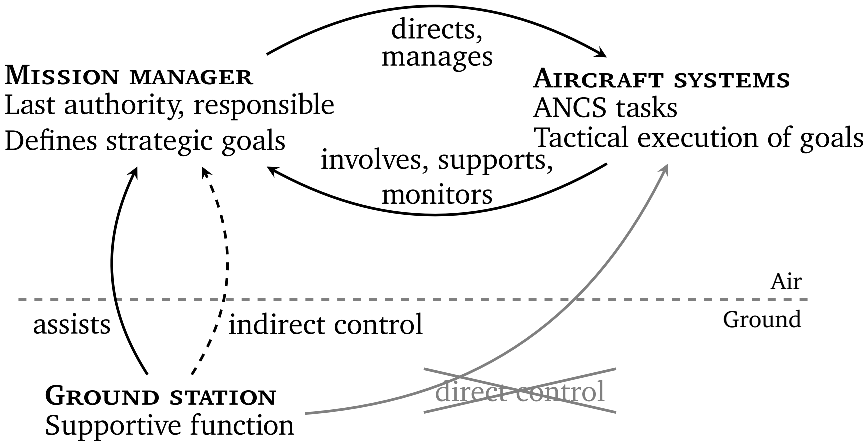


Figure 3.1.: Agents considered in the future flight deck concept and their relations. On the airborne side the mission manager and aircraft systems are shown. Located on the ground is the supporting ground station.

Airborne

On the airborne side of operations the human operator (mission manager) and the aircraft systems are located. Those agents operate the aircraft in a cooperative and self-sustaining work environment. Self-sustaining in this context refers to the fact that all mission-, and safety-relevant decisions can *only* be made from *inside* the aircraft under normal conditions.

Mission manager The human operator on the flight deck will henceforth be called *mission manager*. This renaming is done to highlight envisioned significant role shift for the human operator. In the role of a mission manager, the human operator will focus more on the strategic management of the current mission. Their primary role will be the definition of a strategic mission goal, represented through the definition time or location dependent mission states. As such, goals can be defined in absolutes: "be docked at the gate in Denver at 3pm local time". They can also be defined relational to other goals and can contain constraints: "start the engines after boarding finished, but do no start boarding before fueling is completed”. Mission goals will be further discussed in subsection 4.2.2 and subsection 5.2.1. The notion of setting strategic goals for the automation to execute is a differentiator to other concepts. Task-based approaches are one example where specific tasks are commanded to an automated system and, if necessary, intermediate tasks are automatically identified and supplemented [US11; CS14]. A more closely related example is the concept of directives given to the aircraft systems, or the “H-system” as described in [Sch+07]. Eventually, the mission managers *primary* job is to communicate the goal to the aircraft systems, which will then derive a plan on how to achieve the goal. Planning and executing the required tactical actions to reach the goal is left to the aircraft systems. While doing so they are supervised by the mission manager and supported when necessary, mainly through the provision of context [Har07]. This role is not to be confused with what is described as “supervisory control” [She92]. Instead, higher cognitive planning and decision tasks are assigned away from the human operator. Comparable approaches are present in the relevant research domain and are described in subsection 2.4.4. Micromanagement tasks and direct interaction with the aircraft controls is reduced to adjusting the defined goal or issuing autopilot instructions, for example a temporary heading change, if imminently required. In addition to that occasional interactions may be required when dictated by regulation or by design to keep the operator involved, as it will be described in subsection 4.2.3. Handling of non-normal situations will be touched on later, such as in subsection 4.2.2 or subsection 3.3.4.

In regards to the other agents, the mission manager primarily acts as a bridge between the environment and the aircraft as a system. No command with impact on the mission and/or flight safety can be passed to the aircraft systems from outside the aircraft without the mission manager’s prior approval. Similar to that, the mission manager is also required to pass any information to the aircraft that is outside the scope of the system’s capabilities. Onboard the aircraft the human operator is the highest authority and held responsible for safe execution of the flight as it will be described in subsection 3.3.5.

Aircraft systems Depicted as aircraft systems is a number of computerized systems inside the aircraft. The systems are expected to provide the ability to conduct a defined flight mission fully automatic, not requiring human involvement. The sequence of tactical actions to be performed during the execution, the planning, is performed by the aircraft systems. The plan is derived from the strategic goal as defined by the mission manager. In this role the systems exert control over the following functions:

- Flight controls, primary / secondary (ailerons, elevators, rudder / flaps, trimming)
- Energy management (power systems)
- Supporting functions (landing gear, lights, communication)

The Aviate Navigate Communicate Manage Systems (ANCS) tasks are executed by the automated aircraft systems: Keep the aircraft flying precisely (*aviate*) along a predefined path (*navigate*) while maintaining contact with other airspace users, Air Navigation Service Provider (ANSP) units, and the mission manager (*communicate*). While doing this they are expected to supervise themselves (*manage systems*) [Har07]. In addition to that the monitoring of the human operator concerning physical and psychological health falls into their responsibility. During normal operations the aircraft systems cooperate with the mission manager. After setting the mission goal, they keep the human operator involved and updated on the progress towards the goal. If required, the level of involvement is increased, as it will be outlined in subsection 3.3.4 and later in subsection 4.2.3. Throughout a mission the aircraft systems will furthermore support the mission manager by providing task-related information, decision support tools, or Electronic Flight Bag (EFB)-like applications. There is no direct *control* over the aircraft systems from outside the aircraft. In an emergency situation, however, the systems can be *commanded* to safely terminate a mission, for example when the human operator is incapacitated.

Ground

Supporting functions are placed on the ground, represented by a generic ground station. The ground station can be best described as a single point of service entity, which resembles today's Airline Operation Centers (AOCs). The mission manager (or the aircraft systems) can establish contact with the ground station in case any support is needed. From there on the ground station is able to coordinate with the entities most beneficial to the occurring situation and direct the required help and support.

As previously stated, no direct control of the aircraft is possible. Only indirect control is possible through sending commands or flight routes to the mission manager, allowing execution at their discretion.

Others (not shown)

Aside from the three agents described previously there are other agents to be considered. Included in this category are other airspace users, Air Traffic Control (ATC), ANSP, and data or infrastructure providers. Their influence on the operation is considered to be comparable to today's influence on an airlines individual operational business.

3.3.4. Description of the mission manager's tasks

To get a better understanding of the mission manager's job a description of the tasks they perform is given in section A.4. The definition of the mission manager's tasks is performed in positively (what is a part of the role) and negatively (what is not part of the role) fashion. With each abstract description an example is provided. The examples provided in there are meant to denote the key areas in which the mission manager will have to perform his or her role on the future flight deck.

3.3.5. Distribution of responsibility and authority

The underlying paradigm in the future flight deck concept of operations is *no responsibility without authority and vice versa*. To elaborate on this: no actor can or should be made responsible for something that they are unable to influence by any means possible; e.g. a human operator cannot be responsible for a parameter that is solely controlled by an automated system without being provided an interface to manipulate the parameter. The other way around, an operator is responsible for any consequences that arise from their actions or the omission of these. A similar statement is made by [Bil97]. In this role the human operator, or the mission manager for that matter, will act as the highest authority onboard the aircraft and at the same time be a membrane between the environment and the aircraft systems.

Mission manager as highest authority The mission manager is considered to be the final authority onboard the aircraft and thus is also the person being responsible for the aircraft's safety. This approach reflects the common understanding of human vs. machine responsibility [SWB97]. Responsible in this context refers to the safety of the aircraft, including its passengers, crew, and onboard cargo, and the efficient conduction of the overall flight mission. Hence, the mission manager is granted every authority to influence these values. Especially in regards to the previously mentioned advanced automated systems, the mission manager is able to override any decision or action made by those if deemed necessary. This regulation is compliant with today's practice for commercial aviation under Code of Federal Regulations (CFR) 91.3 [Fed], as described in subsection 2.1.2. The assignment is chosen to avoid a complex structure of responsibilities on the flight deck. Having a human operator on the flight deck would be pointless if there was no way they could exercise authority, for which in turn they have to accept the responsibility.

Mission manager as membrane The relation of the mission manager to the aircraft systems is best described as a (selective) barrier, or in analogy to biological systems: a membrane. As depicted in Figure 3.1, the mission manager is able to filter anything received from external parties, and is able to select which information or commands are propagated to the aircraft systems. For example, if an instruction from ATC is received such as a tactical heading change, the mission manager is able to accept or reject it. Trajectory-related changes require explicit approval from the mission manager. Otherwise, an unwanted external *control* over the aircraft could be exercised by third parties. The aircraft systems will only execute instructions from the outside after approval by the mission manager under normal operating conditions.

3.4. Summary

In this chapter the context in which the to be developed solution will be used was analyzed. An outline of the performed work to better understand today's operating context of pilots on the flight deck is given. The expected changes to the greater context of the GATS as a consequence of research initiatives such as SESAR or NextGen are summarized. The concept of RCO is discussed. In a first step working towards the objective of this thesis the envisioned role of the human operator is described. A new concept of operations for the operator working together with highly automated aircraft systems is described in section 3.3. The term mission manager is introduced. A foundation for the following development process is provided.

4. Information management concept

In this chapter the core development process of the information management concept is continued. Following the process outlined in DIN EN ISO 9241-210 [DIN11] the requirements towards the developed system are specified. A solution for a new information management concept to be used on the flight deck is proposed based on the requirements. In both development steps the findings from the previously performed context analysis, such as the mission manager concept described in section 3.3, are utilized.

Considerations towards the interaction between the human operator and automated aircraft systems are described. The theoretical description of the concept is followed by the depiction of an exemplary implementation of the concept in chapter 5. This application is designed to be used during the evaluation of the concept as described in chapter 6, where it is used in the simulator described in section 5.1.

4.1. Objectives, requirements, and assumptions

First, the objectives for the developed solution are summarized. Requirements for the development process are gathered afterwards. The underlying assumptions are referenced.

4.1.1. Objectives

The objectives and goals of the development are aligned with the overarching objective of this thesis as stated in section 1.2. The objectives set forth in the concept of operations described in subsection 3.3.1 are considered. Two of the therein stated objectives are particularly emphasized:

Safety of operations will and should remain the primary goal when performing any safety critical development. A concept violating basic principles of safety cannot be acceptable and will not be viable. A full formal safety evaluation of this concept must happen later. In this early conceptual stage good judgement should be applied when making safety-relevant decisions. This should not be confused with *doing everything to increase safety*, measure should be taken when it improves safety, while still keeping reliability, comfort, and economy in mind [Bil91].

Put the human operator in the center is the second objective to be emphasized during the development. This objective naturally follows the set forth objective to develop a human-centered information management concept with a high degree of usability as initially described in section 1.2. The abilities of both the human and the automated systems should be used where they fit best in a cooperative environment, following the concept of operations as described in section 3.3. The goal is that the developed concept must be *acceptable* from a human user perspective. The users should not feel constrained or forced into a role by the concept. Instead they should be offered an acceptable level of comfort with expected effects on user satisfaction, safety, and efficiency.

4.1.2. Requirements

The requirements to be met by the developed solution are summarized. The requirements are derived from the literature summarized in chapter 2, focussing on the work referenced in subsection 2.4.3. Whenever possible, direct requirements were adopted from the existing body of research. During the review a majority of the reviewed literature did not state explicit requirements but rather design recommendations. In these cases the recommendations were translated back to solution-neutral requirements, in total over 120 requirements were identified. All requirements were separately cataloged along with a brief summary and their reference. Following that, the requirements were categorized in the following categories:

- Human-Autonomy Teaming (HAT)
- Situation Awareness (SA)
- Information and Interaction Design (IID)
- Others (O)

For each category a set of selection criteria was established to assign the requirements. The selection criteria are listed in Appendix B. A requirement can be assigned to multiple categories. After the requirements were categorized similar requirements were merged. The requirements were finally rewritten to establish a common linguistic style. Eventually, the high level requirements presented in Table 4.1 are derived. A complete listing of the requirements including the more specific sub requirements is presented in Appendix B.

Table 4.1.: Overview of high level requirements towards the information management concept summarized across all categories.

ID	Requirement
HAT-1-0	Provide adequate levels of feedback about the system (automation, aircraft systems) state to the human operator.
HAT-2-0	Make automated systems observable in their actions and the decisions they make.
HAT-3-0	Clearly indicate the ownership of a function (which agent is responsible for doing what).
HAT-4-0	The human operator must always be in command.
	[SW91; SW92; SW95; SW97; Bil91; Bil97; ASS96; Pal+95; Fun+99; CW02; Har07; Let+12; AMR13; End16; Com+13; You+16a; Bai+17]
SA-1-0	Provide effective means of task management for the human operator.
SA-2-0	Support the prediction of mission and system states.
SA-3-0	Support buildup and maintenance of a proper mental model.
SA-4-0	Support and reinforce situation awareness at all times.
	[Bil91; Fun91; SW91; SW92; SW95; SWB97; SMW07; CW02; BBR07; BM04; Let+12; Yeh+13; Fli14; End16; You+16a]
IID-1-0	Support efficient and effective input by the human operator into aircraft systems.
IID-2-0	Provide effective attention guidance principles.
IID-3-0	Support effective and efficient information transfer between aircraft systems and the human operator.
IID-4-0	Establish a standard of information representation across all systems.
IID-5-0	Use a visual design aligned with established design standards and guidelines.
	[Bil91; Fun91; SW92; SW95; SWB97; Pal+95; CW02; WPR02; Wic+03; Wic08; BBR07; Sch+07; Let+12; Com+13; DBS13; Dua+15; Yeh+13; AMR13; End16]
O-1-0	Put the human operator and their role in the center of the design.
O-2-0	Establish trust between automated systems and the human operator.
O-3-0	Consider human limitations (memory, computation, attention, decision making biases, task timesharing, cultural differences).
	[Bil91; Pal+95; Fun+99; Har07; Wic08; AMR13]

4.1.3. Assumptions

Due to the exploratory and academic nature of this thesis, certain boundary conditions that exist in the real world cannot be properly considered during the development process. To not loose focus of the objective of this work the assumptions stated in subsection 3.3.2 are carried forward into the development process.

4.2. Goal oriented Information Management

In this section a new concept for managing information on the future flight deck is presented as the next step of DIN EN ISO 9241-210 [DIN11]. Under the name of Goal oriented Information Management (GoIM) a new way of managing information and exchanging it between human operators and automated aircraft systems is presented. The described concept is the logical continuation of the future concept of operations in the expected context as described in chapter 3. During the conceptualization phase the requirements derived in section 4.1 are considered.

Goal oriented The term *goal oriented* conveys a twofold meaning. On one hand, the term is chosen to highlight the fact that all actors, humans and machines, on the flight deck should be following the same goals when working together in a cooperative environment. On the other hand, the term is chosen to highlight the fact that the human operator will act as a goal setter in the proposed concept of operations as it is described in section 3.3. Human operators, in the role of a mission manager, will be responsible to set high level strategic goals which are then translated into smaller tactical actions by systems supporting them. As such, the principles of information management should follow the overall notion of working goal oriented in the best possible way. Consequently, a *central idea* of GoIM is that information required from, and provided to the human operator is related to reaching the defined goal as visualized in Figure 4.1. This approach is chosen to overcome the challenges related to information management for a human operator, potentially operating under Reduced Crew Operation (RCO), as discussed in section 1.1.

The relation between goals and actions as depicted in Figure 4.1 is described with an example: The underlying assumption is that an organization would always work towards a common goal (1) which is set by the company's leadership (*Always be on time.*). Eventually, the operating elements in an organization, in this case the operators of the aircraft of an airline, are supporting these goals (*They pay higher attention to the on-time performance of their flights.*). They derive their own goals (2). The mission manager, as one operating element, is consequently responsible to share such a goal with the automated aircraft

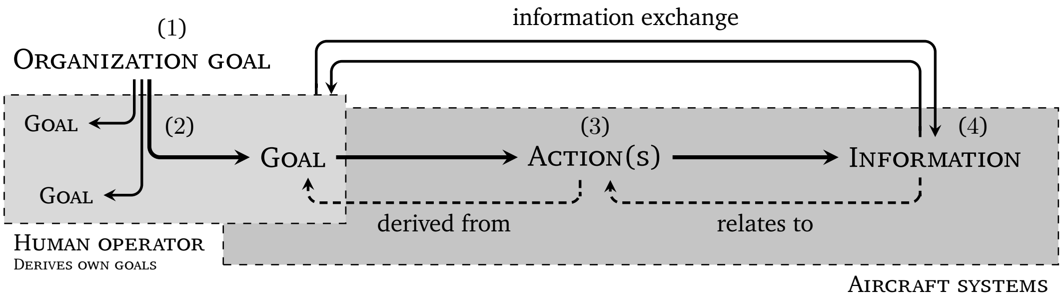


Figure 4.1.: Central idea of GoIM, information must trace back to the defined goal through the actions required to reach the goal.

systems (*Explicit time constraints are set for the arrival at the target airport.*). Based on this input the aircraft systems will derive and plan the required actions (3) to meet the defined goal (*Ensuring on-time departure and preferring speed over economical benefits when calculating the speed profile.*) The planned actions then drive the information (4) which is exchanged with the operator (*Request flight level change (action) to FL340 to benefit from a 20 kn increased tailwind (information).*) Through this information exchange the operator is able to make the required decision to meet the communicated goal or to perform goal adjustments as outlined in section 3.3.

Glossary Throughout the description of the concept, the following terms state, action, plan are centrally used. The *state* of a system is a momentary description of all system variables and their current values. Once a system variable changes, the system is in a different state. Changes to system variables can be deliberately caused by performing actions, or be a result of external influences on the system. Deliberately performed activities with the objective of changing the current system state are called *actions*. Actions require certain preconditions to be met, as reflected in the system state, and on the other hand have certain effects on the state itself. Lastly, a *plan* describes a sequence of actions which is performed to achieve a defined goal state. The terms are based on the terminology used in section 2.3. In the following section the established terminology will be used to describe the GoIM concept and the involved components in detail.

4.2.1. Components

As part of GoIM three major components are considered: the human operator, the aircraft systems, and the flight deck interfaces. The flight deck interfaces are considered to be an enabling component, which allow the communication between the human operator and aircraft systems. As part of the exemplary implementation of the GoIM concept the interfaces are the subject of section 5.2 and chapter 6.

Shown in Figure 4.2 are the three components and their interaction. From left to right the diagram shows three stages of information processing: perception, planning, and execution. The stages are based on the model of human behavior by [Ras83] and the situation awareness model by [End88] as described in subsection 2.2.3 and subsection 2.2.2. On the left side of this depiction the available input channels for the human operator and the aircraft systems are shown. On the right side, in the execution stage, the actions available to either are depicted. In this initial representation no details about the planning stage are shown.

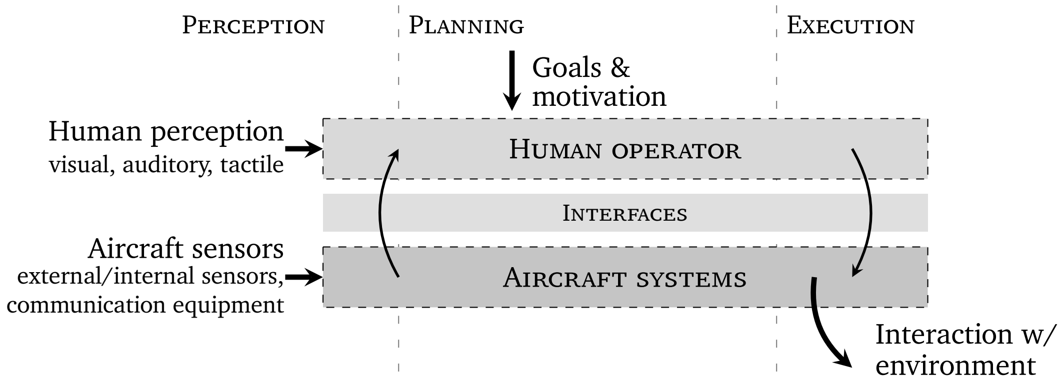


Figure 4.2.: The three main components of the GoIM concept: human operator, flight deck interfaces, and the aircraft systems. The diagram is divided into three stages of perception, planning, and execution (left to right).

Human operator The top component depicted in Figure 4.2 is the human operator. The human operator is able to perceive information about the environment through the human sensory system. Mainly visual, auditory, and tactile cues are perceived for further processing. Input cues are either perceived from the outside environment, such as weather conditions, or aircraft system outputs via the interface layer. Naturally, the human operator can interact with the interface layer through different modalities, such as using a touchscreen or voice commands to communicate with the aircraft systems. As the main goal setter in the flight deck environment the human operator is also assumed to be aware of the overarching goals and motivation.

Flight deck interfaces The flight deck interfaces are the interfaces between the human operator and the aircraft systems. As previously stated they do not fulfill a specific role on their own other than conveying information between the aircraft systems and the human operator, or vice-versa. In the scope of this work the input capabilities are assumed to be of tactile nature only, which includes buttons, switches, touchscreens, or cursor devices. The output from the aircraft systems on the other hand is limited to the visual and auditory channels. As such, the interfaces are a key aspect when considering the requirements of the IID category such as proper alerting as required by IID-2-0. The interfaces are further focused on when the exemplary implementation of GoIM is described in section 5.2.

Aircraft systems Summarized under the term aircraft systems is the combination of aircraft hardware, and the automated/autonomous systems onboard the aircraft as described in subsection 3.3.3. As such, the aircraft systems are the counterpart of the human operator, they are responsible for executing the tactical actions which are required to meet the defined goals. Similar to the human operator the aircraft systems are able to perceive information about the environment through their sensory system. A suite of external sensors is assumed to provide the system with information about the external environment. The internal status of the aircraft itself is assessed through self-diagnosing systems, for example. Input from the human operator is received via the flight deck interface layer, through which direct feedback can be provided as well. During a mission, in contrast to the human operator, the aircraft systems are able to interact with the external environment. A direct interaction can happen through system components such as engines or aerodynamic control surfaces, supported by lower level autoflight systems as they are being commonly used today. In addition to that, the aircraft systems provide the ability to communicate with the external world.

4.2.2. Cohesive description

After having described the high-level components of GoIM a closer look at the interactions and internal processes is taken. The interaction of the human operator with the aircraft systems under GoIM in normal operations is described first. After that, the role of the human operator is reviewed more closely.

Normal operations

Shown in Figure 4.3 is a graphical depiction of the process of how a human operator cooperates with the aircraft systems under GoIM. The diagram is divided into the human operator's domain on the top, and the aircraft systems' domain on the bottom. Both domains are overlapped by the interface layer through which all information exchanged between the agents must flow. The gray arrows traversing the *interface* layer indicate the central information flows to the human operator.

In the *perception* stage the human operator and the aircraft systems are receiving data through their sensory systems. The *environment* from which they receive data includes the external environment, the "world", and the system environment which includes internal aircraft system states. The human operator's perception is supported with supplemental information provided by the aircraft systems, for example the aircraft's exact position when flying over the open ocean or at night or the current airspeed. During the identification and integration of the perceived information both agents are creating a state model of the environment, the current *mission state* (1), highlighted in the upper left of Figure 4.3. Although based on the same factual data, the state models are likely to differ: The human operator's model will be largely based on a mental model, comparable to what is described for situation awareness in subsection 2.2.2, which is influenced by previous experience and training. The aircraft systems' model on the other hand will be based on the underlying system model and pre-trained models used to interpret the received sensory input. However, in both mission state models information about the current, past, and predicted future states are expected to be present.

During the *planning* stage the human operator plans required actions based on the current mission state and cues provided by the aircraft systems. The operator could be required to provide input to an action currently executed by the aircraft systems, or could determine that adjustments to the current mission goal are necessary. In either case, the decision of which action to take and how to take it is driven by the goals and motivations of the human operator, influenced by the goals communicated from higher organizational levels. During the decision making process the human operator can be supported by the aircraft systems. The outcome of this decision is a formulated action plan. In the *execution* stage this action plan is executed (2) by using the interface layer and the therein provided interaction modalities. Notionally, this process reflects the *knowledge-based* level of human behavior by Rasmussen as described in subsection 2.2.3. The planning process is only limited to the immediate action of the human operator such as adjusting the mission goal. The task of planning a sequence action to reach the overarching mission goal is delegated to the aircraft systems.

On the aircraft systems' side a formulated goal is translated into a *target state* (3), which acts as a proxy for the defined goal. The target state is in a format compatible to the perceived state of the mission and the environment as perceived by the aircraft's sensory system. Consequently, a comparison between the target state and a *predicted* future mission state under consideration of any planned actions (4) can be performed. When performing a prediction the aircraft systems are expected to account for the inherent

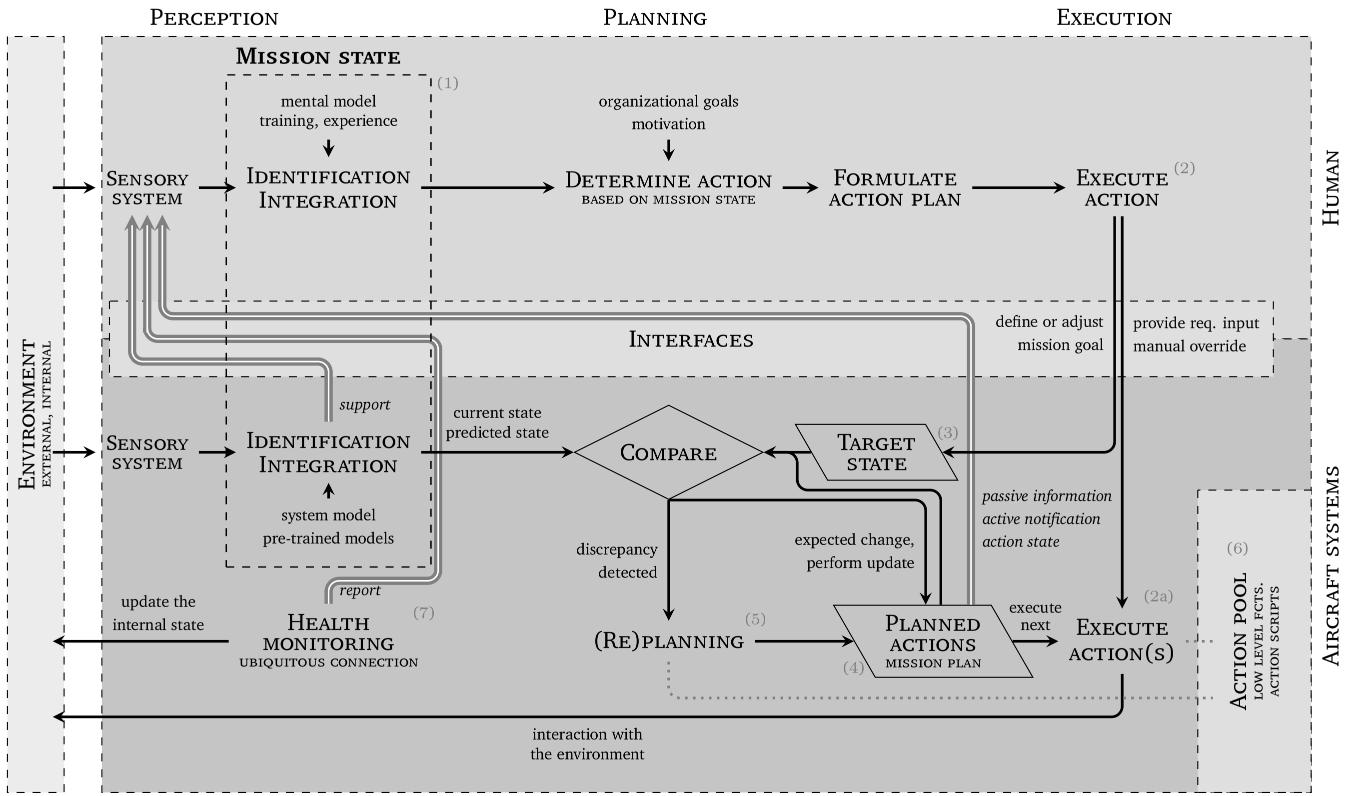


Figure 4.3.: Diagram showing the notional interaction between the human operator (top) and the aircraft systems (bottom). The central information flow to the human operator as a result of the defined goal is depicted by the gray double-lined arrows. Numbers in parentheses are text references.

uncertainties when extrapolating from the current state.

Discrepancies between the predicted future mission state and the goal state are identified by the aircraft systems. If the discrepancies exceed a certain threshold the aircraft systems must plan (5) corrective actions. This will be the case at the start of a mission when a new goal is defined. The aircraft systems will attempt to plan a sequence of actions that cause a transition from the current mission state into the designated target state. During the planning process the actions available in the *action pool* (6), shown in the lower right of Figure 4.3, are considered. Summarized in here are actions and action scripts performed by both the human operator and the aircraft systems. The actions available to the aircraft systems are determined based on the aircraft capabilities as described in subsection 3.3.3. Exemplary actions requiring the involvement of the human operator are discussed in subsection 4.2.3. Based on the current mission state, a defined target state, and the set of available actions a planning method (cf. section 2.3), is then utilized to plan a sequence of actions to reach the target state: the *mission plan* (4). Once a valid solution for a newly defined goal is discovered it is at the discretion of the human operator to review and accept the solution. Aside from the initial planning process after defining a goal, a replanning (5) of actions can happen anytime during a mission. This is triggered by either one of the following conditions:

- An unexpected change of a mission state variable which is not caused by a currently performed action
- A growing discrepancy between the predicted future mission state and the goal state

Depending on the extent to which actions are replanned, it may also be acceptable to inform the human operator as opposed to again requesting authorization from them.

Through the mission plan (4) the human operator and the aircraft systems are able to cooperatively work towards the defined goal. While the operator can retrieve information about performed actions, the plan also indicates when information is required from the operator. Through the mission plan the information management is related to the defined goal and the derived actions, representing the concept of GoIM as outlined in section 4.2.

Over the course of the mission the mission plan is sequentially executed. Every executed action (2a) interacts with the environment and its effects are observable through changing state variables. This change is monitored by the aircraft systems and continuously evaluated against the target state and the currently executed action(s). Once an action is determined to be completed (when all its expected effects are observed) the next action or the next actions are executed. Actions can either be automatically performed by the aircraft systems or with support of the human operator, this depends on the action. In the first case, the operator's involvement is limited to passive observance of the action's status. In the second case, the operator is actively notified that active involvement is required. On the human operator's side this notification is perceived as previously described, followed by determining a course of action and eventually executing it (2). The input is received by the aircraft systems and the currently executed action is continued. Additionally, the human operator would be able to use the same path to manually modify an executed action or override it. This path is retained to account for the distribution of authority and responsibility as discussed in subsection 3.3.5.

Lastly, the aircraft systems provide self-monitoring capabilities, as described in subsection 3.3.3, depicted as an ubiquitously connected *health monitoring* function (7) in the lower left of Figure 4.3. Through this function the human operator is informed about the current status of the aircraft system. Knowing the health status of their cooperation partner is considered essential as it allows the human operator to increase involvement when necessary. Cases in which the human operator is required to support the aircraft systems are described in the following.

4.2.3. Human role in Goal oriented Information Management

Based on the previous description of the nominal operations under GoIM, the role of the human operator in this context is reviewed more closely. One might argue that the human operator is only a passive observer of the automated systems performing the current mission, and is only required to define a goal once. This is not the case. Different situations are envisioned, in which the human operator, in the role of the mission manager, is required to directly support the aircraft systems.

Tasks of the human operator under normal conditions

Aside from the tasks described in the previous section and subsection 3.3.3, a number of tasks is additionally assigned to the human operator by design of the concept. This can be required due to regulatory constraints or operating procedures of the airline. In Figure 4.4 *exemplary* tasks of the operator over the course of a mission are highlighted. (1): Prior to pushback a briefing of mission data, including planned routes, is required. Allowing the passengers onboard the aircraft is also coordinated by the operator. Pushback from the gate must be initialized by the operator after that. (2): The next critical step during the mission is the initiation of takeoff to be performed by the operator. (3): The operator can be required to react to external commands or instructions from Air Traffic Control (ATC) (cf. membrane analogy in subsection 3.3.5). This can happen at any point during the mission as indicated by the gray outline in Figure 4.4. (4): After descent towards the destination airport the operator must allow the aircraft to initiate the final approach and landing.

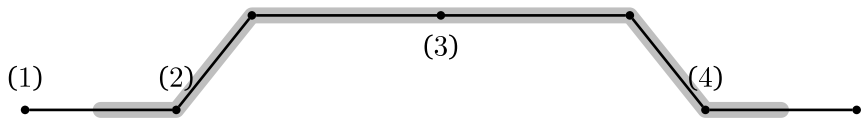


Figure 4.4.: Exemplary points during a flight mission where actions from the human operator under GoIM are required.

The listed actions are chosen as they represent critical checkpoints at which an increased awareness should be expected from an operational and safety point of view. In extension of that, the operator must also be able to exercise tactical interventions when necessary:

- Aborting the mission to the nearest safest location
- Pausing mission execution (flight phase dependent)
- Commanding manual alterations of the plan (e.g., a temporary heading change)

Under normal operating conditions the assignment of tasks and responsibilities is *fixed*, no tasks from the aircraft systems are dynamically reallocated to the human operator. The concept of adaptive automation is not pursued under normal operating conditions, following the recommendations given in [Sch+07]. Instead, the goal is to maintain an evened out level of involvement across a mission, allowing the human operator to work on secondary tasks supporting the overarching operation of an airline. The latter concept is further discussed in [Nei20].

Supporting aircraft systems

Aside from the tasks of the human operator under normal conditions, there are also situations in which the aircraft systems must be *supported*. Support can be required in either one of the three processing stages. In these situations the assignment of responsibilities and abilities, as described in subsection 4.2.2 and subsection 3.3.3, between the human operator and the aircraft systems can be shifted. Even though shifts may occur it has to be noted that this may not happen in an all-or-nothing fashion. The aircraft systems are expected to gradually request support and to maintain supporting functions as long as possible. Derived from the description in section 3.3, an exemplary situation for each of the three processing stages is outlined in the following. A central element of all these situations is that the human operator must provide insight and directions to the automated systems. Quoting Harris [Har07], the role of the human operator is the provision of *context (!)* to the automated systems, as it is reflected in requirement HAT-4-1 and discussed in section 3.3.

State determination support Just as the correctness of planned actions bears uncertainty, the state determination process can be prone to errors and can lead to uncertainties. When aircraft sensors temporally deteriorate through system malfunctions or external factors, the state of the mission might be wrongly assessed, leading to potentially faulty system behavior. In the case that the aircraft systems are unable to correctly determine the current state, or the associated uncertainty rises above a certain limit, the human operator is required for assistance. Without the human operator's support this would lead to an incorrect or incomplete state model on the aircraft's side in Figure 4.3. Such a scenario can be detected by either the operator or the aircraft systems. The operator would be required to now support the aircraft in building up an accurate state model of the environment, contrary to the depiction in Figure 4.3. This way the aircraft systems would still be able to carry out the downstream functions to support mission execution. Exemplary for this would be equipment failures, or situations which are unrecognizable by aircraft sensors. An example of the latter would be a sick passenger which would require the aircraft to divert, as a full supervision of passengers is unlikely to happen.

Planning support The second example considers a need for support during the planning process, as shown in Figure 4.3, which is normally performed by the aircraft systems. Presumably, there can be different alternative plans on how to reach a defined goal state. Each plan will have advantages, disadvantages, and an uncertainty of success associated. Under normal conditions the aircraft systems would weigh the plans against each other and decide on a solution, comparable to the human behavior described in subsection 2.2.3. In case the uncertainties of success grow beyond a pre-defined limit the human operator, as the final authority and responsibility onboard, must be involved. Depending on the root cause the human operator's involvement is different. In case multiple potential plans to reach the defined goal state could be formulated but no optimal solution can be identified, the human operator would be required to choose a solution. Alternatively, the aircraft systems could indicate that no solution can be calculated to reach the currently defined goal state. In this case, the human operator would have to understand the constraining factors and eventually adjust the defined goal. In either case, a partial shift-back of the planning activity originally delegated to the aircraft systems to the human operator would occur. While supporting this activity, the operator is still expected to consider and represent the higher organizational goals. From the aircraft systems the continued provision of supporting information is expected.

Execution support Lastly, the human operator can be required to support the aircraft systems during the execution stage. In such a situation a degraded performance of the tactical actions available to the aircraft systems is expected. The origin of the degraded performance can either be a defect in the aircraft systems, or caused by external influences such as unavailable ground infrastructure. The first step in resolving this situation is diagnosing the problem and attempting to fix it. If the situation cannot be resolved, in terms of restoring functionality, the human operator is now required to directly command lower level aircraft systems. The extent of this mode of operation can range from today's interactions of a pilot with the Mode Control Panel (MCP) up to manually flying the aircraft and safely terminating the mission, which is considered a *last resort* measure. In doing so it is expected that varying levels of support are still provided by the aircraft systems. Again, a shift of responsibilities and tasks back to the human operator would be observed in this case.

4.3. Summary

In this section the development process was continued. Based on the context analysis performed in the previous chapter and a literature review, requirements towards the information management concept were derived. With GoIM a new concept for how the human operator and the automated aircraft systems interact is described. The concept is based on the idea of managing information based on a common goal to be followed by the human operator and the supporting aircraft systems. The information is managed based on the actions to be performed to reach the defined goal. The role of the human operator under GoIM is described, the division of tasks between human operator and aircraft systems in different scenarios is outlined.

5. Implementation

In this chapter the implementation work done in relation to the Goal oriented Information Management (GoIM) and the mission manager concept is described. Requirements and assumptions stated in subsection 3.3.2 and subsection 4.1.2 are translated into design solutions. The development and construction of a purpose-built Reduced Crew Operation (RCO) research simulator is described. The implementation of GoIM in a software application is outlined.

5.1. Reduced Crew Operation simulator

To adequately represent, or evaluate, future cockpit operations a physical environment is needed. At the time this work was started the Institute of Flight Systems and Automatic Control (FSR) at Technische Universität Darmstadt (TUDA) was already in possession of two fixed-base flight simulators: An Airbus A320-like simulator, and a Diamond DA40 simulator. Both simulators were deemed unsuitable to represent *future* operations, an inadvertent association with today's operations cannot be ruled out. It was decided to create a new part-task simulator especially for RCO, as required by O-1-2, without any relations to today's operations. The goal was to create a platform on which future concepts can be *evaluated* and *demonstrated*. The prevailing concept of operations at the time the simulator was conceptualized was the mission manager concept as described in section 3.3.

5.1.1. Design process

The simulator was conceptualized, designed and built during this work. It was chosen to construct the simulator in multiple iteration stages. This was done because there are no existing examples of RCO cockpits in the commercial aviation domain, for which the mission manager concept was developed. Eventually, the design and construction process can be divided into the following stages:

1. Initial wood and cardboard construction
2. Metal frame and installation of visual system

3. Final interior design

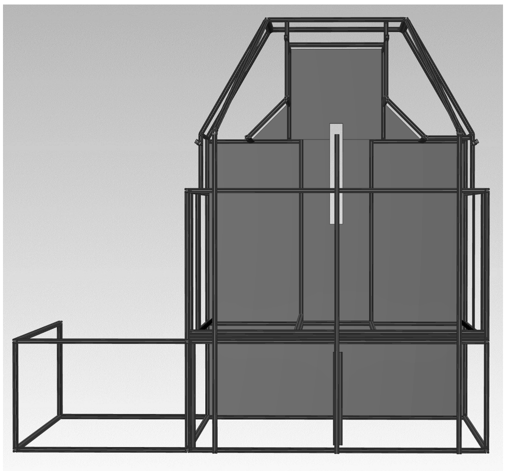
During the initial development stage the interior dimensions were determined. The first version of the simulator consisted of a wooden frame with cardboard and paper applications on the inside. This approach was chosen as a cost-effective way of determining the position of windows, the operator's chair, or displays based on IID-5-1. Guidelines for workspace ergonomics [SLB10] were used to determine the positions of touch-capable screens in relation to the operator's eye reference point. The Federal Aviation Administration (FAA) Advisory Circular (AC) 773-1 [Fed93] was referenced. Shown in Figure 5.1 is a view of the interior during this construction stage while screen positions were determined. An evaluation of screen positions and window cutouts in relation to the eye reference point was performed [Ste17; Sch+18].

Figure 5.1.: Inside view of the simulator during the initial construction stage.

Additionally, the visual system was designed during this development stage to be installed later. A cylindrical 210° projection canvas with a diameter of 5 m is used to display the outside view using three short throw projectors. The canvas is 2.3 m high and starts 0.3 m above the floor. Based on the eye reference point the cockpit windows where then designed to limit the operator's view to only see the projection canvas and no other external references. In contrast to a typical flight deck an undivided front window was chosen to accommodate the centrally placed chair. An additional window segment on each side of the central window provides the necessary peripheral vision, which for example can be required during taxi operations.

During the following development stage, the wooden construction was transitioned to a metal-based frame. The load-bearing structure of the simulator was replaced by metal profiles to which the interior casing was attached. During the last construction stage the interior casing of the simulator was constructed using high pressure laminate sheets. Matte and dark surfaces were preferred to minimize reflections on the displays and windows following IID-5-1 and AC 25-11B [Fed14].

Eventually, the back of the simulator was closed using the same sheet material. A top down view of the simulator frame is provided on the left side of Figure 5.2. The area behind the operator's seat is left empty and allows future expansions. The cockpit area is accessed through a sliding door in the back of the simulator. Located to the left of the simulator is the simulator control station. On the right side of Figure 5.2 a picture of the final interior of the simulator is shown.



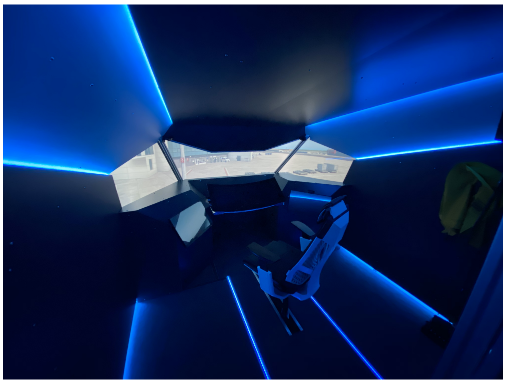


Figure 5.2.: Left: Top-down view of the simulator interior. Right: Picture of the final simulator inside after the inside casings are attached.

As shown in Figure 5.2, large screen devices are chosen to represent all interfaces inside the simulator. Touchscreens were chosen as they allow a more flexible way of reconfiguring the interface and allow a faster iteration of interaction concepts than physical devices would. This decision is supported by IID-1-0, IID-3-0, and the recent introduction of touchscreens to the flight deck for the Boeing 777X [Boe16b]. Centrally located is a curved touchscreen with a resolution of 3440 by 1440 pixels. Left and right of the center screen two 16:9 monitors offer additional display space. While all monitors are within the operator's reach, only the center one is designed to be used for constant monitoring and interactions, it is placed in the operator's central field of view. It was decided that *no* devices for manual flight path control will be constantly present in the cockpit. This decision is directly derived from the operational concept as described in section 3.3, where the usage of flight controls is reduced to a necessary minimum. Although the option to temporarily place a joystick and a thrust lever inside the cockpit exists, in the normal mode of operation none are provided.

As previously described, the simulator serves a dual purpose: as a demonstration and an evaluation environment. To accommodate for the latter one the possibility to attach multiple webcams, three at the time of this thesis, on the inside of the cockpit is given. Additionally, a bio-sensor kit from biosignals PLUX[1] can be used to assess physiological data from participants. Universal Serial Bus (USB) ports inside the cockpit provide power and data connections to the control station. If necessary, a 5 V and 12 V grid is available to power additional equipment. The environmental conditions inside the simulator are controlled through an air conditioning system. Adjustable Light-Emitting Diode (LED) stripes provide indirect lighting inside the cockpit and allow the simulation of different outside lighting conditions. The simulator can be controlled from the previously mentioned control station. From this station a supervisor is able to observe the participants and adjust experiment parameters when needed.

5.1.2. Software architecture

The simulator software is based on a combination of Commercial off-the-shelf (COTS) software and a set of developed microservices which are used to simulate highly automated aircraft operations. Through these microservices the simulated aircraft is capable of performing a full gate to gate flight mission without any human interaction (cf. subsection 3.3.2), including mid-flight flight plan changes. A hardware setup of three servers is used for the simulation environment, all servers are connected through a centralized Ethernet switch. Network storage capacity is provided inside the simulator network to store experiment data. External connectivity with the internet exists, additional devices can be connected to the network via WiFi or Ethernet. All connections between the different components are realized through network communication. Depicted in Figure 5.3 is the software architecture of the simulation core, through which the fully automated flight execution is realized.

[1] https://biosignalsplux.com/ - last accessed on Sunday March 29, 2020.

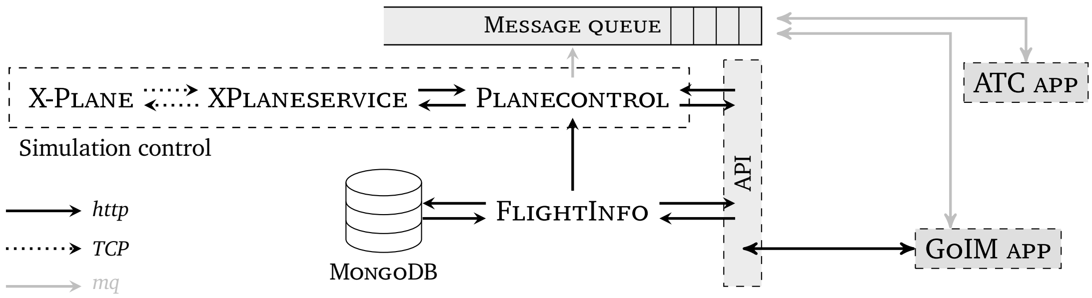


Figure 5.3.: Simulator architecture showing the connection between X-Plane and the developed microservices. Externally facing interfaces are shaded in gray. Physical machine allocations are omitted in this depiction.

The flight simulation software is the commercially available software X-PLANE 11 professional version[2]. To the extent possible, the built-in autopilot provided by X-Plane is used to guide the aircraft during cruise flight, available Instrument Landing System (ILS) capabilities are used to allow automated landings. The X-PLANE server is dedicated to running the simulation itself.

Control over the simulation is performed via the *CONTROL* server, on which a suite of microservices resides: The XPLANESERVICE service provides a web Application Programming Interface (API) through which the access to the so called “dataref” layer of X-Plane is abstracted. The service allows alteration of the internal Flight Management System (FMS) of an aircraft and the activation of autopilot modes. Through the implementation of customized routines it also allows moving the aircraft along a predefined taxi path and performing an automated takeoff without any human intervention.

A modified version of the open source plugin *ExtPlane*[3] plugin is used to connect to X-Plane via Ethernet through a Transmission Control Protocol (TCP) connection. The FLIGHTINFO service provides access to a *MongoDB*[4] database in which flight plans, taxi routes, and other aeronautical information can be stored and retrieved from. Actual control over the aircraft is performed via the PLANECONTROL service. The service then aggregates all required information for the flight phase, for example the taxi route, from FlightInfo and sends it to XPlaneservice where it is eventually translated into commands for X-Plane. Planecontrol, XPlaneservice and X-Plane are forming the central simulation control.

[2]https://x-plane.com/pro/ - last accessed on Sunday March 1, 2020.
[3]https://github.com/vranki/ExtPlane - last accessed on Sunday March 1, 2020.
[4]https://mongodb.com/ - last accessed on Thursday April 2, 2020.

All microservices are accessible via web API. As illustrated in Figure 5.3, FlightInfo and Planecontrol are the primary endpoints to which other software can connect to. Alternatively, a message queue[5] can be used to subscribe to aircraft status updates, such as position, speed, or the currently active flight plan. Depicted in Figure 5.3 are two applications using the provided endpoints: An Air Traffic Control (ATC) application with which communication with ATC can be simulated. The applications implements the International Civil Aviation Organization (ICAO) Doc 4444 [Int16] Controller-Pilot Data Link Communications (CPDLC) message set. The ATC application only uses the provided message queue to retrieve aircraft positions and to exchange CPDLC messages with other applications. Secondly, the application in which the GoIM concept is implemented is using the web API endpoints and the message queue. The user interface of the application is displayed on the touchscreens inside the cockpit, the application is hosted on the third server *DISPLAY*. The internal architecture of the application is further described in the following section.

5.2. GoIM implementation

In this section the implementation of the GoIM concept is described. The concept is implemented in an application designed to be used in the previously described simulator to control the aircraft. Selected technical details of the application and how the concept of GoIM is translated into the application are outlined. An iterative development approach was chosen. During the development process feedback was provided by direct peers of the author and industry experts of the Boeing Global Services (BGS) lab in Frankfurt, Germany. Expert guidance was provided in the areas of User Experience (UX)/User Interface (UI) design and operational knowledge.

Scope It has to be noted that the application does not support all use cases described in subsection 4.2.3 or section 3.3. The developed application is only functional to the extent required by the evaluation scenario described in the following chapter. For example, no means of issuing any sort of tactical commands while in flight, such as a heading of speed change, are provided. Beyond the scope required for the evaluation only limited functionality is provided. Some of the presented data may be mocked or static.

[5] At the time of this work the message queue was based on the Apache ActiveMQ protocol https://activemq.apache.org/, now it is based on RabbitMQ https://rabbitmq.com/ - last accessed on Sunday March 1, 2020.

5.2.1. Mission goal definition and planning

The process of translating the defined goal into a sequence of actions is performed by the aircraft systems, following the description given in subsection 4.2.2. Technical details on goal definition, the planning process, and action execution are provided in the following. Additional details are provided in [Spr19].

Goal definition Only one goal is active at a time, while a second goal may be in a planning state. A mission goal can consist of one to many *goal parameters*, through which the goal is defined. Goal parameters are divided into two categories, spatial and operational parameters. *Spatial parameters* are used to define a geographical location to be reached during mission execution, such as an airport, a waypoint, or a set of geographical coordinates. Optionally, the action to be performed at the target may be defined, e.g. flying over, or landing. *Operational parameters* are used to define the state of a discrete state variable. Goal parameters are represented through state variables, forming the goal state G (cf. section 2.3). Parameters used for goal definition are highlighted in section C.2.

Goal parameters can be supplemented with *constraints* such as temporal constraints on when a location shall be reached or when a state change may occur. Temporal constraints can be used to sequence goal parameters. Similar to temporal constraints goal parameters can also be used to constrain the operational execution of a mission, e.g. to not cross a certain altitude or exceed a certain speed. Such constraints can be issued by ATC or defined by the user. Comprehensive support for constraints is not implemented. The only constraint considered is the ability to restrict the taxi speed of the aircraft as required for the evaluation scenario.

Planning process Once a goal is defined, the planning process is started. The planning process is based on the Goal Oriented Action Planning (GOAP) algorithm, as described in section 2.3. GOAP is considered to be a suitable solution, for the initial implementation of a planner in the scope of this thesis, through its original use case [Neu+19]. The A* search algorithm (A*) [HNR68] is used for plan calculation. The *mission state* (S_m) (i.e. world state in section 2.3) is abstracted by a set of boolean state variables, it contains a representation of all available state variables. In addition to their boolean state, each variable can contain additional object information (e.g., the variable "Mission_Route_Airport_destinationSet" contains the respective airport identifer). Listed in section C.2 are all implemented state variables. A previously defined set of operators (actions) exists in the application memory. The condition described in Equation 5.1 applies to the preconditions (P_{pc}) and effects (P_e) for all available operators ($[O_0 \ldots O_n]$) and all possible mission states ($[S_0 \dots S_n] = [P_0 \dots P_n]$).

$$[P_{pc,1} \dots P_{pc,n}] \subset [S_0 \dots S_n] \wedge [P_{e,1} \dots P_{e,n}] \subset [S_0 \dots S_n] \tag{5.1}$$

The implemented actions are focused on the scope of the evaluation scenario, but allow planning of a full gate-to-gate flight mission. A list of the implemented actions ($[a_0 \dots a_n]$) is provided in section C.3. With each action a dimensionless cost value is associated which is based on the number of effects caused by the action. A heuristics function based on the number of state variables left to change until the goal state is reached is implemented. Mission states are represented as nodes for the A* planning process. Edges are represented through actions fulfilling the condition described in Equation 2.2. Once a path from the initial node of the current mission state to the goal node is discovered it is backtracked, a list of actions is generated. The list of actions is called the *mission plan*.

Action execution By activating the mission plan the planned mission goal becomes the active mission goal and the associated list of action replaces all previously activated actions. Already completed actions are kept in memory to provide traceability for the operator (cf. subsubsection 5.2.2). Upon execution of a plan the first action is executed. Every time an action (a_i) execution is attempted the associated preconditions ($P_{pc,i}$) are compared to the current mission state ($P_{pc,i} \subseteq S_m$). After successful execution of the first action it attempts to execute the subsequent action if the required preconditions are met. This behavior is repeated until an action cannot be executed due to a mismatch in preconditions. Action effects are not instantaneous through the connection to the simulation environment described in section 5.1. Parallel action execution is enabled through this behavior. As long as an active mission plan exists, the planner monitors all changes to mission state variables. As shown in subsection 5.2.4 the central mission state repository notifies the planner about changes. The process depicted in Figure 5.4 is executed.

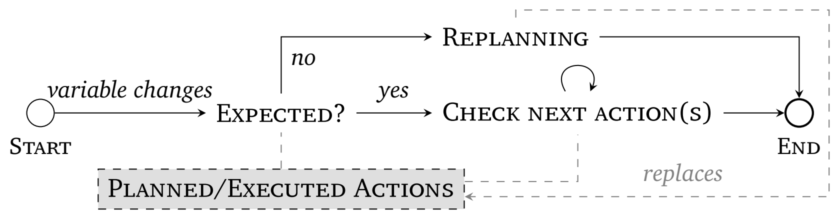


Figure 5.4.: Monitoring of state variable changes through the planner.

Once a state change is registered, the planner checks the observed change against the set of effects of all currently executed actions (a_n). A state change is considered expected when the changed variable is part of the set of effects. Next, it is evaluated if a currently executed action (a_i) is completed. An action is completed when $P_{e,i} \subset S_m$ is true. Analogous to the initial execution of actions, an attempt is made to execute the next planned action. If the changed state variable represents and unexpected state change a replanning is automatically triggered. The planning result replaces the previously planned actions as depicted in Figure 5.4. An example of an unexpected state change is described in the evaluation scenario in section 6.2.

5.2.2. Application layout and components

During the development process it was chosen to follow a widely accepted design guideline, as required per IID-5-0. They benefit from the availability of tested standard libraries and design elements, as well as making users feel more comfortable while interacting with the application, thus reducing the required training effort. Design standards that were taken into consideration were Google's Material Design[6], Apple's Mac OS/iOS design[7], and Microsoft's Fluent Design[8]. Eventually, Microsoft's Fluent Design was chosen. Microsoft Windows is widely known and utilized and other design libraries are strongly focused on mobile devices with different requirements regarding the management of screen space and interactions. Additionally, the application is developed for Windows and thus follows the herein established design and interaction principles. A dark theme is chosen to minimize potential reflections of the application in the windows of the simulator during simulated night operations. The language of the application interface is English.

Shown in Figure 5.5 is an overview of the workspace layout of the developed application. To improve readability of this section only selected parts of the application are shown here. Additional screenshots of the application are shown in section C.1.

As shown in Figure 5.5 the application is divided into four areas: a header bar (**A**), the mission plan on the left (**B**), a central working area (**C**), and an overview of the aircraft systems status on the right side (**D**). The central working area allows the users to dynamically display different information, while the other areas always represent a notionally fixed set of information. This is in line with requirement IID-4-4. The overall workflow inside the application happens between (**B**), where actions to be performed are displayed, and (**C**) where information is effectively exchanged. Areas (**A**) and (**D**) are primarily designed to provide constant awareness of the current mission to the operator. Further details about the interaction concept and the relation between GoIM and the

[6]https://material.io/ - last accessed on Sunday January 12, 2020
[7]https://developer.apple.com/design/human-interface-guidelines/ - last accessed on Sunday January 12, 2020
[8]https://microsoft.com/design/fluent/ - last accessed on Sunday January 12, 2020

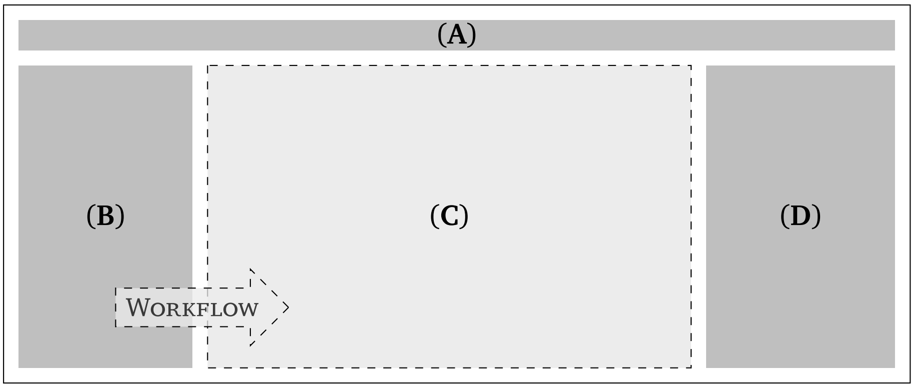


Figure 5.5.: Main screen of the developed application. The application is divided into four areas: A header area (**A**), the active mission plan (**B**), a central working area (**C**), and an overview of the aircraft system status (**D**).

implementation are provided in the following. Shown in Figure 5.6 is a screenshot of the implemented application showing the main screen with no work item being displayed. In this configuration the work area is showing a central navigation menu (1) which allows the user to open different windows inside this area. At position (2) a secondary workload task is displayed which is used for the evaluation and will be referenced in section 6.2.

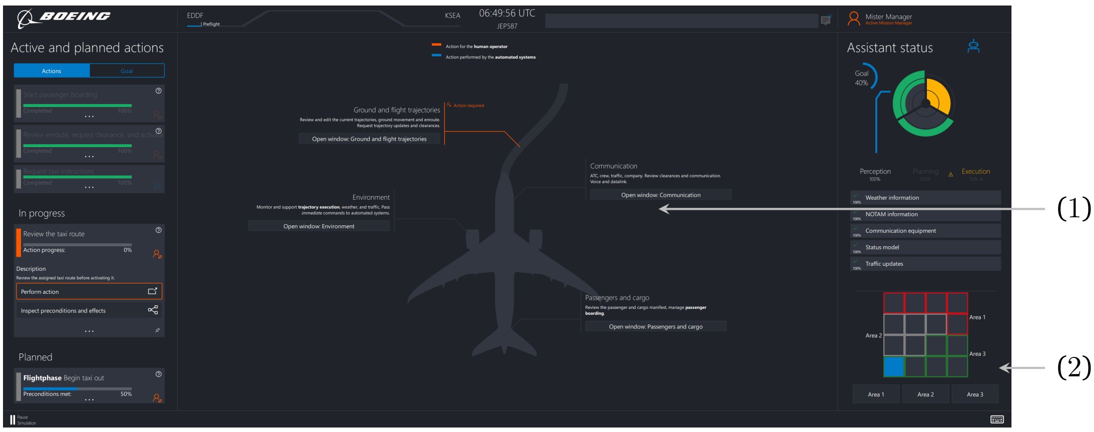


Figure 5.6.: Overview of the application, showing the configuration used for the scenario-based evaluation.

A: Header bar

Analogous to desktop and mobile applications a constantly visible header bar is implemented to provide a quick overview of the current mission state and access to commonly required information. A squeezed representation of the header bar is shown in Figure 5.7. The header bar and the therein contained components are derived from the requirements HAT-1-5, SA-2-0, SA-4-1, IID-2-5.

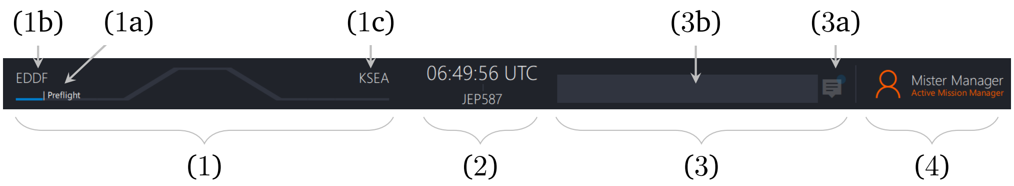


Figure 5.7.: Header bar of the application, components (1) and (3) are shown with a reduced width.

The following information components are contained in the header bar:

(1) Current mission progress, depicted as a flight profile with active flight phase (1a), including departure (1b) and destination airport (1c) with departure times once the aircraft is airborne. An indication about missing route data can be displayed here (not depicted).

(2) Current Coordinated Universal Time (UTC) time and the aircraft's callsign.

(3) Access to an extending message center allowing the human operator to review the last notifications and received messages. An indication about unread messages and unacknowledged notifications is displayed here (3a). A preview of incoming messages is displayed in field (3b).

(4) An indication of whether the human operator is currently considered to be in command of the aircraft. The human operator is assumed to be in command after the preflight procedure is completed and the aircraft is accepted.

B: Active mission plan

On the left side of the application the active mission plan is shown. Through the active mission plan two central elements of GoIM are visualized: First, the sequence of actions, the action list, which is calculated by the aircraft systems is shown here. Secondly, information about the currently active mission goal and its goal parameters is displayed and can be modified here.

Action list The action list is the visualization of the sequence of actions which are required to meet the defined mission goal. The list is the visualization of the result of the planning process performed by the aircraft systems as described in subsection 4.2.2. An example is given in Figure 5.8. Displayed in here are planned (1), currently performed (2), and already executed actions (3). Control (4) allows the user to switch between the view of planned actions and the active mission goal. Through the action list the following requirements are considered: HAT-1-3, HAT-2-0, HAT-3-0, SA-1-0, SA-1-2. All actions

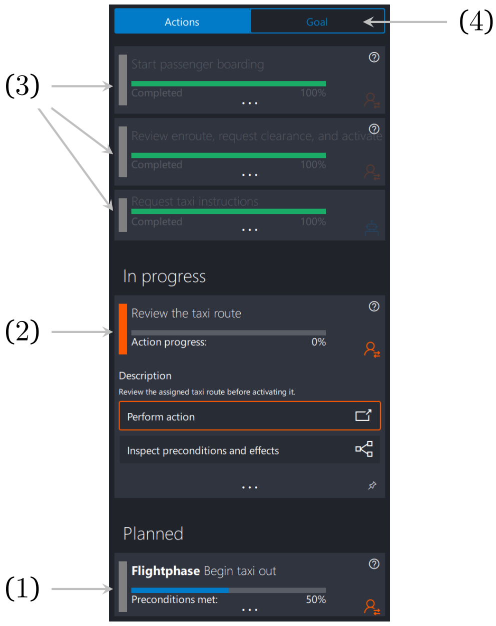


Figure 5.8.: Action list showing completed, active, and planned actions.

are represented as cards. Shown in Figure 5.9 is such a card for the action “Review the taxi route”. The title (5) of the action describes the action to be performed, in this case to review the taxi route. Ownership of an action is displayed by using iconography (6) and coloring as described in subsection 5.2.3. In this example the human operator is responsible to perform the action as described in subsection 4.2.3. With each action a brief description (7) of the work to be performed is provided. A progress bar (8) indicates the current progress of the action towards completeness. The method after which progress is measured is action-dependent. Actions are displayed in the chronological order in which they are executed: Already executed actions are displayed on top of the list, outside the normal view area. Currently being executed actions are highlighted in the center of the view area. Planned actions are placed in chronological order below that, potentially extending beyond the visible area.

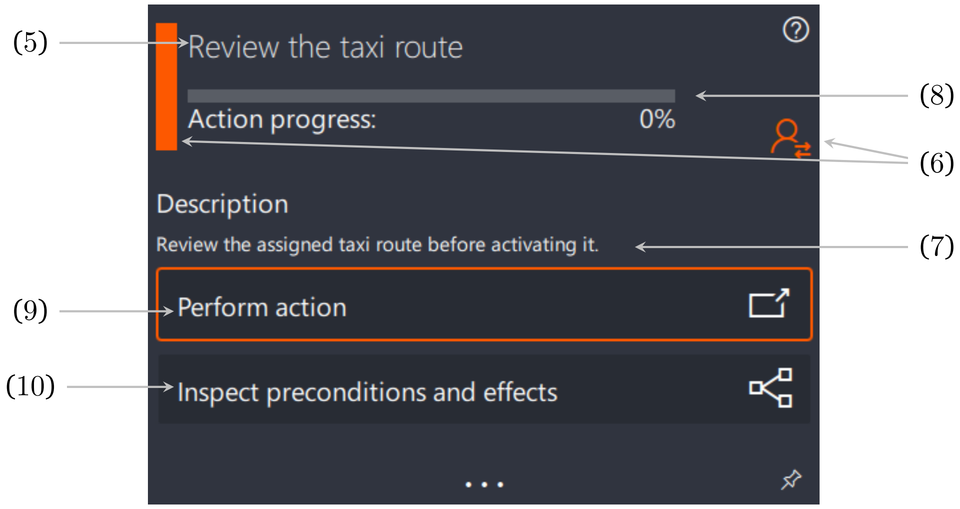


Figure 5.9.: Example of an action card for the action of “Review the taxi route”, which requires human operator input.

The action list is used to guide the operator through the mission and manage the information exchange with the aircraft systems. As such it represent a workflow-management system which prompts the human operator to perform an action when operationally required. This interaction is chosen as it allows the human operator to *work synchronized* with the aircraft systems, based on the plan derived from the defined mission goal. Second to that it also allows the human operator to follow the actions automatically performed by the aircraft systems, and if necessary to intervene. In summary, the action list view is designed to satisfy the requirements HAT-2-0 through HAT-4-0.

Once an action changes from the *planned* state (Figure 5.8, (1)) to the *in progress* state it expands to the appearance shown in Figure 5.9. Through the upper button (9) the operator can now access the details related to the action. Every action which requires some sort of operator intervention highlights this button with the color attributed to the human operator as it will be described in subsection 5.2.3. When clicking on the button a new window will be opened in the central working area. Which window will be opened is based on the action required from the human operator. Within this window the operator is then able to complete the required action, information related to the task is provided. In the chosen example the operator would be presented with a view of the route and the ability to review its parameters. Eventually, after the review is completed the action would change into the *completed* state and, if required, the next action to be performed by the human operator would be highlighted.

Actions which are automatically performed by the aircraft systems are displayed in an analogous manner. Intervention with these actions is possible in the same way as actions that are performed by the human operator. To understand *why* an action is required to meet the defined goal, or how it relates to other actions, a second view of the action list is provided. By clicking the lower button (10) depicted in Figure 5.9 the *action explorer* view is opened. In this view a visualization of the relations between actions is shown, based on the preconditions and effects of the actions. Shown in Figure 5.10 is the example for the operator-initiated action “begin taxi out”. Actions that were executed as preconditions for the selected actions are shown on the left, actions that are (in)directly enabled through the selected actions are displayed on the right.

Figure 5.10.: Action explorer view, showing the relation between actions. Based on the concept of preconditions and effects as described in section 2.3.

Through this view the human operator is supported in understanding *why* certain actions are being planned or performed and how they lead to the defined goal. This supports the requirement HAT-2-0, in which an understanding of why actions are being performed is requested. Secondly, this representation illustrates an important aspect of GoIM: the human operator is able to understand how the defined goal translates into a set of actions to be performed, with the actions themselves requiring and producing information serving the overarching goal. In conjunction with the action list this drives the overall concept of GoIM, allowing the human operator and the aircraft systems to cooperatively work towards a common goal.

Active goal Through a toggle switch the operator can switch from the action list view to the active goal view. In this view the currently active goal is summarized in a textual form, using the parameter classes described in subsection 5.2.1. The progression towards the defined goal is visualized through a progress bar. In addition to that, a button is provided through which the goal can be edited. The process of editing the active goal, and initiating the calculation of a new mission plan, is carried out in a separate window in the central working area. When editing a goal the operator is able to add or remove goal parameters. When adding goal parameters their values must be defined, for example when a spatial parameter describing an airport to land at is added, the user must specify the airport identifier. During this process supporting information is displayed to the operator, in this case for example a map of the airport to verify the correct one is selected. Once all parameters are defined to the operators satisfaction, a planning process can be manually triggered. The goal is sent to the aircraft systems which initiate the process described in subsection 5.2.1. After completion of the initial planning process, the operator can review the solution and subsequently activate the new mission plan, consisting of an action list and the defined goal. This will replace the content previously visible in the active mission plan area.

C: Central working area

The largest portion of the screen is designated to be the central working area. In this area different windows are opened through which actions can be completed and information is exchanged. Windows can be opened in two ways: through the action list, or through a centrally available menu which is visible in Figure 5.6. As previously described the preferred workflow is through interaction with the action list. However, for the human operator to be able to fulfill their task of supporting the aircraft systems, which cannot always be planned for, they require a way of independently accessing different functions. To enable this, the central menu is made available. The following windows are available:

Environment Showing a map-centric view of the environment around the aircraft. The begin of new flight phases can be triggered here. Also allows the human operator to define immediate constraints for the automation as required for the evaluation scenario described in chapter 6.

Ground and flight trajectories Allows the human operator to edit the individual trajectory elements of a mission. The mission is divided into *taxi-out*, *departure*, *enroute flight plan*, *arrival*, and *taxi-in* trajectories.

Communication Provides access to communication channels for example to ATC through Very High Frequency (VHF) radio or CPDLC.

Goal editing Through this window the currently active goal can be edited or a new goal can be defined.

Passenger management Provides an overview of the passengers onboard the aircraft, allows initiation of the passenger boarding process.

In Figure 5.11 the *Ground and flight trajectories* window is shown. Inside the different windows the workflow is always designed from left to right, accounting for requirement IID-4-5. This is a direct continuation of the overall application workflow as previously described and depicted in Figure 5.5. After the window has been opened the operator is presented with a list of trajectory elements (1), the card representations provide a first overview of their status. If an element is selected the middle part of the window shows a graphical representation of the trajectory (2), in this case a taxi route depiction with ownship position (3) (IID-3-3, IID-3-4). On the top right side a textual summary of the route data is presented (4). Actions available to the human operator are grouped in the lower right (5). The action which needs to be performed by the human operator is highlighted (cf. subsection 4.2.3 and subsection 5.2.3).

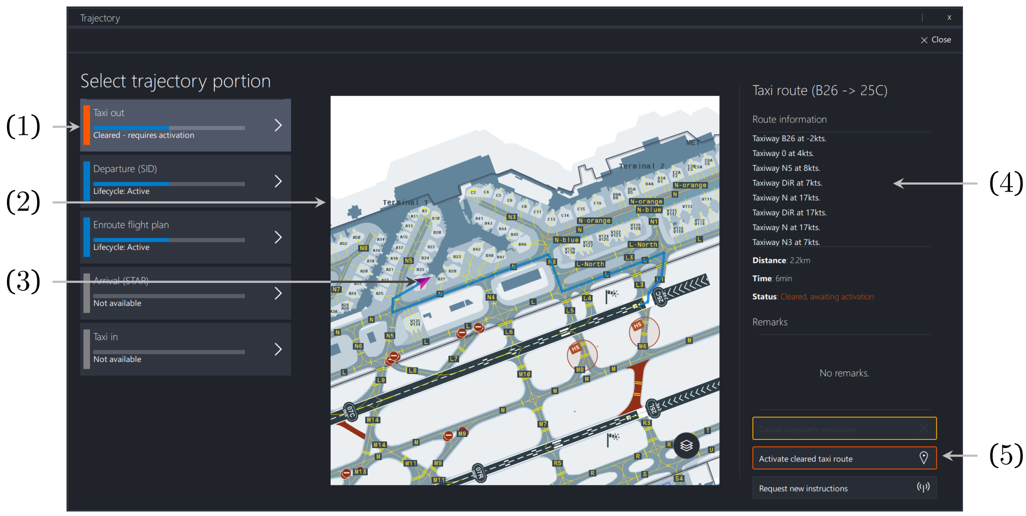


Figure 5.11.: Ground and flight trajectories window in which existing trajectories can be edited or new ones generated.

The content presented within each window is tailored for the evaluation of the GoIM concept and consequently may lack functionality required to perform a *full* mission. A more complete representation would require more sub-configurations of the provided windows.

D: Aircraft system status - the assistant

The aircraft system status representation is introduced to provide a basic overview of the current state of the aircraft systems. This overview is considered to be the primary way of informing the human operator about when they are expected to support the aircraft systems as described in subsection 4.2.3. In the application the less abstract term of *assistant* is chosen to represent the aircraft systems. The constantly visible view on the right side of the application provides an overview of the current system status, divided into the three stages of information processing as described in subsection 4.2.1. Aligned with the textual description of the three areas in Figure 5.12 a graphical representation to quickly assess the state of it is provided. A segmented pie chart (1) is used to visualize the different areas. Larger areas correlate with higher health values. Gray areas, or lines, indicate the moving average of the health value and can be used to detect trends. The chart is color coded in accordance with subsection 5.2.3.

The concentric dividers (2) represent the caution and warning thresholds. The operator is notified through visual and aural cues when a value drops below either threshold. In the provided example the value for *execution* (3) is below the caution threshold, the pie chart segment is colored yellow. The detailed view below (4) provides additional information: a trajectory segment for the next phase of flight is missing. System components and functions are grouped based on the stage they are supporting, for example trajectory completeness would be a part of the execution stage. Each subcomponent again has a health status which is dynamically updated based on changes during the mission. The weighted sum of all subcomponents for a processing stage define the overall health status of the stage which is then visualized. In this case the taxi out route is missing, as it has not been reviewed by the operator yet and therefore the aircraft systems are unable to execute it. Additional information to the operator is provided when clicking on the highlighted element (5), from where they are able to access the relevant window to support the aircraft systems. Through the assistant status view the requirements SA-3-3, SA-3-4, and SA-4-1 are considered.

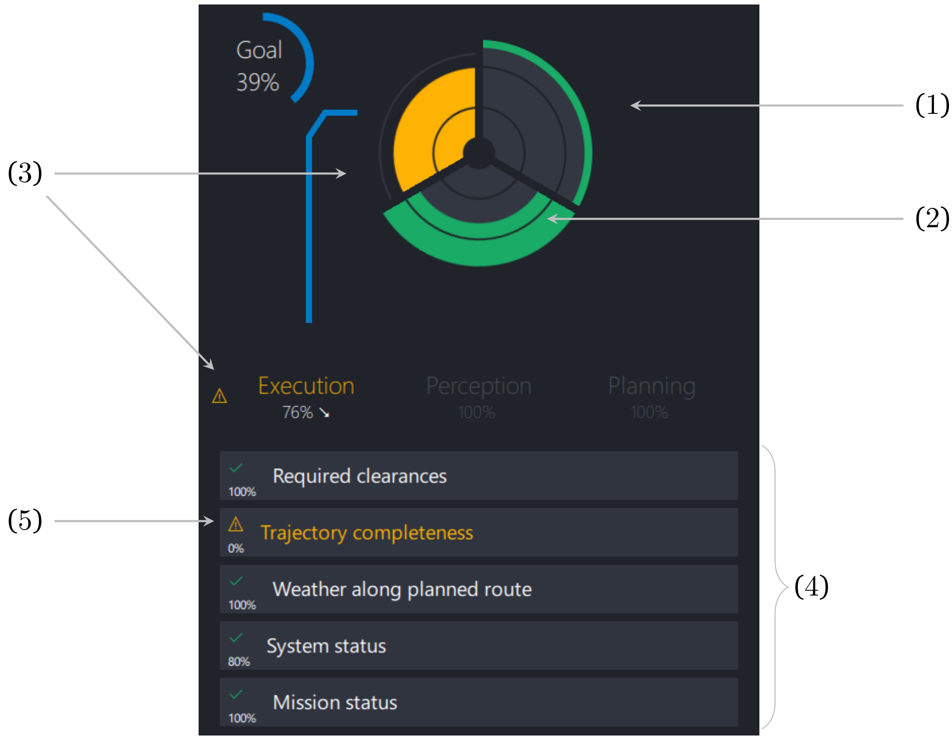


Figure 5.12.: Representation of aircraft system health status using a colored chart. Shown is a decrement in the execution stage due to a missing trajectory component. Note that not all components are required for the evaluation and are only used for filling.

5.2.3. Use of color and alerting

Inside the application, a consistent use of colors in regards to their meaning as per requirements IID-4 and IID-5 is followed. Listed in Appendix C is Table C.3 in which the application's color scheme is described. The use of colors follows the FAA's recommendations as provided in [Fed10].

Two distinct colors are chosen to highlight where either the human operator or the aircraft systems are responsible (HAT-3-0): Items and actions that require support from the human operator are highlighted through an orange color (●). A blue color tone (●) is used to highlight actions and items related to the automated aircraft system. An example of how this coloring scheme is used is present in Figure 5.6 where the upper left menu option is orange colored indicating that an action by the human operator is required here. Analogously, actions that require operator interaction are colored as shown in Figure 5.9, actions performed by the aircraft systems would be colored blue. To account for requirements IID-2-0, IID-2-3, IID-4-3, and SA-3-3, consistent alerting principles were also considered as part of the developed application. The used alerting principles follow the FAA AC 25.1322-1 "Flightcrew alerting" [Fed10].

5.2.4. Architecture

Shown in Figure 5.13 is a notional architecture diagram of the application. Interactions of the core-components are shown. The application back end is developed in C++ 17 utilizing the Qt framework version 5.13, the front end is based on the Qt Modeling Language (QML). Supplementary material is provided in Appendix C. The described architecture is primarily developed to support the GoIM concept. Object-oriented design concepts are used to allow extensibility of the application for future use cases. Objects are derived from the concept description in subsection 4.2.2.

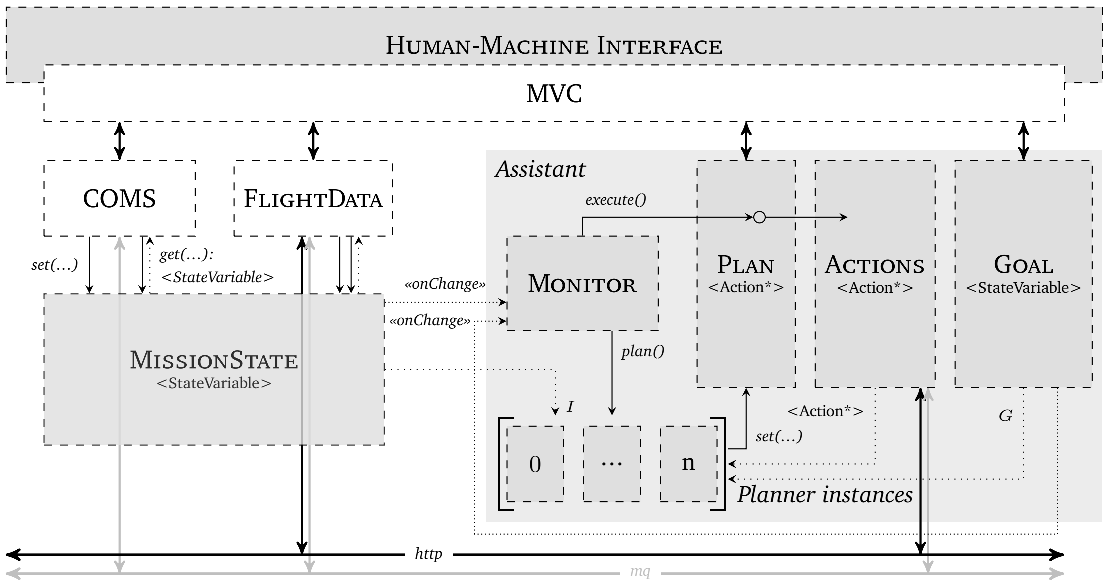


Figure 5.13.: Application architecture showing the core components, simplified view.

Shown at the bottom of Figure 5.13 are the in- and outgoing connections to the simulator as described in section 5.1. Depicted at the top is the Human-Machine Interface (HMI), which is displayed inside the simulator. Information displayed on the application front end is driven by the Model View Controller (MVC) pattern. The respective models and controllers reside within the back end components, the view is managed through custom QML front end components. Inside the application a number of components exist which transform data received from the simulation into information to be used by the application. Two exemplary components are shown in Figure 5.13: Coms and FlightData. Those components play a role in the evaluation scenario, their functionality is outlined. Both components receive and publish data through the message queue and the (http) API. Through the Coms component the communication with ATC is managed. It continuously listens to CPDLC messages on the message queue. The implemented message format follows the specification of [Int16]. All received and sent messages are stored in an internal data structure. Relations between messages are maintained by grouping them into conversations. The conversations are displayed in a dialogue view, an example of such a dialogue is shown in section C.1. Interactions with messages are managed through the view. Exemplary interactions are accepting a clearance or previewing received route data. If required, response or request messages are built using a factory pattern to produce messages following the specifications of [Int16]. The FlightData component is responsible for managing trajectory information for all flight phases (taxi, takeoff, departure, ...). It consumes the aircraft's current position data from the message queue and makes it available to map views. Sending and activating flight- or ground-trajectories is performed via the API interface. Information exchange between components is facilitated through the central MissionState component. All components inside the application have *read and write* access. In this component the current world state is represented (cf. subsection 5.2.1). The component is susceptible to race conditions through its ubiquitous accessibility from different threads. Synchronization across threads is maintained through the Qt signal-slot system[9]. The following example illustrates how different components exchange information through the MissionState component: ATC sends a CPDLC taxi out clearance containing route data to the application through the message queue. After having processed the message, the Coms component notifies the operator about a new message. The operator opens the message view and chooses to accept the taxi route, a positive response is sent to ATC. Next, the Coms component updates a state variable indicating that a cleared taxi route is available (Mission_Route_TaxiOut_availableCleared=*true*). The route data received with the message is stored in the state variable object property (cf. section C.5). The FlightData component is notified about this change through the signal slot system (not shown in Figure 5.13). The route data is displayed to the user in the model of available trajectories in the status "cleared, requiring activation". The user is

[9]https://doc.qt.io/qt-5/signalsandslots.html - last accessed on Sunday March 29, 2020.

required to explicitly activate the taxi route, following the rationale in subsection 3.3.5. The route status changes to "activated". This activation changes another state variable change inside the MissionState (Mission_Route_TaxiOut_activated=*true*).

All state variable changes are monitored by the aircraft systems, summarized under *Assistant*, through the Monitor component. On every state change the process outlined in subsection 5.2.1 is triggered. The planner is realized as its own class. If a (re-)planning is required a new instance is initialized with the current mission state I, the defined goal state G, and references to the available actions (<*Action**>). The planning process happens asynchronously in worker threads. Results of newer (i.e. later launched) planning processes supersede previous ones. This way changes to the current mission state are also considered if they occur while planning. The set of planned actions is stored in the Plan component. Execution of actions occurs as described in subsection 5.2.1. All actions are derived from a base class and must implement an *execute()* function to trigger an associated action script (cf. section C.5). Data requirements are reflected in an action's preconditions and satisfied through access to the central *MissionState* object during execution. The interaction with the user is driven through the Plan component, where planned, currently executed, and completed actions are maintained. The underlying list model serves the views described in subsubsection 5.2.2. Lastly, the user is able to edit and view the current goal state through the HMI. Changes to the goal state require a planning to be performed. In this case the planning result is not directly pushed into the Plan component's model but must first be reviewed by the operator. On acceptance the generated plan replaces the existing plan. Likewise, the defined goal replaces the previously existing one.

5.3. Summary

In this chapter the implementation of the GoIM and the mission manager concept in hard- and software are described: The development and construction of a new RCO research and demonstration simulator is outlined. The implementation of the GoIM into working software is summarized. A description of the application's visual interface and technical details are provided. References to the requirements derived in subsection 4.1.2 are provided for the respective implementations.

6. Evaluation in simulator trials

To continue the development process after DIN EN ISO 9241-210 [DIN11], an evaluation of the previously described Goal oriented Information Management (GoIM) concept is performed. In this chapter the evaluation concept is outlined, the results and subsequent discussion is subject of chapter 7. The objective of the evaluation procedure will be described, the scope is set. The research hypotheses to be tested are derived. Along with the hypotheses the methods to test the hypotheses are provided, acceptable ranges for results are defined.

A two-part evaluation setup is proposed, consisting of an application usability evaluation and a scenario evaluation. The chronological test setup and the evaluation scenario is described. Lastly, the selection criteria for test participants are discussed.

6.1. Objective

The objective of this evaluation is to determine if the the developed information management concept is *a* potential solution to be used on the future flight deck. This reflects the central research objective as stated in section 1.2. The *usability* and the acceptance of the developed *concept* is playing a central role. The selected definition of usability follows DIN EN ISO 9241-11 [DIN18] (highlights by author):

> Usability is the extent to which a system, product or service can be used by *specified users* to achieve *specified goals* with **effectiveness**, **efficiency** and **satisfaction** in a *specified context* of use.

The highlighted terms are further defined in DIN EN ISO 9241-11 [DIN18]:

- Effectiveness is the accuracy and completeness with which the users achieve specified goals.
- Efficiency is the resources used in relation to the results achieved.
- Satisfaction is a person's perceptions and responses that result from use of a system, product or service.

The relevancy of the three different variables in the context of this evaluation is elaborated in subsection 6.1.2. The evaluation of GoIM also poses a challenge that has to be recognized: GoIM is an abstract concept for how information between a human operator and an automated system could be managed in the future. As such, it cannot be tested for usability by a human operator. A suitable interface to make the concept tangible is required. For this matter an exemplary implementation of it was previously performed and described in section 5.2. While this application is neither tied exclusively to GoIM nor is GoIM tied to this application, it will still influence the evaluation. To better understand the impact of the application on the usability of GoIM it must be evaluated for usability by itself.

6.1.1. Scope and evaluation method

The operational scope in which the evaluation will take place is limited to the concept of operations as described in section 3.3. A task-based application usability evaluation and a scenario-based evaluation is chosen to evaluate the usability of GoIM. The experiments are performed in the purpose built Reduced Crew Operation (RCO) research simulator (cf. section 5.1). This evaluation modality is chosen as it is considered a both "useful and efficient" [Abb14] evaluation method for understanding human-machine interaction.

Out of scope for this evaluation is a direct comparison of today's flight deck information management with the herein described concept. To honor the human-centric approach, it was decided that the usability of the concept itself should be evaluated first. A direct comparison is not seen beneficial at this stage, especially under consideration of the extensive changes proposed by GoIM. With the insights gained during the evaluation the concept can then be further matured and a direct comparison can be beneficial.

6.1.2. Hypotheses and testing methods

The evaluation of GoIM is performed by testing three global hypotheses. Two hypotheses are aligned with the concept of usability as previously described. While the first hypothesis is focused on the usability of the application itself, the second hypothesis uses the concept of usability and applies it to the concept of GoIM itself. A third hypothesis is formulated in which the acceptance of GoIM from a user perspective is evaluated. In the following sections the hypotheses are described in detail. With each hypothesis the means of assessing it are described, the practical implementation of the methods in the evaluation setup is then outlined in section 6.2.

Hypothesis I - Usability of the application

Through this hypothesis the usability of the application itself is tested. As described in section 6.1 this is tested to determine the impact of the exemplary implementation on the concept itself. Usability is tested divided into the three components of effectiveness, efficiency, and user satisfaction. The following global hypothesis is formulated:

> **Hypothesis 1:** When using the developed application test participants will show acceptable usability levels which are defined through independent measurements of effectiveness, efficiency, and user satisfaction. Acceptable usability is assumed when effectiveness, efficiency, and user satisfaction are each within acceptable bounds.

Based on this hypothesis the following two-tailed null hypotheses are formulated:

- $H_{0,1}$ When measuring effectiveness during the usability evaluation of the application through combined measures of task completeness and time to complete a task, both expressed through the value λ_I, there is no significant difference from the acceptable level of $\lambda_I = 0.7$.
- $H_{0,2}$ When measuring efficiency through the subjectively reported Modified Cooper-Harper score there is no significant difference from the least acceptable score of 5.
- $H_{0,3}$ When measuring user satisfaction during the usability evaluation of the application through the subjectively reported System Usability Scale there is no significant difference from an acceptable score of 52.

The global hypothesis is evaluated during the first part of the evaluation procedure, the application usability evaluation. During this type of evaluation the users are instructed to perform a series of tasks with the developed application. Their interaction is monitored by the test administrator, errors and correctly performed actions are recorded. According to Cooper et al. [Coo+14], this type of evaluation is best suited for in-development evaluation of user interfaces to get user input on usability. Based on the feedback which is received as part of this evaluation, the usability of the application is determined as described in the following paragraphs.

For the three components of usability *acceptable bounds* and means of assessing them are described. The global hypothesis is rejected if one or more of the components are tested to not be within their respective acceptable bounds.

Effectiveness Effectiveness considers the degree to which a task is completed, as well as the accuracy of completing it. Completeness describes the degree to which user is able to successfully finish a task, including the partial completion of a task. Accuracy is determined by the error- and misstep rate of users while completing the task [Coo+14], as well as the time to completion.

Effectiveness is expressed as the Greek symbol $\lambda\{\lambda \in \mathbb{R}|0 \leq \lambda \leq 1\}$. The initial value of $\lambda_{I,initial}$ is obtained based on the number of correctly completed sub tasks in a given task. Tasks can relate to correctly identifying some information or performing a specific action in the application, e.g. deleting a goal or identifying the current destination airport. A single task may consist of one to many sub tasks, but the overall achievable score will not exceed 1. Partial or erroneous completion of a task results in a degraded completion score down to a minimum score of 0. If a participant indicates that he or she is unable to complete a given task, only the completed sub tasks up to that point are credited.

The time it took the participants to complete a task (t_{task}) is then used to calculate the final score for effectiveness. A correction factor b is calculated as shown in Equation 6.1. Factor b is calculated by putting t_{task} in relation with t_{ref}, the average time it takes an experienced user to complete the task. Time to complete the task is defined by the time span it took the user from finish reading and understanding the task instructions until indicating completeness. Following the recommendation of [Gaw08] start and end of the task are well-defined through the explicit button interactions during the experiment.

$$b = \begin{cases} 1 & \text{if } t_{task} < q \cdot t_{ref} \\ 0 & \text{if } t_{task} > 2q \cdot t_{ref} \\ \frac{2q \cdot t_{ref} - t_{task}}{q \cdot t_{ref}} & \end{cases} \tag{6.1}$$

To account for the fact that users are using the application for the first time during the experiments a multiplication factor of $q = 2$ is applied. Users are given twice the time to complete a task without receiving a score penalty through b. The reference time t_{ref} is calculated based on a series of measurements taken from experienced users and their individual time to complete a task t_{exp}. Experienced users are users that previously completed the experiment, where then specifically trained to use the application, and eventually redid the experiment. The reference time t_{ref} is calculated following Equation 6.2:

$$t_{ref} = \left(\frac{1}{N} \sum_{n=0}^{N} t_{exp} \right) \tag{6.2}$$

Two experienced users were used to provide the reference measurements, each user repeated the procedure three times ($N = 6$). To not skew the reference times, the experienced users were asked to not *rush* through task and avoid a purely pattern-based execution. The final effectiveness score λ_I is then calculated by multiplying the initial score $\lambda_{I,initial}$ with the correction factor b as shown in Equation 6.3:

$$\lambda_I = \lambda_{I,initial} \cdot b \tag{6.3}$$

Acceptable usability in respect to effectiveness is assumed when a final score of $\lambda_I \geq 0.7$ is reached by the participants. This value is chosen as it equals the required passing score defined by the Federal Aviation Administration (FAA) for tests like the Airline Transport Pilot, Aircraft Dispatcher, or Flight Navigator exams [Fed19].

Efficiency Efficiency is defined as the expended human effort while completing a task [DIN18]. In this context the mental workload of the human operator is used to define the expended human effort. For testing this hypothesis self-reported measurements by the participants are used. The rating technique chosen for this use case is the Modified Cooper-Harper (MCH) scale, as introduced in subsubsection 2.2.1.

During the evaluation the participants are asked four times to submit a MCH score. Ratings are requested from the participants after having completed a category of tasks. The acceptance criteria for efficiency is an overall MCH score of ≤ 5.

User satisfaction Lastly, the user satisfaction is assessed. A common method for measuring the subjective user satisfaction when evaluating usability is the System Usability Scale (SUS). The SUS was described by Brooke, and has since evolved into a *de facto* industry standard [Bro13] for usability evaluations [Bro96]. In that regard SUS is considered to be suitable for providing a measurement of a user's satisfaction while using a system, while only consuming little time during an evaluation session [Bro13].

The SUS questionnaire consists of 10 different statements, with alternating positive and negative expressions, about the interaction with the system. Users have to indicate their agreement with the statements on a five point Likert scale. Scoring of the SUS responses follows a formula described by Brooke in the 1996 initial publication [Bro96]. A sample of the utilized SUS rating scale is shown in Appendix D.

Scores obtained through SUS are in a range between 0-100, which does not indicate percentage [Bro13]. Bangor, Kortum, and Miller showed that SUS is applicable over a wide system range and provides an adjective rating scale, allowing a statement about whether an interface is usable or not [BKM08]. An adjective rating scale as described in [BKM08] is shown in Table 6.1. As indicated in Table 6.1 the lower bound for an acceptable user satisfaction is a score of less than 52. It is suggested that a score of 68 is

Table 6.1.: Adjective rating scale for SUS showing score ranges, adjective ratings, and the larger acceptability ranges, as described in [BKM08].

Score	Adjective rating	Acceptability range
<25	Worst imaginable	Not acceptable
25-39	Poor	
39-52	Ok	Marginal
52-73	Good	Acceptable
73-86	Excellent	
>86	Best imaginable	

around the 50th percentile [Bro13]. This coincides with the findings of [BKM08] where a score of 70 is suggested as acceptable.

Hypothesis II - Concept usability in an operational context

While the previous hypothesis is used to test the usability of the exemplary application this hypothesis is used to test the usability of GoIM itself. Testing is performed in an operationally-relevant scenario. Similar to the first hypothesis the term of usability is used to drive the hypothesis formulation:

> **Hypothesis 2:** When using GoIM in an operationally-relevant context the participants will show acceptable usability levels that are defined through independent measurements of effectiveness, efficiency, and user satisfaction. Acceptable usability is assumed when effectiveness, efficiency, and user satisfaction are each within acceptable bounds.

Three two-tailed null hypotheses are formulated:

$H_{0,4}$ When measuring effectiveness during the concept evaluation through combined measures of the user's ability to activate the Very High Frequency (VHF) radio and the position of sending the slow down command, both expressed through the value λ_{II}, there is no significant difference from the acceptable level of $\lambda_{II}=0.7$.

$H_{0,5}$ When measuring efficiency through the subjectively reported Modified Cooper-Harper score there is no significant difference from the least acceptable score of 5.

$H_{0,6}$ When measuring user satisfaction during the usability evaluation of the concept through the subjectively reported System Usability Scale there is no significant difference from an acceptable score of 52.

The operationally-relevant context in which this hypothesis is tested is derived based on the analysis performed in chapter 3. The user will be tasked to work through a scenario which is typically associated with high workload, in this case taxiing in adverse weather conditions under time pressure on a large airport. In this scenario the automated systems will report a condition which requires the user to provide context as described in section 4.2. Through the design of the scenario, which will be subject of section 6.2, the usability components of effectiveness, efficiency, and user satisfaction are evaluated. The acceptable ranges for the three different components of usability are elaborated in the following paragraphs.

Effectiveness Effectiveness is measured through the completion of two tasks during the scenario evaluation. Both tasks are connected to the overarching task of the user to provide context to the automated systems. The tasks act as proxies for potential actions a human operator might be required to perform in situations as described in section 3.3 and subsection 4.2.2. In contrast to the effectiveness measure described for the first hypothesis, the tasks to be performed by the participants are not explicitly provided to them. Instead, the participants are required to detect when an action is required.

The setting in which the tasks occur is described in section 6.2, both tasks are thus only described as far as necessary for the general understanding here: First, the human operator will receive a notification about a Controller-Pilot Data Link Communications (CPDLC) outage. As a reaction, and as suggested by the aircraft systems, the operator is then expected to activate the VHF radio and monitor the audio channel as the scenario progresses. At a later point in time a message will be transmitted via VHF radio stating that the taxi speed should be reduced at a certain taxiway section which is reached later. The human operator is now required to understand what to do and when to act.

Similar to the first hypothesis, effectiveness is measured through task completeness and quality, expressed through the variable λ. In both cases completeness is considered to be binary, no partial completion is assumed. Task quality is measured by the spatial position where the action was triggered. For the first task (switching on VHF radio) the user receives a higher score $\lambda_{II,VHF}$ the earlier it is switched on. For the second task (commanding the automation to reduce speed) a higher score $\lambda_{II,SLOW}$ is attributed the closer the command is issued in relation to the position requested via radio. The score for $\lambda_{II,VHF}$ is calculated by a linear function between the point where the CPDLC outage occurs ($\lambda_{II,VHF} = 1$) and the point where the speed reduction is announced

($\lambda_{II,VHF} = 0$). Similarly, $\lambda_{II,SLOW}$ is calculated as a trapezoid function, increasing from $\lambda_{II,SLOW} = 0$ to a plateau with a value of $\lambda_{II,SLOW} = 1$ around the point where the speed reduction is planned. After that it decreases down to $\lambda_{II,SLOW} = 0$ again.

Both components are individually calculated, the final score λ_{II} is then calculated following Equation 6.4. Analogous to the acceptance criteria defined for the first hypothesis a score of $\lambda_{II} \geq 0.7$ is considered acceptable.

$$\lambda_{II} = \frac{\lambda_{II,VHF} + \lambda_{II,SLOW}}{2} \tag{6.4}$$

Efficiency Two measurements of efficiency are performed. In both measurements the mental workload of the participants is assessed. The MCH rating scale is again utilized to obtain a subjective rating from the participants after having completed the scenario. This measurement is deemed to be the decisive measurement in determining if the null hypothesis can be rejected. An acceptable value for efficiency is assumed for a MCH score of ≤ 5.

A secondary task workload assessment is used to qualitatively describe the experienced workload levels throughout the scenario. The secondary task method as described in subsubsection 2.2.1 is utilized. A secondary task is realized through a separated area on the main application screen where a setup compared to the one presented in Figure 6.1 is visible. Participants are instructed to monitor the matrix and react once one of the squares lights up. Each of the squares can light up at random intervals. Every time a square lights up, the participants must first detect this, then identify the area in which the square lit up and eventually press the related button.

A monitoring task is preferable in conjunction with automated systems [Gaw08], both the primary and the secondary task require the human operator to monitor a system. The average frequency for squares lighting up is 0.5 Hz (min. after 1 s, max. after 3 s). The value was determined empirically during pre-evaluation trials. The time after which an event is considered to be missed is t_{max}=5 s. During the evaluation are the number of correct identifications, erroneous identifications, misses, and the response time between highlighting and reaction of the user (t_{react}) are logged.

The responses are scored as shown in Equation 6.5. No difference between wrong actions and missed events is made, correct reactions are scored individually as a function of reaction time t_{react}.

$$\text{score} = \begin{cases} 0 & \text{if } t_{react} > t_{max} \\ 0 & \text{wrong answer} \\ 1 - \frac{t_{react}}{t_{max}} & \end{cases} \tag{6.5}$$

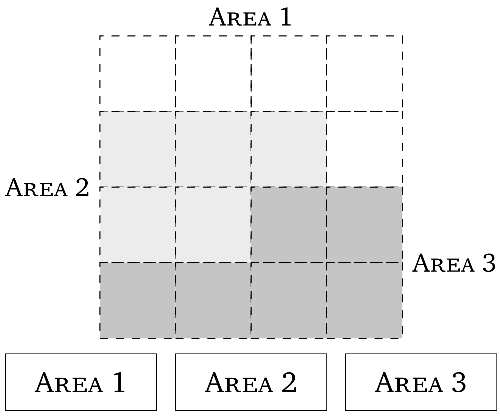


Figure 6.1.: Implementation of the secondary task method for measuring workload. The 4x4 array of squares is divided into three areas. Squares are randomly lighting up. Once a square lights up the user is instructed to push the button of the corresponding area.

The scenario (cf. section 6.2) during which the secondary task is performed is divided into phases. For each phase the secondary task performance is separately assessed. Phases are changed based on trigger actions or events occurring during the scenario. Five logical phases are defined, which are summarized in Table 6.2. The phases are introduced to allow a normalization of the obtained values to a time and location independent measurement to be comparable between all participants.

User satisfaction User satisfaction is again determined through the use of the SUS. In contrast to the previous utilization of the SUS the user is now instructed to think about how they interacted with the automated systems while performing the scenario and less about the interface they were using. The levels for considering the user satisfaction component to be acceptable are unchanged from the first hypothesis.

Hypothesis III - Concept acceptance

As stated in section 1.2 one of the objectives of this work is to put the human operator in the center of the development process, or in other words to consider the human first. It must be assumed that a certain level of acceptance for the GoIM concept exists for it to be a potential information management solution for the future flight deck. Consequently, this user acceptance is focused on in the third global hypothesis. The hypothesis builds

Table 6.2.: Phases into which the evaluation scenario is divided for evaluation of the secondary task workload. Shown are the phase names and the action or event required to trigger that phase. Trigger events relate to the scenario evaluation described in section 6.2.

Phase	Trigger
Preflight	Start of scenario.
Normal taxi	User started pushback.
CPDLC out	Notification about CPDLC outage received.
Event announced	Notification about debris on the taxiway received.
Slowed down	User initiated the slow down of the aircraft.

upon the scenario evaluation performed for the second global hypothesis but focusses more on the individual user's experience with GoIM:

> **Hypothesis 3:** The subjective acceptance rating for GoIM will be in an acceptable range of greater or equal to 3 points on a five point Likert scale.

A single two-tailed null hypothesis is formulated:

$H_{0,7}$ When asked about their experience of working under GoIM the participants will not express overall ratings different from 3 on a five point Likert scale.

To quantify the subjective rating of the GoIM concept a definition and a method of assessing it is required.

As part of a larger questionnaire which accompanies the evaluation a set of statements is created which specifically asks the users about their subjective experience of GoIM. The relevant statements are Q5.12 through Q5.20 as listed in section D.2. For all statements a five point Likert scale is chosen. All statements are formulated positively, e.g. "I always felt in control of the situation.". Participants are able to submit ratings ranging from "Strongly disagree" to "Strongly agree", for which integer point values are assigned from 1 to 5, respectively. A total of nine statements are asked of which the last statement summarizes the previous eight: "Overall, I am willing to cooperate with the automated systems in the demonstrated way.".

The first eight statements touch on different requirements as stated in section 4.1. In particular, the following requirements are implicitly referenced in the statements: HAT-1-0 through HAT-4-0, SA-2-0, SA-3-0, IID-1-0, IID-2-0, O-1-0. For the evaluation all responses are equally weighted. An acceptable user acceptance is assumed for an overall subjectively indicated acceptance score of ≥ 3 on the five point scale. The respective portion of the questionnaire is filled out by the participants after having completed the scenario evaluation and marks the last step in the evaluation as described in section 6.2.

Other aspects to be assessed during the evaluation

In addition to testing the three previously described hypotheses the questionnaire created for the third hypothesis is used to also capture additional insights to be used during the evaluation of the proposed information management concept. The additionally collected data can then be used to support the discussion of the obtained results and provide starting points for future research. The following topics are part of the questionnaire:

Demographics General information about the participants are collected, consisting of: age, gender, proficiency in general and aviation-related English, technology-affinity and experience with automated systems, and the operating systems of their mobile phone and computer. Additionally, participants are asked if they have a professional background in aviation (participants were not required to have piloting experience, cf. section 6.3).

Mission manager concept As this work is the first work in which the in section 3.3 presented mission manager concept is part of a formal evaluation the participants are asked about their opinion of the concept. Participants are asked about how much they agree with the motivation and intent behind the mission manager concept. In addition, the participants are asked to provide their opinion on how challenging or different to a today's pilot they see the mission manager's job and whether they would work as one.

Automated systems Participants are asked about their experience with automated systems, whether they trust them, are able to follow their intents and if they think automation will increase on the flight deck.

Together with the part of the questionnaire used to test the third hypothesis the herein described questions are part of section D.2. Free text answer possibilities are provided alongside with the Likert scale type answers. In section 6.2 it is described when the questionnaire is filled out by the participants.

6.2. Concept and setup

The evaluation sequence is split into two parts. In Figure 6.2 a graphical representation of the procedure is shown. During the first part the application usability evaluation is performed. In the second part the evaluation of GoIM in an operationally-relevant scenario is performed. In this section the evaluation procedure is described in the same chronological order as shown in Figure 6.2.

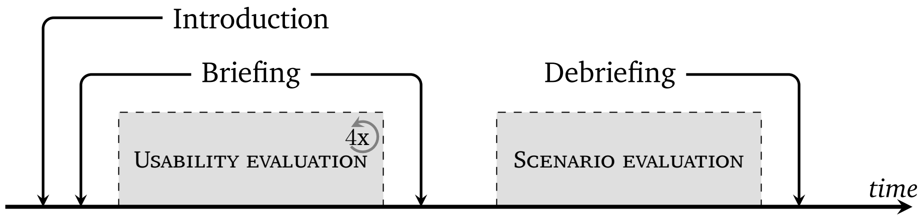


Figure 6.2.: Graphical representation of the procedure followed during the simulator trials. Evaluation runs in the simulator are shown as gray blocks. The usability evaluation is repeated for each group of tasks.

The test procedure is designed for a single participant at a time. However, it is possible to have two participants perform the evaluation session in parallel. In these cases the briefing and debriefing of the participants is performed as a group-exercise, simulation sessions are then performed apart from each other. Communication between participants is limited to before and after the experiment sessions.

Prior to each evaluation session the participants are provided with consent forms and information about data privacy. Information about the participants rights to abort or pause the experimental procedure at any point in time are provided. Briefing material and questionnaires are provided exclusively in English. Data protection consent forms are provided in either English or German. Questions and briefings are answered and held in either English or German. The total duration of a full evaluation session ranges between 2 h and 3 h for each participant, depending on whether only a single participant participates in the procedure or a parallel evaluation is performed. All experiments are performed in rooms of the Institute of Flight Systems and Automatic Control (FSR) at the Technische Universität Darmstadt (TUDA), namely a briefing room and the RCO research simulator. The described test procedure was tested in pre-evaluation test runs during which the participants feedback regarding the overall test setup was gathered and incorporated, resulting in the herein provided description.

6.2.1. Introduction

The evaluation procedure is started by briefing the participants about the motivation of the herein evaluated work and their contribution to the evaluation. A definition of usability and its relation to the evaluation procedure is given as described in section 6.1. Participants are informed that during the first part of the evaluation the usability of the application is tested, whereas during the second part of the evaluation the usability of GoIM as a concept is focused on.

After a general introduction to the experiment procedure and the work at hand, a more specific briefing about the mission manager concept is performed. Based on the concept of operations described in section 3.3 the participants are made aware that they are expected to be in the role of a mission manager during the experiment. This step is chosen as participants might initially expect to work as a today's pilot, this way a common mindset for the job to be performed is created. The following points are emphasized:

- The human operator is responsible for the safety of operations.
- The human operator will define high-level strategic goals which are translated into tactical actions by the automated systems.
- The automated systems will perform all tactical tasks during the flight mission, the human operator will provide context when necessary.

Questions are allowed throughout the whole introduction. During and after the introduction the participants are asked to complete the first two sections of the questionnaire as described in subsubsection 6.1.2. A copy of the utilized questionnaire is shown in section D.2. The questionnaire is handed out as a whole during the initial briefing. Inside the questionnaire well visible *STOP* marks are placed, participants are instructed to not proceed past them unless being told so.

6.2.2. Application usability evaluation

Next, the participants are briefed about the application usability evaluation. This includes an explanation of the performed measurements and how to indicate completeness of a task. The MCH rating scale, as shown in section D.1, is introduced to the participants. Time to familiarize with it is provided, questions are answered. The application usability evaluation itself is performed entirely in the RCO research simulator, using the developed application as described in chapter 5. An initial training is provided to the participants before starting the evaluation procedure.

The application usability evaluation is performed to test the first global hypothesis. Additionally, a training effect when using the application is intended, which will allow the participants to better understand and use the application afterwards. A reduced impact of the application on the subsequent evaluation of GoIM is preferable when trying to mitigate the challenge described in section 6.1.

Application training A standardized training is provided in which participants are taught the basics of the application. To maintain comparable results the entire training is captured in a short video introduction (77 s) and two presentation slides that are only read by the participants without talking by the test administrator. The participants are requested not to ask any further questions after the training started. Aspects highlighted during the training are the general application anatomy, trajectory management nomenclature, and the assistant as described in section 5.2. Eventually, the participants are made aware of the constraints of the utilized application:

- Experimental state of the application, reduced set of functions
- Utilization of operations shortcuts to reduce simulator time
- Simplified terminology to allow for a broader range of participants

Afterwards, the participants are allowed to familiarize themselves with the application in the simulator. The familiarization is performed as a self-paced learning and exploration exercise in which the participant is given 10 min to freely use the application. During the familiarization phase the application was put into a default state with no goals activated, the aircraft is positioned at gate B10 at Frankfurt Airport (EDDF).

Usability tasks After the familiarization, a reset of the simulation environment is performed. For the entire duration of the usability evaluation the aircraft is parked at gate B26 at EDDF. During the evaluation itself the participants are asked to perform a series of tasks with the application. Tasks are grouped in to four logical categories: ASSISTANT AND AUTOMATED SYSTEMS: Working with the assistant, recognizing state changes, supporting the assistant. GOAL DEFINITION AND PLANNING: Defining goals, performing planning, activating planned solutions. ACTION EXECUTION: Understanding the sequencing and ownership of actions, performing actions. MISSION EXECUTION: General tasks that support mission execution and trajectory management.

In total 20 tasks are completed by the participants. An overview of the tasks and the intended solution paths is provided in Appendix D, Table D.1. Some of the tasks resemble actions that the participants are required to perform during the scenario evaluation afterwards. In each group 4-6 tasks are to be completed. Possible tasks are to perform an action, retrieve information, or a combination hereof. The order of tasks within a group is constant. The order of groups is varied between participants, based on a fixed permutation which is seeded by the participant's identification number. Within each group the chronological order of tasks is fixed as some tasks require the previous task to be performed. During the evaluation procedure participants are guided by the application as shown in Figure 6.3. For every task the participants first receive the instructions in a full screen overlay. Initially, the participants are requested to select a new task (1). The task instructions then get displayed (2), in this example task number 1 from the *Assistant* category is displayed. By clicking on button (3) the overlay disappears, the application becomes visible and time recording starts.

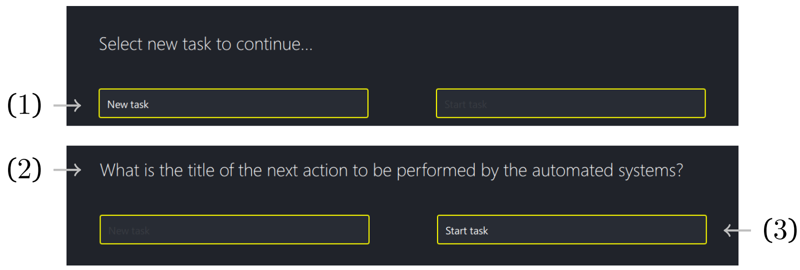


Figure 6.3.: Task instructions as displayed during the usability evaluation.

While performing the task the instructions are constantly displayed in the upper left corner of the application (4) as shown in Figure 6.4. In this example the task would be for the participants to identify the next task in the action list which is allocated to the automated systems “Request taxi instructions” (5) and voice it out to the test administrator. After having done that they must press and hold the task description (4). The participants are made aware of this workflow during the introduction.

The test administrator monitors the participants while performing the assigned tasks. Observations, as well as correctly and incorrectly performed tasks are recorded to calculate $\lambda_{I,initial}$. Participants indicate the completion of a task through interacting with the task description in the upper left corner, the time recording is then automatically stopped. If a participant indicates that they are unable to complete a given task, the test administrator finishes the task and indicates this in the records. The test administrator provides assistance if participants have trouble understanding the task instructions before starting the task.

After having completed a group of tasks an overlay is displayed in which the participants are asked to enter a MCH rating score. The participants are allowed to consult a printed handout of the scale for reference. After completing all tasks and having submitted four MCH ratings the participants return to the briefing room. Upon returning to the briefing room the next part of the questionnaire is filled out, including the SUS. The participants are now allowed to ask questions about the application. The participants are made aware of errors they made while completing the tasks, guidance on how the intended way would have been is given. A summary of the application is given to form a level set understanding

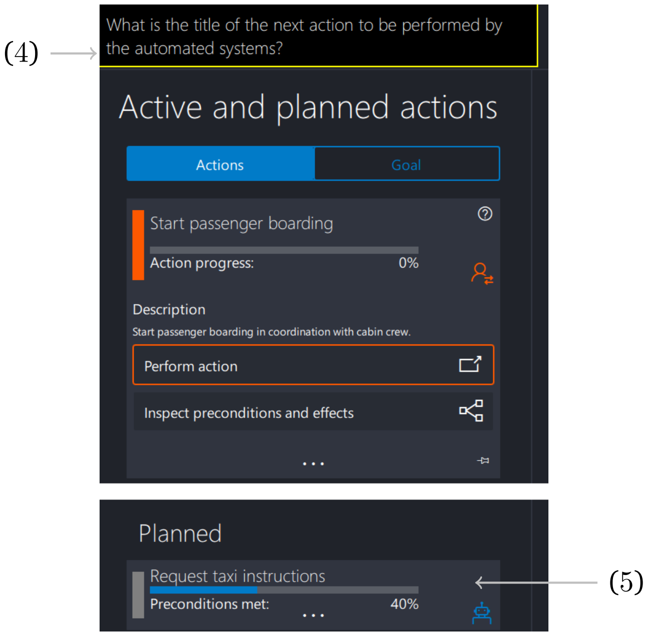


Figure 6.4.: Task instructions as displayed during the usability evaluation. Action list shortened for presentation purposes.

of the application as the interaction with it will no longer be the focus of the second part of the experiment.

6.2.3. Scenario evaluation

The second part of the evaluation is performed as an operationally-relevant scenario which the participants have to complete. The participant's tasks are based on the description of the human operator's tasks as given in section 3.3 and subsection 4.2.2.

Scenario The operational scenario presented to the participants is derived from the researched use cases described in subsection 3.1.3. Maintaining a balance between operational relevancy and realism, high workload, and the skills of the selected participant group, the following scenario is used:

The mission starts in Boston Logan Airport (KBOS) where the aircraft is parked at gate E11 in the late afternoon. The destination for today's flight is Los Angeles International Airport (KLAX). The flight number is JEP587. Preflight is already completed, a flight plan is available, and the passengers are ready to board. Due to deteriorating weather it is expected that the airport might close down soon and it is advised to taxi out and takeoff as soon as practicable, introducing additional time pressure.

The scenario takes place between gate E11 and ends shortly before reaching runway 27. A graphical depiction of the areas relevant to the scenario is provided in Figure 6.5. Larger versions of the screenshots shown in Figure 6.5 are provided in section C.1. The initial parking position of the aircraft at gate E11 is shown at position (0) in Figure 6.5. The scenario starts with the participants defining a mission goal with a parameter of the type "spatial" to reach the destination airport KLAX. A solution is calculated and activated by the participants. The aircraft is then prepared to taxi out. This includes boarding the passengers, requesting a flight plan clearance, activating the flight plan, and eventually, to receive a taxi clearance, activate the route and begin taxi out.

During this part of the scenario Air Traffic Control (ATC) is only contacted via CPDLC, clearances are automatically sent to the application after the participants request them. Upon starting the taxi out phase the aircraft automatically pushes back and begins following the highlighted route in Figure 6.5. Taxiing is performed fully automated, no direct influence of the taxi path is possible. The only exception is a button in the *environment* window which, when activated, causes the aircraft to slow down while following the path.

Upon reaching position (1) the first event is triggered: The assistant will report that CPDLC communication was lost, a health decrease in the perception stage is announced. The participants are required to diagnose the problem by inspecting the detailed information provided through the assistant as described in section 5.2. The recommendation to activate the VHF radio is provided to compensate the CPDLC failure as shown in Figure 6.5. This action is completed by opening the *communication* window either through the central menu or through a button provided next to the recommendation. The distance the aircraft traveled after having received the notification on the communication failure up until activating the VHF radio is reflected in $\lambda_{II,VHF}$ as described in subsection 6.1.2. After having activated the VHF radio an audio message is played that CPDLC failed but all issued clearances remain valid. From this point on generic ATC chatter is being played, referencing different non-visible aircraft on the airport surface. Parallel to that, task is inserted into the action list instructing the pilot to "monitor ground operations", which is shown under (2) in Figure 6.5. Through interaction with the task the *environment* window is opened. Messages intended specifically for the own flight number JEP587 are runway crossing clearances when approaching a runway and future events.

Continuing along the taxi route, the aircraft will then reach position (2) where the next event is triggered: An ATC message is played in which aircraft are advised to slow down on taxiway C between runway 4L and 4R due to reports about debris on the taxiway. The participants are expected to recognize that their assigned taxi route will continue along this taxiway which requires them to slow down when reaching runway 4L. The aircraft is then expected to slow down as close as possible to position (3). Slowing down is performed through the *environment* window, the application screenshot in Figure 6.5 is showing the application after having slowed down. No points are awarded for slowing down prior to position 2 or past position (3a). While slowing down prior to (2) is considered too early and without any reason to do so, (3a) is considered too late and therefore unsafe. Based on the proximity to position (3) $\lambda_{II,SLOW}$ is calculated as described in subsection 6.1.2.

The scenario ends when reaching position (4). Throughout the whole scenario the participants are asked to continuously perform the secondary workload task. The test administrator is present during the whole scenario to supervise the execution.

Briefing A briefing prior to performing the scenario evaluation is conducted. As part of this the briefing package described in section D.4 is handed out to the participants. The mission context is briefly summarized by the test administrator. Participants are advised that there will be a problem during the scenario. The participants are instructed that instructions provided by ATC must be followed. They are also made aware that reading back of ATC received instructions is not necessary and that no steering of the aircraft is required. An introduction to the secondary workload task as described in subsection 6.1.2 is given. The participants are advised to perform the task whenever possible. Participants are reminded that the focus of the scenario evaluation is on evaluating the GoIM *concept*. After the briefing the participants can ask questions.

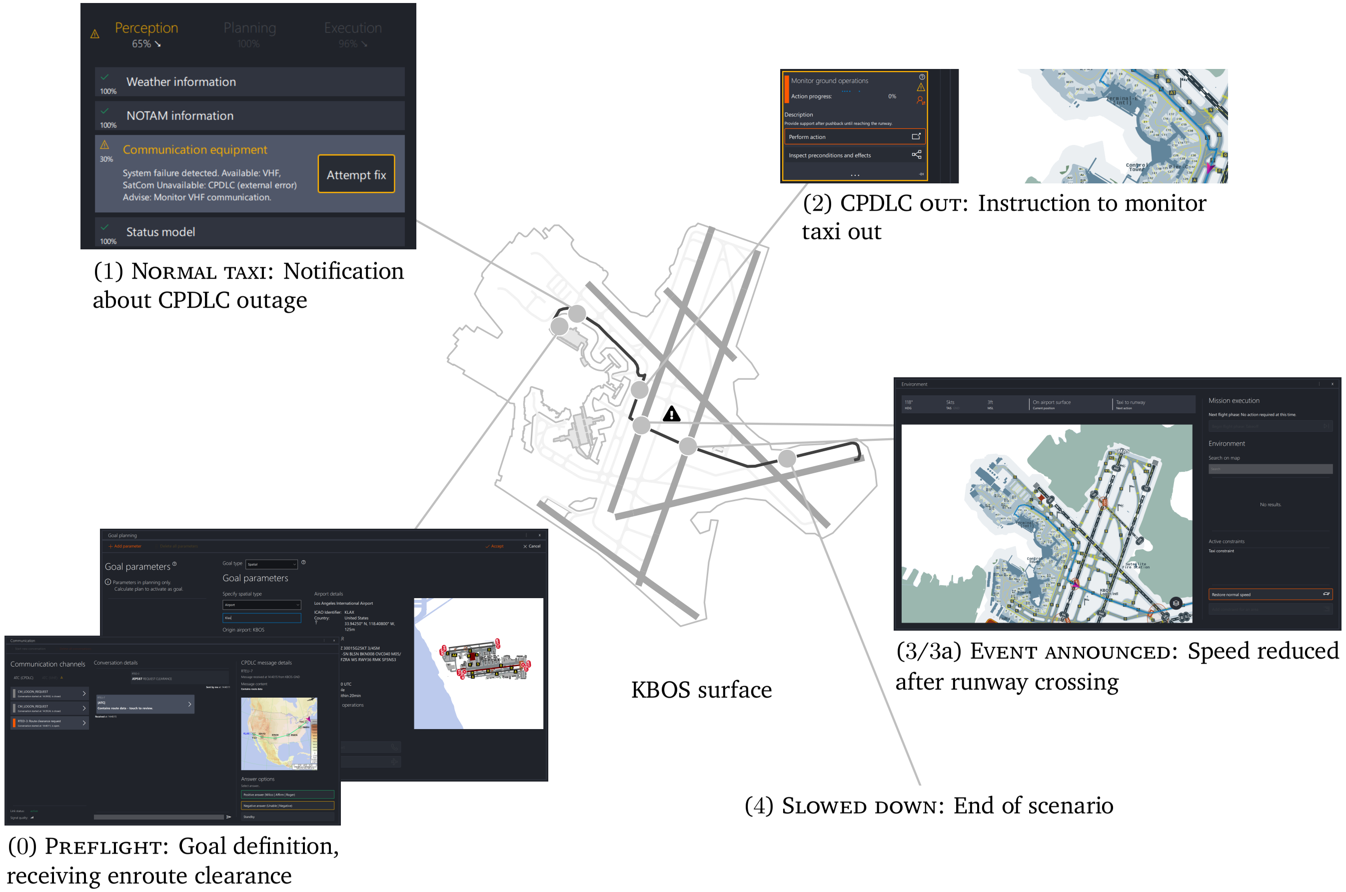


Figure 6.5.: Graphical overview of the events occurring during the scenario evaluation at KBOS. The black line indicates the taxi route which is followed during the evaluation. Events are shown as gray dots. Screenshots of the application related to the events are shown. Caption include scenario phases as defined in Table 6.2.

6.2.4. Debriefing

After having completed the scenario evaluation the participants are guided back to the briefing room. The final part of the questionnaire is filled out. During the debriefing session the participants are able to provide additional feedback and freely ask any questions.

6.2.5. Data recording

During the usability and the scenario parts of the evaluation procedure different data sets are being generated. All data generated during the simulator evaluation is outputted in a JavaScript Object Notation (JSON) text file format. The output from an evaluation session is saved to a secondary hard drive after each completed evaluation procedure. Data recordings are intermittently auto-saved during the evaluation procedure to prevent data loss due to unexpected crashes. This feature was implemented during the already running evaluation campaign as a reaction to such an occurrence.

Anonymization All test data is recorded anonymously. Each participant gets assigned a numeric identification number. All answers and performance measures of the participants are only referenced by this number. A mapping of participant identification number and their names is retained by the test administrator in order to allow participants to withdraw their data from the experiment if requested.

Usability evaluation During the usability evaluation the time from starting a task until clicking the finish button is recorded with ms precision. The time measurements are grouped as JSON objects and are identifiable through a task number. MCH ratings are stored in the same format. No data about whether a task was completed successfully or not is generated by the application. This data is recorded by the test administrator taking notes during the evaluation procedure. The calculated score values are then manually added.

Scenario evaluation During the scenario evaluation a centralized logging system in the application is used to track the currently performed actions. Along with every log message the current timestamp in seconds, the location of the aircraft in geodetic coordinates, and its current air speed are stored. The following events are logged:

- The user opening windows inside the application
- CPDLC messages sent/received from/to ATC
- Events related to the scenario progression

Furthermore, a log file for the secondary task workload assessment is generated. The appearance of a square, the reaction time of the user, and whether the provided answer was right, wrong, or too late is logged.

Others Aside from the data generated by the application during the usability and scenario evaluation, additional test data is recorded: The questionnaire which is progressively completed during the evaluation procedure is exclusively being filled out on paper. A copy of the questionnaire is shown in section D.2. After a completed evaluation procedure the questionnaires are digitized into a Microsoft Excel document.

Eventually, every evaluation session inside the simulator is video recorded. The video recording is top-down on the center main screen with the intent of capturing the user's interaction with the application. A secondary function of the video recording is to serve as a backup solution in case the automated data recording fails during an experiment.

6.3. Participants

Following the human-centered design process as described by Cooper et al. in [Coo+14], the evaluation should be performed with the users that represent the target group for which the development was performed. This, however, is challenging, given the fact that the user, or more so the personality of the user, is unknown, no definite job description exists. The operational concept described in section 3.3 is neither made for pilots today, nor for any other existing job description. No defined user groups exist or have been identified yet. As a consequence the following selection criteria are put in place:

- Must be between 21 and 65 years old
- Must be proficient in English and aviation-related terms in English
- Must have basic knowledge of aircraft operations and terminology
- Must be familiar with state of the art computer systems
- Must be in an acceptable health condition, must not have any conditions that prevent being in a flight simulator

Aircraft flying experience is explicitly *not* part of the selection criteria. Neither during the usability evaluation, nor during the scenario evaluation piloting skills are required. A professional background in aviation is preferred but not considered necessary. No compensation is paid to the participants.

6.4. Summary

In this section the evaluation of the GoIM concept was described. Three test hypotheses aligned with the definition of usability were introduced. The test methods were described, relevant variables were identified. A two-part evaluation method was described consisting of a usability evaluation and a scenario-based evaluation. Eventually, the data recording capabilities are described and the participant selection criteria are outlined.

7. Results

In this chapter the results collected during the evaluation of the Goal oriented Information Management (GoIM) concept are presented. The trials were performed between October and December 2019 at the Institute of Flight Systems and Automatic Control (FSR) at the Technische Universität Darmstadt (TUDA). Statistical tests are applied to determine whether the three global hypotheses are *supported* or *not supported* based on the rejection of the underlying null hypotheses. Following that, the results are discussed.

Assumptions about the test data Statistical test methods are applied to test the hypotheses formulated in subsection 6.1.2. The choice of test methods depends on the sample the data was drawn from, therefore two assumptions about the test data are made: 1. The sample of participants drawn from the population is non-normally distributed. 2. The variance of the participants is unknown.

Both assumptions are made as the number of participants is considered not large enough to assume either. A precondition for parametric tests is therefore not met [FMF13]. Instead only non-parametric tests are used to examine the data. The lower statistical power is accepted in the context of this evaluation. The following test methods are utilized:

- One sample Wilcoxon test in the one and two tailed variants, unpaired, with continuity correction.
- Friedman's Analysis of Variance (ANOVA) wit Bonferroni correction.
- Kendall's rank correlation.

All tests are performed using the statistics computing software environment *R* version 3.6. The base functionality is extended by using packages available via the Comprehensive R Archive Network (CRAN). All utilized packages and their versions are summarized in section D.6. A significance level of 95 % ($p=0.05$ one tailed, $p=0.025$ two tailed) is chosen for all tests. As a measure of effect size Pearson's r is used. Qualitative effect size measures are $|r| > 0.10$ small, $|r| > 0.30$ medium, and $|r| > 0.50$ large [Coh92].

In three cases the recorded test data required post-processing, occurrences are listed in subsection D.9.2. In all cases useable test data was recovered.

7.1. Experiment demographics

The participant demographics are summarized based on the answers provided in the first two sections of the questionnaire. In section D.5 the tabulated results are provided. Overall 13 participants (9 male, 4 female) participated in the evaluation study. Participants were assigned ascending identification numbers starting at 3. All participants were between 25 and 34 years old and considered themselves as tech savvy, having little struggle when interacting with automated systems (Mdn$=4$, IQR$=1$ / Mdn$=2$, IQR$=0$). The majority of the participants indicated a higher than average English proficiency (Mdn$=4$, IQR$=1$). An aviation-related background was reported by 10 participants, piloting experience was reported by 3 participants. A higher than average English proficiency in aviation-related terminology was indicated (Mdn$=4$, IQR$=0$).

7.2. Hypothesis I - Usability of the application

As described in subsection 6.1.2 the first hypothesis is tested through effectiveness, efficiency, and user satisfaction while performing the first part of the evaluation. In the following the results related to each one of the three categories are presented. Statistical tests are performed for each of the related null hypothesis to eventually test the first global hypothesis. In the following only condensed results are presented, the full results are made available in section D.7.

7.2.1. Effectiveness

Effectiveness is the first component of usability to be examined. The acceptance criteria for effectiveness is defined as $\lambda_I \geq 0.7$. The related null hypothesis is $H_{0,1}$. Plotted in Figure 7.1 is the time it took the trial participants to complete the 20 tasks. Times are shown relative to the reference time t_{ref} for each of the tasks as described in section D.3. Indicated are the two relevant limits of $q = 2t_{ref}$ and $2q = 4t_{ref}$ as described in subsection 6.1.2, Equation 6.1.

The relative times (Mdn$=2.30t_{ref}$, IQR$=2.01t_{ref}$) range between $0.65t_{ref}$ and $22.05t_{ref}$ for all tasks. Qualitatively, the *Assistant* category shows the largest variation (IQR$=2.97t_{ref}$) and deviation from q towards $2q$. The *Mission* and *Action* categories show the lowest variation (IQR$=1.56t_{ref}$, IQR$=1.59t_{ref}$). Shown in Figure 7.2 are the relative time values grouped by task category.

In Figure 7.3 the corresponding $\lambda_{I,initial}$ scores reached by the participants for the 20 different tasks are visualized. The score values shown in this depiction are *not* yet

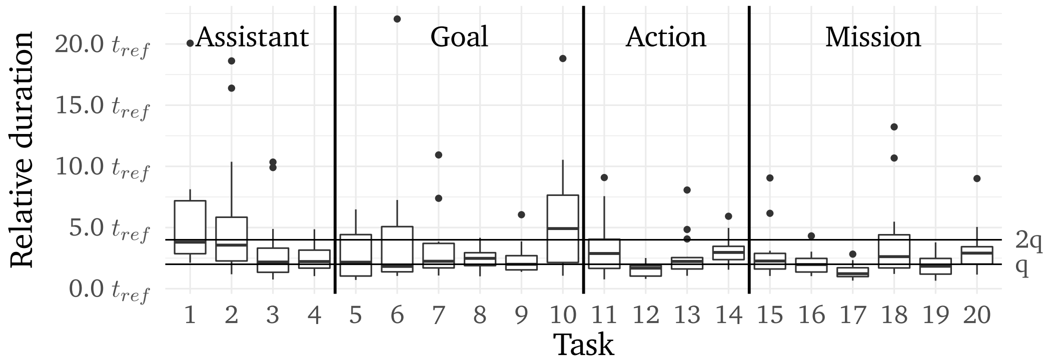


Figure 7.1.: Task completion time relative to the reference time t_{ref}, indicated by the horizontal lines are the scoring-relevant limits of $q = 2t_{ref}$ and $2q = 4t_{ref}$. Tasks numbers reference section D.3.

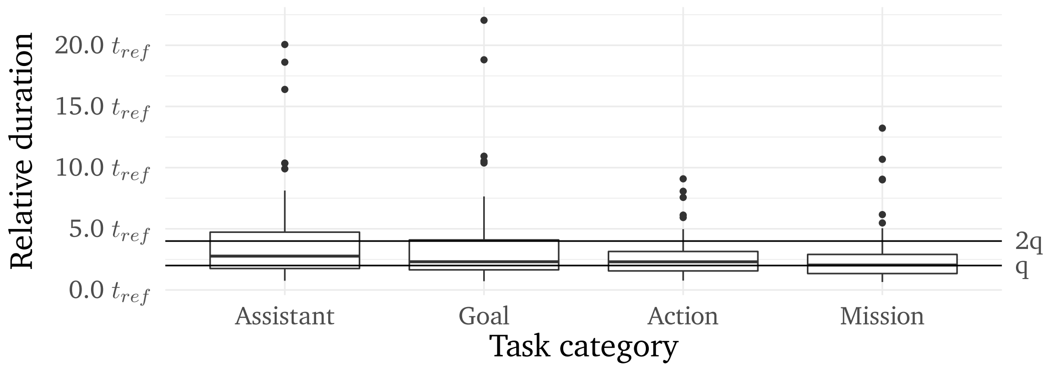


Figure 7.2.: Task completion time relative to the reference time t_{ref}, grouped by task category. Indicated by the horizontal lines are the scoring-relevant limits of $q = 2t_{ref}$ and $2q = 4t_{ref}$.

corrected for completion time. In this depiction the score values are grouped into three levels of completion. Completed indicates that the participant was able to complete the task including all sub tasks. Partially completed describes a task for which not all sub tasks were successfully completed, not completed tasks have no completed sub tasks. Most participants were able to either fully or partially complete all tasks, an average score of $\overline{\lambda}_{I,initial} = 0.89$ is reached for all tasks (Mdn=1.0, IQR=0.0). In the category *Goal* (Mdn=1.0, IQR=0.31) the lowest completion ratings are observed for the tasks 7-9 in the *Goal* category, followed by task 20 in the *Mission* category. For tasks 7-9 the participants

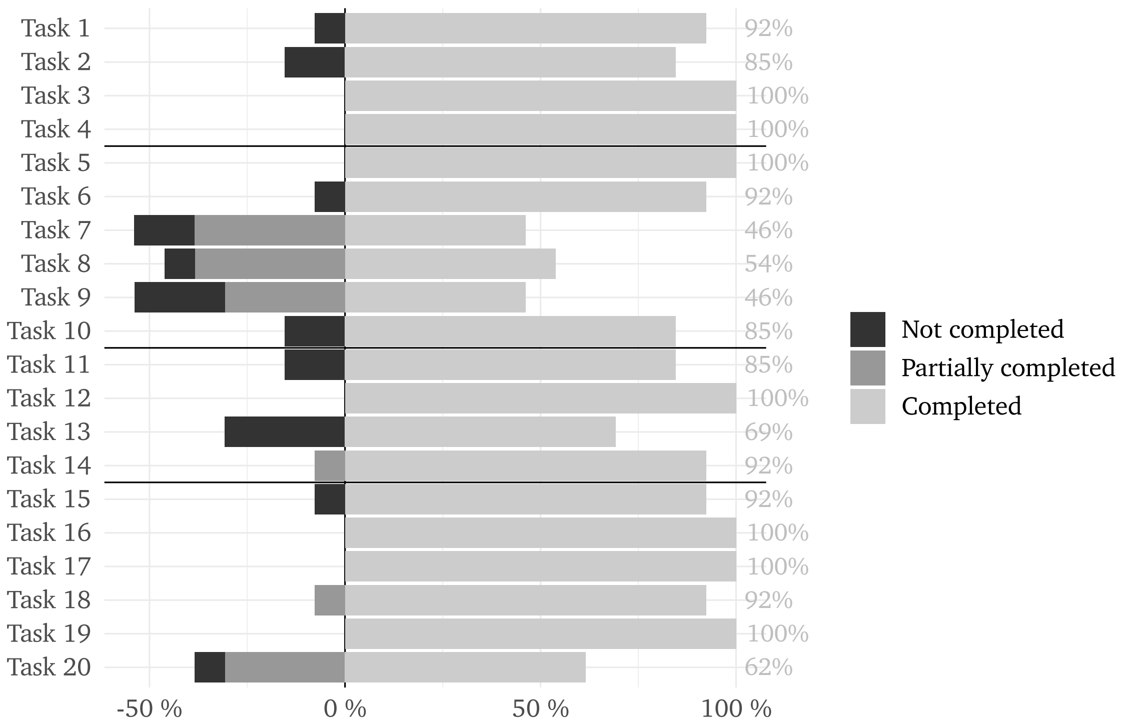


Figure 7.3.: The level of task completeness across all participants based on the value of $\lambda_{I,initial}$ reached during the first part of the evaluation. Task time t_{task} is not accounted for in this scoring. Tasks numbers reference section D.3.

often forgot to complete the last step of the task, such as accepting a goal parameter or activating a solution. While most participants were able to name the correct flight phase for task 20 they were unable to allow execution of it. With four participants unable to complete task 13 "What were the preconditions for the automatically executed action *Request taxi instructions*?" from the *Action* category this task was the not completed one. In this case all four participants were unable to locate the correct window.

The results for t_{task} and $\lambda_{I,initial}$ are combined using Equation 6.3, yielding the normalized effectiveness score λ_I. Shown in Figure 7.4 are the resulting values grouped by task category. A combination of a violin and a box plot is chosen to visualize the distribution shape of λ_I. Across all categories the λ_I (Mdn=0.41, IQR=0.98) scores are spread across the scale. The interquartile ranges for the different categories range between IQR=0.75 for the *Mission* category, 0.82 for the *Action* category, and 1.0 for the *Assistant* or *Goal* categories. An accumulation of score values towards both ends of the scale is visible especially for the *Assistant* and *Goal* category. The median values for all three categories are below the acceptability level of $\lambda_I = 0.7$.

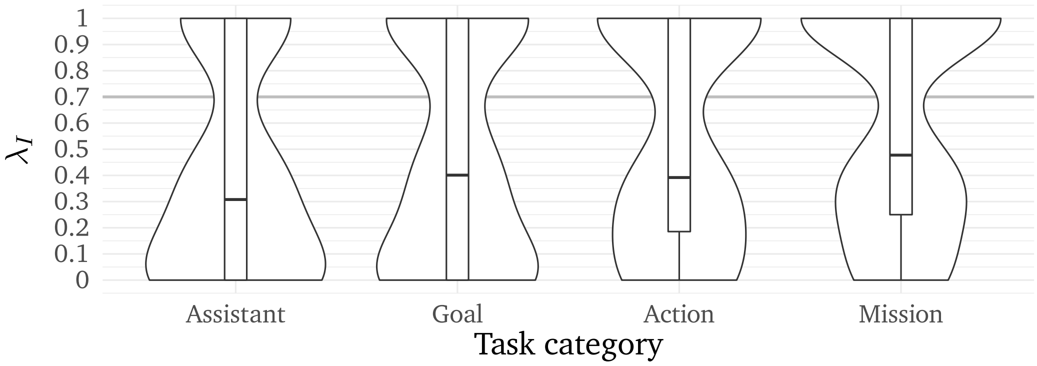


Figure 7.4.: Calculated values for λ_I grouped by task category. The solid gray lines indicate the acceptance level of $\lambda_I = 0.7$ for effectiveness as a component of usability.

Hypothesis $H_{0,1}$ is tested by performing a two-tailed one sample Wilcoxon test with $\mu_\lambda = 0.7$. The calculated value of λ_I (Mdn=0.41, IQR=0.98) differs significantly from the null value of $\mu_\lambda = 0.7$ (W=8418, p=1.18×10^{-12}<0.025, r=-0.44). The null hypothesis $H_{0,1}$ is *rejected*. Given the graphical depiction for λ_I in Figure 7.4 the alternative hypothesis $H_{a,1,<}$ for $\lambda_I < 0.7$ is tested. A significant result is indicated by a one-tailed one sample Wilcoxon test (W=8418, p=5.90×10^{-13}<0.05, r=-0.45) for $\lambda_I < 0.7$. Based on the evidence an *unacceptable* level for effectiveness is assumed.

7.2.2. Efficiency

Next, the results of efficiency are examined. The acceptance criteria is defined as an overall Modified Cooper-Harper (MCH) rating score of less or equal a value of 5. The related null hypothesis is $H_{0,2}$.

Shown in Figure 7.5 is a plot of the MCH scores which were submitted by the test participants during the evaluation procedure. The dashed gray lines indicate the borders of the four major areas present in the MCH scale as described in subsubsection 2.2.1. Indicated by the solid gray line is the threshold for acceptability at MCH=5.

With the exception of two outliers in the *Assistant* and *Action* category all values are below or at a score of 5 points (Mdn=3.0, IQR=1). Based on the plotted interquartile

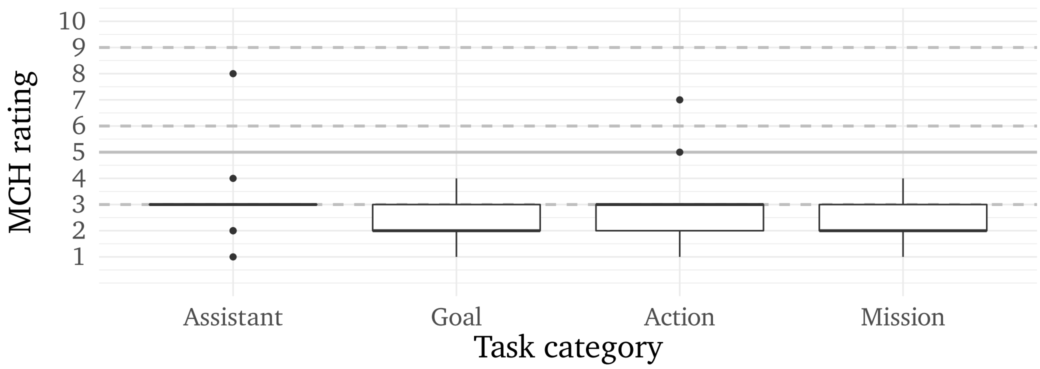


Figure 7.5.: MCH scores grouped by the task category for which the measurements were taken. Dashed gray lines indicate the four different areas of the MCH scale. The solid gray line indicates the acceptance level for efficiency as a component of usability.

ranges more than half of the values are at or below a score of 3 points. The median values alternate between 2 and 3 points.

Given the visual representation of the reported MCH ratings the null hypothesis $H_{0,2}$ is *rejected*. A two-tailed Wilcoxon one sample test confirms a significant difference of the MCH ratings from $\mu_{MCH} = 5$ (W=51.5, p=5.55×10^{-9}<0.025, r=-0.81). The alternative hypothesis $H_{a,2,<}$ of $\mu_{MCH} < 5$ is tested using a one-tailed one sample Wilcoxon test. A significant difference from values $\mu_{MCH} > 5$ is reported (W=51.5, p=2.78×10^{-9}<0.05, r=-0.82), the efficiency level is therefore considered *acceptable*.

7.2.3. User satisfaction

The third component being examined is user satisfaction. User satisfaction is considered to be acceptable for a System Usability Scale (SUS) score ≥ 52. Shown in Figure 7.6 is the summary of the calculated SUS scores (Mdn=75, IQR=12.5). All values are above the acceptability threshold.

Hypothesis $H_{0,3}$ is tested by a two-tailed one sample Wilcoxon test with $\mu_{SUS} = 52$. The SUS score (Mdn=75, IQR=17.5) reported by the participants after having completed the usability evaluation differs significantly from the assumed null level of 52 (W=91, p=0.0016<0.025, r=-0.87). Therefore, the null hypothesis $H_{0,3}$ is *rejected*. A one-tailed one sample Wilcoxon test for a SUS score of ≥ 52 is performed. For the alternative hypothesis $H_{a,3,>}$ a significant result is indicated for $\mu_{SUS} > 52$ (W=91, p=0.0008<0.05, r=-0.93). Therefore, the observed user satisfaction level is considered *acceptable*.

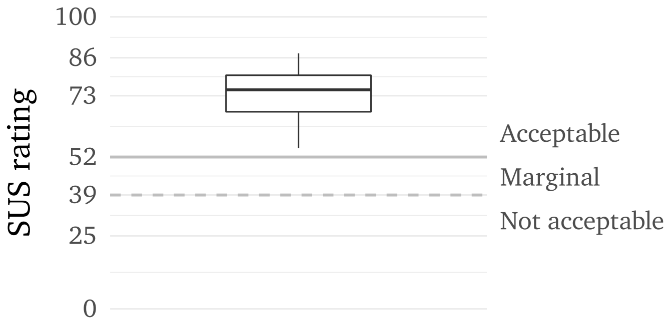


Figure 7.6.: SUS scores reported after the usability evaluation. Indicated by the gray lines are the three acceptability ranges as described in Table 6.1. The solid gray line indicates the acceptance criteria of a SUS score ≥ 52.

7.2.4. User feedback

The digitalized feedback is located in section D.7, no translation and only editorial corrections such as grammar and spelling corrections were performed.

7.2.5. Summary

In Table 7.1 the results of effectiveness, efficiency, and user satisfaction are summarized. As stated in subsection 6.1.2 an overall acceptable level of usability is assumed when all components are considered acceptable. The components efficiency and user satisfaction indicate acceptable results. Effectiveness is unacceptable, the global hypothesis 1 is *not supported*.

Table 7.1.: Summary of the tested usability components for H_1. Shown in column 2 is whether the associated null hypothesis was rejected. Column 3 indicates if the acceptability criteria was met.

Components	$H_{0,1}$, $H_{0,2}$, $H_{0,3}$	Acceptable
Effectiveness	rejected	unacceptable
Efficiency	rejected	acceptable
User satisfaction	rejected	acceptable

7.3. Hypothesis II - Usability of the concept in an operational context

Similar to the previous hypothesis the second hypothesis is also tested through effectiveness, efficiency, and user satisfaction with modified evaluation metrics. In the following paragraphs the obtained test data is described and the corresponding null hypotheses are tested. Applied post-processing steps are described when applicable. Eventually, the second global hypothesis is tested.

7.3.1. Effectiveness

The effectiveness values are examined first. The large majority of the participants were able to fulfill the two tasks of activating the Very High Frequency (VHF) radio and commanding the aircraft to slow down. Shown in Figure 7.7 is a map depiction of the airport surface of Boston Logan Airport (KBOS). The positions at which the participants either activated the VHF radio or sent the slow down command are highlighted.

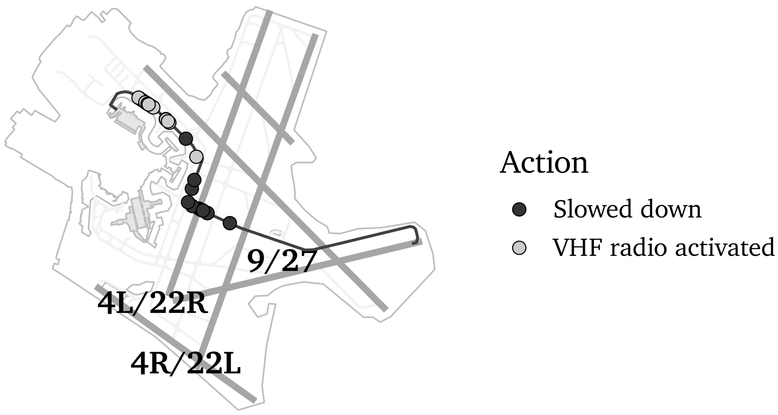


Figure 7.7.: Map positions on which the trial participants performed the required actions of either activating the VHF radio (gray), or commanding the aircraft to slow down (black).

A clustering of positions around the intended positions as they are pointed out in Figure 6.5 is visible. Deviations of when participants sent the slow down commands are observed: One participant (5) accidentally sent the slow down command too early and immediately revoked it, the first erroneous activation was not further considered in the test data. A second participant (14) sent the slow down command after having received the information about the Controller-Pilot Data Link Communications (CPDLC) outage. The participant later realized that the slow down command should have been issued after having crossed runway 4L and did so again. As the first command was not immediately revoked an effective penalty of 50 % to the scoring of $\lambda_{II,SLOW}$ is applied for this participant by also taking the first, too early, activation into account for the total scoring. Participant 11 repeatedly sent and revoked the slow down command and omitted the activation of the VHF radio. A penalty of 75 % for three erroneous activations was applied to $\lambda_{II,SLOW}$ for this participant.

To calculate the scores for λ_{II} the distance of where the event was triggered to where the correct action was performed is calculated along the taxi route. The two dimensional taxi route shown in Figure 7.7 is projected onto a one dimensional representation based on the distance from gate E11 along the route, depicted in Figure 7.8. Shown in here are

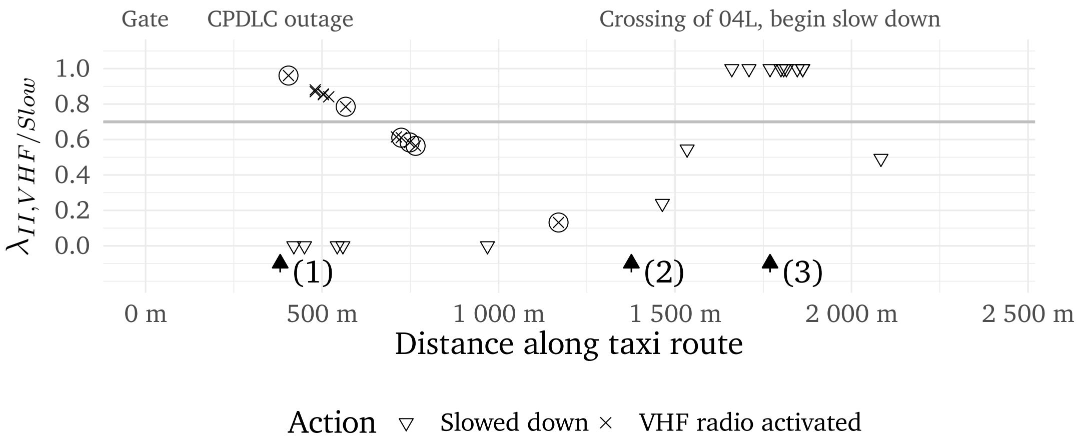


Figure 7.8.: Positions along the taxi route at which the participants performed the required actions during the evaluation scenario, relative to event triggers: (1) CPDLC outage, (2) announcement of debris on the taxiway, (3) crossing of runway 4L, begin of slow down. Circled values indicate where the VHF radio was activated after the participants had first opened the *environment* window.

the positions at which the participants performed the required actions. Indicated at the bottom of the chart are the positions at which the events were triggered. Two groups for when the VHF radio was activated are visible. The circled values highlight participants which first opened the *environment* window after the CPDLC outage was announced. As described in section 6.2 the *communication* window must be opened to activate VHF. With the exception of participant 13, the participants who first opened the incorrect window tended to activate the radio later than other participants.

The vertical position of the markers denote the resulting scores for either $\lambda_{II,VHF}$ or $\lambda_{II,SLOW}$. The horizontal gray line indicates the acceptability threshold of $\lambda_{II,x} = 0.7$. As described in Equation 6.4 the combination of both scores yields λ_{II}. Shown in Figure 7.9 are the summarized score values. Again, the horizontal gray line indicates the acceptability threshold.

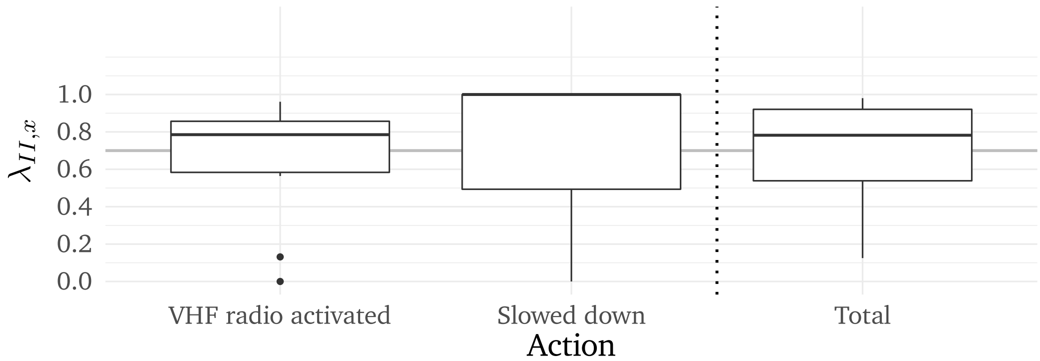


Figure 7.9.: Score values for $\lambda_{II,VHF}$, $\lambda_{II,SLOW}$, and the resulting total score λ_{II}.

For all three values the average values are close to 0.7 (0.66,0.70,0.67, left to right). The values of $\lambda_{II,VHF}$ (Mdn=0.79, IQR=0.27) and $\lambda_{II,SLOW}$ (Mdn=1.0, IQR=0.51) are used to calculate the total value λ_{II} (Mdn=0.78, IQR=0.38) as described in Equation 6.4.

Hypothesis $H_{0,4}$ is tested using the two-tailed one sample Wilcoxon test with $\mu_{\lambda_{II}} = 0.7$, the result is shown in the first row of Table 7.2. The results indicate that the difference

Table 7.2.: Test statistics summary of $H_{0,4}$ and the corresponding alternative hypotheses.

Hypothesis	W	p	r
$H_{0,4}$	404	0.85>0.025	−0.03
$H_{a,4,>}$	404	0.43>0.05	−0.13
$H_{a,4,<}$	404	0.58>0.05	−0.09

in λ_{II} from the assumed null level of 0.7 is not significant. The null hypothesis $H_{0,4}$ is therefore *failed to reject*. Two additional one-tailed one sample Wilcoxon tests testing the alternative hypotheses of $\mu_{\lambda_{II}} > 0.7$ ($H_{a,4,>}$) and $\mu_{\lambda_{II}} < 0.7$ ($H_{a,4,<}$) consequently did not indicate significant results as shown in Table 7.2. It is therefore concluded that no evidence exists that the value of λ_{II} is different from the estimated null value of 0.7. The acceptance criteria is hereby met and the effectiveness is assumed to be *acceptable*.

Based on section 6.1, a correlation of the effectiveness values observed during the usability evaluation and the scenario evaluation is investigated. It is expected that λ_{II} depends on how proficient participants are with the application. Participants that previously reached a higher λ_I can be expected to better understand the application and should consequently show a higher λ_{II} score. Shown in Figure 7.10 are the mean values for λ_I for every participant plotted against the λ_{II} score the participant reached during the scenario evaluation. A linear regression based on the least squares method is plotted along with the data points.

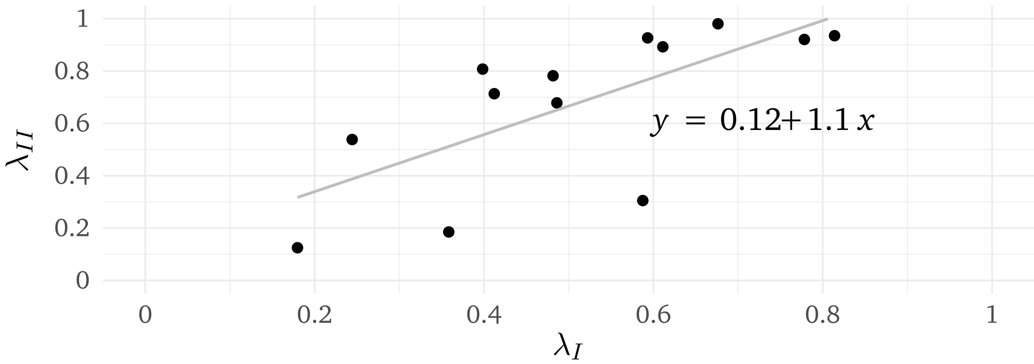


Figure 7.10.: Effectiveness results λ_I and λ_{II} plotted against each other to identify a potential correlation. Shown in gray is a linear regression based on the least squares method.

Kendall's rank correlation τ is calculated to determine whether a correlation between the two variables exists. A significant positive relation between λ_I and λ_{II} is indicated (τ=0.54, p=0.016<0.025).

7.3.2. Efficiency

The assessment of efficiency is primarily performed via the MCH ratings submitted after having completed the scenario. An acceptance threshold of a MCH score less or equal a score of 5 is defined. Secondly, the results of the secondary task workload assessment are used to qualitatively described the workload levels during the scenario.

MCH rating Shown in Figure 7.11 is the result for the post-experiment MCH rating submitted by the participants. At least half of the MCH scores (Mdn=3.0, IQR=1.0)

submitted after the scenario simulator session are at or below a value of 3. None of the MCH scores are exceeding the acceptance level of 5.

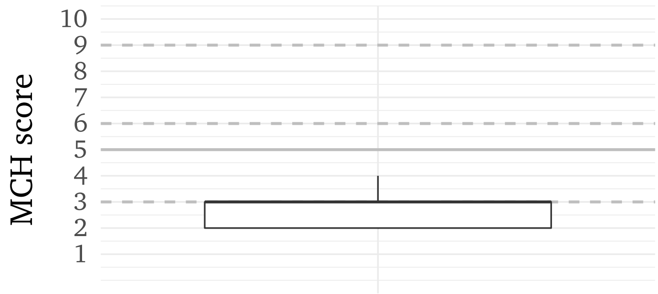


Figure 7.11.: MCH scores obtained after having completed the scenario. Dashed gray lines indicate the four different areas of the MCH scale. The solid gray line indicates the acceptance level for efficiency.

Hypothesis $H_{0,5}$ is tested by a two-tailed one sample Wilcoxon test with $\mu_{MCH} = 5$. The reported MCH score (Mdn=3, IQR=1) differs significantly from the assumed null level of 5 (W=0, p=0.001<0.025, r=−0.89). The null hypothesis $H_{0,5}$ is *rejected*. Considering the plotted results in Figure 7.11 a score of less than 5 is expected, the alternative hypothesis $H_{a,5,<}$ for $\mu_{MCH} < 5$ is tested. A one-tailed one sample Wilcoxon test indicates that the reported MCH score is significantly lower than $\mu_{MCH} = 5$ (W=0, p=0.0007<0.05, r=−0.95), therefore the observed efficiency is considered to be *acceptable*. In comparison to the MCH ratings in section 7.2 a paired Wilcoxon signed-rank test does not indicate any significant differences between the measurements (W=32.5, p=0.64>0.025, r=−0.13).

Qualitative secondary task workload assessment The second measure, for qualitatively describing efficiency, is the secondary task workload measurement. A higher secondary task score indicates a lower workload while exercising the primary task. The recorded scores for the secondary task are shown in Figure 7.12. Scores are divided into the phases as described in Table 6.2. For each participant the mean score in every phase is calculated. The aggregated result is then shown in Figure 7.12. For participant 6 the workload scores that were still recorded after the simulator problem are removed as the participant was unable to effectively continue the task. An increase of the mean score, indicated by the gray line plotted across the phases, after the preflight phase is visible, followed by a drop when the CPDLC outage occurs. Next, an increase towards the end of the scenario is visible. During the preflight phase (Mdn=0.27, IQR=0.21) and after the CPDLC outage event (Mdn=0.48, IQR=0.25) the lowest scores are observed. During the normal taxi

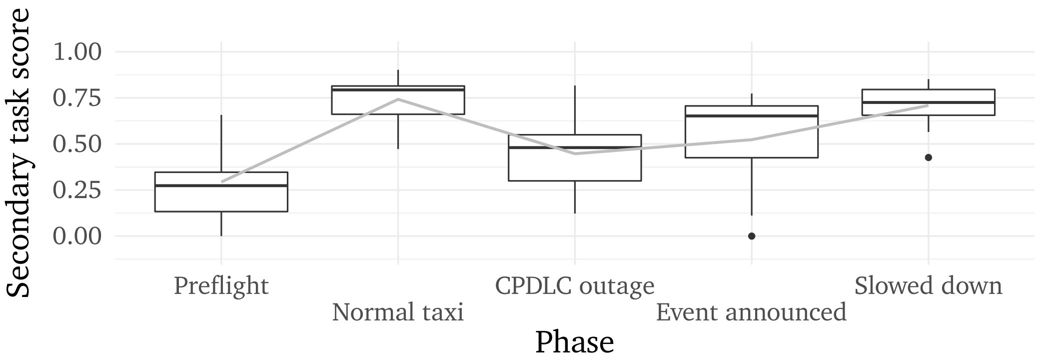


Figure 7.12.: Secondary task workload measurement scores as obtained from the participants while completing the scenario evaluation. Values are categorized in phases as described in Table 6.2. Shown are the aggregated mean values for the participants. The overall mean values are connected through the gray line.

phase (Mdn=0.79, IQR=0.15) the highest mean score is observed. Significant differences between the phases are identified through a Friedman's ANOVA. Post hoc tests are used with Bonferroni correction applied. The test indicates significant differences in secondary task performance between the scenario phases ($\chi^2(4) = 28.0$, p=1.25×10^{-5}<0.05). The critical difference ($\alpha = 0.05$ corrected for the number of tests) for all cases is 22.63. The significant differences are graphically depicted in Figure 7.13.

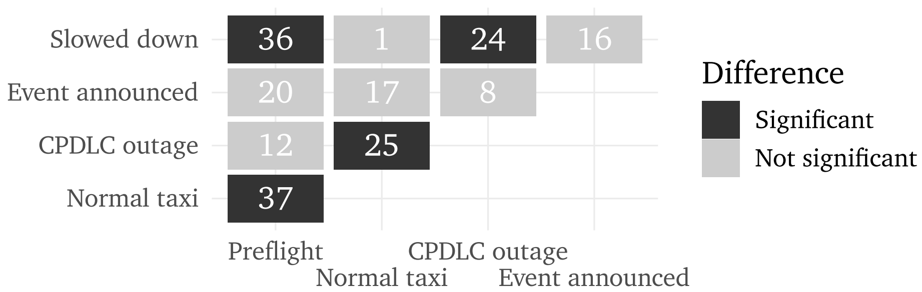


Figure 7.13.: Results of Friedman's ANOVA with Bonferroni correction applied for the secondary task workload measurement. Significant differences for p≤ 0.05 are indicated. Individual differences are shown, critical difference is 22.63.

7.3.3. User satisfaction

The last component of usability in the operationally-relevant scenario is the user satisfaction which is determined through the SUS. As described in section 6.2 the users were instructed to focus their rating on the concept itself this time and less on the application. In Figure 7.14 the summary of the calculated SUS (Mdn=75, IQR=10) scores is shown. The reported values show little spread around the median value (min=$55/45$, max=85) when compared to the user acceptance results reported in section 7.2. One outlier exists at a score of 45 and is generated by participant 11.

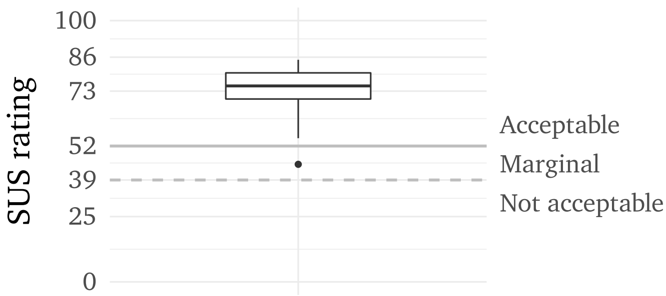


Figure 7.14.: SUS scores reported after the scenario evaluation. Indicated by the gray lines are the three acceptability ranges as described in Table 6.1. The solid gray line indicates the acceptance criteria of a SUS score ≥ 52.

Hypothesis $H_{0,6}$ is tested by a two-tailed one sample Wilcoxon test with $\mu_{SUS} = 52$. The SUS score reported by the participants after having completed the scenario evaluation differs significantly from the assumed null level of 52 (W=88, p=0.003<0.025, r=-0.82). The null hypothesis $H_{0,6}$ is consequently *rejected*. Based on the distribution of the SUS scores in Figure 7.14 the alternative hypothesis $H_{a,6,>}$ of $\mu_{SUS} > 52$ is tested. A one-tailed one sample Wilcoxon test provides an indication that the SUS score is significantly higher than $\mu_{SUS} = 52$ (W=88, p=0.002<0.05, r=-0.87). The user satisfaction level is therefore assumed to be *acceptable*. When compared to the SUS reported in section 7.2 a paired Wilcoxon signed-rank test does not indicate any significant differences between the measurements (W=38, p=0.97>0.025, r=-0.01).

7.3.4. Correlation between subjective ratings and effectiveness

In support of the result discussion in section 7.6 a correlation between the subjective ratings and the effectiveness scores is investigated. Kendall's rank correlation is calculated between λ_{II} and the MCH/SUS scores submitted after the scenario evaluation. No significant correlation between λ_{II} and the MCH scores (τ=-0.48, p=0.04>0.025) or λ_{II} and the SUS scores (τ=0.36, p=0.09>0.025) could be identified.

7.3.5. User feedback

Analogously to the questionnaire answers provided in section 7.2 the answers are digitalized and located in section D.8.

7.3.6. Summary

Based on the results of effectiveness, efficiency, and user satisfaction the decision of whether the second global hypothesis can be rejected or accepted can be made. Summarized in Table 7.3 are the outcomes of the previous paragraphs. All three components are considered to be acceptable. Consequently, the second global hypothesis is *supported*.

Table 7.3.: Summary of the tested usability components for the global hypothesis 2. Shown in column 2 is whether the formulated null hypothesis could be rejected. Column 3 indicates acceptability of usability criteria.

Components	$H_{0,4}$, $H_{0,5}$, $H_{0,6}$	Acceptable
Effectiveness	failed to reject	acceptable
Efficiency	rejected	acceptable
User satisfaction	rejected	acceptable

7.4. Hypothesis III - Concept acceptance

Eventually, the results relevant for determining whether the GoIM concept is acceptable from a user perspective are presented. In doing so, the answers to the questionnaire for the statements Q5.12 through Q5.20 are examined. Shown in Figure 7.15 are the results centered around the least acceptable value of 3. Favorable answers, for concept acceptance, of 4 and 5 are drawn to the right, while less favorable answers are shown on the left side.

The strongest disagreement with $23\,\%$ is reported for statement Q5.16. Here the participants had to assess the statement “I always felt in control of the situation (during the scenario).”. This circumstance was also mentioned as part of the user feedback provided after the scenario evaluation. Apart from that, only the statements Q5.15 (“I

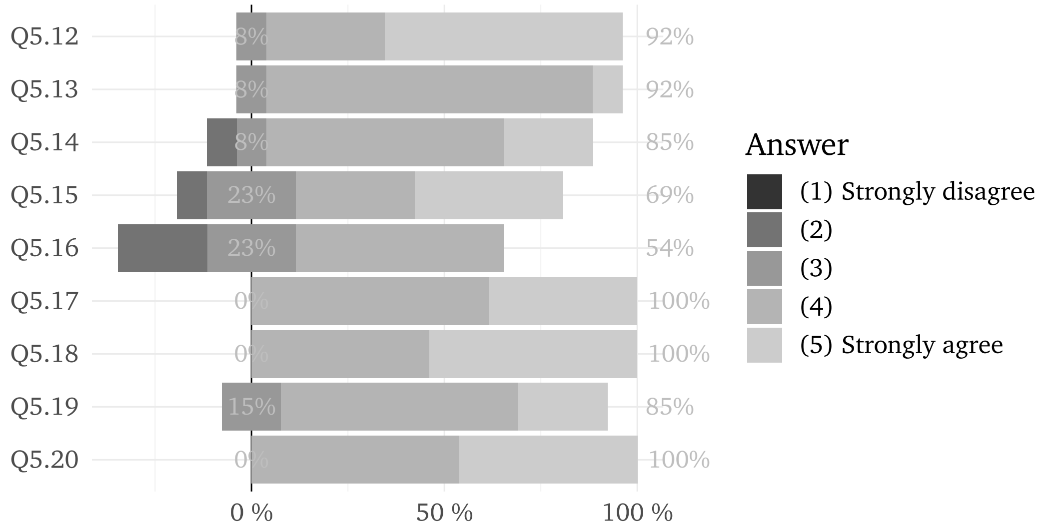


Figure 7.15.: Questionnaire answers regarding concept acceptance. Participants are given positive statements about the GoIM concept, answers are given on a five point Likert scale ranging from "Strongly disagree" (1) to "Strongly agree" (5).

did not feel misplaced or superfluous in my role as a mission manager.") and Q5.14 ("I understood why an action was being performed and what the preconditions, and effects of this action were.") received negative responses. For the closing statement Q5.20, in which the participants had to answer the statement "Overall, I am willing to cooperate with automated systems in the demonstrated way.", no ratings lower than a score of 4 were given.

To test the global third hypothesis all responses given by the participants are summarized. During this process all responses are equally weighted. The total distribution (Mdn=4, IQR=1) of answer values is shown in Figure 7.16. The violin plot in the background illustrates the probability density of the assessments along the Likert scale.

Hypothesis $H_{0,7}$ is tested by a two-tailed one sample Wilcoxon test with $\mu_{acceptance} = 3$. The total concept acceptance score reported by the participants after having completed the scenario evaluation differs significantly from the assumed null level of 3 (W=5498.5, p=2.50×10^{-18}<0.05, r=-0.81). Therefore, the null hypothesis $H_{0,7}$ is *rejected*. Based on Figure 7.16 the alternative hypothesis $H_{a,7,>}$ for $\mu_{acceptance} > 3$ is tested. A one-tailed one sample Wilcoxon test indicates that the true score is significantly above a value of 3 (W=5498.5, p=1.25×10^{-18}<0.05, r=-0.82). The acceptance for the GoIM concept as

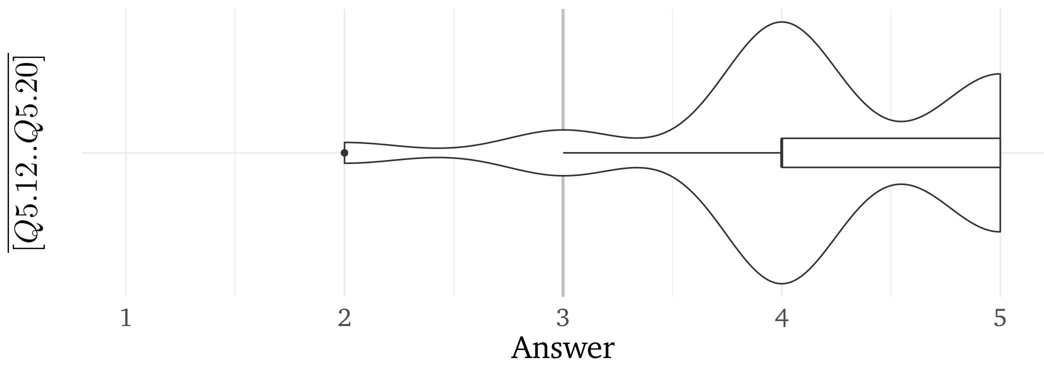


Figure 7.16.: Questionnaire answers regarding concept acceptance for all participants. All questions were phrased as positive statements about the GoIM concept. Assessments are given on a five point Likert scale ranging from "Strongly disagree" (1) to "Strongly agree" (5).

reported by the participants is therefore considered to be *acceptable*. Consequently, the third global hypothesis is *supported*.

7.5. Other results

In addition to the results obtained while testing the hypotheses formulated in subsection 6.1.2 other observations are made. The additionally collected data through the questionnaire is reported in the following. Secondly, subjective observations by the test administrator made during the evaluation are summarized.

7.5.1. Additional questionnaire data

Shown in Figure 7.17 are the answers provided by the participants when asked about the mission manager concept and their experiences with automated systems. On the five point Likert scale the possible ends of the scale varied. For reference review the options as listed in section D.2.

Answers above the horizontal line in Figure 7.17 relate to the mission manager concept. As indicated by the results of Q3.1 the majority (85 %) of the participants agree with the motivation behind the mission manager concept. 54 % consider the job of the mission manager challenging (Q3.2), while 62 % consider the job of the mission manager being different from the job of a today's pilot. With 77 % the majority of participants agrees that

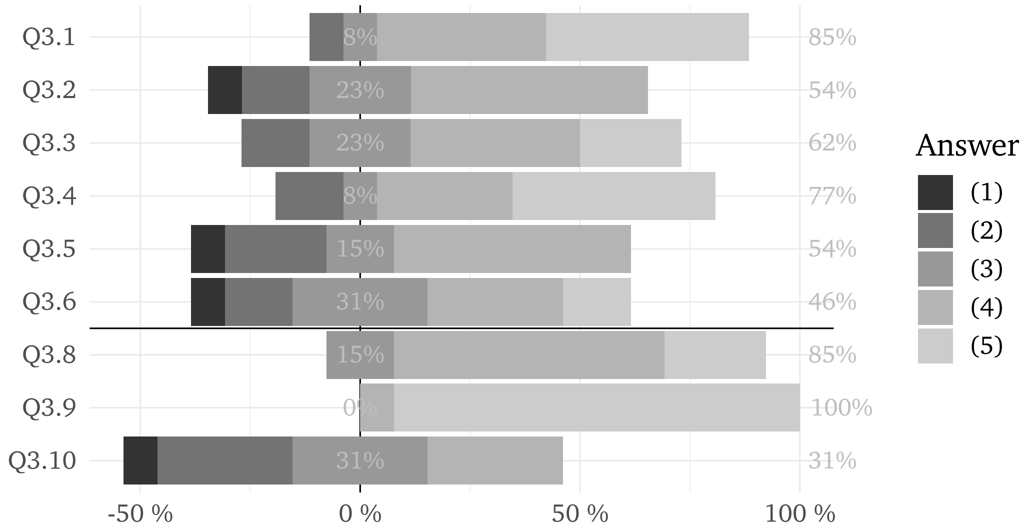


Figure 7.17.: Questionnaire answers of the participants when asked about their opinion of the mission manager concept (Q3.1-Q3.6) and about automated systems (Q3.8-Q3.10). Question Q3.7 is a free text answer field used for feedback on the mission manager concept. Assessments were given on a five point Likert scale. The equivalent answer options are listed in section D.2.

it is important to put the human operator in the center of flight deck operations (Q3.4). When asked about whether they would like to work as a mission manager 46 % answered positively, 31 % neutrally, and 23 % negatively (Q3.6). Free text feedback (Q3.7) on the mission manager concept is summarized in subsection D.9.1.

The majority of the participants indicated a strong trust in automated systems and their capabilities with 85 % and no answers below a neutral score of 3 (Q3.8). All participants indicated that they believe automation will increase on the flight deck (Q3.9). One third (31 %) of the participants agree that actions of automated systems are often hard to follow, another 31 % answered neutral while 38 % indicated disagreement (Q3.10).

7.5.2. General observations

The author of this thesis, who acted as the test administrator, participated in and led all evaluation sessions. He performed the participant briefing, observed the participants in the simulator and was able to collect participant feedback and take notes during the experiments. During all simulator sessions the test administrator was located behind the participants, able to observe the participant's interaction with the application. The subjectively reported notes are summarized in paragraph D.8.

7.5.3. Errors during the evaluation

During the evaluation shortcomings of the experiment design or the utilized application became apparent which did not became apparent during the pre-evaluation testing. With the exception of fixing one reproducible and major application crash no alterations to the evaluation procedure or the utilized application were performed. The summarized list of errors is located in subsection D.9.2.

7.6. Discussion

The usability and user acceptance of GoIM as a potential future information management concept for the flight deck were tested. Three global hypotheses evaluating the usability of the developed application, the usability of GoIM, and eventually the user acceptance were formulated. The evaluation followed the objectives of this work as outlined in section 6.1. The results of the evaluation are reported in section 7.2 to section 7.4. The hypotheses testing the usability and the acceptance of the GoIM concept itself were *supported* with a positive result. The hypothesis used to test the usability of the developed application was *not supported* due to low effectiveness ratings. A discussion of the results is performed in the following. References to the requirements set forth in section 4.1 are given. Eventually, the central objective of this thesis as stated in section 1.2 is revisited once more. An answer to the question if GoIM is a potential solution for information management on the future flight deck is provided.

Hypothesis I The first global hypothesis through which an acceptable level of usability of the application is hypothesized is rejected. A non-acceptable level of effectiveness was reached during the evaluation. As a consequence, an overall non-acceptable level of application usability is assumed following the definition given in subsection 6.1.2. Acceptable levels for user satisfaction and efficiency are indicated. Although, testing this hypothesis is not directly linked to the overarching question if GoIM is an acceptable solution, it should not be left unconsidered. Future iterations of the application can benefit from understanding this outcome.

The majority of the participants was able to either fully or partially complete the tasks. This circumstance is represented in Figure 7.3, as well as the distribution of values for $\lambda_{I,initial}$ (Mdn=1.0, IQR=0.0) across all categories. However, the time it took the participants to complete the tasks negatively impacted the final results for λ_I. This effect is especially observed for the *Assistant* category as depicted in Figure 7.2. Participants often spent a considerable amount of time searching for the requested functionality inside the application. The participants are practically first time users of an application with a wide range of functions. After a short introduction and 10 min time to explore the application on their own (see section 6.2) the observed behavior is not implausible. The possibility of providing a more focused application training to better understand the application was mentioned during the debriefing sessions and in response to question Q4.11 on the questionnaire (see section D.7). It was reported that the application was better understood after the usability evaluation. This effect was intended as described in section 6.2. During the evaluation usability issues were uncovered as documented in section D.7. Reported issues are wording inconsistencies, differentiation between static and interactive parts of the application or similarities between windows. Considerations for future iterations of the application and its evaluation will be outlined in section 8.3.

In contrast to the effectiveness scoring, the participant's subjective feedback is painting a different picture. The MCH ratings submitted by the participants during the evaluation indicate that they experienced no detrimental mental workload levels while using the application. All values were almost consistently below or around a value of 3 as shown in Figure 7.5. The SUS rating submitted after the usability evaluation further supports this. With no rating below the critical score of 52 (min=58) and an average score of 72.5 the mean value is at the border to the *Excellent* rating as described in Table 6.1.

The low effectiveness scores obtained during the usability evaluation are attributed to limited training and experience of the participants. Not meeting the time limits which were set forth by experienced users, as described in subsection 6.1.2, is identified as a cause. In contrast to that, positive feedback through the submitted SUS/MCH answers is received. It is assumed that the application was not perceived as a major obstruction by the participants. Individual influences on other hypotheses are discussed in the following.

Hypothesis II Based on the results presented in section 7.3 the second hypothesis is accepted. An acceptable level of usability is attributed to the GoIM concept in an operationally-relevant context. In all three areas of usability the participants showed acceptable results. They were able to support the automated systems by providing the required context, as described in subsection 4.2.2.

Similar to the application usability evaluation the effectiveness score was the lowest one relative to its acceptability threshold. With no evidence of the score being different from the acceptance level of $\lambda_{II} = 0.7$ an overall acceptable result is reached. The majority of the participants was able to successfully perform both requested tasks as shown in Figure 7.8. Two groups of participants are visible: some activating the VHF radio immediately and a second group requiring more time.

The majority of the participants requiring more time to activate the radio first opened the *environment* window, as indicated by the circled values in Figure 7.8. Once the CPDLC outage event is triggered the status representation of the assistant on the right side of the screen changes which is accompanied by an acoustic signal. At the same time a replanning occurs, caused by the unexpected state variable change (cf. subsection 5.2.1). Caused by the system outage the aircraft systems require the human operator to provide context. An action titled “Monitor ground operations”, shown in Figure 6.5, is inserted into the list of planned actions. This action is highlighted through a pulsating yellow border to draw the user’s attention.

Participants that first investigated the newly added action opened the associated *environment* window. After realizing that no action is required in this window they were able to open the correct *communication* window to activate the radio. The chosen alerting method delayed the participants from taking necessary actions. One participant failed to recognize that no action is required in the *environment* window and repeatedly slowed the aircraft down while failing to recognize the correct action. This observation is reflected in comments made by the participants post-evaluation and in the questionnaire (translated from German, original comment in section D.8):

> Flashing tempts you to do something.
>
> It was irritating that the left field remained orange although no action was required. Hence, I reduced the speed too early.

This indicates that the participants did not understand what was their task in this moment and acted on suspicion. Through this behavior the requirements SA-1-0 and IID-2-0 are violated. Further contributing to this confusion is also the fact that neither the *perception* status value of the assistant nor the contingency action are influenced by the actions of the user. The users are unable to actually *do* something to restore the health of the perception status or to stop the alerting-state of the action card. This behavior was not implemented in the application. The importance of proper alerting and the provision of feedback to the user (HAT-1-0), which is lacking in this case, is underlined by this.

Regardless of whether the participants instantly recognized the situation or not, it introduced a significantly different workload when compared to the previous normal taxi phase. This circumstance is visible in Figure 7.12 and Figure 7.13, where a drop in secondary task performance is visible. After the participants recognized the situation (*identification*), and performed the required actions (*fault isolation*), the performance then started to increase again (*recovery*) until reaching a level with no significant difference to normal taxiing. This behavior can be expected. Based on the submitted MCH ratings, the increase in workload was not detrimental to the participant's perception of GoIM. The submitted SUS ratings are well above the acceptance threshold with an average value of 72.5. No evidence for a significant difference between these values and the ones reported during the usability evaluation was found as shown in section 7.3. The positivity of these results provides a first indication that GoIM in fact represents a potential solution for information management on the future flight deck. Especially with the objective of considering the human operator first, the importance of subjective ratings is an important factor in this development stage.

Lastly, the influence of the developed application on the evaluation of the *concept* is focused on. A connection between the effectiveness scores λ_I and λ_{II} is investigated in section 7.3. The results shown in Figure 7.10 and the associated Kendall's rank correlation indicate a correlation between λ_I and λ_{II}. This provides an indication that an influence of the application exists when assessing the component of effectiveness for GoIM. A relationship, as anticipated in section 6.1, between the effectiveness scores exists. While an influence of the application itself on λ_{II} is comprehensible, a connection between the participant's objective performance during the scenario evaluation and the post-evaluation submitted subjective ratings is investigated. Based on the results of Kendall's rank correlation reported in section 7.3 no evidence for a significant correlation between λ_{II} and the MCH or SUS ratings is identified. The participants were able to experience GoIM in an operationally-relevant scenario and submit their subjective ratings with no significant bias related to their achieved effectiveness score.

In summary, an acceptable level of effectiveness for users operating under GoIM in an operationally-relevant scenario was shown. An influence of the application design on the effectiveness rating could be identified, reducing the weight of this measurement for the evaluation of the GoIM concept. Consequently, more weight is attributed to the subjectively submitted MCH and SUS scores. No evidence for a correlation between effectiveness and the subjectively submitted MCH and SUS scores is identified. Both, the MCH and the SUS, ratings are in favor of GoIM. Eventually, the positive subjective user feedback is seen as an indication that the concept is a potential solution for information management on the future flight deck. It can therefore be concluded that the GoIM concept offers an acceptable level of usability to the extent shown in this thesis.

Hypothesis III Overall, the participants expressed a high acceptance of the GoIM concept after having completed the scenario in the simulator as described in section 7.4. In the following, the answers given in the questionnaire for Q5.12-Q5.20 are reviewed in detail.

The answers given for Q5.12 - Q5.14 indicate that sufficient awareness about the ownership of functions and actions was provided as required by the requirements HAT-3-0, SA-1-0, and SA-1-0. In contrast to that, lower ratings were provided for Q5.15 and Q5.16, those questions are the only ones with ratings below a neutral score. Both questions relate to the role of the human operator. Asked about whether they did not feel superfluous on the flight deck (Q5.15) 31 % of the participants answered neutrally or negated. This is an expected result, considering that with increasing automation the meaningful tasks for the human operator potentially decrease. During the scenario evaluation only little involvement of the participants was requested, essentially two actions after pushback. Naturally, this can lead to the impression of superfluity when working with a highly automated system. Especially under Reduced Crew Operation (RCO) this phenomenon can become an issue (cf. [Nei20]).

Question Q5.16 was related to how much the participants felt in control of the situation while performing the scenario, as required per HAT-4-0. Especially participants with an aviation background or piloting experience mentioned that they had wished for more direct means of interaction (translated from German, taken from section D.8):

> No possibility to brake or stop.
>
> I was unaware how I could brake in case of an emergency (or assume manual control).

It was communicated that this especially applied for when the participants would reach a runway crossing point and have not yet received a crossing clearance and therefore started to feel uncomfortable. Based on HAT-4-0 this capability should exist but was not part of the application scope for the evaluation scenario (cf. section 5.2).

Asked about how well the participants were able to communicate their goals to the automation (Q5.17) or how well the automated systems supported them (Q5.18) only positive answers were given. Together with the answers given when asked about the availability of information (Q5.19) an overall positive result for the GoIM concept is observed. This stance is also visible in the comments provided by the participants (translated from German, taken from section D.8):

> The interaction with the automated system is very well integrated and clearly designed. The system provides sufficient information.
>
> Overall, the collaboration with the system was delightful. All relevant information were available or reachable through a few clicks.

The answers provided to the final question if the participants would like to cooperate with the automated systems in the demonstrated way (Q5.20) were exclusively positive. Ultimately, the previous indications that GoIM is an acceptable solution from the user's perspective is supported through this response.

Mission manager concept As the last part of this discussion a brief review of the feedback to the mission manager concept of operations is given. The majority of the participants agreed with the motivation behind the concept (Q3.2) and considered it important to put the human operator in the center of flight deck operations (Q3.4).

Less predominately is the participants opinion about the job itself. Only about half of the participants see the job of the mission manager as challenging (Q3.2) or potentially satisfying (Q3.5). Consequently, when asked about whether they would like to perform this job (Q3.6) a negative tendency towards not liking to perform this job is observed. This also is reflected in the comments given by the participants (translated from German, original answers in subsection D.9.1):

> Could get boring on a long distance flight.
>
> The pure job of the mission manager as it was presented can be boring (non-challenging).

This again underlines the necessity to ensure that the human operator does not feel superfluous and has a meaningful role on the flight deck. Future research in this area and development of the concept should therefore use the herein presented results to further shape the concept. After incorporating the first feedback received during this work the concept should be further evaluated. A more formal evaluation only focussing on the human operator and the job they have to perform is advisable. For this a full description of the job profile must be derived which extends on the herein described principles. Just as the definition of the full job profile, a more in-depth evaluation of the target users for this job should be performed. Eventually, the identification of a fitting personality profile will also benefit the further development of GoIM, for example by allowing more focused user research and evaluation.

7.7. Summary

The results of the GoIM evaluation study were presented. During the previous discussion an interpretation of the results was performed to determine if GoIM provides an acceptable level of usability to be considered a potential solution for a human-centered information management on the flight deck. An influence of the application on the evaluation of the concept was discussed and found to be present for the measurement of effectiveness. Overall, however, GoIM was backed up by strong subjective user feedback and the expressed willingness to operate under the presented concept. Based on the presented results an acceptable level of usability was attributed to the GoIM concept. Lastly, the feedback for the mission manager concept was summarized.

8. Conclusion and outlook

This chapter is the final chapter of this work. A summarizing retrospective of this work is provided. In the following, the outcome of this work and the answer to the guiding research question is given. Lastly, an outlook providing guidance for future research effort concludes this work.

8.1. Summary

In this work an answer to how a human-centric information management system could be designed for usage on the future flight deck under the assumption of Reduced Crew Operation (RCO) is given. A development process based on the DIN EN 9241-210 [DIN11] was followed during the development of this work.

Initially, the current state of the art is reviewed. An introduction to the Global Air Transportation System (GATS) and its actors is provided. Responsibilities of a two person flight deck crew are defined based on the regulatory definition provided by the International Civil Aviation Organization (ICAO). The concepts of situation awareness and mental workload are introduced. The Modified Cooper-Harper (MCH) scale is subsequently described as a subjective mental workload rating scale. A method for automated action planning is described. A definition of information management in the context of this thesis is derived, commonly used terminology is introduced. Challenges related to flight deck information management are researched. Existing efforts and related work to the topic of this thesis are identified and summarized. A classification of this thesis in relation to the previously identified efforts is done.

The research gap to which this thesis seeks to contribute is the combination of the chosen approach to the human-centric development process and the demonstrated implementation depth. A potential information management system for use on the future flight deck is developed. In this development process the role of the human operator is first redefined, a new concept of operations for the flight deck is described. The developed concept(s) will not remain in the *concept* stage but are implemented in hard- and software, a first evaluation is performed.

Following the development process of DIN EN 9241-210 [DIN11] the context in which such an information management system will be used is analyzed. The results of a preceding pilot task analysis are summarized which is used to form a thorough understanding of a pilot's tasks today. The task analysis is supplemented by research of daily operational challenges a pilot has to work with. High workload use cases are identified based on existing literature and pilot interviews. To form an understanding of the changes to the larger operational context, expected changes to today's GATS as proposed by Single European Sky ATM Research Programme (SESAR) or Next Generation Air Transportation System (NextGen) are reviewed. As no satisfying concept of operations for the future operator under RCO could be identified a new one is created. A forward-looking concept in which the future pilot is called *mission manager* is created. The concept is influenced by existing RCO concepts and the academic body of flight deck research. The operational concept defines the roles of the human operator and their interaction with the highly automated aircraft system. In this context the question of responsibility and authority is answered.

As the next step in the development process the requirements towards a future information management system are summarized. The requirements are derived from the existing body of research and supplementary literature. Assumptions to scope the development process are stated. Based on the previous work a new concept for information management on the future flight deck is introduced. The concept is named Goal oriented Information Management (GoIM). A description of the concept using the three main components of the human operator, flight deck interfaces, and aircraft systems is performed. The concept is based on the idea that both the human operator, and the aircraft systems should be working cooperatively towards the same goal as set by the responsible human operator. Consequently, the information provided to and required from the human operator should revolve around the tasks that are necessary to reach the defined goal. The delineation of tasks between the human operator and the aircraft systems is described, which is based on the mission manager concept of operations. Different operational scenarios highlighting the interaction of the human operator with the aircraft systems are described. Following the conceptual description of the concept is the implementation of GoIM with the goal of evaluating it. First, the development and construction of a new RCO research simulator is described in which the evaluation will take place. The simulator is based on the defined concept of operations. Secondly, the decomposition of the GoIM concept into a software application is outlined.

The overarching development process is continued by evaluating the developed concept in respect to the objectives of this work. An evaluation based on the concept of usability and the user's acceptance for GoIM is planned. In this effort three hypotheses are tested. The first hypothesis is used to test if the developed application itself offers an acceptable level of usability. This is tested by having the users perform a series of tasks with the application while measuring their success rate and time to complete a task. The measurements are supplemented with subjective workload and user satisfaction ratings through the MCH and the System Usability Scale (SUS). Secondly, the usability of GoIM as a concept in an operationally-relevant scenario is evaluated. Based on the use cases identified in the previous context analysis a scenario is generated during which the users must support the automated aircraft systems. Comparable measurements to the first hypothesis are employed. Lastly, the user acceptance for GoIM is determined through a custom questionnaire.

The evaluation study was performed in the previously constructed RCO research simulator at Technische Universität Darmstadt (TUDA) with 13 participants, of which the majority had a background in aviation. The results obtained during the evaluation study were presented. Of the three tested global hypotheses two were accepted, the first hypothesis was rejected. It was found that the application does not meet the previously defined level for acceptable usability. The hypotheses testing the usability and the user acceptance of GoIM were accepted based on the presented evidence. A discussion of the results was performed. Overall, the GoIM concept received positive feedback through the evaluation results. Positive subjective ratings were submitted by the evaluation participants.

8.2. Conclusion

The objective for this work, as set forth in section 1.2, is to answer the question how a human-centric information management system could be designed. This question is asked under the assumption of a fundamentally different operational concept, including a single operator interacting with highly automated aircraft systems. The intention of this work was to look beyond incremental or evolutionary change, to spark future research by introducing a new way of thinking about the collaboration between human operators and automated aircraft systems. As part of reaching the overarching objective three research questions were formulated in section 1.2 and answered throughout this thesis:

The question of what will be the human operator's role was answered through the mission manager concept in section 3.3. Their role will be to set strategic goals, which are then translated to tactical actions by the automation. In this role the human operator will still be the highest authority onboard the aircraft, and will still be responsible for the safe conduction of a flight.

The second question about how a human operator can coordinate with the automated aircraft systems to reach a mission goal was first answered in section 4.2, forming the nucleus of this work. A concept for information management on the flight deck was introduced with GoIM. The idea behind this concept is that the human operator and the automated aircraft systems are cooperatively working towards a common mission goal.

Lastly, the question of how acceptable such a concept is from a user perspective was answered in chapter 6 and chapter 7. The previously developed GoIM concept was implemented in a software application to be used in a purpose-built RCO simulator which was designed and built as part of this thesis. This simulator was designed as a conceptual flight deck for a human operator operating under RCO, it's development is described in chapter 5. A two-part evaluation concept was designed in which the usability of the developed application and the usability as well as the user acceptance of GoIM were assessed. Although the application did not meet the set level for acceptable usability, GoIM as a concept did. Overall, GoIM has received positive feedback through the objective and subjective measurement techniques employed during the evaluation. In particular the strong subjective feedback provided by the participants supports the idea of GoIM, achieving the goal to design a human-centered solution. GoIM has proven to be a viable concept and a starting point for a new way of managing information on the flight deck, potentially supporting future RCO operations.

In this regard, a final answer to the central research question as stated in section 1.2 can be given: Backed up by strong user feedback a first step towards a new way of human-centered information management on the flight deck is taken with GoIM, providing a foundation for future research to build upon.

8.3. Outlook

With this thesis a first step towards a new way of managing information on the flight deck under RCO is taken. At the same time opportunities for future work are identified.

Justified by the positive feedback received for the GoIM concept future work should be focused on broadening and extending the concept. One area to expand upon is the underlying algorithm which allows the aircraft systems to plan the required actions to reach a mission goal. An area potentially worth investigating here is the automated identification of possible system actions including their preconditions and their effects. These for example could be derived from models used in the domain of aircraft safety analysis. Approaches from the machine learning domain should be considered to support this effort. The usage of alternative planning algorithms should be explored. From a user perspective it should be evaluated which level of detail for the performed actions is required and which abstractions can be made. While broadening the scope of the proposed planning algorithm its robustness in abnormal situations should also be investigated. A long-term goal should then be to evaluate potential certification paths of such a system. Synergies from today's efforts to certify pilotless drones for passenger and cargo transport should be used.

Additionally, the application in which GoIM was implemented should be further developed. Expanding the scope of the application beyond the scenario of this work towards a more generalized application should be considered. Required additional functions can be identified based on the function allocation of the mission manager concept of operations and the derived requirements. Concerns shared by the evaluation participants can be addressed through this. Areas of improvement in the User Experience (UX) and User Interface (UI) realm were pointed out during the discussion of the evaluation results in section 7.6. Greater attention must be payed to this aspect because in the end it is ultimately the application with which the user will be interacting. Consequently, future iterations should incorporate the user and additional professional traits, for example UX/UI experts, more closely. The application should further be broken down into smaller components which can be individually designed and evaluated. In faster iteration cycles the application could then be developed closer with the users. Nevertheless, an overall integrated solution should be pursued. Occasional formal human factors evaluation could be used as quality gates during the greater development process.

This again underlines the need for an identified user group or a persona for the mission manager job role to more effectively target the user. Identifying this user group should be the next logical step in defining the future of operations. First, the feedback received during the evaluation should be considered. In doing this, the concept of operations must be extended. A primary question to be answered here is: what is the operator doing when the current mission does not require any actions? Furthermore, it must be ensured that the human operator does not feel superfluous on the flight deck as this was a concern raised by the evaluation participants. Other areas requiring additional work are the ground segment of operations, or the question of how to ensure sufficient vigilance of the human operator when working with the highly automated aircraft systems.

With a refined concept personality traits of potential future operators working as mission managers can be identified. Over time a growing maturity of the concept of operation for the human operator can be assumed. To gain further feedback an involvement of the scientific community should be considered. As early as practical the regulatory authorities should be involved to gain further support.

As mentioned, an iterative approach is desirable to develop a holistic concept for future operations. During this process, however, occasional formal evaluations should be performed. As opposed to the evaluation performed in this thesis future iterations should also consider approaches in which a concept based on GoIM is compared to today's operations. However, this is only advisable once an appropriate maturity stage is reached to avoid running into the previously discussed issues. This will provide insights into areas of improvement to further drive the developed concept. Advanced evaluations will help to better understand how a human operator must be trained to effectively and safely operate under GoIM. For future evaluations an alternative training concept for the application should be considered in response to the results of the application usability evaluation (cf. section 7.6). Providing an extended application training to the users can be acceptable, given the context where the application will be used and the inherent complexity of such an application. The utilization of a proper time limit for future usability evaluations, or alternatives thereof, should be further investigated.

As all of the described future work packages have dependencies on each other, a waterfall development approach will not suffice. Instead, the utilization of agile development methods, as they are being used in the software development domain, should be considered. The aforementioned evaluations during which the concept is tested will be built into this. The outcome of these evaluation should be taken into account for the next iteration, expecting changes to different areas of the overall concept. Eventually, this work can then be seen as the first iteration for future iterations to build upon.

References

[Abb14] Kathy Abbott. “Human Factors Engineering and Flight Deck Design.” In: *Digital avionics handbook*. Ed. by Cary R. Spitzer. 3rd. The electrical engineering handbook series. Boca Raton, Fla. [u.a.]: CRC, 2014.

[AMR13] Kathy Abbott, David McKenney, and Paul Railsback. *Operational Use of Flight Path Management Systems. Final Report of the Performance-based operations Aviation Rulemaking Committee / Commercial Aviation Safety Team Flight Deck Automation Working Group*. Federal Aviation Administration. 2013. URL: http://www.faa.gov/about/office_org/headquarters_offices/avs/offices/afs/afs400/parc/parc_reco/media/2013/130908_parc_fltdawg_final_report_recommendations.pdf (visited on 10/19/2016).

[ASS96] Kathy Abbott, Stephen M. Slotte, and Donald K. Stimson. *The Interfaces Between Flightcrews and Modern Flight Deck Systems*. Ed. by Federal Aviation Administration. Federal Aviation Administration. 1996. URL: http://ocw.mit.edu/courses/aeronautics-and-astronautics/16-422-human-supervisory-control-of-automated-systems-spring-2004/readings/interfac.pdf (visited on 07/04/2016).

[Abb93] Terence S. Abbott. “Functional categories for future flight deck designs.” In: (1993).

[AS18] Fethi Abdelmoula and Marco Scholz. “LNAS - A pilot assistance system for low-nose Approaches with minimal fuel consumption.” In: *ICAS 2018 - CD-ROM proceedings. 31st Congress of the International Council of the Aeronautical Sciences, Belo Horizonte, Brazil, September 09-14, 2018*. Ed. by International Council of the Aeronautical Sciences. International Council of the Aeronautical Sciences. Bonn: The International Council of the Aeronautical Sciences c/o Deutsche Gesellschaft für Luft- und Raumfahrt, 2018. ISBN: 978-3-93218288-4.

[AWG20] Fethi Abdelmoula, Jean Marc Wunderli, and Martin Gerber. *Pilotenassistenzsystem LNAS ermöglicht leisere und treibstoffsparende Anflüge*. Ed. by Deutsches Zentrum für Luft- und Raumfahrt e.V. 2020. URL: https://www.dlr.de/content/de/artikel/news/2020/02/20200622_testfluege-mit-dlr-forschungsflugzeug-a320-atra-in-zuerich.html (visited on 09/27/2020).

[Ack89] Russell Lincoln Ackoff. “From data to wisdom.” In: *Journal of Applied Systems Analysis* 16.1 (1989), pp. 3–9.

[Adl86] Mortimer Jerome Adler. *A guidebook to learning. For a lifelong pursuit of wisdom*. Macmillan Pub Co, 1986. ISBN: 978-0-02-500340-8.

[Aer94] Aeronautical Radio Incorporated, ed. *ARINC Specification 629-3*. 1994.

[ADN14] Aeronautics and Space Engineering Board, Division on Engineering and Physical Sciences, and National Research Council. *Autonomy Research for Civil Aviation*. Washington, D.C.: National Academies Press, 2014. ISBN: 978-0-309-30614-0. DOI: 10.17226/18815. URL: https://www.nap.edu/catalog/18815 (visited on 10/03/2020).

[Air13] Airbus. *A318/A319/A320/A321 Flight Crew Operating Manual*. 2013.

[Air96] Aircraft Accidents Investigation Commission. *Aircraft Accident Investigation Report. Nagoya Airport April 26, 1994*. Ed. by Aircraft Accidents Investigation Commission. Aircraft Accidents Investigation Commission, 1996. URL: http://www.mlit.go.jp/jtsb/eng-air_report/B1816.pdf (visited on 02/17/2019).

[Ann03] John Annett. “Hierarchical task analysis.” In: *Handbook of cognitive task design*. CRC Press, 2003, pp. 41–60.

[AD67] John Annett and Keith D. Duncan. “Task analysis and training design.” In: (1967).

[Avi04] Aviation Accidents Prevention and Investigation Department. *All Engines-out Landing Due to Fuel Exhaustion. Air Transat, Airbus A330-243 marks C-GITS*. Air Transat Airbus A330-243 marks C-GITS Lajes, Azores, Portugal 24 August 2001. Ed. by Aviation Accidents Prevention and Investigation Department. Aviation Accidents Prevention and Investigation Department, 2004.

[Bai+17] Randall E. Bailey et al. “An assessment of reduced crew and single pilot operations in commercial transport aircraft operations.” In: *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*. 2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC) (St. Petersburg, FL, Sept. 17, 2017–Sept. 21, 2019). IEEE, 2017, pp. 1–15. ISBN: 978-1-5386-0365-9. DOI: 10.1109/DASC.2017.8101988.

[Ban+08] Simon Banbury et al. *Intelligent Adaptive Systems. Literature-Research of Design Guidance for Intelligent Adaptive Automation and Interfaces*. Ed. by Defence Research and Development Canada. Defence Research and Development Canada, 2008.

[BKM08] Aaron Bangor, Philip T. Kortum, and James T. Miller. "An Empirical Evaluation of the System Usability Scale." In: *International Journal of Human-Computer Interaction* 24.6 (2008), pp. 574–594. ISSN: 1044-7318. DOI: 10.1080/10447310802205776.

[BL91] S. B. Banks and C. S. Lizza. "Pilot's Associate: a cooperative, knowledge-based system application." In: *IEEE Expert* 6.3 (1991), pp. 18–29. ISSN: 0885-9000. DOI: 10.1109/64.87681.

[Bat18] Tom Batchelor. "Single-Pilot Passenger Planes could soon take to the skies, says Boeing. Aircraft manufacturer also working on hypersonic jet to fly passengers from New York to Tokyo 'for lunch'." In: *The Independent* (Aug. 26, 2018). URL: https://www.independent.co.uk/travel/news-and-advice/single-pilot-plane-boeing-autonomous-jet-technology-cockpit-a8506301.html (visited on 02/23/2020).

[Bat+18] Vernol Battiste et al. "Human-Automation Teaming: Lessons Learned and Future Directions." In: *Human Interface and the Management of Information. Information in Applications and Services*. Ed. by Sakae Yamamoto and Hirohiko Mori. Vol. 10905. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018, pp. 479–493. ISBN: 978-3-319-92045-0. DOI: 10.1007/978-3-319-92046-7_40.

[BBR07] Gordon Baxter, Denis Besnard, and Dominic Riley. "Cognitive mismatches in the cockpit: will they ever be a thing of the past?" eng. In: *Applied ergonomics* 38.4 (2007). Journal Article Research Support, Non-U.S. Gov't, pp. 417–423. ISSN: 0003-6870. DOI: 10.1016/j.apergo.2007.01.005. eprint: 17448437.

[Beg12] Jean-Marie Begis. "Advances in Cockpit Display and Crew Information Management Capabilities." In: *SAE Technical Paper Series*. SAE 2012 Aerospace Electronics and Avionics Systems Conference (Oct. 30, 2010). SAE Technical Paper Series. SAE International400 Commonwealth Drive, Warrendale, PA, United States, 2012. DOI: 10.4271/2012-01-2121.

[BJS14] Karl D. Bilimoria, Walter W. Johnson, and Paul C. Schutte. "Conceptual framework for single pilot operations." In: *Proceedings of the International Conference on Human-Computer Interaction in Aerospace - HCI-Aero '14*. International Conference on Human-Computer Interaction in Aerospace (Santa Clara, California, July 30–Aug. 1, 2014). Ed. by Association for Computing Machinery. Association for Computing Machinery. New York, New York, USA: ACM Press, 2014, pp. 1–8. ISBN: 978-1-4503-2560-8. DOI: 10.1145/2669592.2669647.

[Bil91] Charles E. Billings. *Human-Centered Aircraft Automation: A Concept and Guidelines*. Ed. by National Aeronautics and Space Administration. National Aeronautics and Space Administration, 1991. URL: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19910022821.pdf (visited on 02/20/2019).

[Bil97] Charles E. Billings. *Aviation automation. The search for a human-centered approach*. eng. Human factors in transportation. Mahwah, N.J: Lawrence Erlbaum Associates Publishers, 1997. 355 pp. ISBN: 978-0-8058-2126-0. URL: http://www.loc.gov/catdir/enhancements/fy0738/96020189-d.html.

[Boe+19] Justus Boehm et al. "Entwicklung eines Versuchskonzepts für die Bewertung der Leistungsfähigkeit von Piloten unter hoher Arbeitsbelastung Development of a trial concept to assess the performance of pilots in high workload situations." Flight Systems and Automatic Control. Advanced Research Project. Darmstadt: Technische Universität Darmstadt, 2019.

[Boe14] The Boeing Company. *787 Flight Crew Operations Manual*. 2014.

[Boe16a] The Boeing Company. *2016 Pilot & Technician Outlook*. Support and Services Flight Operations Solutions. Ed. by The Boeing Company. 2016.

[Boe16b] The Boeing Company. *Touchscreens come to 777X flight deck. Bringing today's technology in the hands of pilots*. 2016. URL: https://www.boeing.com/features/2016/07/777x-touchscreen-07-16.page (visited on 09/05/2020).

[Boe18] The Boeing Company. *Commercial Airplanes Price List*. 2018. URL: https://www.boeing.com/company/about-bca/#/prices (visited on 11/14/2018).

[BTF00] Michael Bonner, Robert Taylor, and Keith Fletcher. *Adaptive Automation and Decision Aiding in the Military Fast Jet Domain*. Ed. by Her Majesty's Stationery Office. Defence Evaluation and Research Agency, 2000.

[Boo01] Daniel J. Boorman. “Safety Benefits of Electronic Checklists: An Analysis of Commercial Transport Accidents.” In: *Focusing Attention on Aviation Safety : the 11th International Symposium on Aviation Psychology, March 5-8, 2001, Columbus, Ohio, USA* (Columbus, OH). Ed. by International Symposium on Aviation Psychology et al. Ohio State University Dept. of Aviation and Aeronautical Engineering, 2001.

[BM04] Daniel J. Boorman and Randall J. Mumaw. *A New Autoflight/FMS Interface: Guiding Design Principles*. 2004.

[Boy98] Guy André Boy. “Cognitive function analysis for human-centered automation of safety-critical systems.” In: *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI ’98*. the SIGCHI conference (Los Angeles, California, United States, Apr. 18–23, 1998). Ed. by Michael E. Atwood et al. New York, New York, USA: ACM Press, 1998, pp. 265–272. ISBN: 978-0-201-30987-4. DOI: 10.1145/274644.274682.

[Boy14] Guy André Boy. “Requirements for Single Pilot Operations in Commercial Aviation: A First High-Level Cognitive Function Analysis.” In: (2014).

[BS18] Yannick Brand and Axel Schulte. “Design and Evaluation of a Workload-Adaptive Associate System for Cockpit Crews.” In: *Engineering Psychology and Cognitive Ergonomics* (Cham, 2018). Ed. by Don Harris. Cham: Springer International Publishing, 2018, pp. 3–18. ISBN: 978-3-319-91122-9.

[Bra+18a] Yannick Brand et al. “Design and Experimental Validation of Transparent Behavior for a Workload-Adaptive Cognitive Agent.” In: *Intelligent Human Systems Integration* (Cham, 2018). Ed. by Waldemar Karwowski and Tareq Ahram. Cham: Springer International Publishing, 2018, pp. 173–179. ISBN: 978-3-319-73888-8.

[Bra+15] Summer L. Brandt et al. “Pilot Situation Awareness and its Implications for Single Pilot Operations: Analysis of a Human-in-the-Loop Study.” In: *Procedia Manufacturing* 3 (2015). PII: S2351978915008471, pp. 3017–3024. ISSN: 23519789. DOI: 10.1016/j.promfg.2015.07.846.

[Bra+18b] Summer L. Brandt et al. “A Human-Autonomy Teaming Approach for a Flight-Following Task.” In: *Advances in Neuroergonomics and Cognitive Engineering*. Ed. by Carryl Baldwin. Vol. 586. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2018, pp. 12–22. ISBN: 978-3-319-60641-5. DOI: 10.1007/978-3-319-60642-2_2.

[Bro96] John Brooke. “SUS-A quick and dirty usability scale.” In: *Usability evaluation in industry* 189.194 (1996), pp. 4–7.

[Bro13] John Brooke. “SUS: A Retrospective.” In: *J. Usability Studies* 8.2 (2013), pp. 29–40. ISSN: 1931-3357. URL: http://dl.acm.org/citation.cfm?id=2817912.2817913.

[Bur+13] Barbara K. Burian et al. “Single Pilot Workload Management During Cruise in Entry Level Jets.” In: (2013).

[CJP96] Rodrigo C. Cabrera, Orlando R. Jimenez, and Saul G. Pertuz. *Controlled Flight Into Terrain American Airlines Flight 965, Boeing 757-223, N651AA. Near Cali, Colombia, December 20, 1995*. Ed. by Aeronautica Civil of the Republic of Colombia. Aeronautica Civil of the Republic of Colombia, 1996.

[CML14] Joan Cahill, Nick McDonald, and Gabriel Losa. “A sociotechnical model of the flight crew task.” eng. In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 56.8 (2014). Journal Article Research Support, Non-U.S. Gov’t, pp. 1337–1363. ISSN: 00187208. DOI: 10.1177/0018720814532684. eprint: 25509818.

[Car18] Bill Carey. “Airline Pilots Unions Join To Oppose Single-Pilot Research. Airline pilots’ unions and a business jet pilots association have formed an alliance to oppose a provision of draft FAA reauthorization legislation that calls for research into single-piloted cargo operations.” In: *Aviation Week & Space Technology* (Aug. 1, 2018). URL: https://aviationweek.com/air-transport/airline-pilots-unions-join-oppose-single-pilot-research (visited on 10/03/2020).

[CW83] John G. Casali and Walter W. Wierwille. “A comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications load.” eng. In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 25.6 (1983). Comparative Study Journal Article Research Support, U.S. Gov’t, Non-P.H.S., pp. 623–641. ISSN: 00187208. DOI: 10.1177/001872088302500602. eprint: 6671647.

[CW84] John G. Casali and Walter W. Wierwille. “On the measurement of pilot perceptual workload: a comparison of assessment techniques addressing sensitivity and intrusion issues.” eng. In: *Ergonomics* 27.10 (1984). Comparative Study Journal Article Research Support, Non-U.S. Gov’t Research Support, U.S. Gov’t, Non-P.H.S., pp. 1033–1050. ISSN: 0014-0139. DOI: 10.1080/00140138408963584. eprint: 6510406.

[CS15] Stephen M. Casner and Jonathan W. Schooler. "Vigilance impossible: Diligence, distraction, and daydreaming all lead to failures in a practical monitoring task." eng. In: *Consciousness and Cognition* 35 (2015). Journal Article Research Support, U.S. Gov't, Non-P.H.S., pp. 33–41. ISSN: 1053-8100. DOI: 10.1016/j.concog.2015.04.019. eprint: 25966369. URL: http://www.sciencedirect.com/science/article/pii/S1053810015000963.

[CF17] Jarrod Castle and Celine Fornano. *Flying solo - how far are we down the path towards pilotless planes?* Ed. by UBS Group AG. UBS Group AG, 2017.

[CMF96] Chung-Chi Chou, Das Madhavan, and Kenneth H. Funk. "Studies of Cockpit Task Management Errors." In: *The International Journal of Aviation Psychology* 6.4 (1996), pp. 307–320. ISSN: 1050-8414. DOI: 10.1207/s15327108ijap0604_1.

[CW02] Klaus Christoffersen and David D. Woods. "How to make automated systems team players." In: *Advances in Human Performance and Cognitive Engineering Research*. Ed. by Emerald Insight. 2002, pp. 1–12. DOI: 10.1016/S1479-3601(02)02003-9.

[CS14] Sebastian Clauss and Axel Schulte. "Task delegation in an agent supervisory control relationship capability awareness in a cognitive agent." In: *IEEE International Conference on Systems, Man and Cybernetics (SMC), 2014. 5 - 8 Oct. 2014, San Diego, CA, USA ; proceedings*. 2014 IEEE International Conference on Systems, Man and Cybernetics - SMC (San Diego, CA, USA). Institute of Electrical and Electronics Engineers et al. Piscataway, NJ: IEEE, 2014, pp. 825–830. ISBN: 978-1-4799-3840-7. DOI: 10.1109/SMC.2014.6974013.

[Cla+12] Sebastian Clauss et al. "Design and Evaluation of a UAS combining Cognitive Automation and Optimal Control." In: *Infotech@Aerospace Conferences. Infotech@Aerospace 2012*. Infotech@Aerospace 2012 (Garden Grove, California). [Place of publication not identified]: [publisher not identified], 2012, p. 27. ISBN: 978-1-60086-939-6. DOI: 10.2514/6.2012-2590.

[Coh92] Jacob Cohen. "A power primer." In: *Psychological Bulletin* 112.1 (1992), pp. 155–159. ISSN: 0033-2909. DOI: 10.1037/0033-2909.112.1.155.

[Com+13] Doreen Comerford et al. *NASA's Single-pilot operations technical interchange meeting: Proceedings and findings*. 2013.

[Coo+14] Alan Cooper et al. *About face. The essentials of interaction design*. eng. 4th edition. Indianapolis, IN: Wiley, 2014. 1 p. ISBN: 978-1-118-76657-6. URL: http://proquest.tech.safaribooksonline.de/9781118766583.

[CH69] G. E. Cooper and R. P. Harper JR. *The Use of Pilot Rating in the Evaluation of Aircraft Handling Qualities*. Ed. by North Atlantic Treaty Organization. 1969.

[CR13] Pedro Cruellas and Eric Roelants. *SWIM Concept of Operations*. European Organisation for the Safety of Air Navigation. 2013. URL: http://www.eurocontrol.int/publications/sesar-swim-concept-operations (visited on 05/08/2017).

[CR07] Jerry Crutchfield and Craig Rosenberg. *Predicting Subjective Workload Ratings: A Comparison and Synthesis of Operational and Theoretical Models*. Final Report. Federal Aviation Administration, 2007. URL: https://apps.dtic.mil/dtic/tr/fulltext/u2/a465365.pdf (visited on 02/15/2019).

[CSC16] Mary L. Cummings, Alex Stimpson, and Michael Clamann. "Functional Requirements for Onboard Intelligent Automation in Single Pilot Operations." In: *AIAA Infotech @ Aerospace*. AIAA Infotech @ Aerospace (San Diego, California, USA, Jan. 4–8, 2016). 2016-09-25/2016-09-29. Reston, Virginia: American Institute of Aeronautics and Astronautics, 2016. ISBN: 978-1-62410-388-9. DOI: 10.2514/6.2016-1652.

[Dao+15] Arik-Quang V. Dao et al. "Evaluation of a Recommender System for Single Pilot Operations." In: *Procedia Manufacturing* 3 (2015), pp. 3070–3077. ISSN: 23519789. DOI: 10.1016/j.promfg.2015.07.853.

[DBS13] Asaf Degani, Immanuel Barshi, and Michael G. Shafto. "Information Organization in the Airline Cockpit." In: *Journal of Cognitive Engineering and Decision Making* 7.4 (2013), pp. 330–352. ISSN: 1555-3434. DOI: 10.1177/1555343413492983.

[Dek02] Sidney Dekker. "Frequently asked questions about new technology and human error." In: *Human Factors and Safety in Aviation* (2002), p. 3.

[Dek04] Sidney W. A. Dekker. "On the other side of the promise: What should we automate today?" In: *Human factors for civil flight deck design*. Ed. by Don Harris. Aldershot: Ashgate, 2004, pp. 183–198. ISBN: 978-0-7546-1380-0.

[Det10] Brian Detlor. "Information management." In: *International Journal of Information Management* 30.2 (2010). PII: S0268401209001510, pp. 103–108. ISSN: 02684012. DOI: 10.1016/j.ijinfomgt.2009.12.001.

[DP05] Stephen Deutsch and Richard W. Pew. *Single Pilot Commercial Aircraft Operation*. National Aeronautics and Space Administration Ames Research Center, 2005.

[Deu08] Deutsches Zentrum für Luft- und Raumfahrt e.V. *PILAS (PILoten-ASsistenz-System für Luftfahrzeuge)*. Deutsches Zentrum für Luft- und Raumfahrt e.V. 2008. URL: https://www.dlr.de/fl/desktopdefault.aspx/tabid-1150/1741_read-5015 (visited on 10/05/2020).

[Dia+07] David M. Diamond et al. "The temporal dynamics model of emotional memory processing: a synthesis on the neurobiological basis of stress-induced amnesia, flashbulb and traumatic memories, and the Yerkes-Dodson law." eng. In: *Neural plasticity* 2007 (2007). Journal Article Research Support, N.I.H., Extramural Research Support, U.S. Gov't, Non-P.H.S. Review, p. 60803. DOI: 10.1155/2007/60803. eprint: 17641736.

[DIN11] DIN EN ISO, ed. *Ergonomie der Mensch-System-Interaktion - Teil 210: Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme*. 2011.

[DIN18] DIN EN ISO, ed. *Ergonomie der Mensch-System-Interaktion - Teil 11: Gebrauchstauglichkeit: Begriffe und Konzepte*. Berlin: Beuth Verlag GmbH, 2018. DOI: 10.31030/2757945.

[Dog10] Rigas S. Doganis. *Flying off course. Airline economics and marketing*. eng. 4th. ed. Doganis, Rigas S. (VerfasserIn). London: Routledge, 2010. 336 pp. ISBN: 978-0-415-44736-2.

[Dri+17] Kevin R. Driscoll et al. "Cyber safety and security for reduced crew operations (RCO)." In: *2017 IEEE Aerospace Conference*. 2017 IEEE Aerospace Conference (Big Sky, MT, USA, Mar. 4–11, 2017). IEEE, 2017, pp. 1–15. ISBN: 978-1-5090-1613-6. DOI: 10.1109/AERO.2017.7943785.

[Dua+15] Pengfei Duan et al. "Human-in-the-loop Evaluation of an Information Management and Notification System to Improve Aircraft State Awareness." In: *AIAA Infotech @ Aerospace* (Kissimmee, Florida). 2015. DOI: 10.2514/6.2015-0794.

[Eis19] Stefan Eiselin. "CR929 to become the first large single-pilot aircraft? China and Russia are jointly developing a new long-haul aircraft called CR929. It is supposed to be the first plane that can be operated by a single pilot." In: *aeroTELEGRAPH* 2019 (Dec. 30, 2019). URL: https://www.aerotelegraph.com/en/cr929-to-become-the-first-large-single-pilot-aircraft (visited on 01/03/2020).

[Emp10] Empresa Brasileira de Aeronáutica S.A., collab. *Type Certificate Data Sheet No. A60CE. A60CE*. Ed. by Federal Aviation Administration. Certificate. 2010.

[End88] Mica R. Endsley. "Design and Evaluation for Situation Awareness Enhancement." In: *Proceedings of the Human Factors Society Annual Meeting* 32.2 (1988), pp. 97–101. ISSN: 0163-5182. DOI: 10.1177/154193128803200221.

[End95] Mica R. Endsley. "Toward a Theory of Situation Awareness in Dynamic Systems." In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 37.1 (1995), pp. 32–64. ISSN: 00187208. DOI: 10.1518/001872095779049543.

[End06] Mica R. Endsley. "Situation awareness." In: *Handbook of human factors and ergonomics*. Ed. by Gavriel Salvendy. 3. ed. Hoboken: Wiley, 2006, pp. 528–542. ISBN: 978-0-470-04820-7.

[End16] Mica R. Endsley. *Designing for Situation Awareness. An Approach to User-Centered Design, Second Edition*. 2nd. Endsley, Mica R., (author). CRC Press, 2016. 1 online resource (396). ISBN: 978-1-4200-6358-5.

[EB94] Mica R. Endsley and Cheryl A. Bolstad. "Individual Differences in Pilot Situation Awareness." In: *The International Journal of Aviation Psychology* 4.3 (1994), pp. 241–264. ISSN: 1050-8414. DOI: 10.1207/s15327108ijap0403_3.

[EJ12] Mica R. Endsley and Debra G. Jones. *Designing for situation awareness. An approach to user-centered design*. eng. Second edition. Endsley, Mica R. (VerfasserIn) Jones, Debra G. (VerfasserIn). Boca Raton, London, and New York: CRC Press Taylor & Francis Group, 2012. 370 pp. ISBN: 978-1-4200-6355-4.

[Eth+19] Timothy Etherington et al. "Evaluation of Technology Concepts for Traffic data Management and Relevant audio for Datalink in Commercial Airline Flight Decks." In: *2019 Integrated Communications, Navigation and Surveillance Conference (ICNS)*. 2019 Integrated Communications, Navigation and Surveillance Conference (ICNS) (Herndon, VA, USA, Apr. 9–11, 2019). IEEE, 2019, pp. 1–10. ISBN: 978-1-7281-1893-2. DOI: 10.1109/ICNSURV.2019.8735135.

[Eth+16] Timothy J. Etherington et al. “Quantifying pilot contribution to flight safety for normal and non-normal airline operations.” In: *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC) (Sacramento, CA, USA, Sept. 25–29, 2016). IEEE, 2016, pp. 1–14. ISBN: 978-1-5090-2523-7. DOI: 10.1109/DASC.2016.7778094.

[Eth+18] Timothy J. Etherington et al. “Impact of Advanced Synoptics and Simplified Checklists during Aircraft Systems Failures.” In: *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*. 2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC) (London, United Kingdom, Sept. 23–27, 2018). IEEE, 2018, pp. 1–9. ISBN: 978-1-5386-4112-5. DOI: 10.1109/DASC.2018.8569559.

[Eur18] European Aviation Safety Agency. *CS-25 Large Aeroplanes*. CS-25. Version Amendment 22. 2018. URL: https://www.easa.europa.eu/document-library/certification-specifications/cs-25-amendment-22 (visited on 11/18/2018).

[Eur10] European Organisation for the Safety of Air Navigation. *Long-Term Forecast. Flight Movements 2010-2030*. European Organisation for the Safety of Air Navigation, 2010.

[Eur13a] European Organisation for the Safety of Air Navigation. *Challenges of Growth. Summary Report*. European Organisation for the Safety of Air Navigation, 2013.

[Eur17] European Organisation for the Safety of Air Navigation. *Specifications for System Wide Information Management (SWIM)*. European Organisation for the Safety of Air Navigation. 2017. URL: https://www.eurocontrol.int/publications/eurocontrol-specifications-system-wide-information-management-swim (visited on 02/13/2019).

[Eur04] European Union. *Regulation (EC) No 549/2004 of the European Parliament and of the Council of 10 March 2004 laying down the framework for the creation of the single European sky (the framework Regulation) (Text with EEA relevance) - Statement by the Member States on military issues related to the single European sky*. OJ L 96, 31.3.2004, p. 1–9 (ES, DA, DE, EL, EN, FR, IT, NL, PT, FI, SV). 2004. URL: https://eur-lex.europa.eu/eli/reg/2004/549/oj (visited on 11/19/2018).

[Eur12] European Union. *Standardised European Rules of the Air*. SERA. 2012.

[Eur13b] European Union. *Advanced Cockpit for Reduction Of StreSs and workload. Project Information*. Record number: 106358. European Union. 2013. URL: https://cordis.europa.eu/project/id/314501 (visited on 10/05/2020).

[Eva+16] Emory Evans et al. “Analysis of pilot feedback regarding the use of state awareness technologies during complex situations.” In: *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC) (Sacramento, CA, USA, Sept. 25–29, 2016). IEEE, 2016, pp. 1–9. ISBN: 978-1-5090-2523-7. DOI: 10.1109/DASC.2016.7777995.

[Fed88] Federal Aviation Administration. *AC 25.1309-1 - System Design and Analysis*. AC 25.1309-1. 1988. URL: https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_25_1309-1A.pdf (visited on 05/03/2018).

[Fed93] Federal Aviation Administration. *AC 25.773-1 - Pilot Compartment View Design Considerations*. AC 25.773-1. 1993.

[Fed05] Federal Aviation Administration. *AC 25.1523-1 - Minimum Flightcrew*. FAR/CS 25.1523. 2005. URL: https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/22690 (visited on 10/19/2016).

[Fed10] Federal Aviation Administration. *AC 25.1322-1 - Flightcrew alerting*. AC 25.1322-1. 2010. URL: https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_25.1322-1.pdf (visited on 08/31/2019).

[Fed11] Federal Aviation Administration. *1375.1E - Information/Data Management*. Order 1375.1E. 2011. URL: https://www.faa.gov/documentLibrary/media/Order/1375.1%20E.pdf (visited on 11/10/2018).

[Fed14] Federal Aviation Administration. *AC.25-11B - Electronic Flight Displays*. AC.25-11B. 2014.

[Fed15] Federal Aviation Administration. *Safety Alert for Operators. Roles and Responsibilities for Pilot Flying (PF) and Pilot Monitoring (PM)*. SAFO 15011. Ed. by Federal Aviation Administration. Federal Aviation Administration, 2015. URL: https://www.faa.gov/other_visit/aviation_industry/airline_operators/airline_safety/safo/all_safos/media/2015/SAFO15011.pdf (visited on 05/11/2017).

[Fed16a] Federal Aviation Administration. *Economic Values for FAA Investment and Regulatory Decisions, a guide. Final Report*. Ed. by Federal Aviation Administration. Washington, DC, 2016. URL: https://www.faa.gov/regulations_policies/policy_guidance/benefit_cost/ (visited on 11/14/2018).

[Fed16b] Federal Aviation Administration. *The Future of the NAS*. Federal Aviation Administration, 2016.

[Fed17] Federal Aviation Administration. *AC 120-76D - Authorization for Use of Electronic Flight Bags*. AC 120-76D. 2017.

[Fed19] Federal Aviation Administration. *FAA Airman Knowledge Testing Applicant Identification, Information Verification, & Authorization Requirements Matrix*. 2019. URL: https://www.faa.gov/training_testing/testing/media/testing_matrix.pdf (visited on 11/03/2019).

[Fed20] Federal Aviation Administration. *FAR/AIM Sec. 91 - General Operating and Flight Rules*. 2020.

[Fed] Federal Aviation Administration. *FAR/AIM Sec. 91.3 - Responsibility and authority of the pilot in command*. FAR 91.3.

[FSW10] Thomas Ferris, Nadine Barbara Sarter, and Christopher D. Wickens. "Cockpit Automation: Still Struggling to Catch Up." In: *Human factors in aviation*. Ed. by Eduardo Salas and Daniel E. Maurino. 2nd ed. Amsterdam and Boston: Academic Press/Elsevier, 2010, pp. 479–503. ISBN: 978-0-12-374518-7.

[FMF13] Andy Field, Jeremy Miles, and Zoë Field. *Discovering statistics using R*. eng. Reprint. Los Angeles, Calif.: Sage, 2013. 957 S. ISBN: 978-1-4462-0046-9.

[FHN72] Richard E. Fikes, Peter E. Hart, and Nils J. Nilsson. "Learning and executing generalized robot plans." In: *Artificial Intelligence* 3 (1972). PII: 0004370272900513, pp. 251–288. ISSN: 00043702. DOI: 10.1016/0004-3702(72)90051-3.

[FN71] Richard E. Fikes and Nils J. Nilsson. "Strips: A new approach to the application of theorem proving to problem solving." In: *Artificial Intelligence* 2.3-4 (1971). PII: 0004370271900105, pp. 189–208. ISSN: 00043702. DOI: 10.1016/0004-3702(71)90010-5.

[FN94] Richard E. Fikes and Nils J. Nilsson. "STRIPS, a retrospective." In: *Artificial intelligence in perspective* 227 (1994).

[Fli14] Flight Safety Foundation. *A Practical Guide for Improving Flight Path Monitoring. Final Report of the Active Pilot Monitoring Working Group*. Ed. by Flight Safety Foundation. Flight Safety Foundation, 2014.

[Fri08] Martin Fricke. “The knowledge pyramid. A critique of the DIKW hierarchy.” In: *Journal of Information Science* 35.2 (2008), pp. 131–142. ISSN: 0165-5515. DOI: 10.1177/0165551508094050.

[Fuc06] Tatjana Fuchs. *Was ist gute Arbeit? Anforderungen aus der Sicht von Erwerbstätigen*. Konzeption und Auswertung einer repräsentativen Untersuchung. Ed. by Initiative Neue Qualität der Arbeit. 2006. URL: http://www.inqa.de/SharedDocs/PDFs/DE/Publikationen/inqa-19-was-ist-gute-arbeit.pdf (visited on 02/20/2019).

[Fun91] Kenneth H. Funk. “Cockpit Task Management: Preliminary Definitions, Normative Theory, Error Taxonomy, and Design Recommendations.” In: *The International Journal of Aviation Psychology* 1.4 (1991), pp. 271–285. ISSN: 1050-8414. DOI: 10.1207/s15327108ijap0104_2.

[Fun97] Kenneth H. Funk. *An Agent-Based Cockpit Task Management System*. Ed. by National Aeronautics and Space Administration. National Aeronautics and Space Administration. 1997. URL: https://ntrs.nasa.gov/search.jsp?R=19970026046 (visited on 11/12/2018).

[Fun+97] Kenneth H. Funk et al. “Development and evaluation of an aid to facilitate agenda management.” In: *1997 IEEE International Conference on Systems, Man and Cybernetics*. 1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation (Orlando, FL, USA, Oct. 12–15, 1997). IEEE, Systems, Man and Cybernetics Society Staff. Piscataway: IEEE, 1997, pp. 3120–3125. ISBN: 978-0-7803-4053-4. DOI: 10.1109/ICSMC.1997.633069.

[Fun+99] Kenneth H. Funk et al. “Flight Deck Automation issues.” In: *The International Journal of Aviation Psychology* 9.2 (1999), pp. 109–123. ISSN: 1050-8414. DOI: 10.1207/s15327108ijap0902_2.

[Gaw08] Valerie J. Gawron. *Human performance, workload, and situational awareness measures handbook*. eng. 2nd ed. Boca Raton: CRC Press, 2008. 272 pp. ISBN: 978-1-4200-6449-0. URL: http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10216675.

[Ger+95] Marc Gerlach et al. “The cockpit assistant system cassy - design and in-flight evaluation.” In: *Human-computer interaction. Human and social aspects of human-computer interaction ; proceedings of the sixth International Conference on Human-Computer Interaction, (HCI International ’95), Tokyo, Japan, 9-14 July 1995, volume 2*. Ed. by Yūichirō Anzai, Katsuhiko Ogawa, and

Hirohiko Mori. Vol. 20. Advances in Human Factors/Ergonomics. International Conference on Human-Computer Interaction and ScienceDirect (Online service). Amsterdam and London: Elsevier, 1995, pp. 759–764. ISBN: 978-0-444-81795-2. DOI: 10.1016/S0921-2647(06)80120-6.

[Gra+14a] Jonathan Graham et al. "Design of a single pilot cockpit for airline operations." In: *2014 Systems and Information Engineering Design Symposium (SIEDS)*. 2014 Systems and Information Engineering Design Symposium (SIEDS) (Charlottesville, VA, USA, Apr. 25, 2014). IEEE, 2014, pp. 210–215. ISBN: 978-1-4799-4836-9. DOI: 10.1109/SIEDS.2014.6829879.

[Gra+14b] Oliver Grauers et al. *D10.1-1 Candidate Solutions for the Electronic Standby Pilot. Advanced Cockpit for Reduction Of Stress and Workload (ACROSS)*. British Aerospace Electronic Systems, 2014.

[GT96] Leo J. Gugerty and William C. Tirre. "Situation Awareness: A Validation Study and Investigation of Individual Differences." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 40.12 (1996), pp. 564–568. ISSN: 1541-9312. DOI: 10.1177/154193129604001202.

[Hal+96] James E. Hall et al. *Aircraft Accident Report Collision with Trees on Final Approach American Airlines Flight 1572. McDonnell Douglas MD-83, N566AA*. East Gransby, Connecticut. National Transportation Safety Board, 1996.

[Han13] Peter Adrian Hancock. "In search of vigilance: the problem of iatrogenically created psychological phenomena." eng. In: *The American psychologist* 68.2 (2013). Journal Article Research Support, U.S. Gov't, Non-P.H.S. Review, pp. 97–109. DOI: 10.1037/a0030214. eprint: 23088439.

[Han16] Theo Hankers. "Advanced pilot workflow support using integrated mobile devices." eng. Hankers, Theo (VerfasserIn) Hecker, Peter (AkademischeR BetreuerIn) Klingauf, Uwe (AkademischeR BetreuerIn) Schiefele, Jens (AkademischeR BetreuerIn). Dissertation. Technische Universität Braunschweig, 2016.

[Har07] Don Harris. "A human–centred design agenda for the development of single crew operated commercial aircraft." In: *Aircraft Engineering and Aerospace Technology* 79.5 (2007), pp. 518–526. ISSN: 0002-2667. DOI: 10.1108/00022660710780650.

[Har13] Don Harris. “Distributed Cognition in Flight Operations.” In: *Engineering Psychology and Cognitive Ergonomics. Applications and Services*. Ed. by David Hutchison et al. Vol. 8020. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 125–133. ISBN: 978-3-642-39353-2. DOI: 10.1007/978-3-642-39354-9_14.

[HNR68] Peter Hart, Nils Nilsson, and Bertram Raphael. “A Formal Basis for the Heuristic Determination of Minimum Cost Paths.” In: *IEEE Transactions on Systems Science and Cybernetics* 4.2 (1968), pp. 100–107. ISSN: 0536-1567. DOI: 10.1109/TSSC.1968.300136.

[Has+17] Pouya Haschemi et al. “Creation of a Recommendation-Catalogue for Future Flight Deck Display Design.” Institute of Flight Systems and Automatic Control. Advanced Research Project. Darmstadt: Technische Universität Darmstadt, 2017.

[Hil+92] Susan G. Hill et al. “Comparison of Four Subjective Workload Rating Scales.” In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 34.4 (1992), pp. 429–439. ISSN: 00187208. DOI: 10.1177/001872089203400405.

[Hir11] Mike Hirst. *The air transport system*. eng. Cambridge: Woodhead, 2011. 341 pp. ISBN: 978-1-84569-325-1. URL: http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=689373.

[Hut95] Edwin Hutchins. “How a Cockpit Remembers Its Speeds.” In: *Cognitive Science* 19.3 (1995), pp. 265–288. ISSN: 03640213. DOI: 10.1207/s15516709cog1903_1.

[Int18a] International Air Transportation Association. *Annual Review 2018*. Ed. by International Air Transportation Association. International Air Transportation Association, 2018. URL: https://www.iata.org/publications/Documents/iata-annual-review-2018.pdf (visited on 11/14/2018).

[Int05] International Civil Aviation Organization. *Annex 2 - Rules of the Air*. Version 10th Edition. 2005.

[Int13a] International Civil Aviation Organization. *Phase of Flight. Definition and Usage Notes*. Ed. by National Transportation Safety Board. International Civil Aviation Organization, 2013. URL: https://www.ntsb.gov/investigations/data/documents/datafiles/PhaseofFlightDefinitions.pdf (visited on 02/13/2019).

[Int13b] International Civil Aviation Organization. *Safety management manual (SMM)*. eng. 3. ed. Vol. 9859. Doc / International Civil Aviation Organization /Englische Ausgabe]. Montreal: ICAO, 2013. ISBN: 978-929249214-4.

[Int15] International Civil Aviation Organization. *IMP-Memo/1*. Ed. by International Civil Aviation Organization. International Civil Aviation Organization. 2015. URL: https://www.icao.int/airnavigation/IMP/Memos/IMP%20Memo%201.pdf (visited on 11/10/2018).

[Int16] International Civil Aviation Organization. *Doc 4444. Procedures for Air Navigation Services Air Traffic Managemnt*. Ed. by International Civil Aviation Organization. International Civil Aviation Organization, 2016.

[Int18b] International Civil Aviation Organization. *Annex 11 - Air Traffic Services*. Version 15th Edition. 2018.

[Int18c] International Civil Aviation Organization. *Annex 8 - Airworthiness of Aircraft*. Annex 8. Version 12th Edition. 2018.

[Int18d] International Civil Aviation Organization. *Manual of Eletronic Flight Bags (EFBs) (Doc 10020)*. Version 2nd Edition. 2018.

[JGW10] Alan R. Jacobsen, David A. Graeber, and John Wiedemann. "Crew Station Design and Integration." In: *Human factors in aviation*. Ed. by Eduardo Salas and Daniel E. Maurino. 2nd ed. Amsterdam and Boston: Academic Press/Elsevier, 2010, pp. 533–562. ISBN: 978-0-12-374518-7.

[Jas04] Ashok Jashapara. *Knowledge management. An integrated approach*. eng. Harlow: Financial Times/Prentice Hall, 2004. 324 pp. ISBN: 978-0273682981.

[Jen+99] Florian Jentsch et al. "Who Is Flying This Plane Anyway? What Mishaps Tell Us about Crew Member Role Assignment and Air Crew Situation Awareness." In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 41.1 (1999), pp. 1–14. ISSN: 00187208. DOI: 10.1518/001872099779577237.

[Joi11] Joint Planning and Development Office. *Concept of Operations for the Next Generation Air Transportation System. Version 3.2*. Joint Planning and Development Office. 2011. URL: http://www.dtic.mil/dtic/tr/fulltext/u2/a535795.pdf (visited on 07/04/2016).

[Joi14] Joint Safety Implementation Team. *Airplane State Awareness Joint Safety Implementation Team. Analysis and Recommendations*. Final Report. Ed. by Joint Safety Implementation Team. Joint Safety Implementation Team, 2014.

[JR95] Jon E. Jonsson and Wendell R. Ricks. *Cognitive Models of Pilot Categorization and Prioritization of Flight-Deck Information*. National Aeronautics and Space Administration Langley Technical Report Server. 1995.

[JCM03] Marcel Adam Just, Patricia A. Carpenter, and Akira Miyake. “Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work.” In: *Theoretical Issues in Ergonomics Science* 4.1-2 (2003), pp. 56–88. DOI: 10.1080/14639220210159735.

[Kin19] Max Kingsley-Jones. “Airbus: Single-pilot freighters a step to airliner operations.” In: *FlightGlobal* (Sept. 27, 2019). URL: https://www.flightglobal.com/systems-and-interiors/airbus-single-pilot-freighters-a-step-to-airliner-operations/134505.article (visited on 01/03/2020).

[Lac+12] Marie Lacabanne et al. “Analysis of the flight task around different types of aircraft.” In: *CLLE-LTC and Octogone-Lordat laboratories & THALES AVIONICS* (2012).

[Lac+14a] Joel Lachter et al. “Toward single pilot operations.” In: *Proceedings of the International Conference on Human-Computer Interaction in Aerospace - HCI-Aero ’14*. International Conference on Human-Computer Interaction in Aerospace (Santa Clara, California, July 30–Aug. 1, 2014). Ed. by Association for Computing Machinery. Association for Computing Machinery. New York, New York, USA: ACM Press, 2014, pp. 1–8. ISBN: 978-1-4503-2560-8. DOI: 10.1145/2669592.2669685.

[Lac+14b] Joel Lachter et al. “Toward single pilot operations.” In: *Proceedings of the International Conference on Human-Computer Interaction in Aerospace - HCI-Aero ’14*. International Conference on Human-Computer Interaction in Aerospace (Santa Clara, California, July 30–Aug. 1, 2014). Ed. by Association for Computing Machinery. Association for Computing Machinery. New York, New York, USA: ACM Press, 2014, pp. 1–8. ISBN: 978-1-4503-2560-8. DOI: 10.1145/2669592.2669695.

[Lac+17] Joel Lachter et al. “Enhanced ground support: lessons from work on reduced crew operations.” In: *Cognition, Technology & Work* 19.2-3 (2017). PII: 422, pp. 279–288. ISSN: 1435-5558. DOI: 10.1007/s10111-017-0422-6.

[LM14] Paola Lanzi and Thierry Maret. *Introduction to ACROSS project. A European Commission Seventh Framework Programme*. Rome, July 31, 2014.

[Let+12] Emmanuel Letsu-Dake et al. “Innovative Flight Deck Function Allocation Concepts for NextGen.” In: *Advances in Human Aspects of Aviation* (2012), p. 301. ISSN: 14398711.

[Lig+15] Sarah V. Ligda et al. “Effectiveness of Advanced Collaboration Tools on Crew Communication in Reduced Crew Operations.” In: *Engineering Psychology and Cognitive Ergonomics*. Ed. by Don Harris. Vol. 9174. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015, pp. 416–427. ISBN: 978-3-319-20372-0. DOI: 10.1007/978-3-319-20373-7_40.

[Lom+15] Thomas Lombaerts et al. “Piloted Simulator Evaluation of Maneuvering Envelope Information for Flight Crew Awareness.” In: *AIAA Guidance, Navigation, and Control Conference*. AIAA Guidance, Navigation, and Control Conference (Kissimmee, Florida). Reston, Virginia: American Institute of Aeronautics and Astronautics, 2015. ISBN: 978-1-62410-339-1. DOI: 10.2514/6.2015-1546.

[LDB09] Loukia D. Loukopoulos, R. Key Dismukes, and Immanuel Barshi. *The multitasking myth. Handling complexity in real-world operations*. Ashgate studies in human factors for flight operations. Farnham: Ashgate Publ, 2009. XIV, 188 S. ISBN: 978-0-7546-7997-4.

[Lov95] Edward Lovesey. “Information flow between cockpit and aircrew.” In: *Ergonomics* 38.3 (1995), pp. 558–564. ISSN: 0014-0139. DOI: 10.1080/00140139508925128.

[Ma12] Lai Ma. “Meanings of information. The assumptions and research consequences of three foundational LIS theories.” In: *Journal of the American Society for Information Science and Technology* 63.4 (2012), pp. 716–723. ISSN: 15322882. DOI: 10.1002/asi.21711.

[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*. 16th AIAA Aviation Technology, Integration, and Operations Conference (Washington, D.C). Reston, Virginia: American Institute of Aeronautics and Astronautics, 2016, p. 1. ISBN: 978-1-62410-440-4. DOI: 10.2514/6.2016-3303.

[Mas18] Massachusetts Institute of Technology. *Definition of Block Hour. Global Airline Industry Program, Airline Data Project, Glossary*. 2018. URL: web.mit.edu/airlinedata/www/Res_Glossary.html (visited on 11/18/2018).

[Mat+17] Michael Matessa et al. *Concept of Operations for RCO/SPO*. National Aeronautics and Space Administration, 2017. URL: https://ntrs.nasa.gov/search.jsp?R=20170007262 (visited on 12/01/2019).

[Mat+18] Michael Matessa et al. "Using Distributed Simulation to Investigate Human-Autonomy Teaming." In: *Human Interface and the Management of Information. Information in Applications and Services*. Ed. by Sakae Yamamoto and Hirohiko Mori. Vol. 10905. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018, pp. 541–550. ISBN: 978-3-319-92045-0. DOI: 10.1007/978-3-319-92046-7_45.

[Men13] Heinrich Mensen. *Handbuch der Luftfahrt*. ger. 2., neu bearb. Aufl. 2013. VDI-Buch. Mensen, Heinrich (author.) Berlin, Heidelberg and s.l.: Springer Berlin Heidelberg, 2013. 286 pp. ISBN: 978-3-642-34402-2. DOI: 10.1007/978-3-642-34402-2. URL: http://dx.doi.org/10.1007/978-3-642-34402-2.

[Mes88] Najmedin Meshkati. "Toward Development of a Cohesive Model of Workload." In: *Human Mental Workload*. Ed. by Peter Adrian Hancock. Vol. 52. Advances in Psychology 52. Elsevier, 1988, pp. 305–314. ISBN: 978-0-444-70388-0. DOI: 10.1016/S0166-4115(08)62394-8.

[ML88] Najmedin Meshkati and Alex Loewenthal. "An eclectic and critical review of four primary mental workload assessment methods: A guide for developing a comprehensive model." In: *Human Mental Workload*. Ed. by Peter Adrian Hancock. Vol. 52. Advances in Psychology 52. Elsevier, 1988, pp. 251–267. ISBN: 978-0-444-70388-0.

[MH99] C. A. Miller and M. D. Hannen. "The Rotorcraft Pilot's Associate: design and evaluation of an intelligent user interface for cockpit information management." In: *Knowledge-Based Systems* 12.8 (1999). PII: S0950705199000350, pp. 443–456. ISSN: 09507051. DOI: 10.1016/S0950-7051(99)00035-0.

[Mor67] N. Moray. "Where is capacity limited? A survey and a model." In: *Acta Psychologica* 27 (1967), pp. 84–92. ISSN: 00016918. DOI: 10.1016/0001-6918(67)90048-0.

[Mor14] Michael J. Morgan. "Boeing B-777 Avionics Architecture." In: *Digital avionics handbook*. Ed. by Cary R. Spitzer. 3rd. The electrical engineering handbook series. Boca Raton, Fla. [u.a.]: CRC, 2014.

[MdL18] Daniel Mosquera Benitez, Antonio del Corte Valiente, and Paola Lanzi. "A novel global operational concept in cockpits under peak workload situations." In: *Safety Science* 102 (2018). PII: S0925753517303971, pp. 38–50. ISSN: 09257535. DOI: 10.1016/j.ssci.2017.09.028.

[Nei20] Stefan Manuel Neis. "Evaluation of Pilot Vigilance during Cruise towards the Implementation of Reduced Crew Operations." Dissertation. Darmstadt: Technische Universität Darmstadt, 2020. URL: 10.25534/TUPRINTS-00011376.

[NKS18] Stefan Manuel Neis, Uwe Klingauf, and Jens Schiefele. "Classification and Review of Conceptual Frameworks for Commercial Single Pilot Operations." In: *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*. 2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC) (London, United Kingdom, Sept. 23–27, 2018). IEEE, 2018, pp. 1–8. ISBN: 978-1-5386-4112-5. DOI: 10.1109/DASC.2018.8569680.

[Neu+19] Xenija Neufeld et al. "Building a Planner: A Survey of Planning Systems Used in Commercial Video Games." In: *IEEE Transactions on Games* 11.2 (2019), pp. 91–108. ISSN: 2475-1502. DOI: 10.1109/TG.2017.2782846.

[Nor93] Donald Arthur Norman. *Things that make us smart. Defending human attributes in the age of the machine*. eng. [16. print.] Norman, Donald A. (VerfasserIn). Reading, Mass.: Perseus Books, 1993. 290 pp. ISBN: 978-0-201-58129-4.

[Nor14] Donald Arthur Norman. *Things That Make Us Smart. Defending Human Attributes in the Age of the Machine*. New York: Diversion Books, 2014. 1 online resource. ISBN: 978-1-62681-537-7. URL: https://books.google.de/books?id=yPKkBQAAQBAJ.

[OE86] Robert D. O'Donnell and F. Thomas Eggemeier. "Workload assessment methodology." In: *Handbook of perception and human performance, Vol. 2: Cognitive processes and performance*. Oxford, England: John Wiley & Sons, 1986, pp. 1–49. ISBN: 978-0-471-82956-0.

[OHa97] D. O'Hare. "Cognitive ability determinants of elite pilot performance." eng. In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 39.4 (1997). Journal Article Research Support, Non-U.S. Gov't Journal Article Research Support, Non-U.S. Gov't, pp. 540–552. ISSN: 00187208. DOI: 10.1518/001872097778668004. eprint: 9473974.

[Obj11] Object Management Group. *Business Process Model and Notation (BPMN). Version 2.0*. Object Management Group. 2011. URL: http://www.omg.org/spec/BPMN/2.0/PDF/ (visited on 05/12/2017).

[Onk96] Reiner Onken. "Human-centered Cockpit Design through the Knowledge-based Cockpit Assistant System (CASSY)." In: *Improving Function Allocation for Integrated Systems Design*. Ed. by David Beevis, Peter Essens, and Herke Schuffel. 1996, pp. 179–198.

[OP94] Reiner Onken and Thomas Prévôt. "Cockpit Assistant System for IFR Operation." In: *19th Congress of the International Council of the Aeronautical Sciences. Anaheim, California, U.S.A., 18 - 23 September 1994*. Ed. by International Council of the Aeronautical Sciences. ICAS proceedings 1994,3. International Council of the Aeronautical Sciences. New York, NY, 1994. ISBN: 1563470845.

[OS10] Reiner Onken and Axel Schulte. *System-ergonomic design of cognitive automation. Dual-mode cognitive design of vehicle guidance and control work systems*. eng. Vol. 235. Studies in computational intelligence. Berlin: Springer, 2010. 383 pp. ISBN: 978-3-64203135-9.

[OW01] Reiner Onken and Anton Walsdorf. "Assistant systems for aircraft guidance: cognitive man-machine cooperation." In: *Aerospace Science and Technology* 5.8 (2001). PII: S1270963801011373, pp. 511–520. ISSN: 1270-9638. DOI: 10.1016/S1270-9638(01)01137-3. URL: http://www.sciencedirect.com/science/article/pii/S1270963801011373.

[Ork03] Jeff Orkin. *Applying Goal-Oriented Action Planning to Games*. 2003.

[Ork05] Jeff Orkin. "Agent architecture considerations for real-time planning in games." In: *Proceedings of the First AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*. 3022492. Marina del Rey, California: AAAI Press, 2005, pp. 105–110.

[Ork06] Jeff Orkin. "Three States and a Plan: The AI of F.E.A.R." In: *Game Developers Conference* (San Jose, Mar. 20–24, 2006). Ed. by CMP Media. 2006.

[Ost17] Jon Ostrower. "The U.S. will face a staggering shortage of pilots." In: *CNN Money* (July 31, 2017). URL: http://money.cnn.com/2017/07/27/news/companies/pilot-shortage-figures/index.html (visited on 09/13/2017).

[Pal+95] Michael T. Palmer et al. *A Crew-Centered Flight Deck Design Philosophy for High-Speed Civil Transport (HSCT) Aircraft*. 1995.

[PSW08] Raja Parasuraman, Thomas B. Sheridan, and Christopher D. Wickens. “Situation Awareness, Mental Workload, and Trust in Automation. Viable, Empirically Supported Cognitive Engineering Constructs.” In: *Journal of Cognitive Engineering and Decision Making* 2.2 (2008), pp. 140–160. ISSN: 1555-3434. DOI: 10.1518/155534308X284417.

[PS13] Keri E. Pearlson and Carol S. Saunders. *Managing and using information systems. A strategic approach*. 5th ed. Hoboken, NJ: John Wiley & Sons, 2013. xiv, 401. ISBN: 978-1-118-28173-4.

[Ped89] Edwin P. D. Pednault. “ADL: Exploring the Middle Ground Between STRIPS and the Situation Calculus.” In: (1989).

[Rad00] Radio Technical Commission for Aeronautics, ed. *Design Assurance Guidance for Airborne Electronic Hardware*. DO. 2000.

[Rad12] Radio Technical Commission for Aeronautics, ed. *Software Considerations in Airborne Systems and Equipment Certification*. DO. 2012.

[Ras83] Jens Rasmussen. “Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models.” In: *IEEE Transactions on Systems, Man, and Cybernetics* SMC-13.3 (1983), pp. 257–266. ISSN: 0018-9472. DOI: 10.1109/TSMC.1983.6313160.

[Ras86] Jens Rasmussen. *Information processing and human-machine interaction. An approach to cognitive engineering*. eng. Vol. 12. North-Holland series in system science and engineering. New York, N.Y.: North-Holland, 1986. 215 pp. ISBN: 978-0-444-00987-6.

[Row07] Jennifer Rowley. “The wisdom hierarchy. Representations of the DIKW hierarchy.” In: *Journal of Information Science* 33.2 (2007), pp. 163–180. ISSN: 0165-5515. DOI: 10.1177/0165551506070706.

[Rus+16] Stuart J. Russell et al. *Artificial intelligence. A modern approach*. eng. Third edition, Global edition. Always learning. Boston et al.: Pearson, 2016. 1132 pp. ISBN: 978-1-292-15396-4.

[S-1a] S-18, Aircraft and Sys Dev and Safety Assessment Committee, ed. *Guidelines and Methods for Conducting the Safety Assessment Process on Civil Airborne Systems and Equipment*. ARP. 400 Commonwealth Drive Warrendale PA United States: SAE International. DOI: 10.4271/ARP4761.

[S-1b] S-18, Aircraft and Sys Dev and Safety Assessment Committee, ed. *Safety Assessment of Transport Airplanes in Commercial Service*. ARP. 400 Commonwealth Drive Warrendale PA United States: SAE International. DOI: 10.4271/ARP5150.

[SMC10] Eduardo Salas, Daniel E. Maurino, and Michael Curtis. “Human Factors in Aviation: An Overview.” In: *Human factors in aviation*. Ed. by Eduardo Salas and Daniel E. Maurino. 2nd ed. Amsterdam and Boston: Academic Press/Elsevier, 2010, pp. 3–22. ISBN: 978-0-12-374518-7.

[SMW07] Nadine B. Sarter, Randall J. Mumaw, and Christopher D. Wickens. “Pilots’ monitoring strategies and performance on automated flight decks: an empirical study combining behavioral and eye-tracking data.” eng. In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 49.3 (2007). Journal Article Research Support, U.S. Gov’t, Non-P.H.S., pp. 347–357. ISSN: 00187208. DOI: 10.1518/001872007X196685. eprint: 17552302.

[SW91] Nadine Barbara Sarter and David D. Woods. “Situation Awareness. A Critical But Ill-Defined Phenomenon.” In: *The International Journal of Aviation Psychology* 1.1 (1991), pp. 45–57. ISSN: 1050-8414. DOI: 10.1207/s15327108ijap0101_4.

[SW92] Nadine Barbara Sarter and David D. Woods. “Pilot Interaction With Cockpit Automation. Operational Experiences With the Flight Management System.” In: *The International Journal of Aviation Psychology* 2.4 (1992), pp. 303–321. ISSN: 1050-8414. DOI: 10.1207/s15327108ijap0204_5.

[SW95] Nadine Barbara Sarter and David D. Woods. “How in the World Did We Ever Get into That Mode? Mode Error and Awareness in Supervisory Control.” In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 37.1 (1995), pp. 5–19. ISSN: 0018-7208. DOI: 10.1518/001872095779049516.

[SW97] Nadine Barbara Sarter and David D. Woods. “Team Play with a Powerful and Independent Agent. Operational Experiences and Automation Surprises on the Airbus A-320.” In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 39.4 (1997), pp. 553–569. ISSN: 00187208. DOI: 10.1518/001872097778667997.

[SWB97] Nadine Barbara Sarter, David D. Woods, and Charles E. Billings. “Automation surprises.” In: *Handbook of human factors and ergonomics*. Ed. by Gavriel Salvendy. 2. ed. Vol. 2. A Wiley Interscience publication. New York, NY: Wiley, 1997, pp. 1926–1943. ISBN: 978-0-471-11690-5.

[Sat93] Paul M. Satchell. *Cockpit Monitoring and Alerting Systems*. eng. Satchell, Paul M. (VerfasserIn). Florence: Taylor and Francis, 1993. 1193 pp. ISBN: 978-1-85742-109-5. URL: http://gbv.eblib.com/patron/FullRecord.aspx?p=4758687.

[SLB10] Christopher M. Schlick, Holger Luczak, and Ralph Bruder. *Arbeitswissenschaft*. ger. Heidelberg: Springer, 2010. 1192 pp. ISBN: 978-3-540-78333-6. URL: http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10361974.

[Sch+18] Jonas Schmelz et al. "Advanced Design Project: Entwicklung und Aufbau des Innenraums und der Infrastruktur eines Cockpitsimulators. Development and Construction of the Interior and Infrastructure of a cockpit simulator." Institute of Flight Systems and Automatic Control. Advanced Design Project. Darmstadt: Technische Universität Darmstadt, 2018.

[SK17] Daniela Schmid and Bernd Korn. "A Tripartite Concept of a Remote-Copilot Center for Commercial Single-Pilot Operations." In: *AIAA Information Systems-AIAA Infotech @ Aerospace*. AIAA Information Systems-AIAA Infotech @ Aerospace (Grapevine, Texas). Reston, Virginia: American Institute of Aeronautics and Astronautics, 2017, p. 3303. ISBN: 978-1-62410-449-7. DOI: 10.2514/6.2017-0064.

[SS20] Daniela Schmid and Neville A. Stanton. "Progressing Toward Airliners' Reduced-Crew Operations: A Systematic Literature Review." In: *The International Journal of Aerospace Psychology* 30.1-2 (2020), pp. 1–24. ISSN: 2472-1840. DOI: 10.1080/24721840.2019.1696196.

[SG16] Dieter Schmitt and Volker Gollnick. *Air Transport System*. eng. 1st ed. 2016. Schmitt, Dieter (author.) Gollnick, Volker (author.) Wien: Springer, 2016. 377 pp. ISBN: 978-3-7091-1880-1. DOI: 10.1007/978-3-7091-1880-1. URL: http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1078897.

[SS99] Alex Schulte and Peter Stütz. "Evaluation of the Cockpit Assistant Military Aircraft CAMA in Simulator Trials." In: *RTO SCI Symposium on "Sensor Data Fusion and Integration of the Human Element"*. (Ottawa, Canada, Sept. 14–16, 1998). 1999.

[Sch13] Axel Schulte. "Kognitive und kooperative Automation zur Führung unbemannter Luftfahrzeuge." de. In: *Kognitive Systeme* 2013.1 (2013). Jul. DOI: 10.17185/duepublico/31350. URL: https://duepublico2.uni-due.de/receive/duepublico_mods_00031350.

[Sch15] Paul C. Schutte. “How to Make the Most of Your Human. Design Considerations for Single Pilot Operations.” In: *Engineering Psychology and Cognitive Ergonomics*. Ed. by Don Harris. Vol. 9174. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015, pp. 480–491. ISBN: 978-3-319-20372-0. DOI: 10.1007/978-3-319-20373-7_46.

[ST96] Paul C. Schutte and Anna C. Trujillo. “Flight crew task management in non-normal situations.” In: *Proceedings of the human factors and ergonomics society annual meeting*. Ed. by SAGE. 40. SAGE Publications Sage CA: Los Angeles, CA, 1996. ISBN: 978-0-945289-06-7.

[Sch+07] Paul C. Schutte et al. *The naturalistic flight deck system: An integrated system concept for improved single-pilot operations*. National Aeronautics and Space Administration Ames Research Center, 2007.

[Sch+12] Roger W. Schvaneveldt et al. *Weather in the Cockpit: Priorities, Sources, Delivery, and Needs in the Next Generation Air Transportation System*. Federal Aviation Administration, 2012.

[She92] Thomas B. Sheridan. *Telerobotics, automation, and human supervisory control*. eng. Cambridge, Mass.: MIT Pr, 1992. 393 pp. ISBN: 978-0-262-19316-0.

[Shi+17] Kimberlee H. Shish et al. “Aircraft Mode and Energy-State Prediction, Assessment, and Alerting.” In: *Journal of Guidance, Control, and Dynamics* 40.4 (2017), pp. 804–816. ISSN: 0731-5090. DOI: 10.2514/1.G001765.

[Sin15a] Single European Sky ATM Research Programme. *European ATM Master Plan. The Roadmap for Delivering High Performing Aviation for Europe*. Executive View. Single European Sky ATM Research Programme. 2015. URL: https://www.atmmasterplan.eu/ (visited on 05/08/2017).

[Sin15b] Single European Sky ATM Research Programme. *SESAR 2020. Multi-annual Work Programme*. Single European Sky ATM Research Programme, 2015. URL: https://ec.europa.eu/research/participants/data/ref/h2020/other/wp/jtis/h2020-wp-multi-annual-sesar-ju_en.pdf (visited on 01/03/2020).

[SRW86] J. H. Skipper, C. A. Rieger, and Walter W. Wierwille. “Evaluation of decision-tree rating scales for mental workload estimation.” eng. In: *Ergonomics* 29.4 (1986). Journal Article Research Support, Non-U.S. Gov’t, pp. 585–599. ISSN: 0014-0139. DOI: 10.1080/00140138608968293. eprint: 3709511.

[SJ98] Paul E. Spector and Steve M. Jex. “Development of four self-report measures of job stressors and strain. Interpersonal Conflict at Work Scale, Organizational Constraints Scale, Quantitative Workload Inventory, and Physical Symptoms Inventory.” eng. In: *Journal of occupational health psychology* 3.4 (1998). Journal Article, pp. 356–367. ISSN: 1076-8998. eprint: 9805281.

[Spe71] J. C. Sperandio. “Variation of operator’s strategies and regulating effects on workload.” eng. In: *Ergonomics* 14.5 (1971). Journal Article, pp. 571–577. ISSN: 0014-0139. DOI: 10.1080/00140137108931277. eprint: 5148233.

[Spr19] Sebastian Michael Sprengart. “Automated aircraft system with goal driven action planning.” The Boeing Company. 2019.

[SNS18] Sebastian Michael Sprengart, Stefan Manuel Neis, and Jens Schiefele. “Role of the human operator in future commercial Reduced Crew Operations.” In: *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*. 2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC) (London, United Kingdom, Sept. 23–27, 2018). IEEE, 2018, pp. 1–10. ISBN: 978-1-5386-4112-5. DOI: 10.1109/DASC.2018.8569803.

[SUT16] Tom Stalnaker, Khalid Usman, and Aaron Taylor. *Airline Economic Analysis. For the Raymond James Global Airline Book*. 2015-2016. Ed. by Oliver Wyman. 2016. URL: http://www.oliverwyman.com/content/dam/oliver-wyman/global/en/2016/jan/oliver-wyman-airline-economic-analysis-2015-2016.pdf (visited on 09/12/2017).

[Sta06] Neville A. Stanton. “Hierarchical task analysis: developments, applications, and extensions.” eng. In: *Applied ergonomics* 37.1 (2006). Journal Article Research Support, Non-U.S. Gov’t Review, pp. 55–79. ISSN: 0003-6870. DOI: 10.1016/j.apergo.2005.06.003. eprint: 16139236.

[SHS14] Neville A. Stanton, Don Harris, and Alison 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*. International Conference on Human-Computer Interaction in Aerospace (Santa Clara, California, July 30–Aug. 1, 2014). Ed. by Association for Computing Machinery. Association for Computing Machinery. New York, New York, USA: ACM Press, 2014, pp. 1–8. ISBN: 978-1-4503-2560-8. DOI: 10.1145/2669592.2669684.

[SHS16] Neville A. Stanton, Don Harris, and Alison Starr. "The future flight deck: Modelling dual, single and distributed crewing options." ENG. In: *Applied ergonomics* 53 Pt B (2016). Journal Article, pp. 331–342. ISSN: 0003-6870. DOI: 10.1016/j.apergo.2015.06.019. eprint: 26141908.

[Ste17] Verena Steidel. "Gestaltung eines Pilotenarbeitsplatzes für einen zukünftigen Reduced Crew Operations Cockpit Simulator. Design of a Pilot Work Station for a Reduced Crew Operations Future Cockpit Simulator." Institute of Flight Systems and Automatic Control. Master Thesis. Darmstadt: Technische Universität Darmstadt, 2017.

[SJE03] Laura D. Strater, Debra G. Jones, and Mica R. Endsley. "Improving Sa: Training Challenges for Infantry Platoon Leaders." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 47.19 (2003), pp. 2045–2049. ISSN: 1541-9312. DOI: 10.1177/154193120304701907. URL: https://www.researchgate.net/publication/275455934_Improving_Sa_Training_Challenges_for_Infantry_Platoon_Leaders.

[Tré19] Bruno Trévidic. "Décarboner le transport aérien est une opportunité exceptionnelle pour Airbus." In: *LesEchos* 2019 (Sept. 17, 2019). URL: https://www.lesechos.fr/industrie-services/air-defense/decarboner-le-transport-aerien-est-une-opportunite-exceptionnelle-pour-airbus-1132133 (visited on 12/04/2019).

[TV06] Pamela S. Tsang and Michael A. Vidulich. "Mental Workload and Situation Awareness." In: *Handbook of human factors and ergonomics*. Ed. by Gavriel Salvendy. 3. ed. Hoboken: Wiley, 2006, pp. 243–268. ISBN: 978-0-470-04820-7.

[US 17] U.S. Department of Transportation. *Beyond Traffic 2045*. U.S. Department of Transportation, 2017.

[US11] Johann Uhrmann and Alex Schulte. "Task-based Guidance of Multiple UAV Using Cognitive Automation." In: *COGNITIVE 2011. The Third International Conference on Advanced Cognitive Technologies and Applications : September 25-30, 2011, Rome, Italy*. Ed. by Terry Bossomaier and Pascal Lorenz. Wilmington, DE, USA: IARIA, 2011. ISBN: 978-1-61208-155-7.

[UD15] Maarten Uijt De Haag and Pengfei Duan. "A Multiple Hypothesis Predictive Alerting (MHPA) Method for Improved Aircraft State Awareness." In: *2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC). 13 - 17 Sept. 2015, Corinthia Hotel, Prague, Czech Republic*. Piscataway, NJ: IEEE, 2015. ISBN: 978-1-4799-8940-9.

[UJ02] Pierre Urlings and L. C. Jain. “Teaming Human and Machine.” In: *Advances in Intelligent Systems for Defence*. Ed. by L. C. Jain, N. S. Ichalkaranje, and G. Tonfoni. WORLD SCIENTIFIC, 2002, pp. 29–60. ISBN: 978-981-238-208-5. DOI: 10.1142/9789812776341_0002.

[Wer17] Johannes Werkman. “Scenario-based context analysis for identification of information requirements on the future flight deck.” Institute of Flight Systems and Automatic Control. Master Thesis. Darmstadt: Technische Universität Darmstadt, 2017.

[Wes15] Jendrick Westphal. “Realization and Evaluation of an Aircraft Onboard Retrofit Trajectory Management System.” eng. Westphal, Jendrick (Verfasser) Klingauf, Uwe (Akademischer Betreuer) Kügler, Dirk (Akademischer Betreuer). Dissertation. Darmstadt: Universitäts- und Landesbibliothek Darmstadt, 2015.

[Wic81] Christopher D. Wickens. *Processing Resources in Attention, Dual Task Performance, and Workload Assessment*. Ed. by Office of Naval Research. 1981. URL: http://www.aviation.illinois.edu/avimain/papers/research/pub_pdfs/techreports/EPL-81-3.pdf (visited on 09/14/2017).

[Wic02] Christopher D. Wickens. “Multiple resources and performance prediction.” In: *Theoretical Issues in Ergonomics Science* 3.2 (2002), pp. 159–177. DOI: 10.1080/14639220210123806.

[Wic08] Christopher D. Wickens. “Multiple Resources and Mental Workload.” In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 50.3 (2008), pp. 449–455. ISSN: 0018-7208. DOI: 10.1518/001872008X288394.

[WGM00] Christopher D. Wickens, Keith Gempler, and M. Ephimia Morphew. “Workload and Reliability of Predictor Displays in Aircraft Traffic Avoidance.” In: *Transportation Human Factors* 2.2 (2000), pp. 99–126. ISSN: 1093-9741. DOI: 10.1207/STHF0202_01.

[Wic+03] Christopher D. Wickens et al. “Attentional models of multitask pilot performance using advanced display technology.” eng. In: *Human Factors: The Journal of the Human Factors and Ergonomics Society* 45.3 (2003). Comparative Study Journal Article Research Support, U.S. Gov’t, Non-P.H.S., pp. 360–380. ISSN: 00187208. DOI: 10.1518/hfes.45.3.360.27250. eprint: 14702989.

[Wie88] Earl L. Wiener. “Cockpit Automation.” In: *Human Factors in Aviation*. Ed. by Earl L. Wiener et al. Cognition and Perception. Burlington: Elsevier Science, 1988, pp. 433–462. ISBN: 978-0-08-057090-7.

[WC83] Walter W. Wierwille and John G. Casali. “A Validated Rating Scale for Global Mental Workload Measurement Applications.” In: *Proceedings of the Human Factors Society Annual Meeting* 27.2 (1983), pp. 129–133. ISSN: 0163-5182. DOI: 10.1177/154193128302700203.

[Wol78] James D. Wolf. *Crew Workload Assessment. Development of a Measure of Operator Workload*. HONEYWELL INC MINNEAPOLIS MINN SYSTEMS AND RESEARCH CENTER, 1978.

[WG15] Cynthia A. Wolter and Brian F. Gore. *A Validated Task Analysis of the Single Pilot Operations Concept*. National Aeronautics and Space Administration Ames Research Center, 2015. URL: https://ntrs.nasa.gov/search.jsp?R=20150001421 (visited on 05/11/2017).

[WPR02] David D. Woods, E. S. Patterson, and Emilie M. Roth. “Can We Ever Escape from Data Overload? A Cognitive Systems Diagnosis.” In: *Cognition, Technology & Work* 4.1 (2002), pp. 22–36. ISSN: 1435-5558. DOI: 10.1007/s101110200002.

[Wri13] Robert Wright. “US airlines are ‘hyper-competitive’, trade group insists.” In: *Financial Times* 2013 (Aug. 14, 2013). URL: https://www.ft.com/content/8839cff8-04fa-11e3-9e71-00144feab7de (visited on 11/14/2018).

[XS00] Bin Xie and Gavriel Salvendy. “Prediction of Mental Workload in Single and Multiple Tasks Environments.” In: *International Journal of Cognitive Ergonomics* 4.3 (2000), pp. 213–242. ISSN: 1088-6362. DOI: 10.1207/S15327566IJCE0403_3.

[Yeh+13] Michelle Yeh et al. *Human Factors Considerations in the Design and Evaluation of Flight Deck Displays and Controls. Version 1.0*. Final Report – November 2013, DOT/FAA/TC-13/44, DOT-VNTSC-FAA-13-09. Ed. by Federal Aviation Administration. Federal Aviation Administration, 2013.

[YD08] Robert M. Yerkes and John D. Dodson. “The relation of strength of stimulus to rapidity of habit-formation.” In: *Journal of Comparative Neurology and Psychology* 18.5 (1908), pp. 459–482. ISSN: 0092-7015. DOI: 10.1002/cne.920180503.

[You+16a] Steven D. Young et al. “Evaluating Technologies for Improved Airplane State Awareness and Prediction.” In: *AIAA Infotech @ Aerospace* (San Diego, California, USA). 2016. DOI: 10.2514/6.2016-2043.

[You+16b] Steven D. Young et al. “Flight simulation study of airplane state awareness and prediction technologies.” In: *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC) (Sacramento, CA, USA, Sept. 25–29, 2016). IEEE, 2016, pp. 1–11. ISBN: 978-1-5090-2523-7. DOI: 10.1109/DASC.2016.7777954.

[Zel87] Milan Zeleny. “Management support systems. Towards integrated knowledge management.” In: *Human systems management* 7.1 (1987), pp. 59–70. ISSN: 0167-2533.

List of figures

List of tables

A. Context analysis

A.1. Pilot interview semi-structured interview questionnaire

The following guideline was used by the author while conducting the pilot interviews. Preceding the interview is a short introduction to the topic of this thesis, its motivation, the objectives, and the underlying assumptions. During the interview the interviewees were motivated to talk while notes were taken.

Setting the focus I want to talk about normal situations that can and do occur in every day operations, no abnormal or emergency situations. I want you to think about situations in which the workload level for the cockpit crew increases, for example through unexpected parallel task execution, lots of interaction with systems, or communication. I am also interested in situations that could have turned into an abnormal but were caught by the flight crew.

Main questions A set of questions was developed to learn about high workload situations from the pilots. The questions were asked during the interview when appropriate. No time limit was imposed on the interviewees while answering the questions. The following questions were used to guide the interview:

Q1 Which situations do require both pilots to be involved in every-day operations? What is the task of each pilot? What are other concurring tasks? When do you think it is mandatory to have two pilots? What helps you making decisions?

Q2 What are high workload situations which are less often occurring than the ones previously discussed? What supports you during these?

Q3 Where do you see the biggest challenge when reducing the crew from two to one pilot? Why do you think this is critical?

Additional questions The majority of the interview was spent with discussing the previous questions. If time allowing the following questions were asked towards the end of the interview to gather additional information from the interviewees:

Q4 What makes a good flight?

Q5 What would you like to avoid during a flight? What actitivities waste your time?

Q6 Think about how you receive and exchange information between you, Air Traffic Control (ATC), the airline, other aircraft, copilot... Which are the steps consuming the most time? What information is hard to get?

Q7 What is the most important thing for you to have available? What information are required at every time? What helps you making good decisions?

Q8 What do you like about information management in today's cockpits? What do you not like? How do you work around issues?

A.2. Identified use cases

Table A.1.: Identified use cases with use case identifier, title, and a short description outlining the situation. Use cases are identified as described in subsection 3.1.3.

ID	Title	Description
UC-1	Initial programing of the FMS during preflight	The flight plan data is entered into the FMS, performance data is calculated. The second pilot is required to review the entered data. This should happen at the gate, but could also happen while taxiing and finishing up checklists.
UC-2	Programing the FMS under time pressure	The flight crew is required to enter flight data into the FMS, this could be a list of waypoints rather than a dedicated SID or STAR. Happening in a situation where time is critical (on the runway, in a high traffic area...), or the area is unknown to the flight crew and they need to orient themselves first.
UC-3	Runway change while taxiing	After being pushed back, during taxiing, the crew is notified that the departure runway is changed. As a result a re-briefing of the departure, gathering of charts, and reprogramming of the flight management system is required while still taxiing to avoid further delays.
UC-4	Taxi operations at unfamiliar airport	The flight crew is taxiing at a larger unknown airport, which is unknown to them, under normal traffic conditions. During taxiing they are finishing last checklists while communicating with ATC.
UC-5	Flight preparation under time pressure	Due to previously built up delay the aircraft arrives late at the destination airport. A delay for the following flight is already forseeable. To minimize the downstream impact of the built up delay the flight crew operates under time pressure while preparing the following flight.
UC-6	Departure under turbulences	During takeoff and initial climb the aircraft encounters strong turbulence, requiring a higher level of alertness by the flight crew.

Table A.1.: Continuation of identified use cases.

ID	Title	Description
UC-7	Delay at departure airport	The flight is delayed at the departure airport after being pushed back. Coordination with ATC and the airline is required to determine fuel requirements or ensuring passenger comfort. Both pilots need to coordinate everything until the delay is cleared up and the flight can continue.
UC-8	ATC issues vectors during departure	After take-off the aircraft would normally follow the programmed departure route. However, ATC instructs the crew to manually change heading, speed, and altitude constraints to vector them out of the airspace.
UC-9	Departure route intercepted by traffic	While departing on their assigned SID the flight crew is notified by ATC that crossing traffic requires a manual alteration of the pre-programmed route.
UC-10	Turbulence cause loss of altitude	The aircraft encounters sudden speed and altitude changes caused by turbulence
UC-11	Change of flight path due to weather	A weather cell is detected by the pilots which must be avoided. While one pilot continues flying the aircraft, the other pilot contacts ATC. A solution is developed by the pilots to minimize excess fuel burn and delay, clearance from ATC is requested.
UC-12	Destination airport closed	Caused by a significant delay, which was built up during enroute cruise, the aircraft is no longer allowed to land at the planned destination airport due to a night curfew. The crew is required to plan for an alternate airport in coordination with the AOC and ATC.
UC-13	New STAR assigned to crew	Being 15-20min out the destination, the crew is notified that their expected STAR has changed, the pilots are required to re-program and re-brief the approach. This happens in a phase where concurrent tasks are present, such as following ATC instructions, or flying the aircraft.

Table A.1.: Continuation of identified use cases.

ID	Title	Description
UC-14	Performing approach briefing	When coming into an airport, the aircraft is descending towards the transition waypoint. A STAR is assigned to the aircraft by ATC. After having received this information one pilot starts preparing the procedure, reviewing charts, entering into the FMS and setting up for execution. After that that the other pilots reviews and prepares the STAR by as well, entered information is verified and familiarized. Finally both pilots talk about what is going to happen.
UC-15	Go-around due to turbulence	The flight crew approaches the runway while encountering heavy turbulence. They decide to abort the landing and perform a go-around maneuver, following the published procedures. After the go-around is performed the flight crew coordinates with ATC on how to proceed.
UC-16	Approach under glideslope, altitude warning issued.	During the descent towards the runway the flight crew misses the glideslope. The aircraft systems issue a warning to the crew.
UC-17	Briefing of destination information	Before reaching the destination airport, the flight deck crew needs to establish contact with the airport to retrieve mission-relevant information. This includes information about the weather, the active runway(s), the company situation, assigned gates etc. This is performed in a division of labor where one pilot is flying (more workload in final phases) and the other one retrieving the information, after that a briefing between the two pilots is performed.
UC-18	Runway change during descent	While being on a level leg during descent, the configuration of the aircraft for landing is normally performed. After initially configuring the aircraft, the crew receives instructions that the runway has changed requiring them to re-configure the aircraft to meet the new ILS etc.

Table A.1.: Continuation of identified use cases.

ID	Title	Description
UC-19	Bad weather approach	Coming in from cruise down to the approach phase is a phase in which the aircraft is being reconfigured. The flight crew has limited control and is required to closely monitor the aircraft. Furthermore the crew needs to look out for weather events, keep in contact with ATC, and compare the charts/procedures to what they have programmed into the FMS.
UC-20	Final approach runway change	While already being on the final phase of the approach the flight crew is notified by ATC that their assigned runway has changed to a parallel runway and they are required to swing over.
UC-21	Unruly passenger after pushback	Shortly after pushback one of the passengers becomes unruly and is considered a threat to flight safety. The flight crew decides to turn around to the gate.
UC-22	Congested destination airport	At the destination airport a weather event prevents landing. The aircraft is kept in a holding pattern with other aircraft and the pilots need to decide whether they should wait or divert to a nearby airport before it is filled with other aircraft. Fuel time and induced delay play important roles, communication with the AOC and ATC is made.

A.3. Rating of use cases

The cohesive model of workload by [Mes88] is used to rank the researched use cases based on the expected workload level. In the following, the considered influence factors and how they were quantified is discussed. In the original work [Mes88] no quantifiable rating scheme is described. The model is extended and a quantifiable rating scale is introduced in [Boe+19] under the supervision of the author. In this rating scheme a weighted sum across all influence variables is calculated, leading to an overall mental workload score assigned to each use case. The equivalent variable to this score is *mental work to be done in certain time and environment* in the cohesive workload model [Mes88].

A.3.1. Assessment of influencing factors

For each influencing factor a set of quantifiable measurements is defined. Points are assigned on a weighted six-point scale (0-5) for each influence factor. The sum of the assigned points then results in an overall relative workload rating for a use case. The following factors and rating methods were used: task criticality, environmental factors, amount of information, task structure, information and task complexity, required performance time and responses.

Assessing the *task criticality* is based on Federal Aviation Administration (FAA) AC 25.1309 [Fed88] in which the probability of a failure condition is weighted against its consequences. This rating scheme is adapted, by assigning each use case a probability of occurrence (ranging from 0 "extremely improbable" to 6 "probable") and an analogous rating scale for severity of consequences, should the flight crew not, or incorrectly, act in the specific use case. Points are assigned based on subjective assessment and subject matter expert interviews and normalized to the overall scale [Boe+19]. The *environmental factors* acting on the flight deck crew are assessed subjectively. The *amount of information* present in each use case is estimated based on the task analysis and the interviews conducted during the use case exploration phase as described in section 3.1. The impact of the *task structure* is rated based on five different components:

1. Tasks performed in the different areas of the Aviate Navigate Communicate Manage Systems (ANCS) scheme.
2. Flight phases of the use case, according to the International Civil Aviation Organization (ICAO) definition [Int13a].
3. Modality of information acquisition by the flight deck crew (visual, haptic, auditory).
4. Relevant stages of Crew Task Management (CTM) (cf. [Fun91; Fun97; Fun+97]).
5. Stages of task execution: starting with building and maintaining situation awareness, planning, deciding, executing, and eventually controlling the outcome.

Rating of the different influence factors is performed subjectively, for each applicable criteria in every category a score could be assigned. Following that, the number of assigned points is normalized to maximum possible number of points and summed up. A higher total score indicated a more complex task structure.

The *information and task complexity* and *required performance time and responses* are determined via an online survey targeted at pilots. During this survey, the participants were asked to perform pairwise comparisons of use cases. For each participant a subset of 20 pairwise comparisons was automatically generated. Participants were tasked to indicate (a) which use case is *more demanding and complex* and (b) which use case has a *higher time pressure*. Survey links were posted to online discussion forums targeted at pilots and distributed via email to known pilots. Eventually, 120 answers with a total of 1155 completed pairwise comparisons were received. Based on that the use cases were ranked according to either task complexity or time pressure. According to their relative position the use cases were assigned points on a six-point scale. In subsection A.3.2 the detailed results of the online survey including participant demographics and distribution channels are listed.

The components of *task rate/frequency, used equipment and physical actions, and task novelty* are not explicitly rated. Task rate/frequency is considered to be a part of the required performance time and responses Used equipment and physical actions are considered to be a part of task structure. Task novelty is completely omitted, as no significant influence is expected, it is assumed that the pilots are proficient in the situations they experience.

A.3.2. Online survey results

During the online survey used to determine the influencing factors *information and task complexity* and *required performance time and responses* the participants had to answer the following questions:

- Which scenario would you describe as more complex and more demanding?
- Which scenario concludes to higher time pressure?

The questions were displayed simultaneously. For each of the question the same two scenarios were displayed and the participants had to choose either one for each question. The results of this pairwise comparison are listed in Table A.2:

Table A.2.: Results of the pairwise comparison of use cases. Indicated are the percentiles of the use cases describing how often a use case was rated as more complex or as having a higher time pressure than another use case.

Position	Complexity	Time pressure	Position	Complexity	Time pressure
UC-1	20.0 %	28.6 %	UC-12	65.7 %	61.9 %
UC-2	41.0 %	47.6 %	UC-13	37.1 %	44.8 %
UC-3	57.1 %	62.9 %	UC-14	12.4 %	24.8 %
UC-4	47.6 %	23.8 %	UC-15	65.7 %	55.2 %
UC-5	40.0 %	71.4 %	UC-16	56.2 %	55.2 %
UC-6	48.6 %	25.7 %	UC-17	24.8 %	34.3 %
UC-7	10.5 %	27.6 %	UC-18	64.8 %	61.0 %
UC-8	49.5 %	51.4 %	UC-19	61.9 %	36.2 %
UC-9	28.6 %	43.8 %	UC-20	61.9 %	66.7 %
UC-10	50.5 %	37.1 %	UC-21	61.0 %	49.5 %
UC-11	28.6 %	31.4 %	UC-22	74.3 %	66.7 %

Participants demographics Of the 120 participants 60.8 % had experience flying Boeing aircraft, 23.3 % indicated experience with Airbus aircraft. The remaining 15.9 % are distributed between Embraer, Bombardier, and other aircraft types. The majority of the participants (62.6 %) had between 1000 h and 10 000 h of flying experience. More than 10 000 h of flying experience was reported by 7.5 %. The remaining participants indicated less than 1000 h of flying experience. Survey distribution channels are summarized in Table A.3:

Table A.3.: Distribution channels of the online survey and relative total return rates.

Channel	Return rate	Channel	Return rate
airliners.net	14.2 %	airfleets.net	0.8 %
airlinepilotsforum.com	82.5 %	private channels	2.5 %
civilaviation.co.uk	0.0 %		

A.3.3. Detailed use case ranking

Summarized in Table A.5 are the top-ten use cases sorted by their relative expected workload levels. The detailed ratings for the different influencing factors of the comprehensive mental workload model by [Mes88] are summarized in Table A.6. The weights applied to the different influenced factors are summarized in Table A.4. As reflected by the chosen weights, task complexity and time pressure are emphasized, followed by the amount of information to be processed. The weighting is performed in accordance with work-influencing factors, which are reported the most stressful by employees in [Fuc06].

Table A.4.: Score ranges and weights applied to the considered influence factors of the comprehensive mental workload model [Mes88].

Influence factor	Range	Weight
Task criticality	0-5	3
Environmental factors	0-5	1
Amount of information	0-5	4
Task structure	0-5	3
Required performance time and responses	0-5	5
Information and task complexity	0-5	5

Table A.5.: Top-ten use cases, after ranking by relative workload levels.

Use Case			Scoring	
Position	ID	Short description	Score	Percentile
1	UC-15	Go-around due to turbulence	71.6	100 %
2	UC-19	Bad weather approach	67.5	94 %
3	UC-20	Final approach runway change	66.4	93 %
4	UC-22	Congested destination airport	63.5	89 %
5	UC-18	Runway change during descent	62.6	87 %
6	UC-12	Destination airport closed	62.1	87 %
7	UC-21	Unruly passenger after pushback	61.3	86 %
8	UC-3	Runway change while taxiing	59.9	84 %
9	UC-9	Departure route intercepted by traffic	57.5	80 %
10	UC-8	ATC issues vectors during departure	57.0	80 %

Table A.6.: Detailed use case ratings for all use cases with applied weights as described in Table A.4.

	UC-1	UC-2	UC-3	UC-4	UC-5	UC-6	UC-7	UC-8
Task structure	4	8.7	10.4	7.5	6.9	10.4	5.8	9.8
Environmental factors	1	1	2	2	2	5	1	3
Amount of information	4	8	12	8	8	12	8	16
Required performance time	7.1	11.9	15.7	6	17.9	6.4	6.9	12.9
Information and task complexity	5	10.2	14.3	11.9	10	12.1	2.6	12.4
Task criticality	7.5	3.5	5.5	5	4.5	8.5	2.5	3

	UC-9	UC-10	UC-11	UC-12	UC-13	UC-14	UC-15	UC-16
Task structure	10.4	7.5	9.8	8.7	8.1	6.3	10.4	8.1
Environmental factors	3	4	2	2	2	2	5	2
Amount of information	20	16	12	16	8	8	16	8
Required performance time	11	9.3	7.9	15.5	11.2	6.2	13.8	13.8
Information and task complexity	7.1	12.6	7.1	16.4	9.3	3.1	16.4	14
Task criticality	6	6.5	8.5	3.5	3.5	7.5	10	3

	UC-17	UC-18	UC-19	UC-20	UC-21	UC-22
Task structure	5.8	8.7	7.5	9.2	8.7	9.2
Environmental factors	1	2	4	2	1	3
Amount of information	8	16	20	16	20	12
Required performance time	8.6	15.2	9	16.7	12.4	16.7
Information and task complexity	6.2	16.2	15.5	15.5	15.2	18.6
Task criticality	6	4.5	11.5	7	4	4

A.4. Exemplary mission manager tasks

The following is a collection of exemplary tasks of a mission manager, denoting the key areas in which they have to exercise their role on the flight deck.

A.4.1. Strategic level decision making and goal setter

The mission manager is responsible for the definition of mission goals on a strategic level. The translation of strategic goals into tactical actions and their execution is performed by automated aircraft systems. No low level, tactical, system-related tasks such as monitoring or micromanagement will be performed by the mission manager. This also includes the setting of discrete autopilot parameters as performed today or the reconfiguration of the aircraft (e.g., setting the flaps). Actions like that are derived by the automated systems and executed accordingly. In the role of a mission manager, the human operator will not be tasked to perform any manual flight maneuvers during normal operations.

Example After entering the flight deck, the mission manager starts configuring the aircraft for the next mission. Instead of manually typing in a previously calculated flight plan, the mission manager communicates the goal of the mission to the automated systems. Now, being aware of the mission goal, the automated systems are able to plan all required tactical steps to achieve the set goal and track it over the course of the mission. This includes retrieval of the flight plan under consideration of the current weather and airspace situation. No interaction of the mission manager with systems comparable to the Flight Management System (FMS) nor Flight Control Unit (FCU) would be required.

A.4.2. Monitor of mission relevant parameters

Strategic decisions made by the mission manager need to be based on an accurate overview over the current situation. One of the tasks associated with the role of being a mission manager is the monitoring of mission relevant parameters. This monitoring task relates to the monitoring of grand-scale external parameters, in order to stay ahead of potentially detrimental effects on the mission. This is not to be confused with the monitoring of discrete system and automation parameters as performed in operations today. Instead, the automated systems are performing that (monotonous) task, they are considered to be better suited for that, than a human with limited system insight would be [Har07].

Example During the cruise phase the mission manager is notified that no ground services can be offered at the destination airport. Being aware of that issue the mission manager would now have to consider the impact on the current and potentially following missions. Eventually, by using the available support tools, an alternative solution leading to an assisted adjustment of mission parameters (see below) could be developed.

A.4.3. Assisted adjustment of mission relevant parameters

Related to the monitoring of mission-relevant parameters is the assisted adjustment of selected parameters by the mission manager. Hereby, it is envisioned that the actual adjustment of goals is performed by the mission manager. While doing so, the mission manager may rely on decision support systems and a proper information management, which provide timely and comprehensible information.

Example Following the previous example, the mission manager might decide to deviate to an alternate airport from which the majority of passengers can still be transported to their final destinations. However, rather than manually changing the flight path, the mission manager would delegate this task to the automated systems. The automated systems would then incorporate up-to-date weather information and the current aircraft performance to calculate a viable route. The mission manager would now have to coordinate with the airline to get all passengers to their destinations and reduce the overall impact on the airline's network. While performing the latter task, the mission manager could elect to request support from the ground station.

A.4.4. Management of corner cases, provision of context

Regardless of how sophisticated the autonomous systems are assumed to be, the chance of experiencing so called corner cases does and will exist. Summarized by the term corner cases are situations in which unforeseen events may trigger false or inappropriate responses from the system, or maneuver the system into a state for which it might not have been designed. Similar to what is described by [Har07], an envisioned key role of the human operator will be the management (prevention or mitigation of consequences) for such corner cases. This very central role of the human operator will later be referenced as the provision of *context* to the automated systems when describing the future information management concept in section 4.2.

Example The passengers are already boarded and the aircraft is about to taxi out when suddenly one of the passengers shows symptoms of fear of flying. Irregardless of technological capability an automated system might not even be allowed to monitor the psychological state of the passengers and would thus not notice the passenger having an anxiety attack. In this situation the mission manager would be required to tell the automated system to pause mission execution to not aggravate the situation. Eventually, the situation could be resolved by requesting support from the airport personnel and help the passenger.

A.4.5. Being a safety net - not a backup

A critical point of failure, and discussion, related to the concept of Reduced Crew Operation (RCO) is the question of what is happening when automation breaks down, for example due to some sort of system failure. Often, the human operator is assumed to take over control and act as a backup for the automation [SWB97; CW02]. However, in the herein described concept of operations the human operator will no longer be a backup. The term backup would imply that all tasks previously performed by automated systems would be continued by the mission manager, which is not possible. Instead, a more precise terminology is chosen, the mission manager will act as a safety net. On a first level, this means that an attempt on restoring the automation capabilities is performed. If all measures fail, the mission manager will have to safely terminate the mission, with or without limited support of the automated systems. As a *last resort* measure this could also involve manually controlling the flight path.

Example After a component failure the automation is unable to resume lateral and vertical navigation. In this case the mission manager would try to restore functionality, which in this case does not remedy the situation. Eventually, limited manual control is seized, meaning that assisted flight path control is exercised and the mission is safely terminated at a nearby airport.

A.4.6. Execution of non-mission related tasks

Aside from the previously described tasks, the mission manager will also perform non-mission related tasks. This is motivated by the fact that during large portions of the flight no active interaction of the mission manager with mission parameters will be required. Hence, this down-time can be used to perform tasks which are beneficial for the airline or the mission manager themselves.

Example During a long trans-oceanic part of a mission, the mission manager would be able to perform mandatory training sessions. Alternatively, tasks delegated from the Airline Operation Center (AOC), for which the mission manager is qualified, can be performed.

B. Requirements

B.1. Categorization criteria

The collected requirements are categorized into four categories. A category is assigned based on whether the requirements relates to a certain aspect of the category. The aspects and their description are summarized in respect to their related categories.

Human-Autonomy Teaming Observability: How can the flight crew observe the automation? What feedback is provided to the flight crew? How is feedback provided? Directability: How is the flight crew interaction with the automation? (Error correction, communication of intentions towards automation, context provision) Involvement: How is the flight crew involved with the automation? (In which situations, for what tasks) Authority: Who is in command?

Situation Awareness Mental model: What is currently being done (automation/crew)? What is the state of my system? Predictability: What is going to happen? (Task (time) management, planned actions, automation intentions, consequences of actions) Automation: What is the automation doing, why is it doing that, what is next?

Information and Interaction Design Presentation/ Organization: Is it related to information presentation? How is information organized? (Frame of reference, ordering of information, modality, consistency of tasks) Selection: What information is displayed to the flight crew? How valid is this information? Input: Is it related to entering information? (Procedural guidance, error correction) Alerting: How is the flight crew alerted, how is attention guided? (Alerting modalities, attention guidance, consistency of alerting)

Other Human performance: Is the requirements related to human performance and human capabilities? Safety: Impact on safety levels, information validity, verification, certification, establishing trust in the system. User: Does the requirements relate to the user of the system?

B.2. Full list of requirements

Table B.1.: Full list of Human-Autonomy Teaming (HAT) requirements.

ID	Requirement
HAT-1-0	Provide adequate levels of feedback about the system (automation, aircraft systems) state to the human operator.
HAT-1-1	Provide feedback about flight parameter changes and state transitions.
HAT-1-2	Provide feedback on system goals, and available options.
HAT-1-3	Provide feedback about the actions of automated systems to maintain the current state.
HAT-1-4	Do not make the human operator constantly monitor system parameters.
HAT-1-5	Provide feedback on entries and actions made (or omitted!) by the human operator (e.g., input error checking).
HAT-2-0	Make automated systems observable in their actions and the decisions they make.
HAT-2-1	Do not make the human operator monitor every single decision to understand the intentions of automated systems.
HAT-2-2	Make automation communicative about actions, goals, and intents without being repetitive or pointless.
HAT-2-3	Keep the human operator involved with meaningful tasks (e.g., do not make them consent to every decision).
HAT-3-0	Clearly indicate the ownership of a function (which agent is responsible for doing what).
HAT-4-0	The human operator must always be in command.
HAT-4-1	The human operator must provide context to automated systems when required.
References	[SW91; SW92; SW95; SW97; Bil91; Bil97; ASS96; Pal+95; Fun+99; CW02; Har07; Let+12; AMR13; End16; Com+13; You+16a; Bai+17]

Table B.2.: Full list of Situation Awareness (SA) requirements.

ID	Requirement
SA-1-0	Provide effective means of task management for the human operator.
SA-1-1	Help the human operator to determine task resource requirements (time, physical resources, cognitive resources).
SA-1-2	Let the human operator know what, when, and how to do something.
SA-2-0	Support the prediction of mission and system states.
SA-2-2	Show consequences of interactions to the human operator, before they occur.
SA-2-2	Increase the time horizons for which certain information apply.
SA-2-2	The future intentions and consequences of automated systems behavior must be known to the human operator.
SA-3-0	Support buildup and maintenance of a proper mental model.
SA-3-1	Support the understanding of the relation between system functions and mission parameters.
SA-3-2	Support the detection of changes and impact factors on the mission.
SA-3-3	Raise awareness about the current situation's criticality.
SA-3-4	The human operator must be aware of the current capabilities of the system (performance, failures).
SA-4-0	Support and reinforce situation awareness at all times.
SA-4-1	Establish a basic set of information available to the human operator to describe the current state of the mission (aircraft systems, automation, environment) *at a glance*.
Refer-ences	[Bil91; Fun91; SW91; SW92; SW95; SWB97; SMW07; CW02; BBR07; BM04; Let+12; Yeh+13; Fli14; End16; You+16a]

Table B.3.: Full list of Information and Interaction Design (IID) requirements.

ID	Requirement
IID-1-0	Support efficient and effective input by the human operator into aircraft systems.
IID-1-1	Present procedural (know how) information to guide the user through a process.
IID-1-2	Prevent input format errors.
IID-1-3	Provide meaningful and intuitive hints and labels.
IID-1-4	Do not make the user to reformulate intentions for the automation to understand - allow natural interaction.
IID-2-0	Provide effective attention guidance principles.
IID-2-1	Use aural alerts for immediate attention capturing.
IID-2-2	Use visual alerts for lower-priority attention capturing.
IID-2-3	Guide the human operator towards occurring problems and changes.
IID-2-4	Use positive selectivity above negative selectivity.
IID-2-5	Explicitly identify missing information.
IID-3-0	Support effective and efficient information transfer between aircraft systems and the human operator.
IID-3-1	Do not present raw data.
IID-3-2	Design for rule-based scanning to gather the most important information about the current system state.
IID-3-3	Always present information in context and conformal with the outside world, provide a frame of reference when available.
IID-3-4	Integrate information as much as possible, offloading the operator's cognitive effort.
IID-3-5	Spread information transmission across several modalities.
IID-3-6	Avoid redundant representation of information and data.
IID-3-7	Only show information deviating from normal conditions when monitoring the automation.
IID-3-8	Primarily present the information required for the current situation to the human operator, do not clutter with not required information.
IID-3-9	Present information around goals, emphasize in accordance with importance.

Table B.3.: Full list of Information and Interaction Design (IID) requirements.

ID	Requirement
IID-4-0	Establish a standard of information representation across all systems.
IID-4-1	Maintain a consistent, unambiguous set of labels and icons.
IID-4-2	Maintain a consistent use of color.
IID-4-3	Maintain a consistent, unambiguous, set of alarms.
IID-4-4	Maintain a consistent set of display components.
IID-4-5	Maintain a consistency of interaction and presentation across goals.
IID-5-0	Use a visual design aligned with established design standards and guidelines.
IID-5-1	Refer to FAA (and other) guidelines.
IID-5-2	Allow for easy training through knowledge- and rule based techniques.
References	[Bil91; Fun91; SW92; SW95; SWB97; Pal+95; CW02; WPR02; Wic+03; Wic08; BBR07; Sch+07; Let+12; Com+13; DBS13; Dua+15; Yeh+13; AMR13; End16]

Table B.4.: Full list of other requirements.

ID	Requirement
O-1-0	Put the human operator and their role in the center of the design.
O-1-1	Consider the changed role requirements on the flight deck.
O-1-2	Design for a single operator.
O-2-0	Establish trust between automated systems and the human operator.
O-2-1	Indicate when imperfect prognoses, alerts or information are to be expected.
O-2-2	Ensure information validity, if not possible let the human operator inspect the underlying data the information was derived from.
O-3-0	Consider human limitations (memory, computation, attention, decision making biases, task timesharing, cultural differences).
O-3-1	Maintain constant adequate workload levels.
References	[Bil91; Pal+95; Fun+99; Har07; Wic08; AMR13]

C. Implementation

C.1. Application screenshots

Listed in the following are screenshots of the implemented software application as described in chapter 5. General application screenshots and screenshots of the application highlighting different parts of the evaluation scenario are shown.

General screenshots Shown are screenshots of the application in the configuration used during the scenario evaluation. In the lower right of the application the secondary workload task is visible.

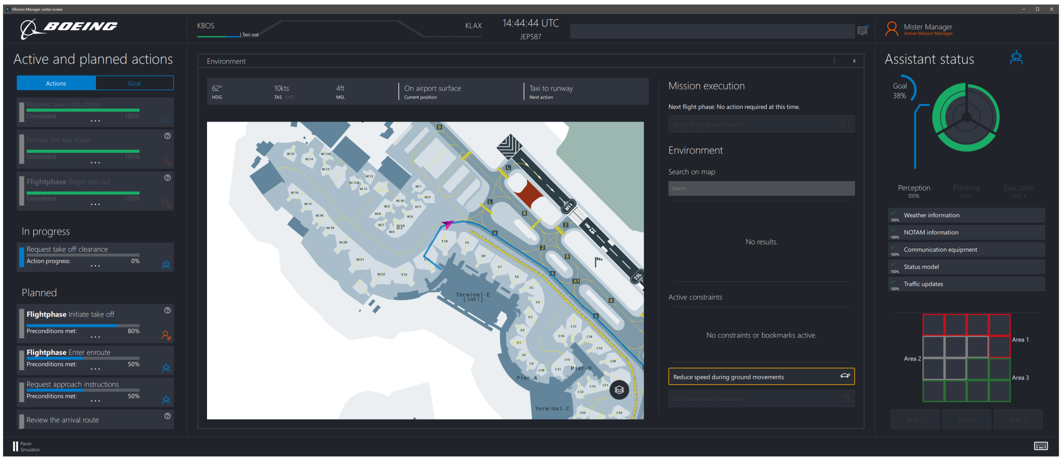


Figure C.1.: Environment window showing the aircraft taxiing on the surface of KBOS during the scenario evaluation.

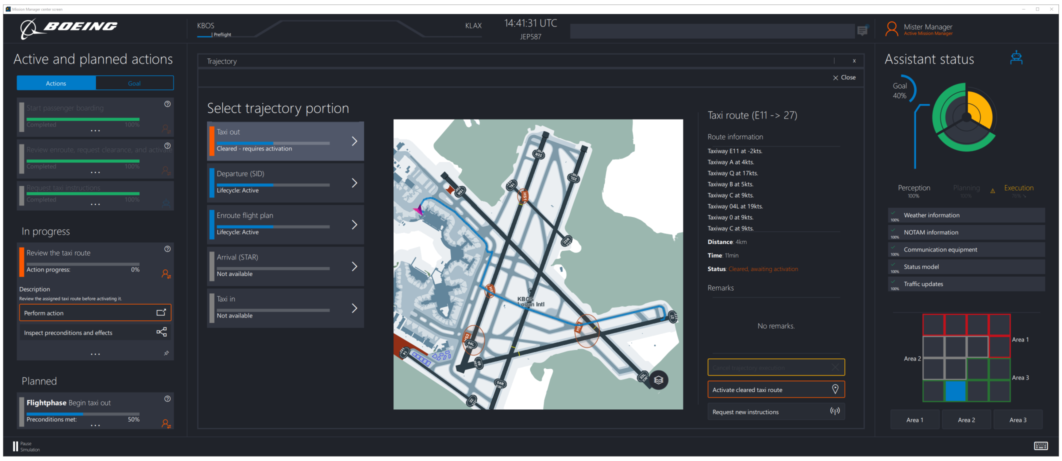


Figure C.2.: Trajectory window showing the review of a taxi route received from ATC and accepted by the operator. Shown route requires activation by the human operator to become executable by the aircraft systems.

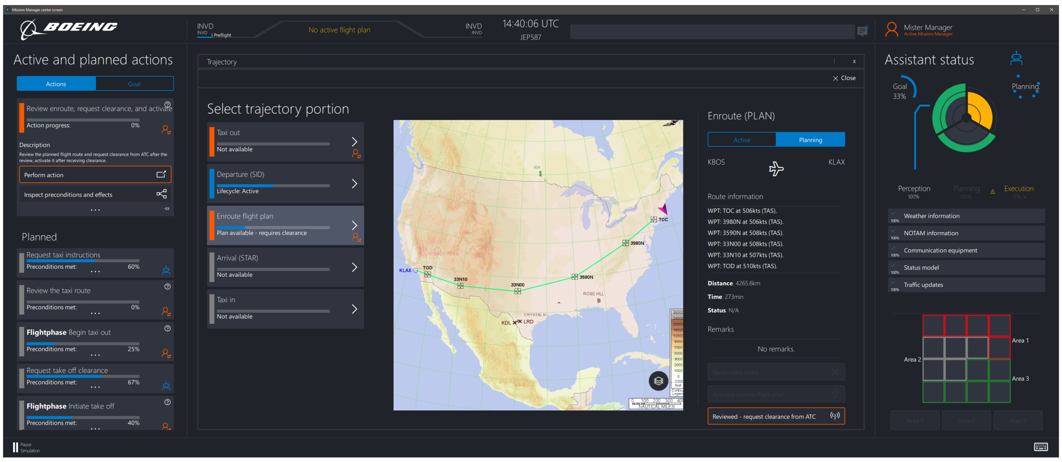


Figure C.3.: Trajectory window showing the review of a flight plan before sending it to ATC for review. Trajectory portion is selected through the overview on the left side of the window.

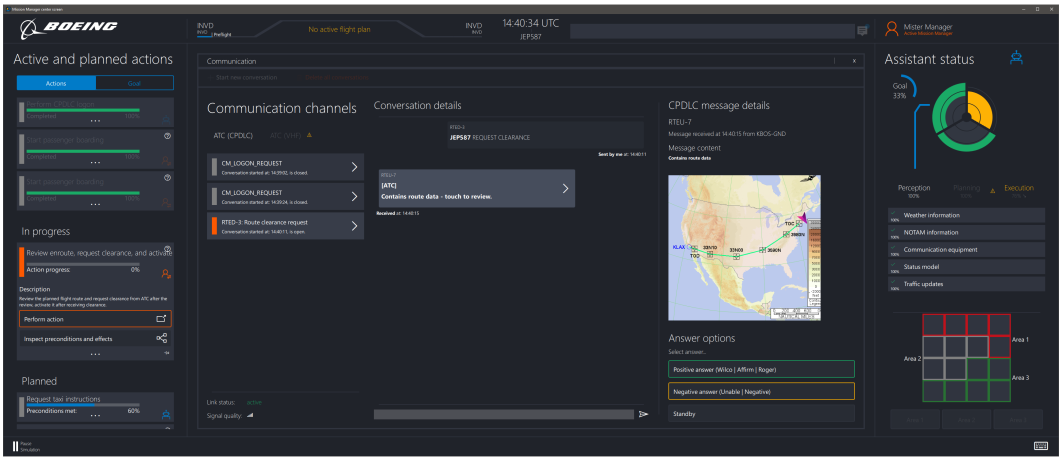


Figure C.4.: Communication window showing an uplinked flight plan. Conversations are selected on the left side of the window, message history shown in the center. Route details are displayed in the detail pane on the right of the window.

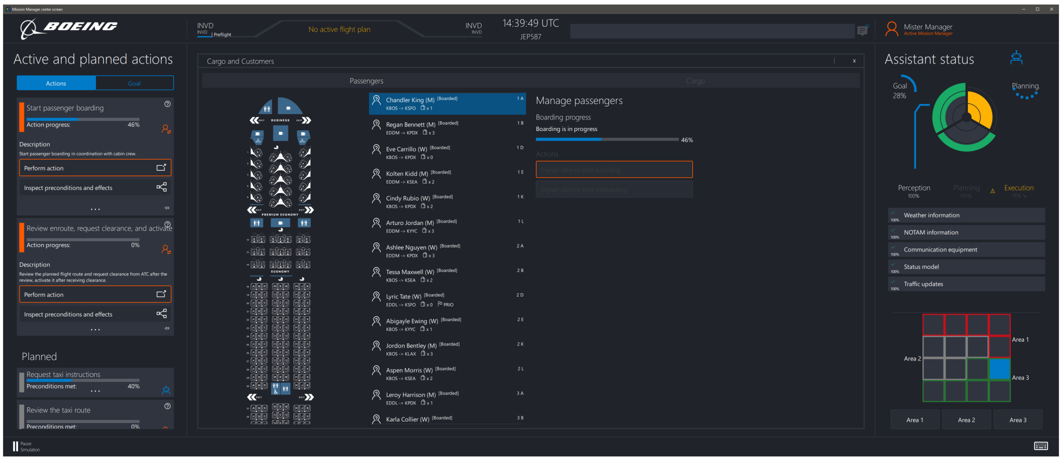


Figure C.5.: Passenger and cargo window showing the boarding of the passengers during the scenario evaluation. Progress bar of the related action on the far left is linked to progress bar inside the window.

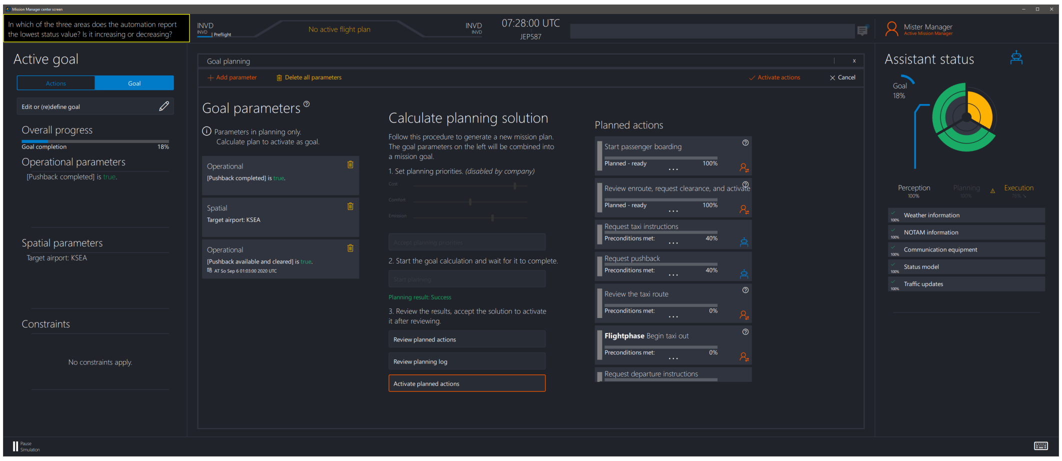


Figure C.6.: Goal planning window showing the result of a manually triggered planning process. List of planned actions requires activation through the operator.

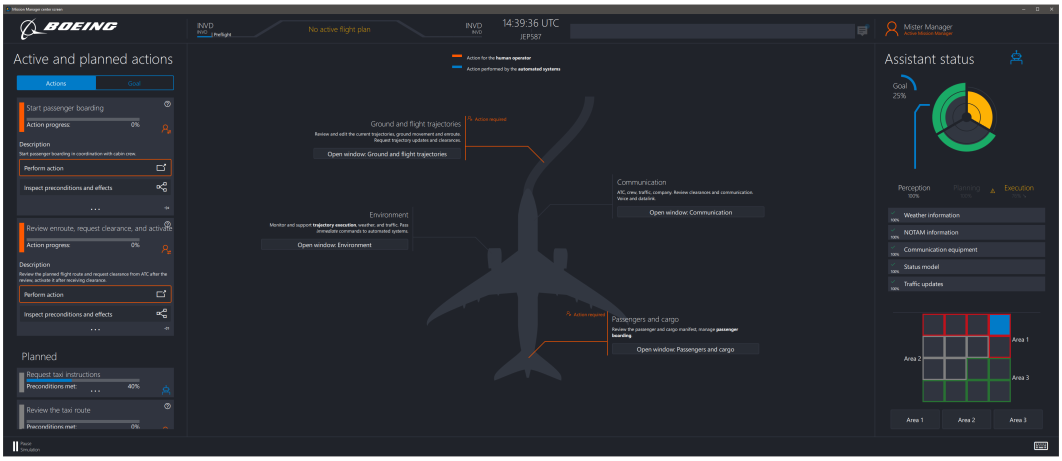


Figure C.7.: Application main menu. Usage of color and iconography guides the operator to windows requiring attention.

Evaluation scenario Shown are screenshots of the application detailing steps of the scenario evaluation. Note that screens are not in the order as they appear in the scenario.

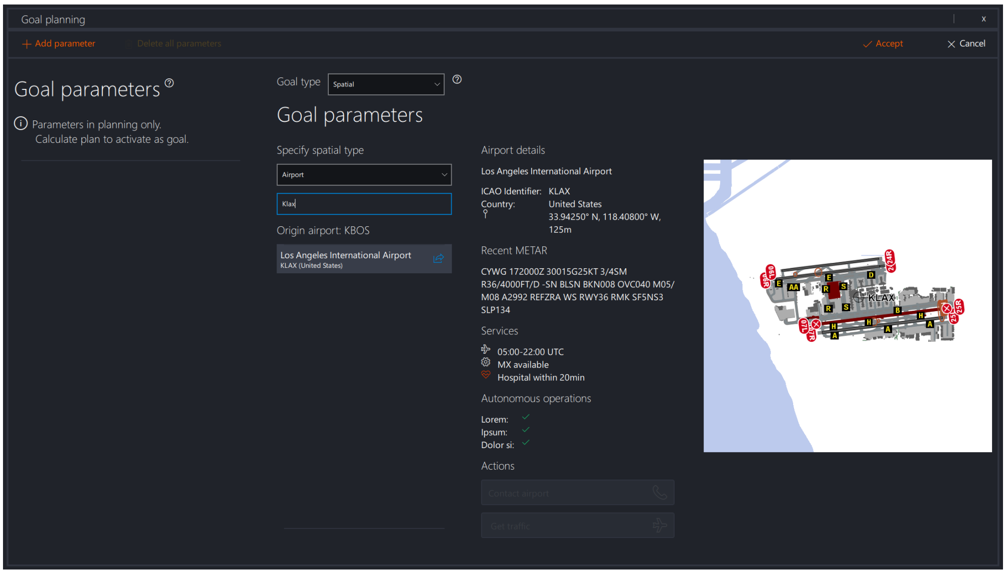


Figure C.8.: Goal planning window is opened. Spatial goal type is selected, KLAX is searched. The origin airport is automatically determined to be KBOS.

Figure C.9.: Instruction to monitor ground operations.

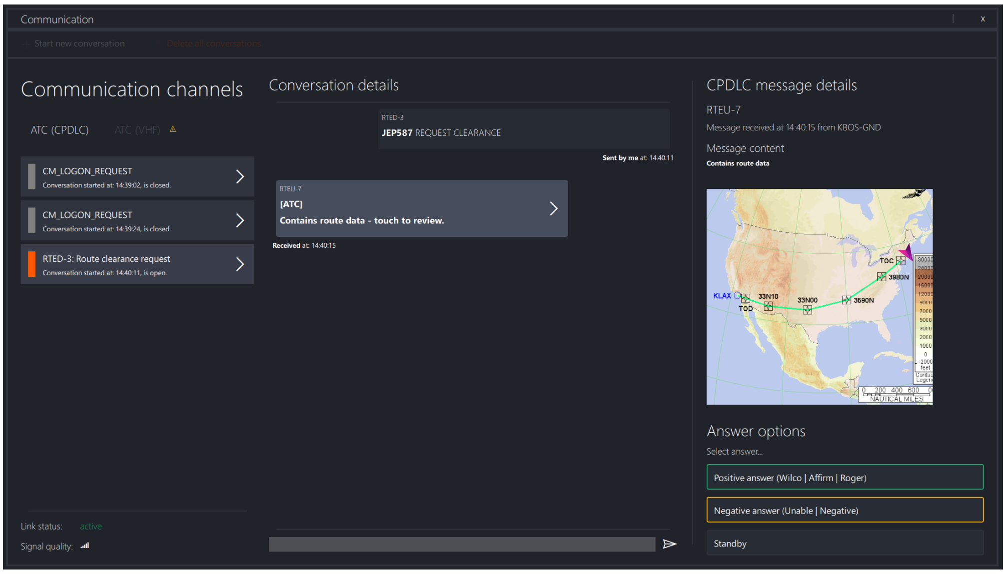


Figure C.10.: Receiving enroute clearance in communication window.

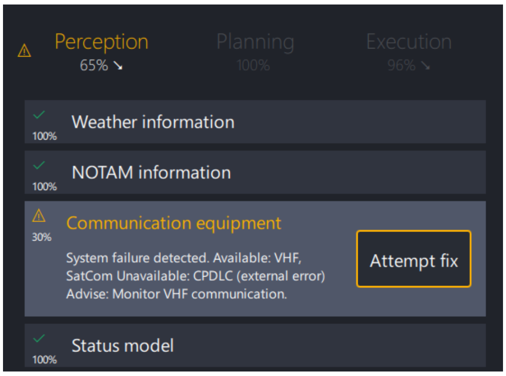


Figure C.11.: Notification about CPDLC outage. Additional information about the failure is provided. Through “Attempt fix” the operator can open the communication window to activate the VHF radio.

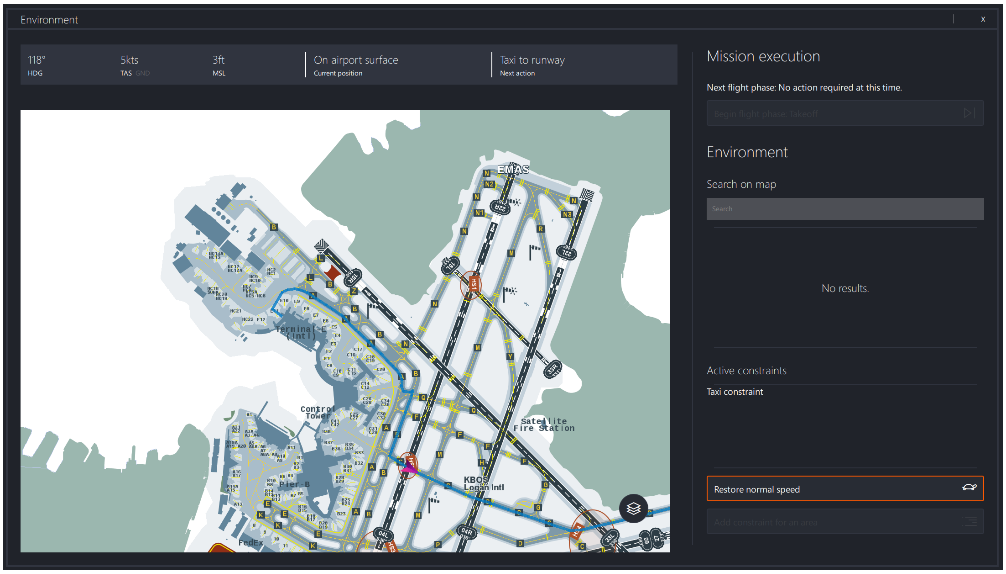


Figure C.12.: Environment window after speed reduction. A constraint “Taxi constraint” is displayed to the operator.

C.2. State variables

Table C.1.: State variables used in mission state modeling. Variables with a ✓ in the *Goal* column can be directly defined as an operational goal parameter. Variables with a (✓) are indirectly set through spatial goal parameters.

ID	Identifier and description	Goal
1	MISSION_ROUTE_AIRPORT_ALTERNATESET Alternate airport set	(✓)
2	MISSION_ROUTE_AIRPORT_DESTINATIONREACHED Destination airport reached	-
3	MISSION_ROUTE_AIRPORT_DESTINATIONSET Destination airport set	(✓)
4	MISSION_ROUTE_AIRPORT_ORIGINSET Origin airport set	(✓)
5	AIRCRAFT_MISSIONMANAGERAPP_READY Application ready	-
6	AIRCRAFT_MISSIONMANAGERAPP_ISLOCKED Application unlocked	-
7	MISSION_PHASE_APPROACH_COMPLETED Approach completed	✓
8	MISSION_ROUTE_ARRIVAL_ACTIVATED Arrival activated	-
9	MISSION_ROUTE_ARRIVAL_AVAILABLECLEARED Arrival route available and cleared	✓
10	AIRCRAFT_SYSTEMS_ENGINES_RUNNING Aircraft engines are running	✓
11	AIRCRAFT_SYSTEMS_ASSISTANT_REQUIRESPERCEPTIONSUPPORT Assistant requires perception support	-
12	AIRCRAFT_SYSTEMS_ASSISTANT_REQUIRESEXECUTIONSUPPORT Assistant requires execution support	-
13	AIRCRAFT_SYSTEMS_ASSISTANT_REQUIRESPLANNINGSUPPORT Assistant requires planning support	-
14	AIRCRAFT_SYSTEMS_CPDLC_ISINOP CPDLC is inoperable	-

Table C.1.: Continuation of state variables used in mission state modeling.

ID	Identifier and description	Goal
15	AIRCRAFT_SYSTEMS_CPDLC_LOGGEDON CPDLC logged on	-
16	AIRCRAFT_CHECKLIST_ACCEPTANCE_COMPLETED Aircraft acceptance	-
17	AIRCRAFT_CHECKLIST_AIRCRAFTDOCUMENTS_COMPLETED Aircraft documents check	-
18	AIRCRAFT_CHECKLIST_EXTERIOR_COMPLETED Aircraft exterior check	-
19	AIRCRAFT_CHECKLIST_INTERIOR_COMPLETED Aircraft interior check	-
20	AIRCRAFT_CHECKLIST_TECHNICALLOG_COMPLETED Aircraft technical log review	-
21	AIRCRAFT_CHECKLIST_PERSONALDOCUMENTS_COMPLETED Personal documents check	-
22	AIRCRAFT_CHECKLIST_PREFLIGHT_COMPLETED Preflight check	-
23	MISSION_ROUTE_DEPARTURE_ACTIVATED Departure activated	-
24	MISSION_ROUTE_DEPARTURE_AVAILABLECLEARED Departure route available and cleared	✓
25	MISSION_ROUTE_DEPARTURERUNWAY_ASSIGNED Departure runway assigned	-
26	MISSION_PHASE_INITIALCLIMB_COMPLETED Initial climb completed	✓
27	MISSION_ROUTE_ENROUTE_CLEARED Enroute cleared	✓
28	MISSION_PHASE_ENROUTE_COMPLETED Enroute phase completed	✓
29	MISSION_ROUTE_ENROUTE_AVAILABLE Enroute route available	✓
30	MISSION_ROUTE_ENROUTE_ACTIVATED Enroute route is activated	-
31	MISSION_ROUTE_LANDING_CLEARED Landing cleared	✓

Table C.1.: Continuation of state variables used in mission state modeling.

ID	Identifier and description	Goal
32	MISSION_PHASE_LANDING_COMPLETED Landing completed	✓
33	MISSION_ROUTE_LANDINGRUNWAY_ASSIGNED Landing runway assigned	✓
34	AIRCRAFT_PASSENGERS_BOARDING_COMPLETED Boarding completed	✓
35	AIRCRAFT_PASSENGERS_BOARDING_INPROGRESS Passenger boarding	-
36	AIRCRAFT_PASSENGERS_HEALTH_ISCRITICAL Passenger health issue	-
37	MISSION_PHASE_POSTFLIGHT_COMPLETED Post flight completed	✓
38	MISSION_PHASE_PREFLIGHT_COMPLETED Preflight completed	✓
39	MISSION_ROUTE_PUSHBACK_AVAILABLECLEARED Pushback available and cleared	✓
40	MISSION_PHASE_PUSHBACK_COMPLETED Pushback completed	✓
41	MISSION_ROUTE_PUSHBACK_ACTIVATED Pushback route activated	-
42	MISSION_ROUTE_TAKEOFF_CLEARED Takeoff cleared	✓
43	MISSION_PHASE_TAKEOFF_COMPLETED Takeoff completed	✓
44	MISSION_ROUTE_DESTINATIONGATE_ASSIGNED Destination gate assigned	-
45	MISSION_ROUTE_TAXIIN_AVAILABLECLEARED Taxi-in available and cleared	✓
46	MISSION_PHASE_TAXIIN_COMPLETED Taxi-in completed	✓
47	MISSION_ROUTE_TAXIIN_ACTIVATED Taxi-in route activated	-
48	MISSION_ROUTE_TAXIOUT_AVAILABLECLEARED Taxi-out available and cleared	✓

Table C.1.: Continuation of state variables used in mission state modeling.

ID	Identifier and description	Goal
49	MISSION_PHASE_TAXIOUT_COMPLETED Taxi-out completed	✓
50	MISSION_ROUTE_TAXIOUT_ACTIVATED Taxi-out route activated	-
51	MISSION_ROUTE_DIRECTTOCOORDINATES_ISSET Direct to coordinates is set	(✓)
52	MISSION_ROUTE_DIRECTTOWAYPOINT_ISSET Direct to waypoint is set	(✓)
53	MISSION_ROUTE_TAXICONSTRAINTS_ISSET Taxi route spatial constraints	-

C.3. Available actions

Table C.2.: List of implemented actions to be used by the aircraft systems during action planning. Preconditions (P_{pc}) and effects (P_e) match the variables listed in Table C.1. (T/F) indicates if the value must be, or becomes, true or false. Actions marked with a ✓ in the *Human* column require human operator involvement during execution.

Name and description	P_{pc}	P_e	Human
UNLOCK THE FLIGHT DECK Authenticate to unlock the flight deck	(6/T)	(5/T), (6/F)	✓
PERFORM ONBOARDING PROCEDURE Perform the onboarding procedure, including preflight checklists and aircraft acceptance.	(16/F), (5/T), (6/F)	(16/T), (22/T), (21/T), (20/T), (19/T), (18/T), (17/T)	✓
START PASSENGER BOARDING Start passenger boarding in coordination with cabin crew.	(16/T)	(34/T)	✓
START PASSENGER DEBOARDING Start passenger deboarding.	(46/T), (34/T)	(34/F)	✓
CALCULATE ENROUTE SOLUTION Calculate a trajectory solution to the specified destination airport.	(22/T), (3/T), (4/T)	(29/T)	-
CALCULATE DIVERSION Calculate a trajectory solution around the area of avoidance.	(22/T), (3/T), (4/T), (5/T)	(29/T)	-
CALCULATE DIRECT TO Calculate a direct to trajectory solution to the specified waypoint or coordinates.	(22/T), (51/T), (5/T)	(29/T)	-
CALCULATE DIRECT TO Calculate a direct to trajectory solution to the specified waypoint or coordinates.	(22/T), (5/T), (52/T)	(29/T)	-

Table C.2.: Continuation of implemented actions to be used by the aircraft systems during action planning.

Name and description	P_{pc}	P_e	Human
MONITOR GROUND OPERATIONS Provide support after pushback until reaching the runway.	(22/T), (0/T), (38/T), (49/F)	(0/F)	✓
PERFORM CPDLC LOGON Request a CPDLC logon at the current ATC facility.	(22/T)	(15/T)	-
REQUEST PUSHBACK Request the clearance to perform pushback from the current gate position.	(48/T), (22/T), (27/T), (15/T), (34/T)	(39/T)	-
REQUEST TAXI INSTRUCTIONS Request a taxi out route to the assigned departure runway.	(22/T), (30/T), (27/T), (15/T), (34/T)	(48/T)	-
REQUEST DEPARTURE INSTRUCTIONS Request a departure route to follow.	(40/T), (15/T)	(25/T), (24/T)	-
REQUEST TAKE OFF CLEARANCE Request the clearance to take off.	(25/T), (23/T), (30/T), (0/F), (40/T), (15/T)	(42/T)	-
REVIEW ENROUTE, REQUEST CLEARANCE, AND ACTIVATE Review the planned flight route and request clearance from ATC after the review, activate it after receiving clearance.	(15/T), (29/T)	(30/T), (27/T)	✓
REQUEST APPROACH INSTRUCTIONS Request approach instructions.	(15/T), (26/T)	(33/T), (9/T)	-
REQUEST LANDING CLEARANCE Request a landing clearance for the assigned landing runway.	(33/T), (15/T), (8/T)	(31/T)	-
REQUEST TAXI INSTRUCTIONS Request a tax in in route to an assigned gate.	(15/T), (7/T)	(45/T), (44/T)	-

Table C.2.: Continuation of implemented actions to be used by the aircraft systems during action planning.

Name and description	P_{pc}	P_e	Human
BEGIN PUSHBACK Starts the pushback of the aircraft.	(30/T), (41/T), (34/T)	(38/T)	✓
BEGIN TAXI OUT Starts taxiing out to the departure runway.	(30/T), (41/T), (34/T), (50/T)	(40/T), (38/T)	✓
INITIATE TAKE OFF Starts the take off roll.	(25/T), (23/T), (30/T), (40/T), (42/T)	(43/T), (49/T)	✓
ENTER ENROUTE Switch to enroute.	(43/T), (30/T)	(26/T)	-
BEGIN APPROACH Starts approaching the destination airport	(33/T), (26/T), (8/T), (9/T)	(28/T)	✓
PERFORM LANDING Starts the final landing phase.	(33/T), (31/T), (28/T)	(7/T)	-
BEGIN TAXI IN Starts taxiing in to the assigned gate.	(45/T), (47/T), (7/T)	(32/T)	-
BEGIN POSTFLIGHT Starts the post flight activities.	(32/T), (44/T)	(46/T), (2/T)	✓
FINISH MISSION Finishes the mission and shuts the aircraft down.	(46/T), (2/T)	(37/T)	✓
REVIEW THE FLIGHT PLAN Review the generated flight plan before submitting it to ATC.	(22/T)	(29/T)	✓
REVIEW THE TAXI ROUTE Review the assigned taxi route before activating it.	(48/T)	(41/T), (50/T)	✓
REVIEW THE TAXI ROUTE Review the assigned taxi route before activating it.	(45/T)	(47/T)	✓

Table C.2.: Continuation of implemented actions to be used by the aircraft systems during action planning.

Name and description	P_{pc}	P_e	Human
REVIEW THE DEPARTURE ROUTE Review the assigned departure route before activating it.	(24/T)	(23/T)	✓
REVIEW THE ARRIVAL ROUTE Review the assigned arrival route before activating it.	(9/T)	(8/T)	✓

C.4. Use of colors

Table C.3.: Highlight colors for use inside the application.

Color	Utilization
	#FD5800, Orange Is used to highlight and indicate tasks that require human operator input or support, can be used in conjunction with the warning and alerting colors.
	#017BC8, Blue Primary accent color, is used in conjunction with the automated aircraft systems.
	#D53600, Red Warning color, used to indicate that immediate attention of the human operator is required.
	#FEB204, Amber Caution color, used to indicate that a value is non-nominal and that attention is required. Also used to highlight important actions.
	#1AAB66, Green Indication of something being as expected or acceptable.

C.5. Class diagrams

Class diagrams used to implement actions (cf. Table C.2) and state variables (cf. Table C.1) are shown in the following. Note that getters, setters, and non-functional relevant variables are omitted in these diagrams.

Action
– identifier : Identifier – uuid : const QUuid – name : QString – description : QString – requiresHumanAction : boolean – isFlightPhaseStart : boolean – isContingency : boolean – progress : double – preconditions : QHash<Identifier, StateVariable> – effects : QHash<Identifier, StateVariable>
+ *execute() : boolean* + *cost() : int* + checkPreconditions(state : MissionState) : boolean + performOnState(stateIn : MissionState, stateOut : *MissionState) # pCurrentMissionState : *const MissionState

StateVariable
– variableIdentifier : Identifier – variableValue : boolean – variableObject : QVariant – variableIsSelectable : boolean

Figure C.13.: Class diagrams of the *Action* and *StateVariable* classes. Implemented actions are listed in Table C.2, implemented state variables in Table C.1.

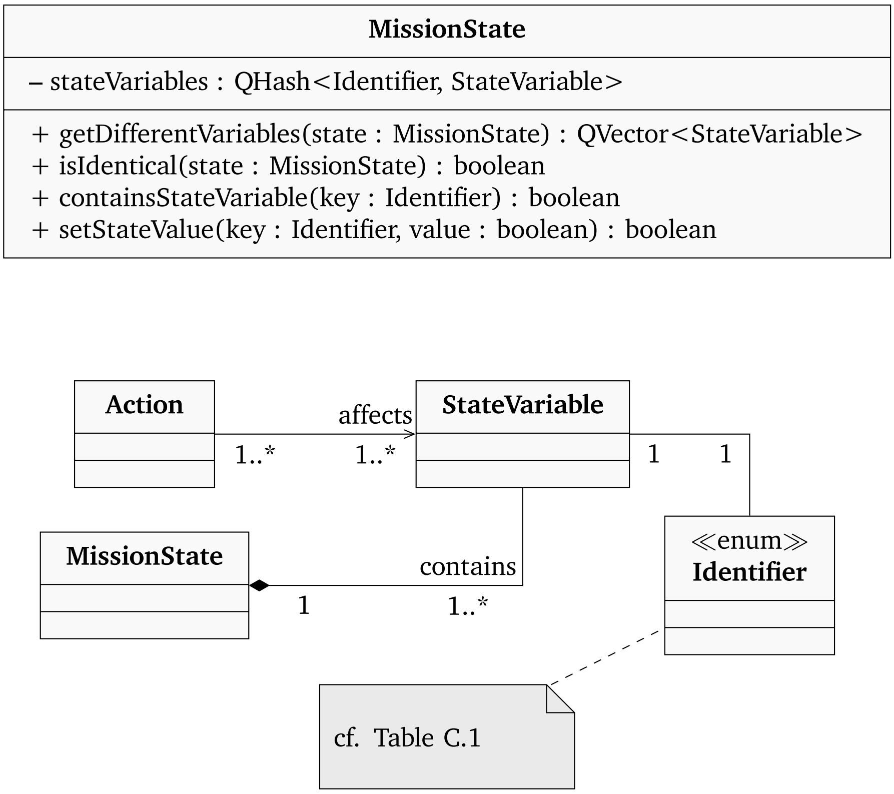


Figure C.14.: Class diagram for *MissionState* class and relationships to other classes outlined in Figure C.13.

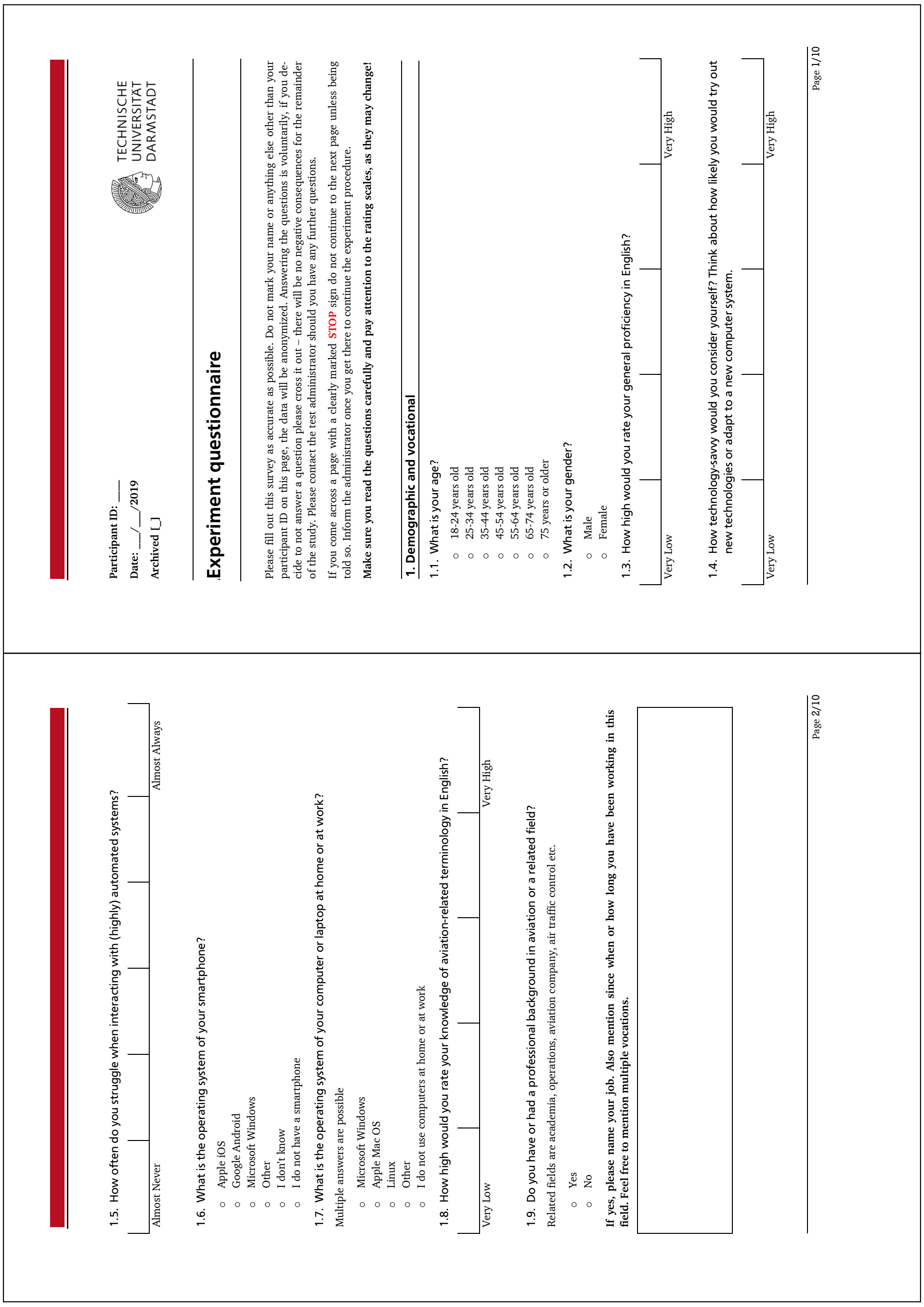
D. Evaluation

D.1. Modified Cooper-Harper scale

The following questions are used for the Modified Cooper-Harper (MCH) rating scale provided to the participants. The layout of the scale is based on the depiction in [Gaw08].

- Even though errors may be large or frequent, can instructed task be accomplished most of the time? *If no...*
  - **(10) Impossible** Instructed task cannot be accomplished reliably.
- *If yes:* Are errors small and inconsequential? *If no...*
  - **(9) Major difficulty** Intense operator mental effort is required to accomplish task, but frequent or numerous errors persist.
  - **(8) Major difficulty** Maximum operator mental effort is required to avoid large or numerous errors.
  - **(7) Major difficulty** Maximum operator mental effort is required to bring errors to moderate level.
- *If yes:* Is mental workload level acceptable? *If no...*
  - **(6) Very objectionable but tolerable difficulty** Maximum operator mental effort is required to attain adequate system performance.
  - **(5) Moderately objectionable difficulty** High operator mental effort is required to attain adequate system performance.
  - **(4) Minor but annoying difficulty** Moderately high operator mental effort is required to attain adequate system performance.
- *If yes:*
  - **(3) Fair, mild difficulty** Acceptable operator mental effort is required to attain adequate system performance.
  - **(2) Easy, desireable** Operator mental effort is low and desired performance is attainable.
  - **(1) Very easy, highly desireable** Operator mental effort is minimal and desired performance is easily attainable.

D.2. Questionnaire



Participant ID: ___
Date: ___/___/2019
Archived [_]

TECHNISCHE UNIVERSITÄT DARMSTADT

Experiment questionnaire

Please fill out this survey as accurate as possible. Do not mark your name or anything else other than your participant ID on this page, the data will be anonymized. Answering the questions is voluntarily, if you decide to not answer a question please cross it out – there will be no negative consequences for the remainder of the study. Please contact the test administrator should you have any further questions.

If you come across a page with a clearly marked **STOP** sign do not continue to the next page unless being told so. Inform the administrator once you get there to continue the experiment procedure.

Make sure you read the questions carefully and pay attention to the rating scales, as they may change!

1. Demographic and vocational

1.1. What is your age?

- 18-24 years old
- 25-34 years old
- 35-44 years old
- 45-54 years old
- 55-64 years old
- 65-74 years old
- 75 years or older

1.2. What is your gender?

- Male
- Female

1.3. How high would you rate your general proficiency in English?

Very Low — Very High

1.4. How technology-savvy would you consider yourself? Think about how likely you would try out new technologies or adapt to a new computer system.

Very Low — Very High

Page 1/10

1.5. How often do you struggle when interacting with (highly) automated systems?

Almost Never — Almost Always

1.6. What is the operating system of your smartphone?

- Apple iOS
- Google Android
- Microsoft Windows
- Other
- I don't know
- I do not have a smartphone

1.7. What is the operating system of your computer or laptop at home or at work?

Multiple answers are possible

- Microsoft Windows
- Apple Mac OS
- Linux
- Other
- I do not use computers at home or at work

1.8. How high would you rate your knowledge of aviation-related terminology in English?

Very Low — Very High

1.9. Do you have or had a professional background in aviation or a related field?

Related fields are academia, operations, aviation company, air traffic control etc.

- Yes
- No

If yes, please name your job. Also mention since when or how long you have been working in this field. Feel free to mention multiple vocations.

Page 2/10

1.10. Do you have piloting experience?

- ○ Yes
- ○ No

If yes, please indicate in which area (GA, CA, BA, Military). Also indicate approximate flight hours, models, seat and certification(s) you obtained.

2. Auxiliary

2.1. Is this the first time you are participating in a flight simulator study?

- ○ Yes
- ○ No

2.2. How are you feeling today?

Think about how awake you are, if you are taking any impairing medication or if you are feeling less concentrated than normal.

		Average		Very Good

Please do not continue unless being told so.

3. Mission manager job and automated systems

After you learned about the concept of the mission manager and their role of the flight deck, please answer the following questions regarding this job description as if you were working as a mission manager.

3.1. How much do you agree with the motivation behind the mission manager concept?

Very Low				Very High

3.2. How challenging do you see the job of a mission manager?

Very Low				Very High

3.3. How much different do you see the job of a mission manager when compared with the job of a "traditional" airline pilot?

No difference				Very different

3.4. I think it is important to put the human operator in the center of flight deck operations.

Strongly disagree				Strongly agree

3.5. I think the job of a mission manager is satisfying.

Strongly disagree				Strongly agree

3.6. I like working as a mission manager. (I would like to!)

Strongly disagree				Strongly agree

3.7. What feedback do you have regarding the mission manager job description (feedback can be in ENG/GER)?

3.8. I have high trust in automated systems and their capabilities.

Strongly disagree | | | | | Strongly agree

3.9. I think that automation will increase on the flight deck.

Strongly disagree | | | | | Strongly agree

3.10. I think that the actions and intents of automated systems are often hard to follow.

Strongly disagree | | | | | Strongly agree

3.11. Think about aviation, in what orders do you think should the following goals be prioritized?

(1 = highest priority, 5 = lowest priority)

___ Reduce travel times
___ Save money
___ Fly as safe as possible
___ Protect the environment
___ Provide comfort

Please do not continue unless being told so.

4. Usability study – Please rate the interaction with the application

4.1. I think that I would like to use this system frequently

1	2	3	4	5
Strongly disagree				Strongly agree

4.2. I found the system unnecessarily complex

1	2	3	4	5
Strongly disagree				Strongly agree

4.3. I thought the system was easy to use

1	2	3	4	5
Strongly disagree				Strongly agree

4.4. I think that I would need the support of a technical person to be able to use the system

1	2	3	4	5
Strongly disagree				Strongly agree

4.5. I found the various functions in this system were well integrated

1	2	3	4	5
Strongly disagree				Strongly agree

4.6. I thought there was too much inconsistency in the system

1	2	3	4	5
Strongly disagree				Strongly agree

4.7. I would imagine that most people would learn to use this system very quickly

1	2	3	4	5
Strongly disagree				Strongly agree

4.8. I found the system very cumbersome to use

1	2	3	4	5
Strongly disagree				Strongly agree

4.9. I felt very confident using the system

1	2	3	4	5
Strongly disagree				Strongly agree

4.10. I need to learn a lot of things before I could get going with this system

1	2	3	4	5
Strongly disagree				Strongly agree

5. Scenario evaluation – Please rate the underlying information management concept

5.1. I think that I would like to use this system frequently

1	2	3	4	5
Strongly disagree				Strongly agree

5.2. I found the system unnecessarily complex

1	2	3	4	5
Strongly disagree				Strongly agree

5.3. I thought the system was easy to use

1	2	3	4	5
Strongly disagree				Strongly agree

5.4. I think that I would need the support of a technical person to be able to use the system

1	2	3	4	5
Strongly disagree				Strongly agree

5.5. I found the various functions in this system were well integrated

1	2	3	4	5
Strongly disagree				Strongly agree

5.6. I thought there was too much inconsistency in the system

1	2	3	4	5
Strongly disagree				Strongly agree

5.7. I would imagine that most people would learn to use this system very quickly

1	2	3	4	5
Strongly disagree				Strongly agree

5.8. I found the system very cumbersome to use

1	2	3	4	5
Strongly disagree				Strongly agree

5.9. I felt very confident using the system

1	2	3	4	5
Strongly disagree				Strongly agree

5.10. I need to learn a lot of things before I could get going with this system

1	2	3	4	5
Strongly disagree				Strongly agree

Page 8/10

4.11. Do you have any further comments at this moment (comments can be in ENG/GER)?

Please do not continue unless being told so.

Page 7/10

5.11. Please enter the Modified Cooper Harper score that you would select after the scenario:

____

You may answer the following questions based on what you experienced during the initial usability evaluation *and* the scenario evaluation.

5.12. I was able to tell which actions were performed by the automated systems and which actions required my attention.

Strongly disagree □ □ □ □ □ Strongly agree

5.13. I was able to tell what actions are going to happen next, I was not surprised by actions.

Strongly disagree □ □ □ □ □ Strongly agree

5.14. I understood why an action was being performed and what the preconditions, and effects of this action were.

Strongly disagree □ □ □ □ □ Strongly agree

5.15. I did not feel misplaced or superfluous in my role as a mission manager.

Strongly disagree □ □ □ □ □ Strongly agree

5.16. I always felt in control of the situation (during the scenario).

Strongly disagree □ □ □ □ □ Strongly agree

5.17. I was able to successfully communicate my goals to the automation.

Strongly disagree □ □ □ □ □ Strongly agree

Page 9/10

5.18. The automated systems were supporting me in reaching the defined goals.

Strongly disagree □ □ □ □ □ Strongly agree

5.19. At all times I had enough information about the mission state to perform my role.

Strongly disagree □ □ □ □ □ Strongly agree

5.20. Overall, I am willing to cooperate with automated systems in the demonstrated way.

Strongly disagree □ □ □ □ □ Strongly agree

5.21. Please leave any comments regarding the overall concept and the scenario in the following field (can be ENG/GER).

Page 10/10

D.3. Usability evaluation tasks and reference times

In Table D.1 the tasks provided to the evaluation participants are summarized along with their reference times. The intended solution path for each task is outlined below the task description.

Table D.1.: Task instructions as provided to the participants. Intended solution paths described below each task.

ID	Instruction	t_{ref}
	ASSISTANT AND AUTOMATED SYSTEMS	
1	What is the title of the next action to be performed by the automated systems? View action list > identify next action marked for automation > END	6.5 s
2	In which of the three areas does the automation report the lowest status value? Is it increasing or decreasing? View aircraft system health > identify area with lowest value > END	5.3 s
3	What is the underlying cause for the status value being low? Aircraft system health > switch to execution stage > END	6.80 s
4	What advice is given by the automated system to fix the issue? Open the recommended window. Aircraft system health > switch to execution stage > open trajectory component > click on "Attempt fix" > END	7.8 s
	GOAL DEFINITION AND PLANNING	
5	Where do you find the current mission goal? What is the active destination airport? Action list > switch to active goal > view spatial parameters > END	11.1 s
6	Edit the current goal in planning, delete all parameters. Action list > switch to active goal > open goal window > click "Delete all parameters" > END	6.5 s
7	Define a new goal parameter, reach the airport with the identifier VIDP, in which country is the airport located? Action list > switch to active goal > open goal window > click "Add parameter" > select spatial / airport > enter airport name > open airport card > click "Accept" > END	2.5 s

Table D.1.: Continuation of task instructions.

ID	Instruction	t_{ref}
8	Delete the previous goal. Define the goal to complete pushback at 12:00 UTC, then fly to EDDB. Action list > switch to active goal > open goal window > click "Delete all parameters" > click "Add parameter" > select operational > select "Pushback completed" > add temporal constraint > click "Accept" > click "Add parameter" > select spatial / airport > enter airport name > open airport card > click "Accept" > END	38.0 s
9	Let the assistant plan a solution for the last defined goal. How many of the planned actions will require your attention? Activate the solution afterwards. Action list > switch to active goal > open goal window > click "Start planning" > wait > click "Review planned actions" > count actions in list > END	34.4 s
10	Tell the automated systems to reduce speed during ground movements. Open environment window > click "Reduce speed during ground movements" > END	10.4 s
	ACTION EXECUTION	
11	What is the title of the next action to be performed by you? Action list > identify next action marked for operator > END	5.3 s
12	Perform the next action that will require your attention. Action list > identify next action marked for operator > click on "Perform action" > start passenger boarding END	17.2 s
13	What were the preconditions for the automatically executed action "Request taxi instructions"? Action list > identify action > click "Inspect preconditions and effects" > read preconditions > END	17.0 s
14	You received a new taxi route from ATC. Accept and afterwards activate it for execution. Main menu > open communication window > select new message > click on message in dialog > click "Positive answer" > END	25.7 s

Table D.1.: Continuation of task instructions.

ID	Instruction	t_{ref}
	MISSION EXECUTION	
15	What is the total cruise duration of the planned flight plan? Main menu > open trajectory window > select enroute flight plan > view time > END	13.3 s
16	Request clearance for the flight plan, accept and activate the flight plan afterwards. Main menu > open trajectory window > select enroute flight plan > click “Reviewed - request clearance from ATC” > END	30.8 s
17	Which taxiways will you follow to get to the runway? Main menu > open trajectory window > select taxi out > read taxi ways > END	15.1 s
18	Activate VHF radio and switch to the ground frequency. Main menu > open communication window > select VHF tab > turn on radio > END	10.7 s
19	What is the current position of the aircraft on the airport surface? Open environment window > view position on map > END	11.2 s
20	What is the next flight phase? Allow execution of it. Action list > identify next action marked as flight phase > click on “Perform action” > in opened environment window click “Begin flight phase” > END	14.7 s

D.4. Briefing package

The following text was handed out to the participants during the scenario briefing:

> **Mission Briefing**
> This is your mission briefing; it will contain all information for the upcoming mission.
>
> **Mission plan**
> Your flight today will leave in the late afternoon from Boston Airport (KBOS). Your destination is Los Angeles Airport (KLAX). Your route will lead you over continental US as shown in the plots on the backside. Your aircraft will be parked at gate E11 in the northern part of the airport. Due to westerly winds, runway 27 is most likely going to be your departure runway. You will have to cross the runways 04L, 04R and 15R. You will operate under the callsign JEP587.
>
> **Weather**
> Severe storms are predicted to hit the area of Boston Airport during the late afternoon. You will have to hurry up to make it in time out of the airport. This is not the first storm in this area during the last weeks. During previous storms intermittent communication outages on the airport surface were reported.
>
> **Maintenance**
> The aircraft was just preflighted, maintenance records are up to date. No flight-impacting faults have been found. The following items were marked as inoperable during the last flight:
>
> **Automated voice recognition** Automated voice recognition capabilities were found to be erroneous; feature was deactivated until scheduled maintenance next week. Voice recognition is considered a secondary system, all communication with ATC will happen via CPDLC.
>
> **Rattling sound of baggage cabinet in row 27** Passengers reported rattling sound, cabinet was sealed.
>
> **Next actions**
>
> 1. Define a goal to reach the destination airport KLAX.
> 2. Let the passengers board the aircraft as soon as possible.
> 3. Begin taxi-out when ready.
>
> END OF BRIEFING SECTION
>
> **Other instructions**
> - Follow ATC instructions when required -listen to your callsign.
> - No manual steering will be required.
> - Perform the secondary task as good as possible.
>
> Coordinate with your dispatcher if you have any questions.

D.5. Participant demographics

Table D.2.: Participant demographics and general questions as reported in sections one and two of the questionnaire. Values marked with an asterisk allowed multiple answers. A total of 13 participants participated in the evaluation.

Question	Answers
Q1.1	13x 25-34 years old
Q1.2	9x Male, 4x Female
Q1.3	Mdn=4, IQR=0
Q1.4	Mdn=4, IQR=1
Q1.5	Mdn=2, IQR=0
Q1.6*	7x Android, 7x Apple
Q1.7*	12x Windows, 2x macOS, 1x Linux
Q1.8	Mdn=4, IQR=0
Q1.9	10x Yes, 3x No
Q1.10	3x Yes, 10x No
Q2.1	9x Yes, 4x No
Q2.2	Mdn=4, IQR=1

D.6. Used R packages

All R packages are downloaded from the Comprehensive R Archive Network (CRAN).

pgirmess 1.6.9
geosphere 1.5-10
tidyverse 1.2.1
fitdistrplus 1.0-14
scales 1.0.0
tikzDevice 0.12.3

D.7. Detailed results - usability evaluation

The results obtained during the usability evaluation are summarized in the following.

Table D.3.: Detailed results for the application usability evaluation. Results are grouped into relative task completion time, score $\lambda_{I,initial}$, and total score λ_I.

	Relative time		$\lambda_{I,initial}$		λ_I	
ID	Mdn	IQR	Mdn	IQR	Mdn	IQR
	Assistant and automated systems					
1	$3.80t_{ref}$	$4.33t_{ref}$	1.0	0.0	0.05	0.29
2	$3.56t_{ref}$	$3.58t_{ref}$	1.0	0.0	0.11	0.43
3	$2.18t_{ref}$	$1.97t_{ref}$	1.0	0.0	0.46	0.83
4	$2.22t_{ref}$	$1.47t_{ref}$	1.0	0.0	0.45	0.79
	Goal definition and planning					
5	$2.17t_{ref}$	$3.39t_{ref}$	1.0	0.0	0.46	1.0
6	$1.82t_{ref}$	$3.69t_{ref}$	1.0	0.0	1.0	1.0
7	$2.24t_{ref}$	$2.00t_{ref}$	0.5	0.5	0.40	0.96
8	$2.48t_{ref}$	$1.06t_{ref}$	1.0	0.25	0.38	0.81
9	$1.99t_{ref}$	$1.16t_{ref}$	0.67	0.67	0.42	0.99
10	$4.92t_{ref}$	$5.50t_{ref}$	1.0	0.0	0.0	0.46
	Action execution					
11	$2.88t_{ref}$	$2.38t_{ref}$	1.0	0.0	0.28	1.0
12	$1.68t_{ref}$	$0.86t_{ref}$	1.0	0.0	1.0	0.0
13	$2.23t_{ref}$	$0.93t_{ref}$	1.0	1.0	0.38	1.0
14	$2.96t_{ref}$	$1.09t_{ref}$	1.0	0.0	0.26	0.27
	Mission conduction					
15	$2.26t_{ref}$	$1.27t_{ref}$	1.0	0.0	0.44	0.72
16	$1.97t_{ref}$	$1.11t_{ref}$	1.0	0.0	1.0	0.62
17	$1.23t_{ref}$	$0.72t_{ref}$	1.0	0.0	1.0	0.0
18	$2.62t_{ref}$	$2.71t_{ref}$	1.0	0.0	0.35	1.0
19	$1.84t_{ref}$	$1.28t_{ref}$	1.0	0.0	1.0	0.62
20	$2.91t_{ref}$	$1.42t_{ref}$	1.0	0.5	0.19	0.30

Free text questionnaire answers (Q4.11)

1. Der Aufbau der Untermenüs war zwischen “Trajectory” und “Environment” fast gleich (bei der Anzeige der Map). Das führte zu Verwirrung.
2. Die Position der Punkte unter dem Assistent (Status Overview) verändert die Position der Drehung - finde ich unglücklich.
3. Flight-Phase relevante Punkte sollte anders gehighlighted werden.
4. Ich hätte mehr Zeit gebraucht, mich besser mit dem System vertraut zu machen 10-20 min. App ist schon gut integriert und es gibt viel verfügbare Fuktionen!
5. Es hat eine Weile gedauert, bis ich die Hauptfunktionen von den 3 Bereichen (Goal(?) Action, Info, Assistant) gut verstanden hatte, aber ich glaube das wurde schon beim Briefing mitgeteilt.
6. Die Szenarien sind gut durchdacht.
7. Die Nutzung von Farben für die Unterscheidung zwischen Mission Manager - Automated System fand ich sehr hilfreich. Auch die Icons.
8. Der Überblick in welcher Action ich mich gerade befinde, fehlt mir ein bisschen. -> Wo ich gerade die Verantwortung habe.
9. Der allgemeine Workflow war für mich auf den ersten Blick nicht intuitiv -> Einige Sachen klappen sehr gut (Passagiere borden, Routenübersicht, ...), an anderen Stellen muss man evtl. zu häufig zwischen den Fenstern (Communication, Environment, Route Planning) wechseln?!
10. Vorherige Einführung in das zugrunde liegende Konzept -> welche Zustände gibt es? aktiv vs cleared vs executing, ...
11. Ich fand die 10 Minuten vor dem eigentlichen Versuch zu kurz, bzw. eine bessere Einführung in den Funktionsumfang hatte einen positiven Einfluss auf die Schnelligkeit und Zufriedenheit bei der erstmaligen Benutzung.
12. Die Grösse des Screens etwas unübersichtlich. nach ein paar “Trockenübungen” sollte es sehr easy sein.
13. Hübsches Design.
14. Somewhat uncomfortable to use the touchscreen - you have to hold your arms in a weird position.

General observations The following observations were made during the usability evaluation:

1. After having performed their first task the participants would often fail to immediately report task completeness. The test administrator then supported the participants by telling them how to use the respective push-and-hold button in the upper left corner. After that no further support was required.

2. In multi-step tasks participants would sometimes forget to complete parts of a task.
3. Participants sometimes mistook flight phases for trajectory portions.
4. The trajectory window was often confused with the environment window.
5. Some participants accessed requested information, for example the current position of the aircraft, through unintended ways such as viewing the aircraft's position on the preview of the taxi route.
6. Static parts of the application were often mistaken as interactive buttons and vice versa.
7. Participants sometimes indicated that they had seen a certain functionality in the application during their 10 min familiarization time, and were able to locate it when required. Other participants mentioned that hey had seen it but were unable to locate it when required.

D.8. Detailed results - scenario evaluation

The results obtained during the scenario evaluation are summarized in the following.

Free text questionnaire answers (Q5.21)

1. Hat mir insgesamt sehr gut gefallen!
2. Wenn "future actions" orange sind, habe ich das Gefühl gehabt, direkt etwas machen zu müssen, selbst wenn das nicht der Fall war.
3. Die Interaktion mit dem automatisierten System ist sehr gut integriert und übersichtlich gestaltet.
4. Das System stellt allzeit ausreichende Informationen bereit.
5. Insgesamt hätte ich mir gewünscht mehr Informationen über die Effekte von meinen Actions angezeigt zu bekommen, aber dann könnte sich die Frage stellen ob nicht dann zu viel angezeigte Informationen werden.
6. An einigen Stellen war es für mich zu schwierig die komplette Kontrolle an das automatisierte System zu überlassen und bei der automatisierten Durchführung wollte ich immer überprüfen ob alles richtig durchgeführt wird.
7. Insgesamt fand ich die Zusammenarbeit mit dem System sehr angenehm. Alle relevanten Infos waren vorhanden oder mit wenigen Klicks erreichbar.
8. Die Lernkurve ist steil. Gegen Ende habe ich die Abläufe so langsam verstanden. -> Der automatisierte Aufgabenplaner funktioniert sehr gut, und hat man einmal verstanden, dass links die Aufgaben (chronologisch sortiert) sind.
9. %-Zahlen haben mich verwirrt.
10. Blinken verleitet dazu etwas zu tun.

11. Keine Möglichkeit. zum Bremsen, Anhalten
12. Keine Möglichkeit bei Systemproblem nach aussen zu kommen.
13. VHF/CPDLC-Seite, Knopf geht unter.
14. Eine intensivere Einweisung würde die Scheue der Probanden bei den Versuchen reduzieren und ein größeres Wohlbefinden bei den Versuchen schaffen.
15. Durch das doch recht große Display kann es passieren das man teilweise am Suchen ist. Elemente könnten vielleicht größer/auffälliger hervorgehoben werden.
16. Ich hatte das Gefühl, dass ich nicht eingreifen kann, falls das Flugzeug eine Runway überfährt obwohl keine Freigabe vorliegt. (Im Szenario kamen die Freigaben früh genug)
17. Mir war nicht bewusst, wie ich im Notfall z.B. bremsen kann (oder manuelle Kontrolle übernehmen).
18. Irritierend, dass linkes Feld orange blieb, obwohl keine Handlung erforderlich war. -> Daher zu schnell die Geschwindigkeit reduziert.
19. Probleme beim verstehen des Radios gehabt. -> zu schnell
20. Irritierend, dass Fehler in der Kommunikation orange war. -> auch keine Handlung erforderlich gewesen.
21. Vielleicht noch eine Farbe einbauen die Aufmerksamkeit erzeugt aber keine Handlung erfordert.

General observations The following observations were made during the scenario-based evaluation:

1. Some participants seemingly have forgotten about the secondary task and were distracted by the outside view of the simulator.
2. All participants were advised that they could close the window containing the Very High Frequency (VHF) radio controls without causing the audio stream to stop as they were reluctant to close it after having activated the radio.
3. Along with the Controller-Pilot Data Link Communications (CPDLC) communication failure the participants were advised to monitor ground operations through the environment window. Some participants were unsure what to do next, and repeatedly opened the environment window.
4. After having received the notification about the upcoming slow down participants would start to look up the location of the taxiway on the map and start waiting to push the slow down button.

D.9. Other

D.9.1. Answers for Q3.7

1. Könnte ziemlich langweilig und einsam auf einem Langstreckenflug sein.
2. Ich denke die Weiterentwicklung des Piloten zum Mission Manager ist der logische Schritt im Kontext der zunehmenden Vernetzung und dem Aufkommen von CPs (cyber-physischen Systemen). Ich vermute durch den Paradigmenwechsel zum Mission Manager wird dessen Rolle generischer und leichter zu Besetzen. Mit Widerstand der Piloten wird zu rechnen sein, da dies einen Teil ihres elitären Status verlieren.
3. A very clear definition of the mission managers responsibilities in emergency scenarios most exist and how his/her preparation should look like for the job.
4. Die Beschreibung klingt sehr logisch und als logische Weiterentwicklung des heutigen Piloten. ich finde die "radikale" Art in Zukunft nicht mehr von Piloten sondern von M.M. zu reden gut. Wie beim autonomen Fahren würde sich mir nur die Frage der Verantwortung stellen.
5. 0 Piloten $>$ 1 Pilot
6. Inwieweit wird der Mission Manager Job in das herkömmliche Fliegen noch eingebunden? (überwacht er das System und dazu Start/Landung?)
7. Ähnliche Aufgaben wie beim konvetionelle Piloten, nur dass die direkte Steuerung der Flugzeuge entfällt.
8. Die reine Tätigkeit des MM, so wie dargestellt, kann sehr langweilig (nicht-challenging) sein.

D.9.2. Errors during the evaluation

- Depending on the order in which the task categories were shuffled downstream tasks could be affected. Participants were for example unable to continue a task because required controls were deactivated as the back end of the application was not reset during the evaluation. If participants noticed this and notified the test administrator the task was marked as completed. All affected participants noticed that the button was not enabled and notified the test administrator, the task was then marked as completed. The impact is considered *minimal*.
- During both the usability evaluation and the scenario evaluation the utilized application could terminate unexpectedly. A *moderate* impact is assumed as the test procedure could either be resumed or reconstructed based on the backup video recordings (participant 3). The application was patched to prevent the reproducible errors.

- Depending on the configuration of the flight simulation software a high framerate could negatively impact the method used for automatic taxiing. The aircraft was then unable to correctly follow the provided taxi path (participant 6). The core part of the scenario was not affected. A note in the test procedures about the correct simulator settings prevented this error in following evaluations. The impact is considered *minimal*.
- Two participants (10 and 11) were unable to continue after the usability evaluation due to technical difficulties. The participants continued with the scenario evaluation at a different day. They were given additional 10 min to re-familiarize with the application, the briefing was repeated. A *moderate* impact is assumed, no effects on the test results are observed.
- Some times the application would not react to touch inputs made by the user. Users were made aware of the issue and were instructed to repeat their inputs when necessary. A *minimal* impact is assumed.
- In one case (participants 14) the VHF radio interaction was double recorded at the same position, one occurrence was removed. A *minimal* impact is assumed.