Exploring Potential Benefits and Challenges of Touch Screens on the Flight Deck

Huseyin Avsar

Thesis submitted to the University of Nottingham for the degree of
Doctor of Philosophy

February 2017

Acknowledgment

I would like to thank my supervisor Tom Rodden for giving me the opportunity to undertake my PhD and guiding me throughout my time as a PhD student. In addition, I would like to thank my supervisor Joel Fischer for his extraordinary support and motivation that always inspired me in seeking new knowledge. Special thanks to our research administrators Samantha Allen Stapleford, Felicia Knowles, Christine Fletcher and Lee Kelsall Edwards for all their help at the University of Nottingham. Last, I would like to thank all lab members of the Mixed Reality Lab, especially my office mates James Colley and Tommy Nilsson.

Many thanks to Nestor Perales Gomez and Joaquin Maceiras who approved and organised my research at the Spanish Maritime Safety Agency (SASEMAR). I would like to thank GE Aviation for supporting my EPSRC ICASE Studentship. This work has partially been funded by EPSRC and GE Aviation.

I am immensely grateful to my parents, Erdogan and Fatma Avsar, for allowing me to go and study in a foreign country and enduring absence of their son. I am thankful to my wife Tugba. I could not have managed this accomplishment without her support and love. Her presence in my life is the biggest source of happiness and strength. I also want to thank my son Erdogan for the happiness and joy that he brought along to our life during this period.

Finally, I would like to thank all my friends and their families around us in Nottingham for their friendships and support.

Abstract

As the avionics industry is seeking to introduce touch screens into most flight decks, it is vital to understand the interactional challenges and benefits of doing so. The potential benefits and challenges of touch screen technology on flight decks was investigated by means of a variety of qualitative and quantitative research methods (mixed methods approach). A number of research questions are addressed, which have been iteratively developed from the literature, interviews with avionics experts and pilots. This work presents one field study, two lab studies, one observational study, one simulation study and one comparative user study, all investigating various factors/variables that could affect touch screen usability on the flight deck.

The first field study investigated interactive displays on the flight deck with search and rescue (SAR) crew members in an operational setting in helicopters. This was the first in-flight experiment where touch screens were evaluated under real conditions. The results showed the impact of target size, device placement and in-flight vibration on targeting accuracy and performance. Presented statistical analyses and observations are essential to understand how to design effective touch screen interfaces for the flight deck.

One of the lab studies evaluated (more in depth) the potential impact of display position of touch screens within a simulated cockpit. This was the first experiment that investigated the impact of various display positions on performance following Fitts' Law experiment. Results revealed that display location has a significant impact on touch screen usability. Qualitative findings from semi-structured interviews and post-experiment questionnaires supported the understanding of interactional issues on a flight deck environment which extended initial design guidelines.

Pilots brought attention to the impact of increased G-force (+Gz) as an additional environmental factor that might affect touch screen usability on agile aircrafts. Therefore, a Fitts' law experiment was conducted to understand the effect of +Gz on touch screen usability. +Gz conditions were simulated with a weight-adjustable wristband, which was the first approach to simulate increased G-force in lab environment. Empirical results and subjective ratings showed a large impact of +Gz on performance and fatigue indices.

An observational study focused on Electronic Flight Bag (EFB) (mobile device) usage on the specific domain of Search and Rescue (SAR) helicopters. The novelty in this study was the focus group in which the aim was to find features, content and functionality that a SAR pilot may wish to see in an EFB. From operational observations and interviews with pilot’s operational requirements were defined. A Digital Human Modelling Software was used to define physical constraints of an EFB and develop interface design guidelines. A scenario and virtual prototype was created and presented to pilots.

A new way of interaction to manipulate radio frequencies of avionics systems was developed based on findings achieved in this work and other relevant studies. A usability experiment simulating departures and approaches to airports was used to evaluate the interface and compare it with the current system (Flight Management System). In addition, interviews with pilots were conducted to find out their personal impressions and to reveal problem areas of the interface. Analyses of task completion time and error rates showed that the touch interface is significantly faster and less prone to user input errors than the conventional input method (via physical or virtual keypad). Potential problem areas were identified and an improved interface is suggested.

Overall, the main contribution of this research is a framework showing the relation between various aspects that could impact the usability of touch screens on the flight deck. Furthermore, design guidelines were developed that should support the usability of interactive displays on the flight deck. This work concludes with a preliminary questionnaire that can help avionic designers to evaluate whether a touch screen is an appropriate user interface for their system.

Table of Contents

Table of Figures

List of Tables

1 Introduction

The first chapter will point the interest and contributing factors of avionics manufacturer which are considering touch screen controls in their future flight deck designs. There are few published studies investigating touch screen usage on the flight deck. Research conducted in other dynamic environments revealed several factors that could affect touch screen usability. The aim of this work is to explore potential benefits and challenges of touch screens on the flight deck. A brief description of the applied research plan will be presented. It will be explained how bigger research questions were operationalised into smaller sub research questions and how they were addressed. The main contributions of this research project can be assigned to the following research areas; Human-Computer-Interaction, Human Factors and Interaction design. A broad overview of novel contributions to relevant research areas are listed. This chapter will be finished with publications and the structure of this thesis.

1.1 Problem definition and objectives

Various input devices such as mouse, trackpad, keyboard and touch screen serve users to input data into (or navigate through) a system. Since each application area has its own specific requirements the performance of input devices may vary across conditions and type of task. One of the remarkable changes of this decade is the transition to touch screen technology in nearly all sort of consumer products. Touch screen technology’s first public appearance was in the early 2000s. Touch screens became a part of the daily life with the invention of smartphones and tablets. Traditionally, cockpit designers relied on hard controls such as knobs, buttons, switches and sliders. Now, this technology has the potential to be the next big change in flight deck design.

The density of air traffic is continuously increasing. New air space concepts like SESAR [2016] (Single European Sky ATM Research) and NextGen (Next Generation Air Transportation System (US)) [2007] are designed to meet future requirements and improve overall operations. To achieve this, new avionic systems and interfaces are required. Avionics industry gained considerable interest and is seeking to understand the challenges and benefits of touch screens on flight decks. Airlines are increasingly interested in the integration of touch screen based Electronic Flight Bags (EFB) into the cockpit in order to benefit from potential reduced operational costs and crew workload [Huguely 2013].

Digital devices have long since started to replace analogue input devices on the flight deck. Considerable changes have consolidated the number of inputs (e.g. buttons, switches and knobs) and outputs (e.g. displays). Touch screen technology could push this trend towards its limits, where majority of interactions are conducted via interactive displays. The extreme case would be that physical input devices completely disappear from the flight deck and interactions with the aircraft system occur exclusively through interactive displays [Bonelli and Napoletona 2013]. An example is the future flight deck concept from Thales [2013] where interactions with the aircrafts system occurs completely through touch screens.

The Federal Aviation Administration (FAA) [2014] advised designers to demonstrate that integration of touch screens should not result in unacceptable levels of workload and error rates. Avionics designers therefore have good reason to seek for ways to reduce cognitive load of pilots with the aim to reduce the potential for human error. The primary goal in designing cockpit displays and controls is to present large amounts of information quickly and in an understandable format to pilots [Read 1996]. Academic research showed that touch screen interfaces reduce cognitive effort and provide an intuitive way of interaction [Albinsson and Zhai 2003]. However, previous studies (e.g. [Kaminani 2011]) also found that the biggest drawbacks of soft buttons (interactive elements) compared to their physical counterparts are unwanted and accidental touches and absence of tactile feedback. The flight deck is a safety critical environment, where errors in operation may result in death or serious injuries to all passengers on board [Knight 2002]. At least, two-thirds of fatal accidents are caused by human error, which makes designing a usable flight deck more important [Boeing 2007; Civil Aviation Authority 2008].

The first academic research that compared touch screen devices with other input devices in a flight deck situation was conducted by Jones [1990]. A simulator was used to compare trackball, touch screen and speech recognition. Results revealed that the touch screen concept was the most effective input method for specific tasks. It took less time to address crew alerting messages, change altitude and navigate through several subsystem menus. Authors concluded that touch screens help pilots to keep their attention, reduce cognitive effort, search time, and motor movement. A similar study was conducted by Stanton et al. [2013] which confirmed these findings. However, subjective impressions revealed an increased discomfort compared to other input devices.

Noyes and Starr [2007] demonstrated that touch screens are not the ultimate solution for input devices within flight decks. An experiment compared speech recognition and touch screen technology for executing checklists. Results showed that control inputs through touch screen are disrupting the flight performance (awareness) more than speech recognition. This is because the need of focusing on the touch screen display while interacting, which is not required for speech recognition.

The primary aim of this PhD project is to investigate potential benefits and challenges of touch screen technology on flight decks by means of a variety of qualitative and quantitative research methods (mixed methods approach). On the basis of this, a framework will be constructed showing the relation between various aspects that could affect touch screen usability on the flight deck. The secondary objective of this work is to address the challenge how to design these touch screens (by developing and recommending design guidelines) so that they are effective (acceptable workload and error rates) and ultimately usable by pilots.

1.2 Motivation

Leading companies like Thales, Honeywell, Rockwell Collins, Boeing, GE Aviation and Gulfstream are working on future flight deck concepts that incorporate touch screen controls. The flight deck evolution shows that usually changes/improvements on the flight deck are made gradually to lower the certification risks [Rogers and Schutte 1997]. An instant change like this would raise many considerations regarding airworthiness, flight integrity and acceptable flight crew performance [Dodd et al. 2014]. So, the main question is; what was the motivation of leading companies to consider this relatively new input device on future flight decks? The following section will list statements of company representatives regarding touch screen integration on flight decks.

- Mark Nikolic, Boeing Flight Deck Human Factors Engineer: “*We want to design a flight deck that pilots are going to be familiar with and that will provide the best interaction experience for them” [Boeing 2016]*
- Brian Gilbert, Boeing Flight Deck Integration Lead: *“We find that touch screens perform as well as or better than current devices in the flight deck for interacting with the displays” [Boeing 2016]*
- Kent Statler, executive vice president and chief operating officer, Commercial Systems for Rockwell Collins: “*A touch-controlled flight deck environment makes it easier for pilots to manage information and do their jobs, and speeds up the process to complete tasks.” “Touch screens are everywhere in our lives” [Rockwell Collins 2016]*
- Bob Feldmann, vice-president and general manager of the 777X programme: *“We think we’re the first [commercial] airplane to really make something that is like all our customers are used to doing in their daily lives”. [Trimble 2016].*
- Project pilots Scott Evans and Scott Martin of G500/G600: *“We have a philosophy of supporting the pilot: What the new design does is simplify the pilot interfaces, including replacing many knobs and switches with touch screen controls and eliminating the massive control yoke in favour of a new type of sidestick control that makes the*

cockpit look much less cluttered, improves the view of the instrument panel displays and helps keep pilots in the control loop." "We were charged with how to design the flight deck and its interface to be more capable and add more functionality and at the same time be more intuitive to the crew." "It's a flexibility for design that physical controls constrain you from [being able] to do." [Thurber 2015].

- Brian Sill, president, Business and General Aviation, Honeywell Aerospace: *"From consumer-like touch functionality in the cockpit to mechanical systems that reduce weight and increase flight efficiency, collectively we are providing customers, pilots and passengers with the best flight experience possible." "… touch screens dramatically reduces the number of switches, thereby enhancing pilot and passenger safety." [Honeywell 2014]*
- Jeff Merdich, director of Product Marketing for Cockpit Systems at Honeywell Aerospace: "*Pilots use touch screens in their daily consumer devices and because of this are much more accustomed to interfacing with machines through interactive screens" [Honeywell 2014]*
- Jean-Noël Perbet, head of scientific relations for Cockpit Engineering and Development at Thales: *"Touch screen interaction revolves around touch, obviously, but sight also plays a key role in optimising eye-hand coordination. Ultimately, the technology offers a much more natural and intuitive way of interacting with the system." [Thales 2015]*
- Joe Razo, principal marketing manager of Pro Line Fusion business and regional systems at Rockwell Collins: *"It's a heads-up eyes forward flying flight deck operating philosophy", "So while you maintain your scan, you can reach up and touch and you can make changes to the avionics system without breaking your concentration and your focus and looking down." [Bellamy 2013]*.
- Mr. Bonnet, the head of cockpit innovation at Thales: *"We want to create an interaction that is more intuitive and that reduces the workload, helping to keep the pilot focused on flying." "The screens*

enable imagery to be rearranged, while maps can be zoomed and manipulated in the same way as an iPhone screen“ [Clark 2013].

- Alain Paul, director of the cockpit competency centre at Thales: *“We are using the multi-touch because that can help to reduce the training burden”, “These movements are very natural, because people are using their smart phones with them, there is no need to introduce a new set of rules for people to relearn.” [Osborne 2013]*

Based on these statements it may be fair to assume that leading avionics manufacturer want to integrate touch screens because they think/found that touch screens; are easy to learn, have a more natural and intuitive way of interaction, reduce crew workload and training time, perform better than current input devices, declutter/tidy up the flight deck, reduce weight, increase flight efficiency and enhance pilot and passenger safety.

The HCI community has extensively investigated various variables that could affect touch screen usability (Chapter 2.3). Potential benefits, which are stated by the manufacturer, can only be achieved if designers understand the flight deck environment and develop design solutions that supports touch screen usability. The oldest statement from a company representative regarding touch screens on the flight deck is from 2013. At the beginning of the PhD project (2012) there were only few research that studied touch screens on flight decks. Research that were conducted in other non-stationary environments (Chapter 2.3.8) showed that this area has many open research questions and opportunities for technical solutions, such as the questions and techniques examined in this thesis.

The motivation of this work is to contribute to the design of future flight decks with touch screens by; identifying potential variables that could affect the usability, investigating the effect of these variables, understanding their relation to each other and developing design solutions that mitigate the drawbacks of this technology.

1.3 Research Questions

Interviews with avionics experts and pilots (Chapter 4) revealed various factors that might affect touch screen usability on the flight deck. It was possible to categorise these factors into four main groups: **environmental, physical, virtual** and **user** factors. Thus, the main over-arching research question **"What are the potential benefits and challenges of touch screens on the flight deck"** will be addressed with these four research questions:

- What are the environmental factors which can cause movements in the flight deck and how much will these factors affect touch screen usability?
- What physical/hardware factors are existing that can influence touch screen usability on a flight deck situation?
- How should be the interface design so it is ultimately usable by pilots in a flight deck environment?
- What are the personal factors between users that can cause a difference in performance?

The logical question resulting from these questions is:

- How are the variables from these groups related to each other and what are the physical and virtual countermeasures to alleviate negative effects of theses variables?

Later, these four main research questions were operationalised into 18 sub research questions which are iteratively developed from the literature (Chapter 2). The effect of various variables was investigated and design solutions were developed that should mitigate the drawbacks of touch screen technology in this type of environment. These sub research questions are explored and addressed in five studies, for which the research questions were;

(Note: the letter E (environmental), P (physical), V (virtual) and U (user) at the end of each sub-research question indicates the contribution to the main research question.)

1. What is the impact of in-flight vibrations on usability? (E)
2. Is there a difference in performance for device placement (display fixed or mobile)? (P)
3. What is an appropriate size for interactive elements (button size) on a touch screen installed on a flight deck? (V)
4. What is the preferred hold strategy in mobile placement? (U)
5. How should be the physical shape of the display, so it supports usability? (P)
6. Which areas on the display have an increased error rate? (V)

A field study (Chapter 5.2) was undertaken with Search and Rescue (SAR) crew members (Spain) in an operational setting in helicopters where the primary aim was to investigate the impact of in-flight vibration, device placement and target size (size of interactive elements on the user interface) on touch screen usability. Participants performed a tapping task (a modified Fitts' Law experiment) on a tablet device in mobile and fixed placement in all possible flight phases.

---

7. Is there a difference in usability for different display positions? (P)
8. Is there a difference for display displacement in vertical and horizontal direction? (P)
9. Does handedness effect the usability? (U)
10. What are physical and interface countermeasures to alleviate negative effects of handedness? (P and V)

A lab experiment (Chapter 5.3) was conducted to investigate the potential impact of display position on touch screen usability. Participants conducted a Fitts' Law experiment (as described in ISO-9241-9 [2007]) in 20 discrete display positions.

---

11. What is the impact of +Gz on error rates and usability? (E)
12. How are fatigue symptoms affected with +Gz? (U)
13. Can experience and fitness influence overall performance? (U)

A weight-adjustable wristband was used to simulate increased G-force (+Gz) conditions in a lab study (Chapter 5.4). Participants conducted a Fitts' Law experiment in three conditions (1Gz, 2Gz and 3Gz) on a fixed display.

---

14. What features, functionality and content are pilots expecting from a mobile device? (V)
15. What are physical expectations from a mobile device? (P)
16. How will pilots use mobile devices on the flight deck? (U)
17. What are interface design guidelines for one handed thumb operation? (V)

The primary aim of the observational study (Chapter 6.1) was to define features and functionalities of a mobile device (user interface) within flight deck environment. A Digital Human Modelling (DHM) software was used to determine physical constraints of an EFB. A prototyping tool was used to mock up an EFB application, which was presented to pilots. Pilots used a scenario to list requested feature and functionalities.

---

18. Which input method provides the best and safest interaction method for radio frequency changes? (P and V)

Based on developed design guidelines a prototype was mocked up that simulated a novel way to manipulate radio frequencies of COM devices. A usability experiment (Chapter 6.2) simulating departures and approaches to airports was used to evaluate the new developed interface and compare it with the current system (Flight Management System).

1.4 Research Areas

Adopted mixed methods approach to explore potential benefits and challenges of touch screens on flight deck contribute to the following research areas:

- Human Computer Interaction (HCI) - Daintith and Wright [2008] defined HCI as: *“The means of communication between a human user and a computer system, referring in particular to the use of input/output devices with supporting software. Devices of increasing sophistication are becoming available to mediate the human-computer interaction. These include graphics devices, touch-sensitive devices, and voice-input devices. HCI is a branch of the science of ergonomics, and is concerned especially with the relationship between workstations and their operators. The aim is to develop acceptable standards for such aspects as display resolution, use of colour, and navigation around an application”.*
- Human Factors (HF) - Stramler defined Human Factors as “… the field which is involved in conducting research regarding human psychological, social, physical, and biological characteristics, maintaining the information obtained from that research, and working to apply that information with respect to the design, operation, or use of products or systems for optimizing human performance, health, safety, and/or habitability”.
- Ethnography - Hammersley and Atkinson [1995] defined Ethnography as *“…a particular method or set of methods which in its most characteristic form involves the ethnographer participating overtly and covertly on people’s daily lives for an extended period of time, watching what happens, listening to what is said, asking questions – in fact, collecting whatever data are available throw light on the issues that are the focus of research”.*
- Interaction Design (IxD) - Cooper et. al [2007] defined interaction design as: *“…the practice of designing interactive digital products, environments, systems, and services”.*

1.5 Novel Contributions

All studies had at least one contribution to the listed research areas. With the aim to visualise key contributions, minor findings are filtered for this part. Figure 1.1 shows a Venn diagram of the key contributions (A-F) that shaped the framework of this research. The points below will give a broad overview of the main contributions, detailed analysis will be provided in the Chapter 8 (Discussion):

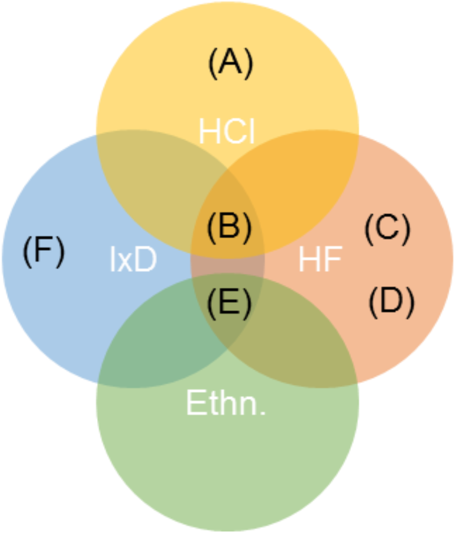


Figure 1.1 Venn Diagram of the Contribution to the Relevant Research Disciplines (A-F)

A. *A modified Fitts' Law experiment for multi touch enabled interactive displays* – Pilot studies demonstrated that the tapping task design as described in ISO 9241-9 is not suitable for devices with multi-touch capability. Participants tended to hover their finger over the next target before clicking the current target with the other hand. This kind of predictability would lead to contrived movement time measurements compared to realistic operational use. This can cause a problem especially, if one of the objectives is to observe how potential users are going to use the device in a real-world situation (Chapter 5.1). A task design was created in which the size and the distance of each target varied dynamically from the previous one.

B. *Target size guidelines for fixed and mobile displays* - This was the first in-flight experiment (Chapter 5.2) that evaluated the effect of in-flight vibrations, device placement and size of interactive elements on touch screen usability. All tested variables have a significant impact on touch screen usability. However, increasing target size (15 mm for mobile devices and 20 mm for fixed devices) eliminates the negative effects of placement and in-flight vibration in most cases. Based on observations initial design guidelines for the physical shape of the displays and user interface are created.

C. *The impact of display position (Lab Study)* - This was the first experiment (Chapter 5.3) that investigated the impact of various display positions on performance following Fitts' Law experiment (ISO 9241-9). 20 discrete display positions were tested. Both quantitative results and semi-structured interviews showed that the location of the display has a large effect on speed and accuracy. Best results were achieved on the display position which was directly in front of participants. Performance results degrade if the display position was moved to the side of participants dominant hand. The worst performance was achieved at participants non-dominant hand side. Participants achieved higher performance values for displays positions at nearer distances than farther distances. Additional design guidelines were developed from the outcome of this study.

D. *Effect of +Gz on touch screen usability (Lab Study)* – The gravitational force was simulated with a weight adjustable wristband. This approach was the first approach that simulated +Gz in a lab environment (Chapter 5.4). Findings suggested that this method reflect ecological valid data in some extent. Empirical results and semi-structured interviews with participants showed that +Gz has a large effect on performance and fatigue development and need to be considered in the design process for agile aircrafts where pilots are frequently exposed to increased G-forces. Statistical results revealed that while the simulated +Gz increased linearly, performance decreased exponentially, and movement time increased exponentially.

E. *Electronic Flight Bags (EFB) in Search and Rescue Operations* – Operational observation, interviews with pilots (Chapter 4), questionnaire and a prototype were used to define expected features and functionalities from an EFB for Search and Rescue operations (Chapter 6.1). Results showed that each domain and type of aircraft (military, commercial or parapublic operations) will have their own specific requirements and expectations. Physical constraints of an EFB with no dedicated mounting device on the flight deck were developed with a Digital Human Modelling (DHM) software. Additional information and feedback received from the pilots extended initial design guidelines that were created during the field trials.

F. *Guidelines for touch screen user interfaces for flight decks* – A usability experiment (Chapter 6.2) comparing a new developed user interface, grounded on developed design guidelines, with the current system (Flight Management System) revealed that the touch interface is significantly faster and less prone to user input errors than the conventional input method (via physical or virtual keypad). Analyses showed that designing user interfaces that represent their real-word counterparts (skeuomorphism) will not improve the usability. User interface and physical factors of the display are playing a key role in performance.

G. *A framework for touch screen integration on the flight deck* – This is the main overarching contribution of this thesis. The outcome of research conducted within this thesis and other relevant studies were used to create a framework showing the relation of various variables with the main four groups (environmental, physical, virtual and user) that could affect touch screen usability on the flight deck. A preliminary questionnaire was created that avionics experts can use to get an initial idea whether a touch screen technology is a suitable interface for their avionics system.

1.6 Publications of this Thesis

Parts of the contents of this thesis have been accepted by or are in submission to peer-review for publication in conference proceedings in Digital Avionics Systems Conference and International Conference on Human Computer Interaction in Aerospace:

- The first set of the results of the field trials (Chapter 5.2) is the content in; Huseyin Avsar, Joel Fischer, and Tom Rodden. 2015. Target size guidelines for interactive displays on the flight deck. In 2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC). Prague: IEEE, 3C4-1-3C4-15. [Avsar et al. 2015]

  DOI:http://dx.doi.org/10.1109/DASC.2015.731140

- Expanded initial results of the field trials, the results of the lab study (Chapter 5.3) investigating the impact of display position and the literature review about HCI research (Chapter 2.3) is submitted; *Huseyin Avsar, Joel Fischer, and Tom Rodden. 2016. Physical and environmental considerations for touchscreen integration on the flight deck. Submitted.* [Avsar et al. 2016e]

- The lab study (Chapter 5.4) that tries to understand the impact of +Gz is presented in Huseyin Avsar, Joel Fischer, and Tom Rodden. 2016c. Future flight decks: impact of +Gz on touchscreen usability. In International Conference on Human Computer Interaction in Aerospace: HCI-Aero. Paris: ACM Press.[Avsar et al. 2016c]

  DOI:http://dx.doi.org/http://10.1145/2950112.2964592

- The ethnographical study (Chapter 6.1) investigating the potential benefits of a mobile device in SAR environment is published in *Huseyin Avsar, Joel Fischer, and Tom Rodden. 2016. Designing touch-enabled electronic flight bags in sar helicopter operations. In International Conference on Human Computer Interaction in Aerospace: HCI-Aero. Paris: ACM Press.* [Avsar et al. 2016a]

  DOI:http://dx.doi.org//10.1145/2950112.2964591

- The final study (Chapter 6.2) comparing touch input with conventional input methods on flight deck is presented in *Huseyin Avsar, Joel Fischer, and Tom Rodden. 2016. Designing touch screen user interfaces for future flight deck operations. In 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC). Sacramento: IEEE. (BEST STUDENT PAPER AWARD)* [Avsar et al. 2016b]

  DOI:http://dx.doi.org/10.1109/DASC.2016.7777976

- The framework (Chapter 7) showing the relation of each variables and the history of flight deck evolution (Chapter 2.1) is presented in *Huseyin Avsar, Joel Fischer, and Tom Rodden. 2016. Mixed method approach in designing flight decks with touchscreens: A framework. In 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC). Sacramento: IEEE* [Avsar et al. 2016d]

  DOI:http://dx.doi.org/10.1109/DASC.2016.7778066

1.7 Structure of the Thesis

This section will describe the structure of the thesis. Chapter 2 is the literature review which is divided into three sections. The flight deck evolution will be introduced with a special focus on how pilots retrieved and input information on the aircraft. The second part of this chapter summarises available touch screen technologies. The last part reviews research that evaluated input devices (including touch screens) in different conditions.

Chapter 3 describes the applied approach and methodology in this thesis. First, the term “usability” will be defined and introduced. A brief review of flight deck design process will be used to justify the adopted “mixed methods approach”. Qualitative and quantitative research methods which were used within this thesis will be discussed regarding their definitions and advantages.

Chapter 4 presents the initial interviews with avionics experts and pilots which were used to identify potential variables that may affect touch screen usability on the flight deck. Operational observations and interviews with pilots were conducted to understand and specify the use of context of touch screen enabled devices.

Chapter 5 is dealing with the experimental research of this thesis, which is divided into 3 parts. A general description of applied task design will be given before the field trials are described. The first part investigates interactive displays on the flight deck with Search and Rescue (SAR) crew members in an operational setting in helicopters. The second study evaluates the potential impact of display position of touch screens within a simulated cockpit in a laboratory study. The last study explores the potential impact of +Gz on touch screen usability.

Chapter 6 includes the two design studies of this thesis. The first part investigates touch screen based Electronic Flight Bags (EFB) on the specific domain of Search and Rescue (SAR) helicopters. A scenario was created that describes how a SAR pilot would use a mobile EFB in the future. Developed interface design guidelines were used to mock up an EFB application for SAR operations. Expected features by pilots are presented. The second research is a user study where a new way of interaction to manipulate radio frequencies of avionics systems is examined. A usability experiment simulating departures and approaches to airports was used to evaluate the interface and compare it with the current system (Flight Management System).

The framework showing the relation between various variables that could impact the usability of touch screens on the flight deck is presented in Chapter 7. This chapter will be concluded with a questionnaire that avionics designers can use to evaluate whether a touch screen interface is suitable for their aircraft system.

Chapter 8 presents the discussion of this thesis which will begin with an analysis of the applied methodology (mixed methods). The discussion will continue with addressing the main research questions that were raised in the Chapter 1. The last chapter concludes this thesis by summarising the thesis’ findings and contributions to relevant research areas.

2 Literature

The literature review will begin with the history of flight deck evolution with a special focus on how pilots retrieved and input information into the aircraft system. The second part of this chapter will summarise available touch screen technologies. The analysis will concentrate on advantages and drawbacks that different technologies might have in a dynamic (non-stationary) environment. The last reviews academic research that evaluated input devices (including touch screens) in different conditions. The literature review will be used to create a set of sub research questions that are essential to understand the potential benefits and challenges of touch screens on the flight deck.

2.1 Flight Deck Evolution

Cambridge dictionary defines the “flight deck” (or cockpit) as the part (located in front) of an aircraft where the pilot sits and where the controls (and instruments) are. It is a safety critical environment where pilots can see various instruments (information output) to monitor the state of the aircraft (e.g., speed, altitude and attitude) and use controls (input) to change the state. To serve the purpose of this thesis the flight deck evolution described in the following sections will largely focus on how pilots retrieved information and interacted with the aircraft system.

In 1903 Wright brothers made the first controlled, sustained powered flights. At that time, there were only three instruments on board and there was no enclosed cockpit. The pilot was only able to control the aircraft for 59 seconds and covered 260 meters [Wright Brothers Aeroplane Company 2010]. The demand for more flight information increased once aircraft were able to fly higher, faster and farther. Avionic systems made it possible to navigate through airspaces and to communicate with other aircraft and ground units. Systems and instrumentation in this period were analogue electro-mechanical or only mechanical designs. Every meter, gauge, indicator and readout provided one particular information from a (in few cases multiple) sensor and needed its own space in the cockpit. The number of instruments grew exponentially, which caused physical constraints on the flight deck. There were significant improvements in performance. For example, Lockheed SR-71 (1966-1998) was able to fly beyond three times the speed of sound at an altitude of 25 000 meters [LockheedMartin 2013]. However, the appearance of instruments and the way of interaction on the flight deck has barely changed between 1930 and 1980.

The number of instrumentation was so enormous that large commercial aircraft like Boeing 314 Clipper (1938-1941) was flown by a crew of five: two pilots, a flight engineer, a navigator and a radio operator. In the following 30 years, automation and advancement in avionics systems reduced the number of crew members from five to three. However, towards the end of the 1970s the number of mechanical instruments and controls in a commercial aircraft was more than one hundred [Wallace 1994]. Computer based technology which could increase the level of automation was available at that time, but they did not meet the safety requirements. This technology required another 10 years until it found its way into the cockpit.

Figure 2.1 Flight Deck of Concorde © C.Kath

The flight deck of the Concorde (1969-2003) can be categorized as a classical or conventional flight deck [Spitzer et al. 2000]. Figure 2.1 shows the flight deck layout of the Concorde [Kath 2006]. This cockpit was packed with analogue instruments and gauges, and compared to current flight decks there was almost no automation, which required more active flying by the pilots. Pilots were overwhelmed with information which result in increased crew workload and attention demand. This prevented a further reduction in the minimum number of crewmembers.

The increase in automation reduced crew workload and the introduction of cathode ray tubes (CRTs) gave the opportunity to “tidy up” the flight deck and to operate it with a two-man crew. CRTs enabled to display of succinct information on a small area. The first generation of “glass cockpit” had a mix of CRTs and analogue instruments. A representative example for the first generation “glass cockpit” is the flight deck of the Airbus A310 (1983), which is shown on Figure 2.2 [Califlier001 2014]. Comparing this with a classic flight deck design, it is noticeable that the newer generation looks less complex. Another significant invention was the Flight Management System (FMS) which was coupled to the map display. The FMS is a small computer that enabled pilots to create their flight plan through a keyboard, which is illustrated on the map display. There were also other avionic systems that had a digital readout, however controls were still mechanical.

Figure 2.2 Flight Deck of A310 © Calflier001

The second generation of “glass cockpit”, which include A320 (1987), had a higher level of automation. The flight deck of the A320 is shown on Figure 2.3 [Curimedia 2011]. Previously pilots had to actively fly and monitor the state of aircraft. Some models of this generation enabled coupling of autopilot with FMS. The majority of the workload was transferred from flying the aircraft to monitoring automatics. CRTs were replaced by active matrix liquid crystal displays (LCDs) that are thinner, generate less heat and consume less power [Harris 2004]. The number of displays were similar to the first generation “glass cockpit”. The reduction of analogue instruments on the dashboard is remarkable. Mechanical gauges and warning lights in previous generation were replaced, although there were some analogue instruments as backups in case of display failure. Significant changes were made on information output. Automation reduced the number of input devices; however, controls (input) were still implemented using hard controls like buttons, switches and sliders.

Figure 2.3 Flight Deck of A320 © Curimedia

The Boeing 777 (1995) was the first commercial aircraft that incorporated “cursor control”, allowing pilots to use a touchpad to interact with “soft buttons” on certain displays [K. H. Abbott 2001]. The Boeing 787 (2011) has one of the newest flight decks (Figure 2.4) [Jetstar Airways 2011]). It has fewer but larger displays and there are few hard controls installed on the dashboard. A significant advancement in terms of information retrieval replaced paper documents with integrated Electronic Flight Bags (EFB). Pilots had access to various paper charts and checklist through the EFB, which reduced the search time for documents significantly (located on the diagonal of both pilots) [Kaminani 2011]. In the area of avionics systems more advances were made in the past two decades than previous 90 years. Comparing this flight deck with its predecessors the consolidation of input and output devices is noticeable.

Touch screen technology offers a new way of intuitive interaction, which can push this trend to its limits where the majority of interaction occurs through interactive displays. All information and input keys can be accessed through the same interface, so there is less physical or space constraints [Bonelli et al. 2013].

Figure 2.4 Flight Deck of B787 © Jetstar Airways

Touch screens are adaptable to any configuration by changing the underlying software, and they do not require removing and reconfiguring physical input devices [Dodd et al. 2014]. Zero displacement between input and output, control and feedback, hand action and eye gaze, make touch screens very intuitive to use. In addition, it helps users to keep their attention, reduce cognitive effort, search time and motor movement [Albinsson and Zhai 2003]. A comparative study between various input devices revealed the touch screen as the most effective input method for navigations through subsystems [Jones 1990]. However, compared to their physical counterparts the biggest drawback of touch screen interaction is unwanted and accidental touches [Degani et al. 1992]. Another significant drawback is the absence of tactile feedback which request users to focus solely on the screen [Kaminani 2011].

More recently, Original Equipment Manufacturers (OEM) have recognized the potential benefits of this technology and started to explore opportunities for the integration of touch screens in and around the flight deck. This applies both for military and commercial aviation. An example for military is the flight deck of the Lockheed Martin F-35 [2014] and for commercial aviation is the flight deck of the Gulfstream G500/600 [Gulfstream 2015].

Advancement in avionics systems cannot prevent that ‘human error’ is the primary cause for fatal accidents. According to Boeing [2007] more than 80% of accidents are caused by the flight crew, which makes reduction in the potential for these errors through good interface design even more important.

2.2 Touch Screen Technology

This section will introduce and compare four different touch screen technologies; *resistive, capacitive, surface acoustic wave and infrared touch screens.* Depending on the purpose, each technology has its own advantages and disadvantages, which will be discussed after a brief explanation of the working principle of all touch screen technologies.

Resistive touch screens - use two layers of flexible sheets coated with a resistive material which is separated by a thin gap of air. A touch is recognised once someone (finger) or something (stylus) touches the screen and close this gap. *Surface acoustic wave (SAW) touch screens* - produce acoustic waves on the surface of the display. A part of the wave is absorbed once a solid object touches the screen. Receivers use this to estimate where the solid object interfere with the wave and set the position. *Capacitive touch screens -* consist of an insulator such as glass, coated with a transparent conductor. Since the human body is also an electrical conductor, touching the screen with a bare finger results in a distortion of the screen's electrostatic field which is measurable as change in capacitance. This will be used to determine the location of the touch on the screen. *Infrared touch screens -* have an array of infrared LED and photodetectors that are positioned around the edges. Photodetectors sense visual hulls in the LED beam once an object enters the interactive area [Dhir 2004].

Gaspar [2011] compared these technologies for an in-vehicle touch screen device. Strengths and weaknesses regarding; image quality, way and type of interaction, durability, costs were compared. In the following sections, this comparison will be performed from the perspective of flight deck design.

Ideally, touch screens on the flight deck should be usable with any object because some operations (like SAR) request pilots to wear heat resistant gloves. A significant drawback of capacitive touch screens against other technologies is that users cannot use any object to trigger the interaction. Users have to use their fingers, a special treated glove or a stylus.

Some future flight deck concepts (e.g. Thales [2013]) have only one large touch screen integrated. By taking into account that commercial flights are conducted with two pilots using a technology without multitouch capability would be a significant drawback. Previously, capacitive touch screens were the only touch screen technology that enabled multitouch functionality. Nowadays, there are different type of resistive touch screens with multitouch capability. For an intuitive operation, multitouch screen offers the possibility to design a wide range of gestures including drag, swipe, pinch and pan.

Users have to apply a certain amount of force on a resistive touch screen before it can be detected. This is an advantage for a safety critical environment because it can decrease the amount of accidental touches. Some SAW and capacitive touch screens have also the capability to measure the force applied on the screen. This would enable different actions for the same interactive element depending on the amount of pressure. However, using a resistive touch screen can be frustrating if the user has to repeat the same action on the device until it detects the touch, which would consequently increase the task completion time.

The durability (life span) is a very important topic in aviation. SAW touch screens can be damaged by outside elements. Contaminants on the surface can also interfere with the functionality of the screen. Resistive and capacitive (even longer than resistive touch screen if protective layers are integrated) touch screens have a longer live span.

Another point worth discussing is the image quality. Due the two layers on top of the screen, resistive touch screens have the worst visibility and the least amount of emitted light compared to other touch screen technologies. SAW and capacitive touch screens have the advantage that they need only one layer which means they offer a better image quality and resolution. Infrared touch screens technology may offer the best visual quality because the surface area of the screen is free.

A logical research question resulting from this section would be: *“Which touch screen technology is the most suitable one for the flight deck environment?”.* This was originally one of the research questions at the beginning of the project, however Dodd et. al [2014] published a study comparing resistive and capacitive touch screen technologies in a simulator. Results revealed that pilots committed more errors on the capacitive touch technology compared to the resistive touch technology. Authors suggested that some of these errors were due to inadvertent touches, as capacitive screens are more sensitive to touch than resistive screens.

This drawback can be compensated with a pressure sensing capacitive touch screen (e.g. Apple 3D Touch [2016]). This is a relatively new feature and there is no existing research for the flight deck environment investigating a pressure threshold that designers can use to determine whether a touch was intendent or inadvertent. Another possible solution could be a camera based eye tracking system, where the system can check whether the pilot is looking to the area where he is touching. Both potential solutions are subject to future work. This problem can also be addressed with interface design. Related academic work, which will be presented in following section, revealed that performance degrading factors can minimised by using an appropriately large target size.

Capacitive touch screens have a longer life span and a better image quality. Solving the problem with accidental touches (e.g. by setting a pressure level as activation threshold) could make this a suitable technology for the flight deck.

2.3 Related Work in HCI

The HCI literature reports a host of studies of interaction with touch screens that are reviewed in the following. Independent variables that have been studied include activity (walking or standing), mobility (mobile devices or fixed devices), usage (one handed thumb, index finger or both hands), feedback modality (auditory and haptic), target population (younger adults, elderly people, people with disease), task (alphanumeric text entry, numeric text entry, tapping task context related tasks) and environment (dynamic, in-vehicle usage). The majority of the experiments compared larger targets (or buttons) versus smaller targets and investigated if padding (small space) between targets would have a significant effect on the overall performance. Common results show that larger targets result in better accuracy than smaller targets, and that “small” padding between targets does not have a significant impact.

Related work consists of eight subsections. After summarising recommendations and design guidelines from mobile device suppliers and organisations, it will be explained in which way mentioned studies are related to this work.

2.3.1 Mobile Device Suppliers and Organisations

Mobile device suppliers have their own recommendations for target sizes, which are in general a trade-off between acceptable error rate and available screen area [Henze et al. 2011]. Apple [2014] advised developers to use 15.5 mm target size in their designs. In addition to that it is recommended to use plenty spacing between interactive elements. Microsoft [2014] recommended minimum target size is 7 mm. It is recommended to use 9 mm targets for more frequent used actions and critical tasks. It is acceptable to apply 5 mm targets if the design does not allow to use larger targets and if a mistake can be corrected within few seconds. Expected error rates for 5, 7 and 9 mm targets are 3%, 1% and 0.5%, respectively. A standard padding of 2 mm between targets is recommended for all mentioned target sizes. Google [2014] recommended a minimum target size of about 7 mm. Similar to Microsoft it is recommended to use larger targets for frequently used tasks.

Others rely on anthropometric measures to suggest appropriate target size. Ubuntu [2008] takes the size of an adult finger as a base to determine the size of interactive elements. At this point, Ubuntu is referring to research that found that the average index fingertip width is between 16 mm and 20 mm [Dandekar et al. 2003]. Targets smaller than 10 mm should be avoided. International Organisation for Standardization (ISO) [2007] has a similar view and recommend a target size equal to the breadth of the distal finger joint of a 95th percentile male (approx. 22 mm). In addition, the American National Standard Institute / Human Factors and Ergonomics society ANSI/HFES 100– [2007] standard states that there is no improvement in accuracy for target sizes larger than 22 mm.

Mobile device supplier's recommendation for target sizes produce an error rate which might be acceptable for daily usage but not for safety critical tasks. For flight deck interfaces, an appropriate target size should be selected which provides the best accuracy even in worst case situations (e.g. high vibration, turbulence and bad weather conditions) The first research question developed from this section is;

***Sub RQ**: "What is an appropriate size for interactive elements (button size) on touch screens installed on a flight deck?"*

2.3.2 Keypad (Numeric Text Input)

Gauci et. al [2015] designed a touch screen interface which was connected to a flight simulator. Pilots were able to control the aircraft system through the touch screen interface. One of the features was changing the heading, altitude and speed of the aircraft via a virtual keypad. Currently, this kind of interactions will occur through rotating buttons or a physical keyboard. Novel flight deck designs, which is already discussed in Chapter 2.2, have reduced number of physical input devices. In the following section, virtual keypad related research will be summarized. Primary independent variable in these studies were the size of interactive elements, difference between single and serial tasks and the difference between various input devices.

Schedlbauer [2007] evaluated the performance and accuracy of data input on keypads by using a fixed experimental apparatus, where the task was to type ten-digit GPS coordinates. Trackball, stylus and touch input were studied and compared. His results showed that a key size of 15 mm appears to be sufficiently large to provide acceptable accuracy for touch input (error rate: 1.9%). Padding between target sizes had no measurable effect. This value was confirmed by Tsang et. al [2013] who performed a similar experiment, and defined 15 mm targets as a cut-off point where target sizes below should be avoided. Another finding was that 20 mm targets yielded lower error rates. This outcome is supported by Colle and Hiszem [2004], who tested target sizes between 10 mm and 25 mm. Subjective and empirical measurement showed no significant difference between 20 mm and 25 mm target sizes. Spacing between targets did not show a significant effect.

Parhi and Karlson [2006] performed an experiment and evaluated the differences between discrete (single) task and serial task (input four-digit number). Participants operated a mobile device, one handed with their thumb. For discrete tasks, the authors recommended to use 7.7 mm targets and for serials task 9.6 mm. These values had error rates of 5% which is acceptable for daily usage.

Feedback modality is another independent variable which can influence touch screen performance. Lee and Zhai [2009] compared physical buttons with virtual buttons (finger and stylus use) and investigated whether audio and tactile feedback would have a significant effect on error rates and performance. The task used in this experiment was a simple multiplication operation (four digits multiplied by four digits). Results revealed that either audio or tactile feedback improves soft button performance, but no further improvement is made when both are combined. Accuracy was similar for all conditions.

In this section studies were conducted in mobile or fixed display placement and the results showed that the device placement might have a significant effect on error rates. Future flight deck incorporate mobile as well as fixed displays [Bonelli and Barsotti 2014] therefore the question is;

> ***Sub RQ:*** *“Does the device placement (display fixed or mobile) have an effect on performance on the flight deck”?*

2.3.3 Keyboard (Alphanumeric Text Input)

Creating flight plans require alphanumeric text or only text input. ICAO code of an airport has four letters, codes of navigational aids and waypoints are 4 or 5 letter alphanumeric text. Wang et al. [2015] investigated the effect of target size and shape of interactive elements. The task was to create a flight plan through a simulated Flight Management System (FMS). Usability increased with increasing target size up to 19 mm where the error rates as well as subjective rating reached asymptotes (error rate < 1%). In addition, to that, results revealed that square keys provided a better usability than rectangular keys. Keyboard studies below concentrated on the effect of touch target size on typing speed and comfort values.

Despite the fact that typing performance on a virtual keyboard is 60% slower than a conventional keyboard [Kim et al. 2012], virtual keyboards are replacing conventional keyboards. Early research conducted by Sears et. al [1993] investigated four different keyboard sizes. The target size ranged from 5.7 mm to 22.7 mm. Experienced users were able to type 21 words per minute on the smallest keyboard and 32 words per minute on the largest keyboard. In another research, Sears [1991] compared mouse, touch screen and conventional keyboard to input strings. In this experiment, he observed touch biases. Shifting touch positions allowed target size to be reduced from 26.1 mm to 22.7 mm while maintaining an error rate of less than 1%. Typing performance was similar to the results achieved by Kim. Later, he performed a study with a handheld device, where participant input strings and alphanumeric data via a stylus. Results show that keyboard size does not affect neither entry rates nor error rates. Alphanumeric tasks which requires to switch between alphabetic keyboard and the numeric keyboard do result in significantly slower data entry rates [Sears and Zha 2003].

More recent research [J. H. Kim et al. 2014] investigated typing force, muscle activities, posture and comfort during keyboard usage. Tested keyboard had square keys ranging from 13 mm to 22 mm with 2 mm padding between keys. Findings indicate that virtual keyboards with a key size of 16 mm and smaller, result in slower entry speed, high static muscle activity and lowest subjective preference. In addition to that it was demonstrated that participants with wider finger width ended with reduced typing accuracy and data entry speed. The relation between finger width and error rate was also found by Mac Kenzie [2015].

Keypad and keyboard studies showed that user interfaces representing their real-word counterparts (skeuomorphism) will worsen the usability (speed and accuracy). This is a logical outcome because these interfaces are designed for physical input devices (e.g. keyboard) Therefore, the interaction design of the user interface should be optimised for touch interaction. Thus, we can ask the question

> ***Sub RQ:*** *"Which input method provides the best and safest interaction method for flight decks?".*

2.3.4 Tapping Task and Effect of Touch Location

Tapping is one of the simplest gestures on multi-touch enabled devices. In aviation context, a single tap can trigger on-off functions, select waypoints on map, execute checklists, put landing gears or flaps up and down, activate or disable functions. The Pro Line Fusion Cockpit [Rockwell Collins 2015] design is one of the first available cockpits for retrofitting. The design has screens with single touch and all interactions occur through tapping the screen.

Henze and colleagues [2011] developed a tapping task game for smartphones. Participant's task was to touch circles appearing on the screen. This was an unsupervised experiment, which found that targets below 15 mm had an increased error rate. The error rate increased to over 40% for targets smaller than 8 mm. Over 120 million touch events were recorded which enabled to show that touch positions are systematically skewed towards a position in the lower-right of the screen. Authors assumed that the way of how participants hold the device may cause this shift. Since this experiment was uncontrolled authors cannot say exactly whether this played a role. A compensation function that shifts touch areas showed improvement in error rates. Another finding was that error rate at the border of the screen is much higher than in the centre.

Previously, Park and Han [2010] performed a tapping task with a mobile device and defined the lower right area of the screen for one handed thumb usage as inappropriate. It was demonstrated statistically that it is possible to reduce the error rates by shifting touch regions. Avrahami [2015] compared targets that appear on the centre of a tablet with targets that appear on the edge. Controversially to mouse, targets appearing on edges of the screen have a significant negative effect on reaction time.

These studies demonstrated that the target location on the screen has a significant effect on error rates. The question: ***Sub RQ**: "Which areas on the display have an increased error rate?"* should be reinvestigated in a flight deck environment before an appropriate target size recommendation can be made.

Another physiological factor that could have an impact on touch screen usage is the grip and used finger. Trudeau et al. [2016] measured the difference of one handed thumb usage and two-handed thumb usage. Tapping with a two-handed grip revealed faster and more accurate interaction than one-handed grip. Perry and Hourcade [2008] found that participants performing a tapping task with their dominant hand completed tasks more quickly and accurately than participants who used their non-preferred hand. Tested targets ranged from 3.8 mm to 11.5 mm. The difference resulting from dominant and non-dominant hand usage disappears with increasing target size. The error rate for both conditions at 11.5 mm target is around 5%. Later, Kim and Jo [2015] showed that used finger has also an impact to the usability. One-handed thumb input compared to the cradled finger-based input, revealed a significant reduction in speed and accuracy.

These studies were more focused on grip and how users use touch screens. There is no study existing that investigated the following questions in a flight deck environment:

Sub RQ: *What is the preferred hold strategy in mobile placement?*

Sub RQ: *Does the handedness effect the usability?*

Sub RQ: *What are interface design guidelines for one handed thumb operation?*

2.3.5 Age Related Differences

The minimum age to start a flight training is 16. Future pilots can have their exams with 17 (private pilot certificate) and 18 (commercial pilot certificate). Private pilots can fly an aircraft as long as they pass the medical examinations. Commercial pilots retire with the age of 65 [Federal Aviation Administration (FAA) 2015]. The potential age difference is approximately 40 years, which makes research about age-related differences on touch screen usage (in this context) important. Again, the studies reported below concentrated on the effect of target size and investigated whether padding between adjacent buttons would improve the accuracy.

Leitao and Silva [2012], published interface design guidelines for older people. Participants performed tapping and swiping tasks on a handheld device. Tested targets ranged from 7 mm to 21 mm. In their study, 14 mm (for tapping task) 17.5 mm (for swiping task) could be considered as a break-even point since there was no significant improvement for larger targets in terms of accuracy and speed. Spacing between targets did not show significant effects in either of the tasks. Xiong et al. [2014] investigated age-related difference on touch screen usability by asking participants to press (serial) square number buttons on a fixed touch screen. Tested target size ranged from 6 mm – 16 mm. Results indicated that independently from the target size elderly people (mean age 68) required approximately twice the time to complete the task with respect to young adults (mean age 22) (also stated in [Bakaev 2008]). In terms of errors, there was a significant effect only for targets below 10 mm. Wulf et al. [2015] confirmed these results and added that device orientation has a significant effect on error rates. Participants made more errors in portrait orientation than for landscape orientation.

Gao and Sun [2015] demonstrated that spacing between targets decreased the number of errors for elderly people. Findlater et al. [2013] investigated age-related performance with touch screen compared to traditional mouse input. Participants performed various tasks including pointing, dragging, crossing and steering. As expected, findings showed that elderly people (mean age 74) were significantly slower than younger adults (mean age 28). However, the gap between touch was smaller to the mouse. By elderly people, the movement time on a touch screen was 35% over the mouse. This value was 16% by younger adults. In general touch input was faster than mouse. The review showed that age difference is a significant factor that can affect movement speed and accuracy on touch screens. However, the difference in accuracy can be compensated by accommodating appropriately large targets. This shows the importance of previously stated research question:

> ***Sub RQ:*** *What is an appropriate size for interactive elements (button size) on touch screens installed on a flight deck?*

2.3.6 Impact of Disabilities

At the first glance, this subtitle seems to be irrelevant for this research area, because pilots cannot have a limited motor ability or a disability. However, the only research that investigated the impact of display position to touch screen usability was found in this area. This was another research question that was addressed in this thesis.

People using wheelchair often have to approach ATM or kiosk from the side. Participants performed a four-digit entry task on a fixed touch screen. Tested target size ranged from 10 mm – 30 mm. Sitting in a parallel orientation (screen on side) in front of a touch screen reduced the performance up to 48%. Authors recommended to use targets larger than 20 mm to compensate the adverse effects of sitting orientation on performance [Chourasia et al. 2013]. The flight deck is an environment where pilots cannot adjust their posture with respect to the systems they are interacting. One of the reason is limited mobility since pilots are usually strapped to the seat and there could be the case where they have to monitor different screens and systems parallel. In the following section, further research in this topic area is briefly provided.

Guerreiro et. al [2010] conducted a study with motor-impaired users and evaluated various touch gestures. Tapping was the most preferred technique by participants. It was recommended to use targets greater than 12 mm on mobile devices. Chen et. al [2013] performed a study where participants with motor control disability completed a 4-digit entry task. Tested target size ranged 10 mm - 30 mm. As stated by previous studies, participants without disabilities reached their asymptotes in error rates at 20 mm targets. In comparison, disabled participant performance continued to improve as target size increased. There was no significant effect found for padding between targets. Bertucco and Sanger [2014] evaluated whether Fitts' law prediction model held for different user groups. The user groups were tested; adults, children and children with dystonia (a disorder that causes muscles in the body to contract and spasm involuntarily). The linear relationship by Fitts' law detained for all groups, adults had the fastest movement time and children with dystonia had the slowest movement time.

The initial idea to investigate the impact of various display positions on the flight deck came during the initial interviews with avionics experts (Chapter 4.1) aiming to understand the context of use and to identify important variables that could affect touch screen usability on the flight deck. In a modern flight deck pilots are surrounded with displays. In example, Gulfstream 500/600 has displays in front, on diagonal, on side and above. The first question to investigate is:

Sub RQ: *"Is there a difference in performance for different display positions?".*

The distance between the displays and pilots should be optimized for direct manipulation. In a Agusta Westland 139 the distance between the pilot sitting position to the head down display is 65 cm. According to Pheasant [2005] this is outside the "zone of convenient reach". Therefore, the following question should be investigated as well

Sub RQ: *Is there a difference for display displacement in vertical and horizontal direction?*

2.3.7 Effect of Walking (Divided Attention)

There is a significant body of research that investigated the impact of walking to mobile device usage. Operating a mobile device while walking requires people to split their attention. In this context, the activity walking can be classified as primary task and using a mobile device as secondary. A similar situation applies to pilots flying an aircraft. Their primary task is to fly the aircraft safely. Interacting with aircraft system has a secondary order, which need dividing their attention. In some studies, researchers controlled the path (pre-defined road) and speed (treadmill) and observed how participants used the mobile device while walking.

A study [Schildbach and Rukzio 2010] with mobile devices found that walking (on a pre-defined test track) degrades the performance and increases cognitive load significantly. While standing, users performing a two-dimensional tapping task (as described in ISO 9241-9) made on average 6.77% fewer errors and time on task was reduced by 30%. The largest tested target size was 9.5 mm (error rate 16%). The authors claim that increasing the target size by 40% would compensate the negative effects of walking.

Bergstrom-Lehtovirta et al. [2011] performed target selection while walking on a treadmill, and conclude that all types of walking, regardless of speed, causes a noticeable decrease in accuracy. A different research showed that holding mobile devices with both hands does not provide additional stability or input accuracy [Nicolau and Jorge 2012]. In real world situation users has to be attentive to the environment to avoid obstacles and collisions. Conradi et al. [2015] added a virtual scene in front of a treadmill. The primary task was to navigate through a hierarchical menu structure (5 touches/task) on a smartphone and the secondary task was to report distractors as soon as they show up on screen. Tested target size (square) ranged from 5 mm to 14 mm. 14 mm target showed low error rates while walking as well as while standing.

Hayes et. al [2014] conducted a user evaluation using a tablet to present a target selection task within a map-based interface. Participants performed the experiment while seated or while walking in an uncontrolled indoor environment. Investigators requested to hit the centre of the targets. Results showed that participants had a higher deviation from the centre while walking. 7 mm targets while seated and 9 mm while walking result in 4% error rate. Mizobuchi et. al [2005] recorded walking speed of participants to see whether it has a significant impact to performance, which showed no significant interaction on text entry speed or accuracy. This can be supported by Lin et. al [2007], that compared stylus input while sitting, standing, walking on a treadmill and walking on an obstacle course. Analogue to previous mentioned studies error rate was highest on the obstacle course and lowest in seated position. An observational study should be conducted to see and understand how pilot’s interacting with the aircraft system currently. Further, an in-flight experiment with touch screens can provide an idea about:

***Sub RQ**: “How would pilots use touch screens on the flight deck?”*

2.3.8 Dynamic Environments (In Vehicle Usage)

This is the part that is most relevant to this work. Pilots have to interact with the aircraft system in a dynamic/vibrating environment. Relatively to air vehicles, there are lots of research published in the recent years for ground vehicles. This subtitle consists of two sections where the first part deals with ground vehicles and the second with aircrafts.

Lin et. al [2010] evaluated touch screen, mouse and trackball on a motion platform where a vehicle vibration was simulated. Results indicated that vibrations had a significant impact on all devices where performance, error rates and end point variation are degraded. Baldus and Patterson [2008] evaluated the usability of mouse, touchpad and touch screen while moving in a tractor on an off-road environment. Mouse and touch screen received the best performance results. For this setting, the mouse received the best subjective usability ratings. Authors assume that using a larger screen with larger targets would improve the subjective ratings of the touch screen. In addition to that it was proofed that using input devices in a moving vehicle has a significant negative effect.

Hong et al. [2011] compared touch screen with thumbstick and keyboard for pointing, dragging and text entry tasks in a military vehicle context. Results indicate that thumbstick has better performance in dragging, touch screen in pointing and keyboard in text entry tasks. The study revealed that participants preferred a handheld device which they can hold in their hands as they would be less affected by the vibration of the vehicle. Increased error rate discomfort on the arms and the obstruction of the screen by hands are disadvantages that appeared during touch screen operation. Authors recommend not to perform dragging operation with a touch screen in a moving vehicle. Wearing gloves reduce tactile feedback and consequently the performance. For applications in vehicles or with the potential use of gloves, the Department of Defense (DOD) [2012] recommended target sizes are between 10 mm and 25 mm.

More and more cars have integrated touch screens as in-vehicle information systems. Kim et al. [2014] investigated the effect of target size with respect to safety issues besides its usability. Participants entered 5-digit numbers with various target sizes while performing simulated driving. Tested target size ranged from 7.5 mm – 27.5 mm. Driving safety and the usability of in vehicle information system increased as the target size increased up to 17.5 mm (error rate 1 %) at which it reached asymptotes. Conti et. al [2015] investigated additionally age related differences and padding between targets. The mean age for younger adults were 25 and the mean age for older adults was 56. Results did not reveal any significant difference between the age groups. Additionally, there was a small effect on performance for the largest tested (10 mm) spacing. However, authors mentioned that this factor needs additional investigations. Ahmad et al. [2015] performed a tapping task study while driving in a real car on roads with different conditions. The experiment was conducted on three different road conditions; well-maintained motorway, road with mild pave, manhole covers raised depressed and minor bends and a road which has rutted and potholed surface with sever pave, milan blocks, rover bumps, random pitch and manhole covers raised-sunken. The speed of the car was adjusted according to the road condition. Depending on the road condition in-vehicle accelerations changed. Increased vibrations in the worst road condition result in high error rates. 7 mm target were used in this study. The number of errors can be minimized by increasing the target size by 3 mm, 4 mm, and 7 mm when on road type 1,2 and 3 respectively.

The flight deck is an environment, in which errors need to be minimized. The Federal Aviation Administration (FAA) (2011) advised designers to demonstrate that integration of touch screens should not result in unacceptable levels of workload and error rates. There was no explicit guidance on minimum target size or acceptable error rate under high-vibration conditions that are particularly likely in helicopter operations.

However, there is little research about the impact of dynamic (e.g. vibrating, turbulent) environments. During a flight, pilots could face particular difficulties operating touch screen devices when the display is moving or vibrating independently from the body. Recently, Dodd et al. [2014] published research performed in a flight simulator, and found that turbulence has a significant effect on error rates. Their experimental design suggests that this research was focused on commercial aircraft (above 8000 feet, at an airspeed of approximately 250 knots). Since general aviation aircraft and helicopters are smaller, lighter and operating at lower altitudes, pilots are likely to feel higher vibrations/turbulences. Thus, results from a commercial aircraft setting may not be transferrable. Therefore, the following research question should be revaluated from the perspective of an light aircraft;

Sub RQ *"What is the impact of in-flight vibrations on usability?*

Increased G-force (+Gz) is another environmental factor that can change dynamically during agile flight manoeuvres. Pilots stated (Chapter 4.3) that +Gz might have an decremental effect on touch screen usability. The first and only study that investigated the impact of +Gz on touch screen usability is performed by Le Pape and Vatrapu [2009]. Participants performed button selection and letter selection tasks on a mobile device that was attached on the thigh of participants in an aerobatic aircraft. The experiments were performed in 5 alternating Gz levels (+1Gz, +2Gz, +3Gz, -1Gz and -2Gz). Results revealed that, performance on both the button selection and letter selection tasks worsened under altered ±Gz acceleration conditions compared to the +1-Gz condition. The difference in time latency between +1-Gz and +3-Gz was approximately 20%.

In this experiment the mobile device was inside the zone of convenient reach [Pheasant and Haslegrave 2005] and the participant's hand was always at the same height. Future flight deck concepts incorporate fixed as well as mobile touch screens. For fixed displays, pilots have to extent and raise or lower their arms to interact with the aircraft system; this could be a further degrading factor (assuming no hand support is provided) on usability which needs further investigation. This raised the following research questions;

Sub RQ: *What is the impact of +Gz on usability (on fixed displays)?*

Sub RQ: *"How are fatigue symptoms affected by +Gz?"*

Sub RQ: *"Can experience & fitness influence overall performance?"*

3 Methodology

This chapter describes applied approach and methodology in this thesis. Aside from exploring the potential benefits and challenges of touch screens on the flight deck, the secondary aim was to develop design guidelines and recommendations for touch screens so that they are effective and ultimately usable by pilots. First, the term “usability” will be defined and introduced. A brief review of flight deck design process will be used to justify the adopted “mixed methods approach”. After describing available mixed methods approaches, selected “exploratory sequential mixed methods design” will be discussed. Qualitative and quantitative research methods which were used within this thesis will be listed with their definitions, and advantages will be considered.

3.1 What is Usability?

Usability is the core psychological and physiological construct in this thesis. International Standard Organisation (ISO) defines usability as “*… the effectiveness, efficiency and satisfaction with which specific users can achieve specific goals in particular environments”*. [International Standard Organisation 2015]. Jordan [1998] described effectiveness as the extent to which a goal is achieved, efficiency as the amount of effort required to accomplish a goal and satisfaction as the level of comfort and acceptability that users feel when using a product. Satisfaction is the most important aspect for consumer products whose use is voluntary.

However, the flight deck is a safety critical environment where effective and efficient operation has a higher priority than user satisfaction. Failing to operate a safety critical system may result in loss of life, significant property damage, or damage to the environment [Knight 2002]. The majority of fatal accidents are caused by human error, which makes designing a usable flight deck more important [Boeing 2012; Civil Aviation Authority 2008]. Bad interfaces are slow or error-prone to use [Dix et al. 2004]. There are various measures of usability for effectiveness and efficiency but they are supposed to test a complete system. Input via touch screen is a new way of interaction on the flight deck. At the beginning of this project there were less research about how to design usable avionics systems with touch screen interfaces. This required a reinvestigation of potential factors within the flight deck that could affect human-computer-interaction. Therefore, it was worth to consider flight deck design and other product design processes to create the approach that should investigate potential benefits and challenges of touch screens on the flight deck.

3.2 Flight Deck Design Process

Designing a flight deck is a complex, largely unwritten, variable and nonstandard process that requires simultaneous and cooperative work from a number of people with different expertise [Palmer et al. 1995]. Developing a new aircraft today takes five years from the program launch to entry into service [Reuzeau and Nibbelke 2012]. The average life cycles of military and commercial aircrafts are more than 30 years (e.g. Grumman F-14 1974-2006, Lockheed C-130 Hercules 1954-present, Boeing 737 1966-present and Airbus A320 1986-present). Douglas et. al [1998] stated that typically no change in a flight deck will be made unless there are new requirements or new objectives. Gradual changes/ improvements on a new flight deck which is similar to a previous type, lower the certification risks [Rogers and Schutte 1997]. Accepted designs (precedence) are used as a basis for certifying many of the human factors aspects on flight decks [Abbott et al. 1996]. A radical change in flight deck design would also have disadvantages for the customers in form of increased training costs.

Palmer et al. [1995] created a simplified representation of user centred flight deck design process. One of the very first steps is to define external requirements about mission, customer, flight crew, environment and regulations. This initial step applies to other product design models like sequential design process [Benington 1983], concurrent engineering [Parsaei and Sullivan 2012], “Vee” development cycle [Forsberg and Mooz 1994], DoD development cycle [Department of Defense 1988] and spiral model [Boehm 1988]. User-centred design requires designers to shape the system around the capabilities and needs of the users.

Potential users are involved from the beginning of the project and are an incremental part of each development stage [Endsley 2016]. Abbott and Rogers [1993] combined user-centred design principles with a systems-oriented approach to design a new flight deck which meet overall missions requirements. As well in this approach, designing of the flight deck or other aircraft systems will be conducted after mission requirements are defined.

Development of the Boeing 777 was one of the first projects that involved representatives from subcontractors and customer airlines. This was driven by the fact that competitors like Airbus and McDonnell Douglas were developing their own products for an emerging segment of passenger aircrafts (which was between the companies largest (B747) and second largest aircraft (B767)) and they were far ahead in the development phase [Sabbagh 1996]. Applying concurrent engineering methods, cross-departmental cooperation and transition from physical to virtual mock-ups shortened development time and reduced life cycle costs for the Boeing 777 [Sharma and Bowonder 2004; Jørgensen 2006].

Touch screen technology is a relatively new technology for the flight deck environment, which needs investigation about potential benefits and challenges in respect to current system. For this thesis, the implication from this review is to involve potential users (organisations, airlines, and pilots) and manufacturer from the begin on to identify potential factors that could affect touch screen usability on the flight deck. Based on that the extent of the impact of various factors can be examined. Moreover, in order to address all research questions stated in the introduction (Chapter 1.3) and literature review (Chapter 2.3) qualitative as well as quantitative research methods need to be applied. Therefore, a mixed methods approach was adopted in this research.

3.3 Mixed Methods Approach

This section will focus at the methodology that underlies the research presented in this thesis. Applying one particular research methodology did not suffice to address the research questions that were required to understand potential benefits and challenges of touch screens on the flight deck. Therefore, mixed methods [Creswell and Clark 2007] approach was adopted where qualitative and quantitative data collection is integrated. There are a number of definitions for “Mixed Methods Approach” which were summarised and analysed by Johnson et al. [2007]. As a result, a general definition is proposed as:

> *“Mixed methods research is the type of research in which a researcher or team of researchers combines elements of qualitative and quantitative research approaches (e.g. use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding and corroboration”.*

Mixed Methods Approach is a new methodology based on work, conducted around the late 1980s and early 1990s, from researchers with various backgrounds such as evaluation, education, management, sociology and health sciences [Creswell 2013]. It has gone through several periods of development including the formative stage, the philosophical debates, and the procedural developments which are described in detail by Creswell and Clark [2007], Teddlie and Tashakkori [1998], Johnson et al. [2007] and Symonds and Gorard [2010].

Bryman [2006] reviewed 232 social science mixed methods papers and identified 16 reasons for conducting mixed methods studies. The reason that motivated researchers to adapt/develop mixed methods approach is coincident with our motivation. It is very difficult (or not possible) to address all research questions using only qualitative or quantitative research methods since each methodology has its specific strengths and limitations (which will be discussed in the following sections). Mixed methods approach combines the strength of qualitative and quantitative data collection and minimize its limitations [Kurosu 2013]. Creswell [2013] stated that at practical level mixed methods could be an ideal approach if the researcher has access to both quantitative and qualitative data. The technique of using multiple sources to generate new knowledge (triangulation) will answer research questions from a number of perspectives [Lazar et al. 2010]. Qualitative and quantitative data are integrated in the design analyses through merging, connecting or embedding the data which will provide a more complete understanding of the research questions.

The three basic forms of mixed methods design are: *Convergent Parallel* where both methods are conducted concurrently*, Explanatory Sequential* where first quantitative method is performed than the qualitative method is performed and *Exploratory Sequential Mixed Methods* where first the qualitative method is completed before the quantitative method. Qualitative and quantitative methods can be weighted, prioritized or emphasized equal when both methods are equally important to address the research question. This applies often in convergent parallel mixed methods design. Exploratory and Explanatory sequential mixed methods design have often an unequal weighting where one method (quantitative or qualitative) is emphasized over the other method within the study [Hesse-Biber and Leavy 2010]

3.3.1 Convergent Parallel Mixed Methods Design

In convergent parallel mixed methods design both quantitative and qualitative data collection is done concurrently. The quantitative and qualitative methods are often prioritized equally. This is the only mixed method approach that enables simultaneous data collection [Stentz et al. 2012]. Therefore, it is suitable for researchers who have limited time and opportunity to collect data. First, data analyses is conducted separately, and then findings are compared whether they confirm or disconfirm each other [Watkins et al. 2015]. The key assumption of this approach is to gather information from different sources (qualitative and/or quantitative) that yield to the same result [Campbell and Fiske 1959]. To analyse and compare the results it is required to collect both forms of data using the same or parallel variables. The basic idea is merging both forms of data into a single picture [Creswell 2013].

3.3.2 Explanatory Sequential Mixed Methods Design

The explanatory sequential mixed methods design is a two-phase approach in which the researcher conducts a quantitative study in the first phase, analyses the results, and then uses the results to create the second qualitative study. Basically, qualitative data collection builds directly on the quantitative results. As it can be derived from the name the overall intention is to use qualitative data to explain and understand more in-depth initial quantitative results, which is the key idea of this design. It is useful especially if unexpected results arise from a quantitative study [Morse 1991]. Quantitative results can shape the types of qualitative questions in the second phase. Quantitative and qualitative data are analysed separately in this approach. Researchers report first the quantitative results and then qualitative findings to expand or explain the quantitative results [Creswell 2013].

3.3.3 Exploratory Sequential Mixed Methods Design

Exploratory sequential mixed methods design is the complete opposite of explanatory sequential mixed methods design where researchers first begin with a qualitative study and then conduct a quantitative study that builds on findings from the first qualitative study. The intention is to explore new variables or factors during the qualitative study that can be evaluated more in depth during quantitative study. This approach is especially useful if researchers cannot begin with a quantitative study because specific theories, variables, and measures are not known at the beginning [Hesse-Biber 2011]. Therefore, the qualitative part can be seen as a pre-study to the actual quantitative research [Baumgarten and Lahusen 2006]. The aim of qualitative study is to clarify concepts, gather explanations, gain insight, refine problems and ideas, and form hypotheses which can be used as the underlying construct for the quantitative phase [Andrew et al. 2011]. Qualitative findings and its use to build the quantitative study will be reported before quantitative results of the final phase [Creswell 2013].

3.3.4 Justification of Selected Mixed Method Design

In this thesis, a two-phase ‘Exploratory Design’ was selected where the results of the first method (qualitative) were used to develop the second method (quantitative) [Greene et al. 1989]. The “instrument development model” and the “taxonomy development model” are two kinds of exploratory model [Doyle et al. 2009]. Starting with an initial qualitative study and finishing with a quantitative study apply to both models. The difference is how the researcher connects the two phases. In instrument development model qualitative findings provide guidance of elements and scales that are needed to develop and implement a quantitative survey instrument [Beerbaum 2016].

At the beginning of the project it was unknown which variables could affect touch screen usability on the flight deck. Experienced researchers in Human Factors or in Human Computer Interaction are often able to hypothesise whether an independent variable can cause a significant effect on a dependent variable. The more interesting challenge is to find the ‘effect size’ that shows the strength of the difference between the levels of independent variables [Green et al. 1997]. Thus, a ‘taxonomy development model’ was applied where initial qualitative study is conducted to identify important variables and relations, and the following quantitative phase to test these results more in detail [Tashakkori and Teddlie 1998; Morgan 1998].

This model was applied twice in this thesis. In Chapter 4.1 (Interviews with Avionics Experts, qualitative), Chapter 5.2 (Field Trial, quantitative) and Chapter 5.3 (Lab Experiment (display position), qualitative) in the first instance, and in Chapter 4.3 (Interviews with Pilots, qualitative) and Chapter 5.4 (Lab Experiment (+Gz), quantitative) in the second instance. Both started with qualitative research where identified variables are tested in an empirical work (quantitative). Chapter 6.2 is a user study where all findings (qualitative and quantitative) from previous research were used to create the study. In the following sections applied qualitative and quantitative research methods will be introduced. Each method will be introduced with a set of definitions. Different types, structures or categories of research methods and their potential advantages and disadvantages will be listed. Finally, justification of selected methods will be given.

3.4 What is Qualitative Research?

> Denzin and Lincoln [2000] defined qualitative research as: *"… multi method in focus, involving an interpretive, naturalistic approach to its subject matter. This means that qualitative researchers study things in their natural settings, attempting to make sense of or interpret phenomena in terms of the meaning people bring them. Qualitative research involves the studied use and collection of a variety of empirical materials – case study, personal experience, introspective, life story, interview, observational, historical, interactional, and visual texts – that describe routine and problematic moments and meaning in individuals' lives."*

The most common method used to generate data in qualitative research is interview [Savin-Baden and Major 2012]. Other frequent used techniques are observations, field notes, reflexive journals and analyses of documents and materials [Marshall and Rossman 2011; Bogdan and Ksander 1980]. Qualitative methods, such as interviews, provide a better understanding of a phenomena that could not be achieved from purely quantitative methods, such as questionnaires [Silverman 2009]. In a qualitative interview, good questions should be open-ended (require more than a yes/no answer), neutral, sensitive and understandable [Britten 1999].

3.4.1 Type of Questions (Closed and Open Ended)

The way of information transfer in interviews is done by asking closed or/and open-ended questions to interviewee/s.

There are two types of closed ended questions. One type has ordered response categories, and the other type does not [Lazar et al. 2010]. In ordered closed ended questions interviewees have to select one item from a list of choices, which have a logical order [Dillman et al. 2011]. An example is Likert scale [1932] questions, where interviewees rate whether they would “strongly agree” or “strongly disagree” with a statement on a scale ranging from 1 to 5, 7 or 9. In unordered closed ended questions there is no logical order, which can be designed where respondents select one or more items. For questions designed for single selection, interviewees could answer with one or two words (like “yes” or “no”) or select a single item from a number of choices (similar to ordered questions). “How old are you?”, “Do you use a smartphone or tablet?”, “How many hours do you spent on these devices per day?” are examples which were used in this thesis where participants replied with a single word. “Which application do you use most during flight preparation?” is a question where the interviewee replied by saying an application from a number of available applications. “Which features do you want to see on an Electronic Flight Bag in the future?” was a question where interviewees selected multiple items from a list of features that could be incorporated on a mobile device.

Open ended questions cannot be answered with a simple “Yes” or “No”. Typically, open questions begin with what, how, why, or could [Ivey et al. 2011]. For example, “What are your opinions about future flight deck designs with touch screens”. Open questions allow respondents to express themselves in their own words [Foddy 1994].

MacKay and Weinstein [1998] stated that closed ended questions are helpful to verify information and open ended questions provide valuable information, greater insights, and more understanding. Fink [2003] developed a checklists to help researchers whether to use open or closed questions. Generally, it is recommended to start with easy to answer questions and then proceed to more difficult or sensitive topics. This supports to build up confidence by interviewees and create rich data that subsequently develops the interview further [Britten 1999; Gill et al. 2008].

3.4.2 Type of Interviews

> Kvale [1996] defined interview as: *"…an interchange of views between two or more people on a topic of mutual interest, sees the centrality of human interaction for knowledge production, and emphasizes the social situatedness of research data".*

Lazar et al. [2010] argued that the ability to "go deep" is the strongest argument in preferring interviews. In an interview, there are two parts; an interviewer (investigator) who is seeking for information about a specific topic, and an or several interviewee/s (participants) who has the potential to provide this information. There are three categories of interviews; fully, semi and unstructured interviews.

In a fully structured interview, the investigator uses a well-defined order of questions [Love 2005]. It is possible to skip questions based on previous questions. Questions could include both closed and open-ended questions. However, the investigator has not the freedom to add questions during the interview. The advantage of this method is that the results may be relatively easy to analyse. This kind of interview will be used to test specific hypothesis which is normally not the aim in other interview structures [David and Sutton 2004].

Semi-structured interviews give the freedom to interviewers to ask for clarification and follow up interviewees statements. New paths of views and opinions which were not initially considered can be explored [Gray 2004]. The challenge is to analyse these answers which may take ten times longer than the interview itself [Robson 2002]. Bless et al. [2006] stated that semi-structured interviews are very helpful in exploratory research. In a semi-structured interview, interviewer prepares questions as in a fully structured interview. However, the interviewer has the freedom to change the order of the questions. The questionnaire consist almost entirely of open-ended questions with probing instructions [Brace 2008].

An unstructured interview is based on a list of topics or simple questions known as an interview guide [Robson 2002]. The interviewer may ask a simple question to an interviewee at the beginning and leave the discussion go into the direction where it goes. The questions are designed to be as open as possible [Bailey 2008]. Semi-structured and unstructured interviews are considered as qualitative research method. [David and Sutton 2004].

Applied taxonomy development model requires a qualitative research method at the beginning to identify important variables. A structured interview was not suitable because the interviewer has not the freedom to add question to clarify or go deeper (with the aim to identify factors that could impede touch screen usability on the flight deck). There was the risk that valuable questions could not be considered initially because “flight decks with touch screens” was a relatively new research area and structured interviews are considered mainly for quantitative research which would conflict with the applied research methodology. The complete opposite interview strategy (unstructured) was not suitable as well because it was possible to create some questions based on previous studies (discussed in Chapter 2.3) that evaluated touch screen performance under various conditions. Therefore, semi-structured interviews were applied.

Semi-structured interviews were conducted during initial conversations with avionic experts (Chapter 4.1), after each experiment and the study that explored features, content and functionality of mobile devices (Chapter 6.1). The interviews served the function of defining important variables, creating scenarios and questionnaires. Except post experiment interviews, interviews were conducted with a focus group (experts or pilots). A set of questions were used to start and guide the interviews, the aim was to transform this to a discussion between participants to receive valuable information. If there was a statement made by a participant which was not considered initially, was asked to the following participants whether they would agree with this statement. This also helped to spot the point for data saturation. For post-experiment interviews, questions were about the experience and observations that participants made during the experiment. The output data of interviews (verbal communication) was qualitative. Quantitative data was collected with more closed questions written on questionnaires, which will be discussed in the next section.

3.4.3 Type of Questionnaires

Many people use the terms survey and questionnaire for the same purpose. However, the “questionnaire” is a list of questions and the “survey” is the entire methodological approach. Dillman [2000] stated that the questionnaire is only one element of a well-done survey. Brace [2008] described questionnaires as remote conversation between researcher and respondent.

Analogue to previous chapter questionnaires might have open as well as closed questions. Open questions are rarely used in questionnaires because they are more difficult to analyse [Gillham 2008]. In addition, the researcher will not have an immediate possibility to ask for clarification and follow up respondent thoughts. A key advantage of questionnaires compared to semi-structured (or unstructured) interviews is low cost in time and money. The investigator can send thousands of questionnaires with one click. However, there is a typically low response rate in questionnaires [Mathers et al. 2009]. There is also a known problem with motivating respondents. Initially, it was intended to distribute a questionnaire to pilots to figure out features that they would like to see on an Electronic Flight Bag. However, the response rate was very low which motivated to conduct semi-structured interviews instead.

In addition to empirical measurements, an independent rating scale was used to assess subjective impressions in the lab studies (Chapter 5.3 and 5.4). The independent rating scale taken from ISO-9241-9 [International Standard Organisation 2007] have two group of indices; general and fatigue indices. On a 7-point scale the questionnaire was formatted in a positive direction, with the highest values being associated with the most positive impressions. These data were used to understand and support quantitative data.

After the experiments, the investigator conducted a semi-structured interview with participants about their experience and observations. After all participants finished the experiment, all statements were collected and a post-experiment questionnaire was created. On five-point Likert scale participants rated if they would agree with the issues that other participants mentioned.

A similar approach was also applied during interviews with pilots where the aim was to explore features, content and functionality of mobile devices on flight decks (Chapter 6.1). The investigator took note of statements that pilots made from the previous interview. These statements were asked to other pilots whether they would agree with their colleagues. Information gained from these interviews were used to create a scenario. The scenario describes the daily life and routine of a pilot and how he uses his tablet device to complete various tasks. Participants task was to tick the features and functionality that they would like to see on a mobile device in the future.

3.4.4 Observation

Observation is a widely used method in ethnographic studies which investigates broadly the human behaviour [Angrosino 2007]. Erlandson et al. [1993] defined observation as a method that enable researchers to describe existing situation using their five senses, providing a “written photograph” of the situation under study. Marshall and Rossman [1989] defined observation as *"… the systematic description of events, behaviours, and artefacts in the social setting chosen for study"*. DeMunk and Sobo [1998] listed several advantages of applying participant observation. This include the access to the “backstage cultures” which allows detailed description of behaviours, intentions, situations and events which cannot be captured with other data collection methods. DeWalt and DeWalt [2002] stated that observation improves the quality of data collection and interpretation and facilitates the development of new research questions or hypotheses.

Bailey [2008] described four distinct forms of observation methods determined by the type/level of environment and structure. The observation can be conducted in a natural environment or in a laboratory setting. An observation is structured if the researcher counts the frequency of particular events. In an unstructured observation, the researcher records current observations and events and does not look for specific events. Observations can be conducted either as a participant observation study or a non-participant observation study [Sears and Jacko 2012]. In a participant observation study the researcher is a part of the team and act as a team member, which is not the case in a non-participant observation. Another variable is whether participants know that they are being observed or not [Karwowski 2006]. McLeod [2015] summarised three methods for data sampling; event sampling, time sampling and instantaneous sampling. In “event sampling” the researcher records only pre-defined events of interest. All other types of events are ignored. In “time sampling” the research defines a specific time period and record events occurred within this time period. In “instantaneous sampling” the research defines event which will trigger the observation and events are recorded. Everything happening before or after is ignored.

Observations were conducted during the field study (Chapter 5.2) in a natural environment to see how crew members are using mobile and fixed devices during the operation and to understand the process of operations. This was a non-participant observation where data collection was done via “event sampling method”. Participants were aware that a research was conducted that investigates the potential benefits and challenges of touch screens on the flight deck. However, the specific details the investigator was looking for was not given. These notes were also used to cross-check in which flight mode (cruise, transition and hover) the aircraft was, while participants conducted the experiment.

3.5 What is Quantitative Research?

> Given [2008] defined quantitative research as: *"… the systematic empirical investigation of observable phenomena via statistical, mathematical or computational techniques. It provides fundamental connection between empirical observation and mathematical expression of quantitative relationships. Quantitative data is any data that is in numerical form such as statistics, percentages, etc."*

According to Balnaves and Caputi [2001] measuring observations is the key task of quantitative research methods. The aim of quantitative research methods is to test pre-determined hypotheses and produce generalizable results that can be used to describe variables, examine relationships among variables and to determine cause-and-effect interactions between variables [Grove and Burns 2005; Marshall 1996]. Harwell [2011] said that quantitative research methods attempt to maximize objectivity, replicability, and generalizability of findings, and are typically interested in prediction. There are three types of research categorise; library, field, laboratory and simulation research [Kothari 2004].

Library research can be referred to the classical literature review process which needs to be done at the beginning of each research project. Analysing previous work can produce quantifiable results however in this thesis the literature was largely used to understand the problem area, to define questions that can be asked to avionics experts and to create hypothesis which need to be tested. All other mentioned research categories were incorporated in this thesis. Feasibility of laboratory and field trials were evaluated and optimised using pilot studies. In the following subsection, applied quantitative research methods will be introduced, if applicable different categories and their advantages will be described. Each subsection will be concluded with the justification of the applied method.

3.5.1 Pilot Studies (Preliminary Studies)

Van Teijlingen and Hundley [1998] describe pilot studies as mini versions of a full-scale study. Preliminary studies increase the likelihood of success during the main study. The aim of a pilot study is to identify potential problem areas that may affect the quality and validity of results [Blessing and Chakrabarti 2009]. Factors like feasibility, time, cost, adverse events and effect size are evaluated during this phase [Hulley 2007]. The setup should be as close as possible to the setup of the intended study. Testing, changing or developing new hypotheses is another advantage of pilot studies. It provides researchers with novel ideas and approaches that cannot be foreseen before the pilot study is conducted. Pilot studies provide sufficient evidence for researcher who have to decide whether to proceed with the main study. It is possible to test various approaches to collect data and to decide which approach would provide the clearest results. These advantages were summarised by Woken [2013].

With the aim to identify and correct problem areas, to evaluate the feasibility of task, to improve the experimental design and to adjust levels of independent variables pilot studies were conducted with at least three participants.

A major contribution of pilot studies was the modification of task design in the field study (Chapter 5.2). Two-dimensional Fitts' Law Experiment (as stated in ISO 9241-9 [2007]) is one of the common methods to evaluate (or compare) input device in various conditions. The task is to tap targets located around a circle in a sequential order. Since the location of the next target was predictable, participants tended to hover over the next target with one hand while tapping the current target with the other hand. Restricting participants to use only one hand would have conflicted with the goal of seeing how participants would use the device in a real-world situation. Thus, it was decided to modify the task in which the size and the position of the targets changed dynamically after each tap.

Another benefit was shaping the levels that defined display positions in the lab experiment reported in Chapter 5.3. Initially, it was envisioned to have more distinct display positions, however the pilot study revealed that participants cannot cope with this experimental setting. Therefore, levels of various independent variables were reduced so it was possible to conduct the experiments within two days (per participant).

In the lab study described in Chapter 5.4 which explored the potential impact of +Gz on touch screen usability. Participants who piloted this study determined the level of simulated G-forces to be tested in the main study.

The pilot study investigating the potential of free-air interaction described in Chapter 10.1 revealed that this kind of interaction method is not suitable for the flight deck. Thus, it was decided to cancel the main experiment which saved time and effort during the research period.

3.5.2 Empirical Methods (Lab and Field Study)

A variety of laboratory and non-laboratory research methods are available for human-computer-interaction. The most frequently used include observations, field studies, survey, usability studies, interviews, focus groups, and lab experiment. The majority of this methods are applied within this thesis, which will be discussed in the following section. This section will concentrate on field studies and lab experiments.

The key difference between field and laboratory experiments is the environment in which the intended study is conducted. The location of the experiment affects also the controllability of the study [Preece et al. 2002]. A field study is conducted in a natural environment providing ecological valid data. However, experimental manipulations can be best controlled under laboratory conditions [Lehner 1998]. In general, a lab experiment makes it easier to assign people to random conditions [Gilbert et al. 1998] and it is easier to replicate the results by a different researcher. However, being observed can cause participants to make short-term improvements which would not be the case in a real world situation (Hawthorne effect) [Landsberger 1958]. Sun and May [2013] recommended to conduct lab experiments for usability experiments and field experiments for investigating factors affecting the overall acceptability of the system.

In a real-world setting (Search and Rescue helicopters) the impact of in-flight vibrations on touch screen usability was investigated (Chapter 5.2). The investigator controlled the order of the experiment and recorded his observations. This was a semi-controlled task where the crew conducted the tapping task experiment at their own discretion, in periods of downtime from their primary activities. If participants exceed a certain amount of time on task the investigator asked to stop the task to avoid fatigue effects.

The majority of reviewed studies that compared or evaluated touch screen usability was conducted in a lab environment. This type of experiment can be easily controlled and more accurate measurements can be achieved. Research questions about the impact of display position (Chapter 5.3) and increased G-force on touch screen usability (Chapter 5.4) were addressed with data collected and analysed from lab experiments.

3.5.3 Type of Simulation Methods

> The Department of Defense (DoD) [1994] defined modelling and simulation as: *"… the use of models, including emulators, prototypes, and stimulators, either strategically or over time, to develop data as a basis for making managerial or technical decisions".* A simplified description is provided by Banks et al. [2001] who described simulation as *"... the imitation of the operation of a real-world process or system over time".*

Simulations are used to gain insight of functioning of human and natural systems [Smith 1998]. Simulations are used if real systems are not accessible, dangerous to use, designed but not yet built, or the real system itself does not exist [Sokolowski and Banks 2011]. Potential advantages and disadvantages of simulation methods are summarised by Hancock et al. [2008]. Similar to surveys the key advantage of simulation is cost and time effectiveness. Orlansky and String [1977] estimated that commercial air carriers could pay off the cost of a simulator after 9 month and the entire training facility in fewer than 2 years. Patenaude [1996] summarised time savings during the design process from 9 organisations who applied modelling and simulation methods. Another advantage is the availability of simulators, which do not require the physical presence of the object simulated. It gives the opportunity to provide training in non-existent aircraft or in aircraft in which first performance in a new system is critical [Jones 1967]. Simulators provide experience for normal and abnormal conditions in a safe and non-threatening environment. Consequently, the number of hours on vehicles are reduced which means reduced mechanical wear and tear, maintenance cost and infrastructure load on the national airspace system. The fact that simulators are environmental friendly compared to real vehicles is another point voting for modelling and simulation [Hancock et al. 2008]. There are four different simulation methods; live, virtual, constructive and hybrid [Andrews et al. 1998].

Live simulations involve live people using real systems. In example, field trial described in Chapter 5.2. The lab studies described in Chapter 5.3 and 5.4 are examples for virtual simulation where live people use a simulated system. In constructive simulation both people and system are simulated. In Chapter 6.1, pilots were asked about their physical expectations from a mobile device. A Digital Human Modelling software was used to determine the optimal size of a mobile device which can be used by the majority of pilots. A hybrid simulation is a combination of these simulation methods, where real people use proposed operational equipment in a simulated operational environment. Chapter 10.2 is discussing the envisioned human-centrifuge project where pilots will use the same equipment as pilots do in a fast jet aircraft.

3.5.4 Usability testing

In usability testing, users (target population) perform representative tasks in representative environments on early prototypes of computer interfaces [Lewis 2006]. It involves a systematic observation under controlled conditions that provides feedback on how users use the system [Nielsen 1994]. Lazar et al. [2010] stated that the basic goal of usability testing is: *"… to improve the quality of an interface by finding flaws in it"*. Usability testing can be conducted on any device ranging from desktop or laptop computers to mobile device such as tablets and smartphones [Schusteritsch et al. 2007]. Usability testing could be as simple as paper prototypes or high-fidelity prototypes that simulate real interfaces. Low fidelity prototypes or paper prototypes are used during the early design stage [Dumas and Fox 2009]. This is a cost and time effective way to present and evaluate interfaces with potential users where users may feel more comfortable giving feedback and criticize the interface [Snyder 2003]. Usability experiments are conducted later in the design stage as well when high level design choices have been made. The goal is to evaluate the effectiveness of specific design choices.

Findings from previous research was used to create a new user interface (presented in Chapter 6.2), that pilots could use to manipulate radio frequencies. The aim was to compare input methods and to figure out flaws in the initial design solution.

4 Approach

This chapter describes the approach that was applied to identify important variables that could affect touch screen usability on the flight deck. As prescribed in the adopted mixed method approach qualitative research methods were applied during the first stage of the research. Interviews and operational observations were performed with avionics experts and pilots. Questions and answers (direct quotes) will be presented alongside with supporting references.

The results were used to create the foundations of the framework. Variables were sorted into four categories (virtual, environmental, user and physical) which created the foundations of the intended framework. This framework served as a guide for further quantitative (Chapter 5) and qualitative (Chapter 6) research.

4.1 Interviews with Avionics Experts

Two unrecorded semi structured interviews were performed with avionics experts from GE Aviation Ltd. and National Police Air Support Unit (NAPS). Interviews (qualitative method) were held before the data collection process (quantitative method). A set of questions were prepared to guide the interviews. The investigator had the freedom to ask follow-up questions and to ask for clarification. Interviews were conducted with focus groups. The aim was always to turn the interview into a discussion to gain valuable information. The interviews with avionics experts revealed their intention and motivation to integrate touch screen technology into future flight decks. The primary aim of the interviews was to identify important variables that might affect touch screen usability within the flight deck. Four themes were identified from the statements that avionics experts made in the interviews;

- Touch screen - an alternative input device.
- Influence of air carriers and other customers.
- Motivation for touch screen integration.
- Factors that may affect touch screen usability

4.1.1 Touch Screen - an Alternative Input Device

The interviews started with background information asking about when and why avionics experts had the idea to consider touch screen displays as an alternative input device to hard controls.

> *Q1: "When did you had the initial idea to consider touch screen displays as an alternative input device to current available devices?"*
>
> *Engineer 1: "The idea of integrating touch screens on the flight deck existed longer. However, at the beginning computing power and response time rate did not meet the (operational) requirements. Nowadays, the current state of technology motivated us to reconsider this technology as an additional (or alternative) input device"*
>
> *Engineer 2: "Once (touch enabled) tablet devices were available we observed that significant number of pilots found their own ways to use them…"*

Early research [Albinsson and Zhai 2003; Degani et al. 1992; Noyes and Starr 2007] on touch screens stated poor computing power, response time and display update rate, which can be neglected by the current state of technology. In 2012, many avionics systems manufacturer worked on future design solutions with touch screen interfaces. This motivation may be triggered by general aviation and commercial pilots who used touch enabled mobile devices to execute a host of tasks [Barstow 2012].

> *Engineer 2: "Basically, current technological capabilities and projects initiated by SESAR and NextGen motivated us to consider touch screen technology in future flight deck concepts"*

In addition engineers mentioned SESAR [2016] (Single European Sky ATM Research) and NextGen (Next Generation Air Transportation System (US)) [2007] which are new air space concept that have common goals like to improve overall aviation system performance, to meet expected demands for increased capacity and to maintain the highest levels of safety [Coordination Committee 2014]. To achieve this, new avionic systems and interfaces are required.

The research was accelerated with the beginning of ALICIA [2014] (www.alicia-project.eu) project in 2009. The project lasted for four years where the primary aim was to extent aircraft operations in degraded visibility conditions. New technologies and applications were investigated which included touch screen controls [Bonelli et al. 2013]. ODICIS (One Display for Cockpit Interactive Solution) project was a different project that aimed to develop a single touch enabled display cockpit that will offer more space and a larger adaptability to display new functions required by SESAR and NextGen ([Kenterlis 2012]. The outcome of this project is the future flight deck design concept of Thales [Porcu 2013].

The questions about why avionics manufacturer wants this change has produced similar statements as listed in Chapter 1.2.

Q2: Why do you want this change/transition in the flight deck?

Engineer 3: "…touch screens offer an intuitive way of interaction"

Engineer 1: "I think they (touch screen interfaces) are easy to learn."

Engineers believe that touch screens are easy to learn, have a more natural and intuitive way of interaction compared to other input device. Comparisons and measurements with other input devices demonstrated reduced cognitive effort, workload, search time, motor movement and hand-eye coordination problems [Shamo et al. 1998; Kaminani 2011; Shneiderman 1997]. Since the input and output (zero displacement) occur in the same location, interaction with touch screens has been found to be intuitive [Jones 1990; Albinsson and Zhai 2003].

As we can see from these statements we can say that technological advancements in recent years, new airspace concepts and operational benefits are the main contributing factors that triggered/accelerated touch screen integration.

4.1.2 Influence of air carriers and other customers

During the second part of the interview it was asked if air carriers and other operators (e.g. military, police, search and rescue organisations) requested this integration.

> *Q2: "Is this change also requested from air carriers and other customers?"*

> *Engineer 3: "Air carriers can be seen as early adopters of touch screen technology in commercial aviation. … they saw that replacing the 15-16 kg flight bag with a tablet is a cost-effective integration. "*

Air carries recognized the potential benefit of reduced operational costs and crew workload and started their own Electronic Flight Bag (EFB) program. In 2011, FAA has authorized to use of the Apple iPad as EFB [Murphy 2011] [Paur 2011]. This was a further benefit that pilots appreciated. Approximately two years later, American Airlines was the first major air carrier that successfully integrated its EFB program [Huguely 2013].

> *Q3: "What benefits motivated air carriers to deploy tablets?"*

> *Engineer 2: "Common benefits are weight saving by replacing the traditional flight bag (saving fuel), reducing cost, and increasing operational efficiency by reducing (or eliminating) paper processes."*

> *Engineer 1: "…it offers several safety advantages (like completeness of the paperwork). For example, paper chart revisions are issued every two weeks and it is a known problem that pilots misfile a paper chart (or remove the wrong one). Pilots are able to update the revisions on a tablet within seconds."*

Searching documents, performing performance calculations, and updating documents and weather reports is significantly faster and safer with tablets.

> *Engineer 3:" ... another advantage is that personal injuries which are related to carrying the conventional flight back are completely eliminated."*

Patrick O'Keeffe, American Airline's vice president of Airline Operations Technology said that American Airlines has reduced the single biggest source of pilot injuries that are caused by carrying flight bags by using mobile EFBs" [Frost 2013].

> *Engineer 2: "…in future (air) carriers and organisations are expecting more functionality from these (mobile) devices… connectivity to aircraft system and other units is one of them"*

In future, air carriers and other customers are expecting more functionality from these devices. One common request is that mobile devices can communicate with the aircraft system. Uploading flight plans or flight plan modification using the tablet is a requested feature. Another feature that air carriers request is enabling communication with ground units (air carriers) through the tablets.

Basically, reduced (physical and cognitive) workload by crew members was the main benefit that enabled the integration of mobile touch-enabled devices into the cockpit.

4.1.3 Motivation for touch screen integration

The interview followed with questions about the potential benefits manufactures and pilots can expect from touch screen integration.

> *Q4: "What is your main motivation (as manufacturer) in this integration process?"*

> *Engineer: 1:" Touch screen technology will provide the flexibility to change the interfaces without removing (or reconfiguring) physical input devices. The interface can be customized so each part of the aircraft system has the same look and feel."*

Changing the interfaces without removing and reconfiguring physical input devices is the key advantage from the perspective of the manufacturer because after each step of the design process (e.g. requirements, analysis, design, and production) the flexibility of making design changes is reduced. This was also stated by Dodd et. al [2014]. Anderson [2014] predicted that it costs 10 times more to make a change at the next design stage. For example, spotting an error in the design stage would cost $10, however missing and detecting the error in the production will cost $100 to fix it. The increased cost is largely caused by undoing things and replacing tools or fixtures. Conventional aircraft system interfaces have hard controls (e.g. buttons, sliders and switches). A human factor related issue can be hidden until the product is launched and the device is used by many pilots.

> *Q4: "Why pilots are using mobile devices? What benefits can pilots expect from touch screen interfaces."*
>
> *Engineer 3: "Pilots are able to carry all paperwork (e.g. navigation charts, taxi procedures, weather maps, minimum equipment list, company policy manual, federal aviation regulations) on a single (mobile) device." Previously, pilots had to carry all the paperwork and the mobile device was considered as a supplement."*
>
> *Engineer 2: "Touch screen devices (smartphones and tables) are available since a decade and future flight deck concept will be available after 2020. Therefore, the pilots who will operate aircraft with touch screen flight decks will not have an adaptation problem, because they grew up with this technology."*

The main motivation why pilots used a touch enabled mobile device was the practicality of the product. Pilots were able to execute a host of tasks in all possible flight phases. Pre-flight tasks include flight planning and whether checking, in-flight tasks include checklist execution and post-flight tasks include logbook filling. From manufactures perspective, the main benefit pilots can expect from flight decks with touch screen is the familiarity of the technology they are going to use.

4.1.4 Factors that may affect touch screen usability

The last part of the interview focused on the main objective of this research. Engineers were asked what factors they would expect to have significant effect on touch screen usability on the flight deck.

> *Q5: What factors (variables) would you expect to have significant effect on touch screen usability?*

> *Engineer 3: "… usage will be in a non-stationary environment, therefore the movements within the aircraft can degrade the interaction speed and accuracy."*

The most mentioned factor was the movements within the aircraft. In-flight vibrations, turbulences and weather can cause these movements. Type of aircraft, speed and operation altitude can determine the total amount of movements felled by the pilots on the flight deck. The HCI Literature (Chapter 2) showed that the target size (size of interactive elements) should be appropriately large in a non-stationary environment to minimise errors. A small target size would increase the errors and completion time of specific tasks, which may be not acceptable for a safety critical environment such as the flight deck. A very big target size would reduce the area which can be used to display information. Based on current design prototypes from leading avionics manufacturer we can assume that touch screen displays will be significantly larger than current cockpit displays.

> *Engineer 2: The impact of various display positions (dashboard, pedestal and overhead) should be evaluated ...touch screens cannot provide tactile feedback"*

Another physical factor is the position of the display on the flight deck. As stated before; future flight deck designs incorporate mobile as well as fixed interactive displays. Beneath interaction speed and accuracy, it may have an impact on fatigue development, because the distance between the displays and pilots are not designed for touch interaction. Especially, if pilots use their non-dominant hand in particular display positions.

Available studies (e.g. [Degani et al. 1992; Kaminani 2011]) which were performed in a flight deck situation revealed that unwanted and accidental touches and the absence of tactile feedback are the biggest drawback against conventional hard controls (e.g. switch, button and slider).

> *Engineer 1: "…more important is to understand how the flight crew will operate these devices during the operation… observations can influence the interface design…. interaction strategy and interface design may influence the usability (of touch screens)".*

The touchable area (target size) is only one part of the interface. The arrangement of touchable area, used font size and icons are additional factors of the interface which could affect the usability. Touch screens offer the ability to make gestures (drag, swipe, pinch and pan). A new interaction strategy can be created for a particular task, which can be used to investigate the acceptability of pilots, the extent to which the task is achieved, completion time and accuracy.

> *Engineer 3: "… it is interesting to see what the operational differences, requirements and expectations of commercial aircrafts and other operations are (police, SAR and air ambulances) are… this area is currently unexplored""*

Commercial flights are conducted under instrument flight rules (IFR). Para public operations are usually conducted under visual flight rules (VFR) which requires actively looking outside. Touch screens request users to focus solely on the display which may be acceptable for IFR flights. Except at take-off and landing (2% of the entire flight [Boeing 2012]) pilots are not relying on looking outside. However, it is likely that this fact will be a significant trade-off against the potential benefits of touch screens. The effect of vibration and turbulence could be significantly higher in a helicopter, which would make interacting with touch screens more difficult. Engineers were interested in such operations since this was "unexplored" at this time. This motivated us to approach the Spanish Maritime Safety Agency (SASEMAR) with the aim to investigate how beneficial interactive displays would be in their operations. Identified variables in this section were the first set of variables that were identified in this research. These are listed at the end of this chapter.

4.2 Operational Observation of SAR units

Spanish Maritime Safety Agency (SASEMAR) was one of the main collaboration partner in this research project. It is essential to observe how pilots are currently interacting with aircraft system. Air bases were visited to understand how interactive displays might be used within this context. On the basis of operational observations and interviews with pilots a scenario was developed to understand how pilots wish to benefit from an EFB. This scenario is presented in Chapter 6.1.

SASEMAR have 11 helicopter bases alongside the Spanish coast. Each Search and Rescue (SAR) group consist of air and ground units. Air units conduct the operations and ground units maintain the helicopters for safe operation. Crews are operating on 12-hour shifts. The shift change occurs at 12 pm. There are 4 crew members operating the helicopter: 2 pilots, one hoist operator and one rescue swimmer. Before the current crew hand over the shift to the new crew, crew members have an informal chat about the state of the aircraft and whether they were faced with any problems during flight.

Apart from scheduled training and patrol flights, crews do not know when and where they are going. Because of the nature of rescue missions, response time is critical. Once a distress call is received, the crew is ready to take off within 15 minutes. In the air (1500-2000 feet above ground level), the crew flies with maximum cruise speed (120-130 knots) to the target location. Targets could be small and moving objects such as a person over board or small watercraft. Helicopters may have to operate in challenging areas (sea or cliffs) and weather conditions. During training flights, the crew is simulating possible scenarios. Variables for such operations are search required or not required, target type, rescue procedure, and rescue equipment used. For each training flight, two or three possible scenarios will be trained. This kind of training flight takes on average 2:15 hours. Each crew member has separate responsibilities, and they are interacting with each other continually. In real rescue missions, the pilot is usually the on-scene coordinator (OSC), who coordinates all other units. Detailed information about SAR operations are available in in the IAMSAR (International Aeronautical and Maritime Search and Rescue) Manual [2013]. In the following sections a detailed description of pre-, in- and post-flight activities will be given.

4.2.1 Pre-Flight Activities

The first thing that pilots are doing is to check the weather and NOTAM's in their responsibility area. If the crew does not have a scheduled training flight they are on standby until they are called for a mission. If a distress message reaches the responsible maritime rescue coordination centre (MRCC), pilots will be contacted via mobile phone.

After a distress call is received pilots start with mission preparation and ground crews prepare the helicopter for the flight (refuelling, loading required rescue equipment, pulling out the helicopter from the hanger). If the location of the target is known the MRCC will provide the coordinates. If there is an uncertainty about the exact location of the target, the crew have to search the estimated area. The search area and pattern is determined by MRCC which uses a simulation program that estimate the area where the target could be. If search is required, the MRCC send the search plan via email to the pilots. Previously, the MRCC provided the corner points of the search area and pilots had to calculate the waypoints by hand. Nowadays, pilots receive the parameters and they have to put this information into the Flight Management System (FMS).

Pilots check different weather reports from the area. If they are searching for a vessel and they know its name, they look for its picture online. It was noticeable that pilots have to visit various websites to gather all required information. In addition, they decide what kind of SAR equipment they plan on using during the operation. After the flight plan is created and the amount of required fuel is calculated, pilots perform the weight and balance calculation.

Once the mission preparation is finished the captain of the flight performs a mission briefing to all crew members. First the pilot describes the nature of the operation, the area (if the exact location of the target is known) and the time of the incident. If the target is a vessel, the length, structure colour and identifiable beacon light are given. In addition to that speed and heading of the vessel and the number of persons on board will be given. Secondly, the mission plan is explained, the pilot reports on the state of the sea, swell and direction and the height of waves, wind speed, and visibility on scene. After that the weather, wind speed (METAR, TAF) at the destination and an alternative return airport are given.

After that the pilot reports on the kind of SAR equipment to be used during the operation and required medical equipment. Weight and balance calculation will be presented. If search is required, the type of search pattern, the area, and the wind speed at the search area are presented. Finally, the emergency procedures are reviewed.

4.2.2 In-Flight Activities

After the briefing crew members require approximately 5 minutes to prepare themselves. In the meantime, ground units pull out the helicopter and if necessary refuel the aircraft. In a real mission, the time between first call and take-off is approximately 15 minutes.

While pilots perform pre-flight checklist, the hoist operator checks the winch and the rescue swimmer his equipment. Once the engine runs pilots require approximately 4-5 minutes to take-off. Before take-off the co-pilot uses the FMS to create the flight plan and requests clearance for take-off from the Air Traffic Controller (ATC).

As soon as the aircraft is in the air (1500-2000 feet above ground level), the crew flies with maximum cruise (120-130 IAS) speed to the target location. The co-pilot performs the after take/off checklist. On scene, targets could be small and moving objects, such as a missing person or vessel. It could be the case that helicopters have to operate in challenging areas (sea or forest) and weather conditions. If the mission involves several rescue units, the captain of the aircraft is usually the “On-Scene-Coordinator (OSC)” who coordinates all other units. OSC’s are determined by the responsible MRCC.

The captain informs the cabin crew approximately 10 minutes before they arrive at the target location. If the position is known, the helicopter will fly directly to the target and contact the vessel; if not, the pilot will head to the first waypoint of the search pattern and the search will start. The search is conducted visually. Additionally, the cabin crew can use and control the FLIR camera. Pilots can mirror the imagery on their centre display. Once the target is spotted, the co-pilot initiates the appropriate checklist. The captain will slow down and transits from cruise to hover. Once the aircraft is in hover, pilots require in average 3 minutes to position the aircraft close to the target. The hoist operator opens the door and talks with the pilot to make fine adjustments. It is also possible that the hoist operator takes full control over the aircraft and positions the aircraft by using his controller. The rescue swimmer may be connected to the winch and lowered to the target. After that the rescue equipment will be lowered. The rescue swimmer uses this equipment to secure the person to be rescued. If a belt is used, the hoist operator will pull up both in one go. If they use a basket (or a stretcher) the person to be rescued will be pulled up first, then the rescue swimmer. In training missions 2 or 3 possible scenarios will be simulated.

4.2.3 Post-Flight Activities

After the rescue mission is completed the pilot transits to cruise and head directly to target destination. Before they approach the airport, the co-pilot initiates the approach checklist and contacts the air traffic controller to request clearance to land. The approach chart of the airport is reviewed before landing. The helicopter lands on the airport and taxis towards the hanger. In a real mission, the crew transport the person into an ambulance.

After the mission, there is a debriefing session where the crew discuss the mission. Crew members share their ideas and provide constructive criticism of the mission procedure. Unusual circumstances during operation, operations which do not confirm to the manuals and procedures, and potential improvements are discussed.

After that, pilots have to do some paperwork for at least 40 minutes. They have to fill out reports for INAER (provider of aerial emergency service and aircraft maintenance), SASEMAR, aircraft, engine and personal logbook. Required information is similar and will be duplicated in different documents. Pilots have to enter the time to start engine, take-off, on-scene, rescue operation starts and end, landing, and shut down of the engines.

4.3 Interviews with Pilots

4 semi structured interviews were performed with pilots from the Spanish Maritime Safety Agency. Eight male pilots participated in the interviews. There were always two pilots on duty and interviews were conducted with both pilots at the same time. At that time SASEMAR had 3 female pilots (out of 110), which were not on duty. Participants age ranged from 32 to 47 (M=40, SD=6.2). Logged flight hours ranged from 3500 to 6000 (M=4500, SD=1200) (Participant information sheet - Appendix III). Interviews were performed after the in-flight experiments (Chapter 5.2) was completed. Interviews with pilots revealed their opinions about future flight decks with touch screens. The main objective was to define pilot expectations and requirements from a touch screens interface with a special focus on mobile devices. Four themes were identified from the statements that pilots made in the interviews;

- Thoughts about future designs
- Factors that may affect touch screen usability
- Physical and design requirements for mobile devices
- Preferred features and functionality from an EFB

4.3.1 Thoughts about future designs

Future flight deck concepts (e.g. [Thales 2014], [Rockwell Collins 2012] and [Honeywell 2015]) with touch screens were exposed to pilots and their opinions were asked whether this type of flight deck is suitable for SAR operations. The majority of pilots were sceptical about general (fixed and mobile displays) touch screen integration and pointed out a potential threat that was mentioned during the introduction.

> *Q1: What are your opinions about future flight deck designs with touch screens? Do you think they are suitable for SAR operations?*
>
> *Pilot 1: "I flew previously a (Eurocopter) Super Puma with an analogue system for COM. I was able to operate it without looking on it. Digital systems are lot easier in design but less efficient in use compared to the analogue system…"*

Touch screen interaction require users to focus solely on the screen. Observations showed that controlling through touch screen disrupted the primary flying task [Noyes and Starr 2007]. SAR pilots perform search visually and looking at the touch screen inside the flight deck would decrease the search performance.

Pilots were able to learn the patterns of an analogue interface (hard controls like, buttons and switches). Pilots are able to interact with the device without looking at it, which is not possible with a touch screen interface. At the beginning of the research there were few academic research (case studies), which are mentioned in the literature (e.g. [Jones 1990; Stanton et al. 2013; Noyes and Starr 2007]), that evaluated or compared touch screen usage in a flight deck environment. Therefore, it should be thoroughly investigated whether a touch screen interface is suitable for a particular avionics system or mission.

It was observed that some pilots use mobile devices on the ground and during the operation. Therefore, the question was asked why they are using mobile device and what sort of task they performing.

Q2: Do yo)u use a mobile device on the ground or during operation? If yes, why are you using a mobile device and what sort of tasks are you performing? If not, would you like to use one?

Pilot 2: "…keeping all important information in one place and having fast access to desired information is my main reason why I use a mobile device."

SASEMAR did not initiated an EFB program yet however two pilots (interviewees) use a tablet device to conduct various tasks. These are; checking weather and NOTAMs, executing checklists and searching approach charts. Both pilots reported that they have few colleagues who use a mobile device, as well. Pilots who do not use currently a mobile device would prefer to use a mobile device in the future.

EFB's could remove hard copies from the flight deck, which means savings in space, weight and costs. In addition, it is reported that searching, updating of documents, checklist completion and performance calculations can be done quickly and more accurately [Noyes and Starr 2007; Hamblin C 2003; Shamo et al. 1999]. Using a mobile device has the flexibility to adjust the position and view angle to achieve maximum usability. Software may provide intuitive zoom interaction and the possibility to de-clutter charts [Chandra et al. 2003].

4.3.2 Factors that may affect usability

Pilots were asked what factors they may imagine to have significant effect on touch screen operation during the flight.

Q3: What factors (variables) would you expect to have significant effect on touch screen usability?

Pilot 3: "…it could be very dangerous if I touch a different button due to vibrations… during thunder storms the vibrations are very high."

Pilots stated that in-flight vibrations and weather could impede touch screen usability. This was also mentioned by avionics experts during the initial interviews aiming to identify important variable that might affect touch screen usability.

Pilots categorized in-flight vibrations in helicopters in three categories; cruise, transition and hover. Transition down to hover phases generate the highest vibrations on the aircraft. In comparison, vibrations during cruise and hover are smaller. Especially, in winter months' pilots have to operate in challenging environments (e.g. turbulences, thunder storms). Sudden movements within the aircraft can cause accidental and unwanted touches.

> *Pilot 2: "… it would be better if I have to press harder (apply more force on the screen for activation… like I put my finger on the screen and then press harder.*

To avoid unwanted touches or touch by accident due to in-flight vibrations, pilots recommended a pressure sensitive touch screen, where pilots have to apply a certain amount of force on the interactive element to activate it.

> *Pilot 4: "I think I have to lean forward to reach the screen and if I have to repeat this each time it is fatiguing… we have to be strapped during the flight"*

Discussions between pilots revealed that the display position might also influence the performance. Pilots said that it would be more difficult in a helicopter to interact with a fixed display where the pilot has to extend his arm to reach the display.

The majority of SASEMAR pilots have a military background. Two pilots stated another environmental factor which rarely occurs in a helicopter but more frequently in fast jet aircrafts. Pilots identified increased G-force that occur during steep turns as a potential threat that could impede touch screen usability. Pilots recommended to investigate these environmental factors and consider it in the design process.

4.3.3 Physical and design requirements for mobile devices

Since, some pilots are using mobile device as EFBs and everybody would like to use one in the future the following questions was about EFB usage on the flight deck. First set of questions were about the physical aspects.

Q4: "What should be the physical size of the EFB on the flight deck, so it does not disrupt your primary task?"

Pilot 5: "There are periods where we experience high vibrations in the aircraft, especially in transition to hover phases. Thus, retrieving information from the head down displays is difficult… so the display should be large enough"

Pilot 6: "This one (10-inch) is ok for me… but I think it would be too big and heavy for smaller pilots who want to use it on the knee"

The size of the devices used by pilots range from 8 to 10-inch. The investigator showed 7, 8 and 10-inch tablets to pilots not using a mobile device and asked which device they would prefer during the flight and why. Majority of pilots' opinion was that a 7-inch tablet could be too small to see/read information in a helicopter. Since, the device is relatively small, consequently information (font size) will be small as well. Small screens have been shown to increase information retrieval time and workload significantly [Hamblin C 2003].

A 10-inch tablet would be good for information retrieval however some pilots pointed out that this device might be too large and heavy for use in a cockpit, especially when pilots would use it on their knee. Pilots predicted that the optimal screen size will be between 8 and 10-inch.

Q5: How are you using the EFB currently?

Pilot 6: "We are flying like this (imitating the posture as shown on Figure 3.1), so the tablet should not be larger than my leg and I should have place on leg where I can put my arms".

There is no dedicated mounting device for EFBs on the flight deck to which pilots can attach the tablet. Pilots who use a device, strap their EFBs to their knee. Both pilots who already use a mobile EFB and pilots who said they would like to use one stated a common requirement. They expected that a portable EFB maximises screen area while minimising overall weight. It should also fit properly onto the knee, while there should be room on the thigh to rest the arms.

Figure 4.1 Cockpit view of AW139.

As shown on Figure 4.1 the captain (yellow helmet) holds the stick with his right hand while resting both arms on his thighs. The cyclic control stick is between the feet of the pilot. The tablet must not reduce the controllability of the cyclic.

Another observation which was made and stated by pilots was that pilots interacting with the aircraft system (e.g. Flight Management System (FMS)) rest (or stabilise) their hands while inputting data. This can be also seen on *Figure 4.1*; the co-pilot is interacting with FMS. To minimize the effect of vibration and turbulence, pilots may hold/stabilise the EFB with their hand and operate it with their thumb.

> *Q6: "What problems are you facing with EFBs and how can be these addressed?"*
>
> *Pilot 7: "If I use my tablet a lot on my knee it heats up and I start to sweat on my knee. If want to remove my kneeboard it. It would be better if I have magnetic attachment so I can take it off more easily"*

Pilots who use a tablet during the operation mentioned that heat generated by the tablet causes discomfort. Mobile EFBs are mostly attached to the kneeboard. Generated heat by the device could has a negative impact on comfort [Chandra et al. 2003].

> *Pilot 8: "…it is hard to read the tablet if the sun lights hits the screen."*

Another common mention was that the angle of tablets strapped directly to the leg is not ideal, and that sun light can produce glare. They recommended the design of a kneeboard that pilots are able to tilt up the tablet, while preventing heat transformation. Some pilots requested that the tablet should be easily removable if the device is not used or if the pilot wants to show something on the EFB to his co-pilot. The captain is likely to strap the EFB to his left knee, because he is the flying pilot and he keeps his right hand on the cyclic stick. So, if parallel usage is required pilots are likely to strap it to their left knee. The co-pilot has a little bit more freedom because he is not interacting with aircraft controls as much as the flying pilot. It was predicted by avionics experts that pilots would strap the EFB to the left knee, since the left hand would be used infrequently. However, considering that approximately 10% of the population is left-handed [Hardyck and Petrinovich 1977] there will be pilots who will prefer the right knee, to facilitate usage with their preferred hand.

> *Pilot 6: "The EFB (Application) should be easy to use. For instance, if I want to perform a checklist or want to look something on the map it should be available after a few clicks"*

All pilots expressed the desire for an easy to use and intuitive interface design. The EFB must not distract pilots. Colours and animations should be thoroughly investigated. The number of buttons on display area should be minimised to avoid clutter. Navigation through the app should be intuitive and the number of control inputs required to get to the required command should be minimised.

> *Pilot 2: "… do not forget to use big letters. We had this problem previously with the checklists. Later we created our own checklist with larger letters."*

The font size and the size of interactive elements should be appropriately large because vibrations in a helicopter could be higher compared to a fixed wing aircraft. Another pilot stated that they created the checklist using 14 pt font because they could not read the checklist in high turbulent environments. This is substantially larger than the recommended font size, which is about 8 pt [Tinker 1963]. In high vibration and turbulence phases pilots face difficulties in retrieving data from head down displays.

This section will be completed with a brief description of EFB regulation. The FAA categorised EFBs (Hardware) in three different groups [Federal Aviation Administration 2012]:

- An EFB Class 1 is a portable device that is not attached to any aircraft-mounted device. Any data connectivity to the aircraft system is forbidden, and it is not a part of the aircraft configuration. Therefore, a Class 1 device does not require airworthiness approval.
- EFB Class 2 is also portable. However, it requires a dedicated mounting device. This kind of equipment may have limited data connectivity. Airworthiness approval is needed for some physical aspects (e.g. mounting, connections and antennae).
- EFB Class 3 is fully integrated (fixed) into the aircraft flight compartments and systems. It requires an airworthiness approval via a type certification.

Applications (or software) that run on EFBs are defined by their functionality. The three levels of functionality are summarised below:

- Type A software are static applications such as document viewer for aeronautical data (maps, charts, manuals, checklists and NOTAM)

- Type B software include dynamic interactive applications which, could perform various calculations and are able to zoom, pan, and scroll approach charts (to display own-ship position requires further approvals). It has the permission to receive (or update) weather information. An authorised person should validate such applications.
- Type C software can display own-ship position on charts. This kind of application must run on EFB Class 3, therefore a type certification via airworthiness approval is required.

Most airlines prefer class 1 or 2 devices because they are cheaper and easier to deploy. American Airlines (AA) was the first major commercial air carrier that integrated mobile EFBs. The software [Pschierer et al. 2012], used by AA, has the following features: Enroute charts and airport diagrams (displays own-ship position), arrival, departure and approach procedures and change notifications (terminal and enroute).

4.3.4 Preferred features and functionality from an EFB

The last questions were about features and functionality pilots would prefer in an EFB. Some available tablet applications were demonstrated to pilots. We asked pilots to list features and functionality they would like to have on an EFB. The most wanted features were i) performing checklist, ii) weight and balance calculations, iii) download mission related information, iv) upload the flight plan to aircraft system, v) searching approach plates, and vi) to use the tablet to fill the paperwork after the mission.

The last part of the interview was separated into three sections; pre-flight, in-flight and post-flight. It was requested to describe the pre-flight tasks they have to complete on a daily basis, then, to list the tasks that can be done via a mobile device. This part of the interview was mostly a conversation between pilots where they discussed the features and functionalities they would like to see on an EFB. The investigator asked additional questions to clarify their thoughts. This was repeated for in-flight and post-flight tasks. The outcome of these interviews was used to create a scenario describing the daily routine of a pilot who use a mobile EFB. This scenario is presented in Chapter 6.1.

4.4 Parts of the Framework

Based on the interviews, we categorised the emergent variables into four groups; **environmental, physical, virtual** and **user**. As stated at the beginning of the chapter; usability is the core psychological and physiological construct in this thesis. Based on ISO DIS 9241-11 [2015] there are three separate aspects of usability; effectiveness, efficiency and satisfaction. Jordan [1998] described effectiveness as the extent to which a goal is achieved, efficiency as the amount of effort required to accomplish a goal and satisfaction as the level of comfort and acceptability that users feel when using a product.

During the interviews, avionic experts used the terms interaction speed, task completion, accuracy and fatigue. It can be seen that task completion corresponds to effectiveness, efficiency to interaction speed and accuracy and fatigue to satisfaction. Avionics experts were largely concerned about which variables (environmental, physical and virtual) could affect the usability (user - speed, accuracy and fatigue). In the following section user factors, will be listed and defined (all general definitions at the beginning of the description are from Oxford Dictionaries):

4.4.1 User Factors

- **Speed** – "*The rate at which someone or something moves or operates or is able to move or operate"*. This term was used in this thesis as the movement time between two targets (button) in Fitts' Law Experiments, completion time of frequency manipulation task and recognition speed of icons/symbols.
- **Accuracy** – *"The degree to which the result of a measurement, calculation, or specification conforms to the correct value or a standard".* This term was used to reveal error rates for particular target size and specific positions in Fitts' Law Experiments and the number of errors during the frequency manipulation task.

- **Fatigue** – "*A reduction in the efficiency of a muscle or organ after prolonged activity"*. Unstructured interviews and questionnaires (mainly conducted after experiments) were used to rate general (e.g. effort and comfort) and fatigue (e.g. wrist, arm and shoulder) symptoms.

During the interviews experts mentioned additional factors that can be assigned to user factors, these are:

- **Hold Strategy** – Hold – "Grasp, carry, or support with one's arms or hands" Strategy – "A plan of action designed to achieve a long-term or overall aim". In this project, this term was used to describe strategy how participants hold/support the touch screen device in mobile as well as fixed placement.
- **Handedness** – "*The tendency to use either the right or the left hand more naturally than the other."* This was a variable in the lab study that evaluated the impact of display position on usability. The effect of handedness on speed and accuracy was evaluated.

Empirical and qualitative findings revealed further user factors that can affect usability, these are:

- **Experience** – "The knowledge or skill acquired by a period of practical experience of something, especially that gained in a particular profession". In this project, this term described the impact of familiarity of touch screen usage and icons on interaction/recognition speed."
- **Vision** – *"The faculty or state of being able to see"*. This term was used in two different meanings. First, whether the selected font size has an impact on readability. Second, whether touch screen usage can cause occlusion on the display.
- **Finger** – "*Each of the four slender jointed parts attached to either hand (or five, if the thumb is included)"*. Touch screen operations are conducted usually with the thumb or the index finger. This variable showed what variable caused participants to use which finger.

4.4.2 Environmental Factors

The first and one of the most mentioned variable that could affect touch screen usability is in-flight vibrations. As stated before there are many factors like weather and domain (type of aircraft and operation) that can determine the total amount of vibration experienced by pilots on the flight deck. Increased G-force (+Gz) is another factor which came not initially during the interviews. Interviews with pilots which was conducted at a later stage of the project, revealed that this phenomenon is an additional environmental factor that needs to be investigated.

- **In-flight vibration** – *in-flight* – "*Occurring or provided during an aircraft flight" vibration – "An instance of vibrating"*. In this project this term describes the total vibration that was measured during the flight at various phases.
- **Domain** – *"A specified sphere of activity or knowledge"*. This term was used to describe the impact of type of aircraft and operation on touch screen usability.
- **G-Force (+Gz)** – "A form of acceleration that causes the accelerating object to experience a force acting in the opposite direction to the acceleration". One of the aim of the project was to understand whether +Gz, occurring during steep turns, has a significant impact on usability.

4.4.3 Virtual Factors

A significant part of the interviews was focused on the interface design. Another frequent stated variable was the target size. However, experts pointed that addressing this issue will not sort the entire problem. There pointed to other factors like interface layout, font size, icons and interaction strategy. An additional factor which was not mentioned initially was the impact of target location. The last variable of virtual factors is the content, features and functionality of interfaces requested by pilots.

- **Target Size** – Target – “An objective or result towards which efforts are directed” Size – “The relative extent of something; a thing's overall dimensions or magnitude; how big something is”. In Computer Science, this term is the size of interactive elements (button size) on the interface.
- **Target Location** – *Location – “A particular place or position”.* This is the particular position of buttons on the interface.
- **Layout** – “*The way in which the parts of something are arranged or laid out”.* This the arrangement of text, icons, button and other information on the interface.
- **Content** – *“The things that are held or included in something”.* Features, content and functionality that pilots would like have in an aircraft system.
- **Icons** – “A symbol or graphic representation on a screen of a program, option, or window”. Symbols which were used on the touch interface.
- **Font** – *“A set of type of one particular face and size”.* In this context, this is the size of fonts on the interface.
- **Interaction Strategy** – Interaction – “action or influence” Strategy – “A plan of action designed to achieve a long-term or overall aim”, The way how users will interact with the interface.

4.4.4 Physical Factors

The last most frequent stated variable that could have a significant impact on touch screen usability was the display position on the flight deck. There are two types of displays envisioned in future flight deck concepts; mobile and fixed. The position of the display in mobile placement is similar (within the zone of convenient reach) for all users. However, there are various opportunities on the flight deck to install a touch screen display. The effect of used touch screen technology was also mentioned by experts. Physical variables which has also a significant effect to another variable, but were not stated during the interviews are; the size and shape of the display.

- **Placement** – *“The action of placing someone or something somewhere”*. In this thesis, this term described whether a touch screen is fixed or mobile.
- **Position** – *“A place where someone or something is located or has been put.”* This describe the position of fixed displays on the flight deck.
- **Shape –** “*The external form, contours, or outline of someone or something.”* This is the shape of the touch screen for both placements.
- **Size –** “The relative extent of something; a thing's overall dimensions or magnitude; how big something is.” This is the size of mobile and fixed touch screens.
- **Technology –** *“Machinery and devices developed from scientific knowledge.”* This is the touch screen technology (capacitive or resistive) used in the study.

5 Experimental Research

The experimental work presented here examines fundamental design choices for touch screens with the goal to provide guidelines that enable the design of touch screens that are effective while minimising errors, in order to be ultimately usable by pilots. The contribution of this work are recommendations and design guidelines for touch screens on the flight deck, derived from extensive trials in the field and in the lab. This chapter presents three novel studies: the first in-flight study in which touch screens are evaluated under real conditions, the first experiment that investigated the impact of various display positions on performance following Fitts' Law experiment (ISO 9241-9) and the first study that simulated +Gz using a weight adjustable wristband.

We had the opportunity to conduct experiments in Search and Rescue helicopters in Spain. Conversations with avionics experts revealed that minimizing error rates has a higher priority than fast interaction with aircraft system. Due to time limitations, it was decided to reduce the levels of display placement and increase the levels in target size for the field trials and conduct a separate lab experiment in order to investigate the potential impact of display position on usability. Increased G-force (+Gz) which is another environmental factor could not be investigated during the in-flight experiments. A further lab study was conducted to understand the impact of +Gz on touch screen usability. Following sections will provide a brief description of the studies.

The in-flight experiment investigates the impact of vibration (cruise, transition and hover), device placement (mobile and fixed) and target size (5, 10, 15 and 20 mm) on touch screen usability with Search and Rescue (SAR) crew members in an operational setting in helicopters. The purpose of this research is to establish design guidelines and recommendations for fixed and mobile touchscreens on a helicopter flight deck. Key hypotheses driving this work are:

Hypothesis: *Vibration, placement and target size have a significant negative effect on error rates and performance.*

> *Increasing target size will minimize the negative effects of vibration and placement.*
>
> ***Hypothesis:*** *Participants make fewer errors when the device placement is mobile compared to when it is fixed.*

The second study evaluates the potential impact of display position within a simulated cockpit in a laboratory study. The impact of angular displacement (45° between each 5-discrete position), vertical displacement (near and far) and horizontal displacement (low and high) on throughput, error rate and movement times was investigated. Hypothesis in this work are:

> ***Hypothesis****: The* position of the display has a significant effect on touch screen usability.
>
> ***Hypothesis:*** *Handedness has a significant effect on error rates and performance.*

The last study investigates the potential impact of +Gz on touch screen usability. Pilots flying a fast jet aircraft are frequently exposed to alternating G-forces. A Fitts’ law experiment was conducted to understand the effect of +Gz on touch screen usability. The key hypotheses driving this work are: Increased

> ***Hypothesis:*** *+Gz will have a negative impact on interaction speed and accuracy.*
>
> ***Hypothesis:*** *Participants subjective ratings for their fatigue indices will be affected by increased +Gz..*

Sub-research questions (1-13) stated previously in Chapter 1.3, will be addressed at the end of each study. Questions 1-6 will be addressed with the field trials, questions 7-11 with the lab study investigating the impact of various display positions and last 3 questions with the study aiming to understand the effect of +Gz on touch screen usability.

5.1 Fitts' Law Experiment

The task design is similar in all experimental studies. Before, starting with the method for the field trials a general description of applied task design will be given. Rapid aimed movement tasks modelled after Fitts' Law [1954] (cited nearly 6000 times) is known as a good model to predict pointing performance for various input devices under various conditions. ISO 9241-9 [2007] suggested a two-dimensional tapping task where targets are arranged around a circle (Figure 4.1).

The order of targets is predefined and the sequence finish once the participant tapped all targets. Then the Throughput, which is the index of performance, can be calculated by taking the quotient of Index of Difficulty (ID) and Movement Time. (*Equation 1*)

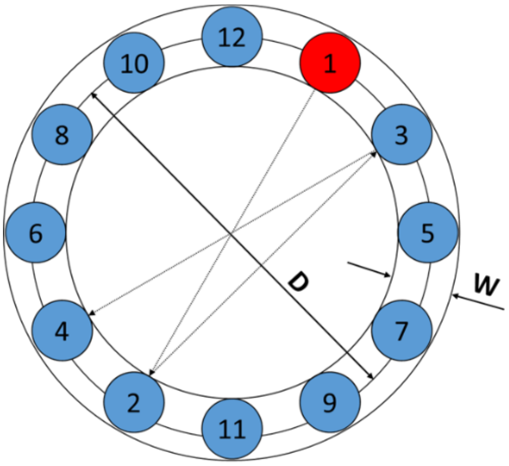


Figure 5.1 ISO 9241-9 Multi Directional Tapping Task.

$$TP = \frac{ID_e}{MT}$$

Equation 1

The Shannon formulation of the index of difficulty (in bits) is calculated by using distance between two targets (D) and the target size (W). Movement Time (Movement Time) is the mean movement time (seconds) between targets during a sequence. (*Equation 2*)

$$ID_e = log_2(\frac{D_e}{W_e} + 1)$$

Equation 2

The subscript e, which is available at ID, D and W is indicating the adjustments for accuracy which is proposed by Grossmann [1960]. W_e is calculated as 4.133 x SD_x, where SD_x is the standard deviation in the selection coordinates and D_e is the mean of the actual movements distances in the sequence of trials. Fitts' Law prediction model can be created by using a series of data generated over a wide range of ID. *Equation 3* shows the required (predicted) movement time to reach a target of size (W) over a distance (D). The two constants a and b are found using regression analyses. impact

$$MT = a + b \times log_2\left(\frac{D_e}{W_e} + 1\right) = a + b \times ID_e$$

Equation 3

5.2 Field Trials; In-Flight Experiment (IFE)

The first part of the research was carried out in a Search and Rescue (SAR) setting. Our site of study was the Spanish Maritime Safety Agency, also known as SASEMAR, between April and May 2015. SASEMAR has eight identical Agusta Westland AW139 Helicopters (Figure 5.2) distributed along the Spanish coast. Data was collected during 12 training flights in four different bases (Reus, Valencia, Almeria and Jerez). The crew conducted the experiments at their own discretion, in periods of downtime from their primary duties.

Figure 5.2 SASEMAR AW139.

5.2.1 IFE - Method

A mixed methods approach was adopted where a series of experiments (described below) were undertaken in a lab setting prior to moving to more open-ended field trials in a real-world setting. Initial experimental results showed significant differences in targeting accuracy and movement time for using touch screens in a static environment compared to a dynamic (vibrating) environment. This motivated the transfer of experiments into a real-world setting to achieve ecologically valid results.

5.2.2 IFE - Participants

The target population are pilots. However, for safety reasons pilots could not directly participate in field trials. Participants were hoist operators and rescue swimmers on board of the helicopter. 14 male crew members conducted the experiment (there were no women on duty at the time of the trials). Their age ranged from 27 to 52 years old (M=35.6, SD=11.8). Two of the participants were left-handed. The number of years on duty ranged from 3 to 25 years (M=9.6, SD=8.6). 13 Participants used a touch-enabled device (smartphone or tablet) and rated their touch screen skills on a 10-point scale (10 means very good) (M=7.9, SD=0.9). (Participant information sheet - Appendix I & Appendix II)

5.2.3 IFE - Apparatus

In the study (Chapter 6.1) aimed at learning about the features, content and functionality that pilots would like to see in an electronic flight bag (EFB), we asked what kind of tablet device they would prefer to use within the cockpit. Qualitative and empirical results suggested that an 8-inch tablet would be sufficiently large to display flight related information. Three pilots already used an iPad Mini as an EFB. Thus, an Apple iPad Mini (7.9” capacitive touch screen) was used for the entire experiment.

During the flight, vibrations were recorded with a Samsung Galaxy S4 (GT-I9505). The on-board accelerometer sensor is a K330 3-axis from STMicroelectronics. The resolution is 0.001m/s^2 and the range is 19.613m/s^2. Minimum delay is 0.01 seconds. Experiments were performed with two different device placements (mobile and fixed). In the mobile condition, participants hold the device while performing the experiment.

Figure 5.3 Experimental Setting (fixed placement).

In the fixed condition (Figure 5.3), the tablet is attached to a suction cup holder mounted on the window. The distance from the screen to seating position is 65 cm, which is approximately the same distance as that between pilots and the main instrument panel. Some double-sided tape was affixed to the window in order to stabilize the tablet in its position and to absorb its vibrations.

5.2.4 IFE - Experimental Design

A 2x3x4 within-subjects design with repeated measures was used for the experiment. Independent variables in this experiment were placement (2 levels - fixed and mobile), in-flight vibration (3 levels – cruise, transition and hover) and target size (4 levels – 5 mm, 10 mm, 15 mm and 20 mm). The minimum target size (5 mm) was determined using Google's Design Guidelines [2014]. The largest target size (20 mm) was adopted from previous work, in which authors achieved almost 100% accuracy. The target was displayed randomly, and the position and size of the target was recorded. Recorded dependent variables were movement time, touch position, distance and error rate. There was no minimum quantity of data that participants had to generate during a flight.

5.2.4.1 Vibration Measurement

An application called "Physics Toolbox Accelerometer" [Vieyra and Vieyra 2015] was used to record vibrations within the aircraft. Measurements were taken in three different locations. The first measurements were collected at the point where the experiment was conducted with fixed device placement. These measurements were compared with another measurement on the dashboard (Figure 5.4). The smartphone was attached between the Multi-Function Display and Central Display Unit. When the placement was mobile, participants held the device in their hand with the aim to see whether and how much the human body is able to compensate vibrations. 50 measurements were recorded per second.

Figure 5.4 In-situ Vibration Measurement.

5.2.4.2 Flight Recording

Another research objective was to understand how pilots interact with the aircraft system; thus, video recordings were made. The camera was positioned at an angle from which it was able to capture the pedestal, dashboard and the outside view from the pilot’s side (Figure 5.5). These recordings were used to verify in which flight mode (cruise, transition, or hover) the aircraft was in while participants commenced the tapping task.

Figure 5.5 Flight Recording.

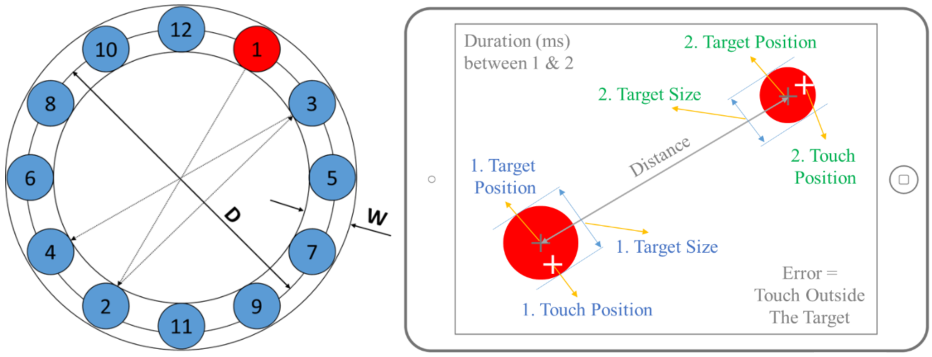
5.2.4.3 IFE - Task Design

The ISO 9241-9 [2007] recommended task design for input devices evaluation is illustrated in Figure 5.6a. In this multi-directional tapping task targets are arranged around a circle. The task is to tap all targets in a consecutive order. Taps outside of the circle are recorded as an error.

The distance (D) between targets and the width (W) (the actual size of targets) changes after the sequence is completed.

This task design was tried out in the lab. Initial results showed that participants tended to hover their finger over the next target before clicking the current target with the other hand. This kind of predictability would lead to contrived movement time measurements compared to realistic operational use.

However, the potential solution of restricting participants to use only one hand would have conflicted with the goal of seeing how participants use the device in a real-world situation. As it was not intended to compare results with prior work that applied the ISO task design, it was decided to modify the task design by creating a task in which the size and the distance of each target varied dynamically from the previous one.



a) b)

Figure 5.6 ISO 9241-9 Task and Tapping Task and Recorded Variables

A tapping task (first contact touch strategy) was created using JavaScript (Figure 5.6b). The task was to tap targets (displayed as red circles) sequentially. Data recording occurs as follows: the first target is displayed and the user taps the target. The position of the target and the actual touch position are recorded. The current target disappears and the next target is displayed, the user taps the next target. Again, the actual target and touch position are recorded. Using time stamps the duration between subsequent targets (movement time in milliseconds) is calculated and stored. In addition, the distance between subsequent targets is recorded. Touching outside the target is recorded as an error. The target remains visible until the user touches the target. The number of errors per task are recorded. The mean errors are calculated by dividing the number of errors by the number of tasks. Since, this task design differs significantly from the two-dimensional task design as proposed in ISO [2007] the effective values for width and distance and consequently the index of difficulty cannot be calculated. Instead, alternative analyses will be performed by using the actual width and distance values.

5.2.5 IFE - Procedure

The aims and objectives were explained to participants. Each participant was notified that the aim was to investigate the impact of in-flight vibration and turbulence to targeting accuracy and movement time on touch-enabled devices. Participants were asked to be as accurate as possible, while performing the task at a normal pace.

The experiment started with a baseline determination, replicating previous work e.g.[MacKenzie 2015]. Participants conducted some trials in both placements on the ground. Figure 5.7 illustrates the default positions of each crew member during take-off. The investigator sat on the seat from which the experiment would be conducted in the fixed placement condition.

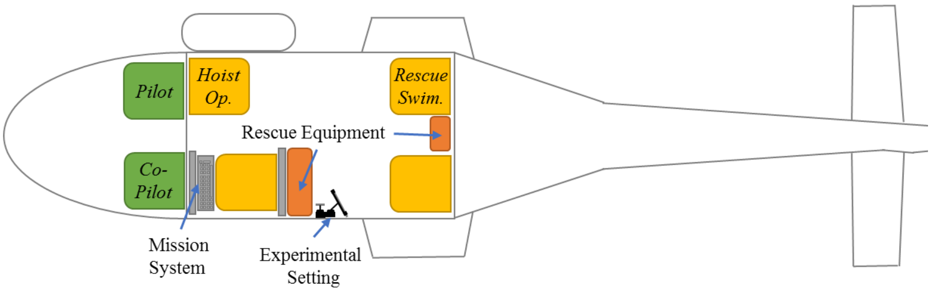


Figure 5.7 Aircraft Layout illustrating the Experimental Setup.

In the following sections, possible time frames are described, in which crew members were able to perform the experiment. To avoid fatigue effects, the investigator asked participants to stop after 5 minutes. Participants took their gloves off during the experiment. Some hoist operators had gloves without index finger; thus, they were able to conduct the experiments while wearing gloves.

Before take-off, the screen of the tablet was cleaned. The experiment started in the mobile placement condition. After take-off, the rescue swimmer started with the tapping task. After approximately 5 minutes, the rescue swimmer handed over the tablet to the hoist operator and he continued the experiment. The pilot notified the persons in the rear cabin approximately 10 minutes before reaching the target. The rescue swimmer started with preparations. The investigator gave the hoist operator a signal when the transition to hover was attempted (around 80 knots).

Once the aircraft was in hover, pilots required on average 3 minutes to position the aircraft close to the target. The hoist operator handed over the tablet to the rescue swimmer. The rescue swimmer continued with the experiments. The hoist operator opened the door and spoke with the pilot to make fine adjustments for the position of the aircraft. It was also possible for the hoist operator to take full control over the aircraft and position the aircraft by using his controller. At this stage, the experiment was done in the mobile condition for all flight modes (cruise, transition and hover).

After the first training was completed and the door was closed, the investigator attached the tablet device to the fixture. From that point, the experiments were conducted in fixed placement conditions. Pilots are strapped to the seat all the time; however, hoist operators and rescue swimmer are connected with a wire to the aircraft, thus they can move freely in the cabin. Participants were asked not to fasten seatbelts to save time and not to lean towards the display.

The helicopter flew away from the target and circled. The investigator swapped his seat with the hoist operator. Once the helicopter approached the target (when transitioning occurred), the hoist operator started with the taps. The hoist operator finished the task once the helicopter was ready for opening doors. He swapped his seat with the rescue swimmer who continued with the task. The rescue swimmer stopped once his duty started.

Once the second training was completed, the hoist operator closed the door and the helicopter took off and turned for the third scenario if there was one, otherwise, the crew returned to base. During this transit flight, the crew performed the experiment again. Approximately 10 minutes before landing, the investigator gave the hoist operator a signal to start the experiments; after 5 minutes, he swapped with the rescue swimmer who performed the experiments until landing.

Data was recorded in nine flights as mentioned above. At this point, it was noticed that more data had been collected in the mobile condition than with the fixed placement. Thus, during the last three flights the experiment was conducted mainly in the fixed placement.

5.2.6 IFE - Results

First, vibration analyses will be performed. The results will reveal that all flight modes (cruise, hover and transition) have different characters. After that it will be described how raw data was treated and sorted into subgroups (determined by the level of placement, in-flight vibration and target size). Furthermore, analyses of the distribution characteristics of subgroups will be presented. The main part of the results is throughput, error rate and movement time analyses which will be presented in the same order.

5.2.6.1 Vibration Analyses

The application recorded the acceleration in x, y, and z directions with a timestamp. The magnitude of the vibration was calculated by using *Equation 4*.

$$M = \sqrt{x^2 + y^2 + z^2}$$

Equation 4

At least 15 measurements are recorded per second. The flight protocol and recordings were used to determine the timeframes for specific flight modes. The data was annotated with a key value describing the flight mode. The key value is the same as described in the next section.

Timelines are added to visualize flight modes. (Note: transition phases are the timeframes between cruise and hover).

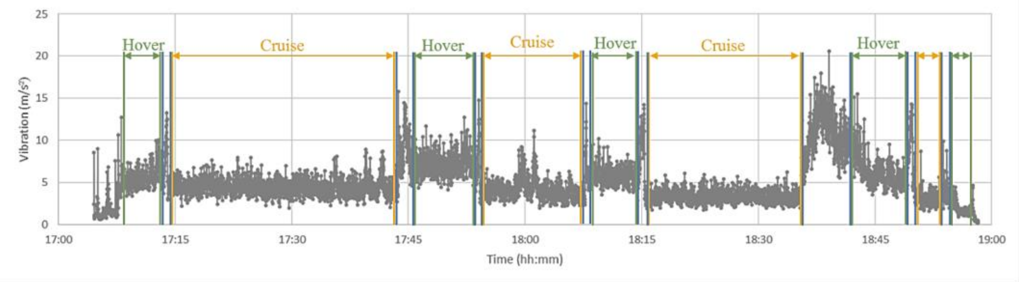


Figure 5.8 Vibration Measurement in Fix Position.

Figure *5.8* shows vibrations recorded during a flight in Valencia. The smartphone was attached to another suction cup holder, which is mounted behind the fixed device placement (see Figure 5.7). For this particular flight, the mean vibration for cruise was around 5 m/s^2, for transition 12 m/s^2 and for hover 7 m/s^2.

However, this does not mean that vibrations always lead to the same values. The airspeed is a significant factor during cruise that can cause high vibrations. During this flight, the cruise speed was always below 120 knots. During a different flight in Reus, the cruise speed was sometimes over 130 knots and the smartphone measured a mean vibration of 6 m/s^2.

Depending on the weather and location, vibrations during hover could be as small as 4 m/s^2. The magnitude of vibrations during transition phases depend on how fast the pilot transitions through the critical speed where the vibrations are highest. Thus, the measurements reflect when the pilot decreased speed during a transition down phase more slowly. In this transition phase, vibrations of more than 15 m/s^2 were measured.

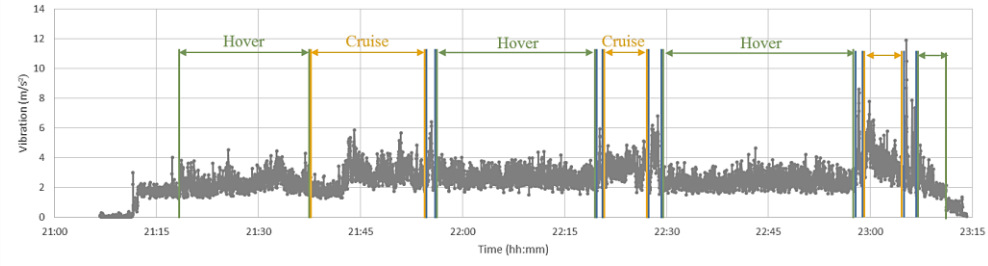


Figure 5.9 Vibration Measurement on the Dashboard.

The data shown in Figure *5.9* was recorded on the main instrument panel during a night flight in Almeria. Vibrations for cruise were around 3 m/s^2, hover were 2.5 m/s^2 and transitions were 5 m/s^2. The second recording in this setting had similar values.

The lastFigure 5.10 is a collection of different vibration measurements, which were taken on the hand of participants, to see whether the human body is able to compensate vibrations. Results show that the majority of measurement for cruise and hover were below 2 m/s^2 where the average was around 1.5 m/s^2. During transition phases, vibrations increased to 3 m/s^2. There are fluctuations in the measurement, which are likely caused by hand movements.

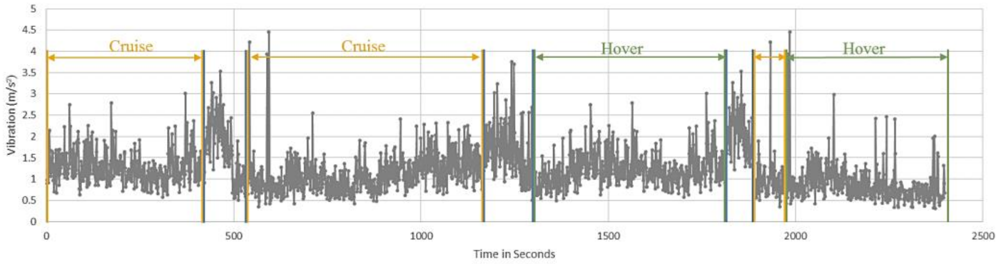


Figure 5.10 Mobile Vibration Measurement.

All measurements were imported to IBM SPSS to test the groups for statistical significance. ANOVA revealed for all cases that the levels of vibration (cruise, hover and transition) are significantly different from each other. The highest vibrations were achieved during transitions phases. The vibrations during hover were in average slightly but significantly higher than vibrations during the cruise. An ANOVA for mobile measurement was not performed because of few and intermittent measurements.

It was expected that vibrations measured in the fixed condition would be more intense than those on the main instrument panel, which is installed on a system, which absorbs a certain amount of vibrations. By contrast, in the fixed placement condition the smartphone and tablet were attached to the window via a suction cup fixture, which transferred the entire airframe vibration to the devices without absorption.

Interviews with pilots showed that there are times, especially during winter months, in which they have to operate in challenging weather conditions. In these times, pilots are exposed to higher vibrations and turbulences. Thus, experiments conducted with higher vibrations resulting from the fixed placement may be considered to emulate a certain amount of realism.

The analysis of vibration measurements gathered in the mobile condition showed that the human body is able to absorb a certain amount of vibration. The peak value was measured as expected during transition phases. In other flight modes, which cover the majority of the flight, vibrations did not increase beyond 3 m/s^2.

Observations showed that pilots performed more ‘manual’ actions during hover compared to cruise. During hover, the wind is pushing the aircraft away from its position and the pilot has to steer manually to keep the aircraft at the desired position. This causes additional unexpected movements in the aircraft. Another factor, which could impede the accuracy, is the downwash wind that blows into the door during hover.

5.2.6.2 IFE - Data Pre-Processing and Manipulation Checks

17,346 data points (14,356 generated in the air) were imported from the app. Each task received a key value describing the placement, vibration and target size. The key value consists of four digits (see Figure 5.11). The first digit describes the placement (1-fixed, 2-mobile), the second digit describes the vibration (1-cruise, 2-transition, 3-hover) and the last two digits describe the target size. For example, 1115 means that the task was performed with a fixed placement, during cruise and the target size was 15 mm.

Data received their key value by using the flight protocol. These values were double-checked with vibration measurements and video recordings. Tables 1 and 2 present the mean and standard deviation on task error rate and Throughput in percent versus several different conditioning factors. A probability value (p) of 0.05 was chosen as a cut-off level for statistical significance.

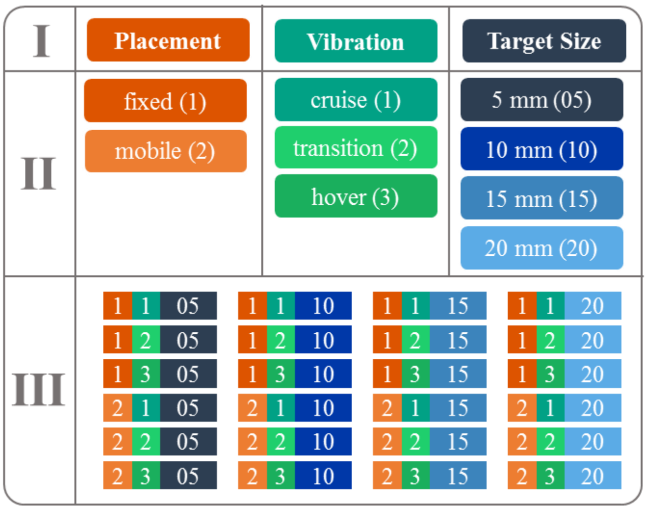


Figure 5.11 Independent Variables.
I-III correspond to different levels of analysis.

Analyses start at top level where all independent variables were considered separately. For throughput analyses, levels of placement and vibration were combined and examined for significant differences. For error rates analyses, target size levels were added and each condition was evaluated for significant differences. Targets appeared on an 8 x 10 array, which enabled the possibility to analyse the error rate by specific target locations.

Soukoreff and MacKenzie [2004] recommended range for Index of Difficulty (ID) is between 2 and 8. Due to small screen area ID values ranging from 1.2 to 6.2 were presented. Due to experimental design, ID values were not distributed evenly. Data was binned into subgroups by the level of placement and vibration. The mean value for all subgroups were calculated. ANOVA was applied (only in-flight data) to ensure that participants were assigned similar task difficulties in each condition. Results showed a mean ID value around 3.7 with a standard deviation of 1.0, which indicated that participants were exposed to the same level of difficulty in each condition ($F_{5,\ 14351} = 1.22,\ p=.293$). The same test was applied to the distance between two targets. The mean distance between two targets was 66 mm with a standard deviation of 32 mm. There was no significant difference for each subgroup ($F_{5,\ 14351} = 1.39,\ p=.223$).

Movement Time (MT) values from field trials were compared with values from the lab study. Skewness values were (100x) higher during field trials. That can be explained by the fact that conducting experiments during the flight had a secondary order for crew members. In addition, it was possible to observe and count the breaks that participants made in some cases. Participants took 2 or 3 breaks per 100 touches during the mobile placement condition. This value increased to 4-5 breaks during the fixed placement condition, which were mainly caused by fatigue. It was decided to use the first 95th percentile for each subgroup, 'cutting off' the long tail. As a result, the skewness for this modified data set as was 3 times higher than in the lab study. Keeping in mind that this task design required extra search time for the next target, this kind of skewness is acceptable. The difference between two ID values was as small as 0.01 and most tasks appeared around the mean ID value. For the Fitts' Law prediction model all ID values were binned into groups with a 0.1 increment and the average Movement Time (MT) was calculated.

The distribution characteristic for Throughput results (95th percentile) were assessed. The mean skewness of the distributions, for subgroups defined by level of placement and vibration, was 0.240. The mean kurtosis was 0.187. Both of these values are low, indicating no overall tendency towards a negative or positive skewness or towards a flat or peaked distribution. A Kolmogorov-Smirnov test with Lilliefors Significance Correction (half of conditions satisfied this criteria) and a visual inspection of their histograms, normal Q-Q plots and box plots showed that Throughput scores were approximately normally distributed.

Since the trials were integrated into the training flights of SAR units, the crew conducted the experiments at their own discretion, in periods of downtime from their primary duties. In this semi controlled experiment, it was not possible to assign the experiments to both participants evenly. Therefore, it is possible that one participant produced more data in a particular condition than his crew member. Thus, it was decided to use the average values (Throughput, Error Rates and Movement Time) per flight in the statistical analysis.

5.2.6.3 IFE - Throughput Results

Throughput is the index of performance, which is calculated by dividing the Index of Difficulty (ID) by Movement Time. Figure 5.12 shows the mean Throughput by placement and vibration with 95% confidence intervals (left), and a matrix that illustrates the significance for pairwise comparisons (right). Table 1 presents the mean, standard deviation and the result of statistical tests on the effects of the independent factors on throughput.

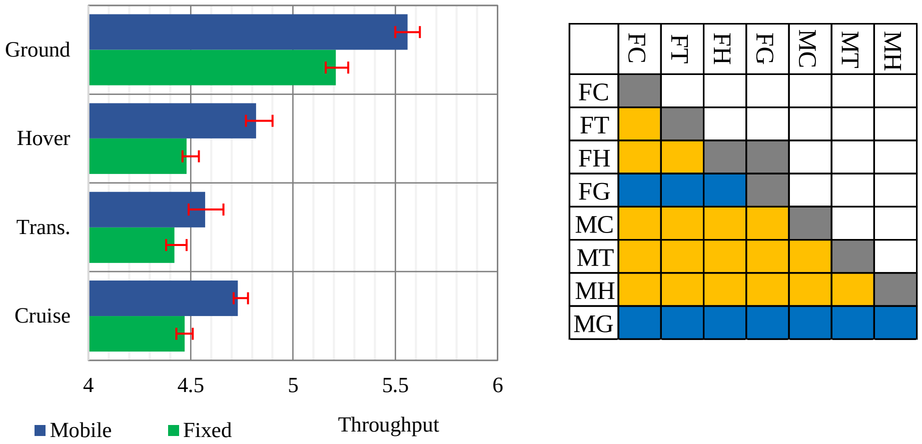


Figure 5.12 Throughput Results for Placement and Vibration (left) – error bars represent 95% confidence interval) and Pairwise Comparison Matrix (right)

Throughput for experiments conducted on the ground was significantly higher (large effect) than results generated in the air (during flight). By taking the literature into consideration this reduction in interaction speed was expected. Average Throughput values on the ground were approximately 18% higher than the average values generated in the air.

Comparing different levels of in-flight vibration did not show a significant difference. Ground and air data were grouped with placement levels (fixed and mobile). ANOVA and Bonferroni post-hoc tests showed that placement has a significant effect on the ground but not in the air. Vibration levels were combined with placement levels. ANOVA indicated a significant difference on combined variables. Bonferroni post hoc test revealed that pairwise condition where ground data were involved produced a significant different effect to air data.

Table 1 Statistical Analyses and Results for Throughput (Field Trials)

Description	Levels	M	SD	Result
Ground and Air	Air	4.55	0.25	F (1,11) =71.7,p<0.001, η_p^2 = 0.87
	Ground	5.36	0.11	
Vibration	Cruise	4.61	0.25	F (3, 9) = 29.69, p <.001, η_p^2 =0.91
	Transition	4.43	0.54	
	Hover	4.60	0.27	
	Ground	5.35	0.11	
Ground/Air & Placement Combination	Air & Fix.	4.48	0.36	F (3, 9) = 37.08, p <.001, η_p^2=0.93
	Air & Mob.	4.75	0.45	
	Gnd. & Fix.	5.17	0.10	
	Gnd. & Mob.	5.52	0.15	
Placement & Vibration Combination	Cruise & Fix.	4.52	0.33	F (5, 7) = 30.49 p <.001, η_p^2 = 0.98
	Trans. & Fix.	4.29	0.39	
	Hover & Fix.	4.47	0.13	
	Gnd. & Fix.	5.17	0.10	
	Cruise & Mob.	4.70	0.32	
	Trans. & Mob.	4.56	0.60	
	Hover & Mob.	4.82	0.46	
	Gnd. & Mob.	5.52	0.15	

5.2.6.4 IFE - Error Rate Results

Error rates are calculated by taking the quotient of number of error and number of trials. Targets appeared on a 8 by 10 array which enabled to investigate the error rates for specific areas (*Figure 5.13*).Figure 5.14 shows error rate development by changing target size for each condition. Figure 5.15 contains a matrix that illustrates whether pairwise comparisons yield significant differences for each condition. Table 2 presents the mean, standard deviation and the results of statistical tests on the effects of independent factors on error rates and results for significant interaction of the independent variables.

Error rate for experiments conducted on ground was smaller than results generated in the air. Average errors generated in the air were 2.8 times higher than the errors on the ground. There was a significant effect of vibration on error rates. Bonferroni post-hoc test compared effects pairwise and showed that errors generated during cruise are significantly lower than the errors generated during the transition phases. Errors generated during the hover did not revealed any significant difference to the other two flight modes. ANOVA detected a significant effect of target size on error rates. Bonferroni post-hoc test found a significant difference for pairwise combinations apart from the combination of target sizes 15 mm and 20 mm. Ground and air data were grouped with placement levels (fixed and mobile). ANOVA and Bonferroni post-hoc tests showed that changing placement on the ground does not have a significant impact on error rates. Other pairwise combinations are significantly different.

A univariate analysis of variance revealed significant interaction effects between placement and target size and also between vibration and target size. There was no significant interaction between placement and vibration. This suggests that the impact of placement and vibration depends on the size of targets.

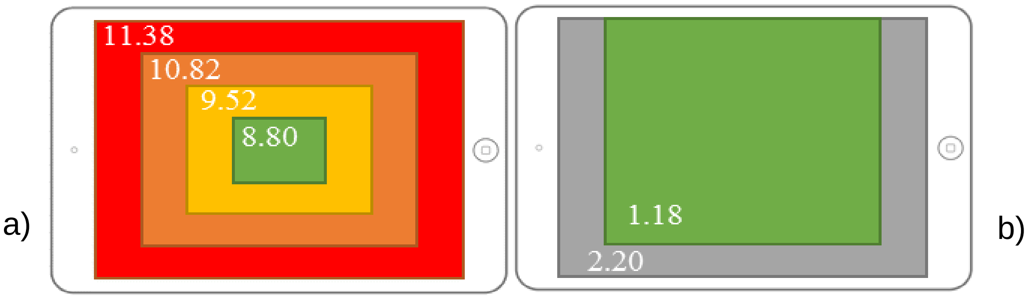


Figure 5.13 Error Rate areas in mobile placement (all Target Size) & Error Rate for input and output areas (15 & 20 mm)

Targets appeared randomly on an 8 by 10 grid. During fixed placement for all target sizes, the error rate was similar (13-14%) for all areas. For mobile placement, it was noted that participants made fewer errors on the centre of the screen and the error rate increased by moving towards the edge (Figure *5.13*a). Recommended target size for fixed and mobile placement is 20 mm and 15 mm, respectively. Usually interactive elements are placed alongside the edges (grey area) and the centre of the screen (green area) is reserved for displaying information (Figure *5.13*b). For fixed placement, the error rate was around 4% for both areas. For mobile placement, the areas where interactive elements are normally placed had higher error rate than the placement where information is displayed.

Error rates for each placement condition are plotted by target size on Figure 4.14;

- The largest difference in error rates occurred in the mobile condition for 5 mm targets. The difference between cruise and transition was 20% (for the fixed placement the difference is 19%). This margin decreases for all vibration levels with increasing target size.
- The largest difference for placement was also found at 5 mm target size. The difference for all vibration levels were around 12-13%. Like before, increasing the target size reduces the effect of the placement.

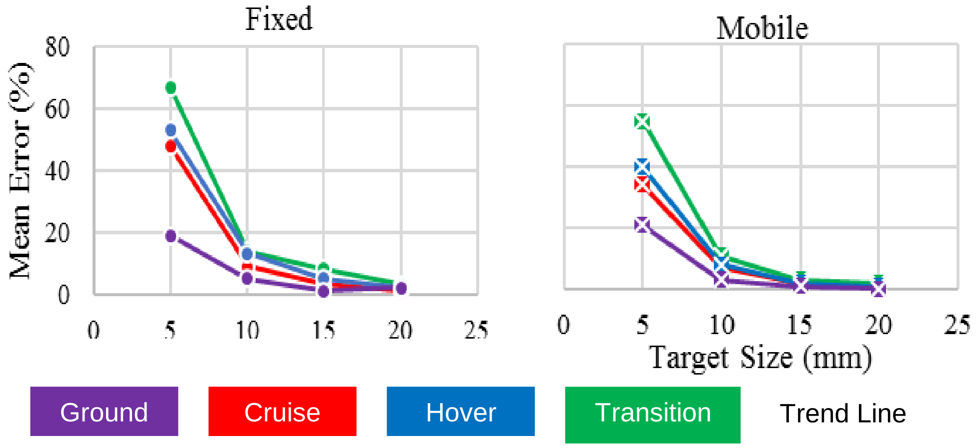


Figure 5.14 Errors by Target Size for Fixed and Mobile Placement.

Bonferroni post-hoc analyses compared all conditions pairwise for significant difference (Figure 5.15), main results are as follows: (The number in brackets are referred to the numbers on Figure 5.15)

- 5 mm target sizes were significantly different to all other target sizes. However, there were a few pairs, which were not significantly different (FH5/MT5, FH5/MH5 and MC5/MH5); amounting to 2% of the comparisons in which 5 mm targets were involved (green - 1);
- Comparing 10 mm targets with the same level and larger target sizes reveal more cases that are not significantly different. 24 % of the pairwise comparisons in which 10 mm targets were involved showed no significant difference (orange - 2);
- The first level of analysis with all factors considered independently showed no significant difference for 15 mm and 20 mm targets. Considering all conditions separately as shown in Figure 5.15 showed that the error rate for 15 mm targets during the transition phase with a fixed placement (FT15) differed significantly from 15 and 20 mm

targets during cruise for both conditions (FC15, FC20, MC15 and MC20). 58 % of the comparisons in which 15 mm targets were involved showed no significant difference (grey - 3);

- Comparing conditions that have 20 mm targets involved did not show any significant difference (violet - 4).

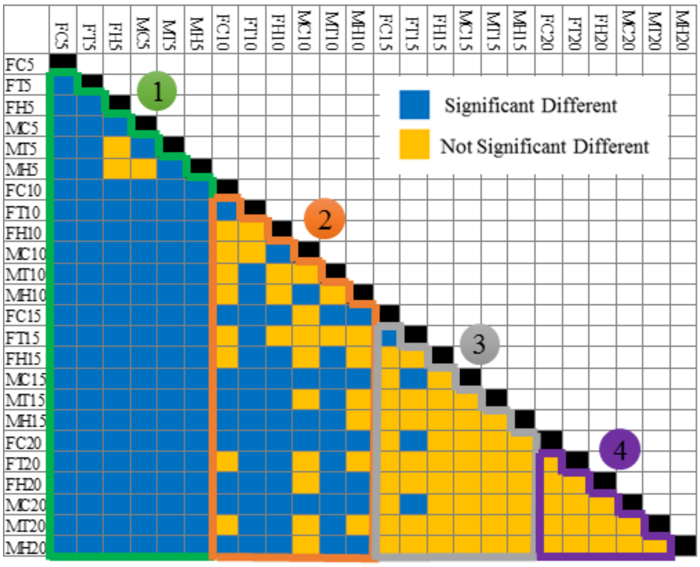


Figure 5.15 Matrix for Significance (Error Rates for Field Trials).

Table 2 Statistical Analyses and Results for Error Rates (Field Trials)

Description	Levels	M	SD	Results
Ground and Air	Air	19.00	4.99	F(1,11) = 61.9, p<0.001, $\eta_p^2 = 0.85$
	Ground	6.84	2.03	
Vibration	Cruise	14.58	2.87	F (3, 9) = 23.35, p <.001, $\eta_p^2 = 0.89$
	Transition	23.08	6.28	
	Hover	19.75	8.03	
	Ground	6.84	2.03	
Target Size	5	47.5	8.98	F (3, 9) = 104.8.1, p <.001, $\eta_p^2 = 0.97$
	10	9.75	3.93	
	15	3.25	1.55	
	20	1.17	1.19	
Placement & Vibration Combination	Air & Fix.	21.58	4.14	F (3, 9) =33.1, p <.001, $\eta_p^2 = 0.92$
	Air & Mob.	14.58	6.01	
	Gnd. & Fix.	7.08	2.27	
	Gnd. & Mob.	6.67	3.20	
Interaction between IV.	Target Size and Placement			F (3,9) = 6.35, p<.001
	Target Size and Vibration			F (6,9) = 22.9, p<.001
	Placement and Vibration			F (3,9) = 2.04, p=.106
All Conditions				F (19,17314) = 101.6, p<.001

5.2.6.5 IFE - Movement Time Results

Movement time (MT) is the required time to point the next target in sequence. Fitts' Law prediction models for both placements were created by binning ID values to subgroups with 0.1 increment and by plotting all data. Figure 5.16 shows the model, equations and R^2 value for all data and subgroups for fixed placement during hover. Figure 5.17 and Figure 5.18 include the models and equations for all data and subgroups for both placement and vibrations levels.

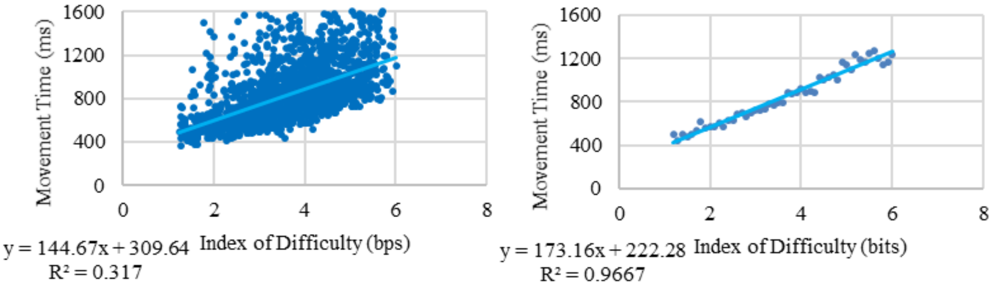


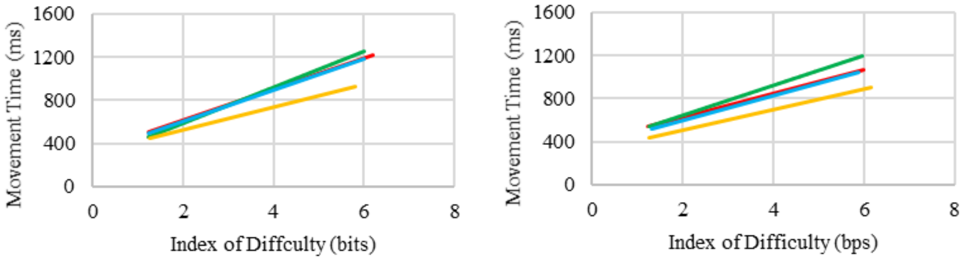
Figure *5.16 All Data and Subgroups for Fixed Placement during Hover (Graph, Equation and Regression)*

Due to "noise" generated from all data which are plotted separately the R^2 values are low. Models created from subgroups have high R^2 values which produced meaningful data. It was noticed that in all conditions, equations for mobile placement had a smaller slope compared to equations for fixed placement, which means that participants were able to point the same target (condition) in mobile placement faster than in fixed placement. Whereas, the off-set in the mobile condition was higher than in the fixed placement condition. This can be explained with occlusion problems which is likely to happen in mobile placement. More details are given in the following sections;

Figure 5.16a showed the linear regression trend line for fixed placement during hover, which is created from all single data points which was conducted during this condition. The longest five percent of movement time were removed with the aim to filter data that were generated after long breaks (for each subgroup defined by vibration and placement). 2060 data point were used to create the trend line. In that case as well in all cases the R^2 values were very low. The reason for this is each single data point for this particular condition was used to create the model. Normally in a standardized Fitts' law model you have a certain number of targets per sequence and the software will calculate the average of ID and movement time values. Another reason might be personal differences of users who created this data. Last but not, least it need to be mentioned that this was a secondary task and there was a divided attention present. Produced prediction models in this way were not interpretable.

Once data were binned into groups with 0.1 increment in ID values (Figure 5.16b). The regression value (R^2) was very high for all 8 models. R^2 value ranged from 0.87 to 0.97. This filtered all degrading factors. According to Soukoreff and MacKenzie [2004] the intercept (a) value should be smaller than 400 ms. The average intercept value for fixed placement was 214 ms and the average intercept for mobile placement was 250 ms. Occlusion could be potentially a contributing factor for this difference. Participants reported that sometimes their hand covered the next target. The majority of participants conducted the study with their preferred hands index finger. If next target appears below participants hand it is likely that the distance between the current and next target is relatively small, which will produce a low ID number. This will potentially increase the search time and consequently the movement time for small ID numbers. This is not the case in the fixed setting. The screen is far away from participants siting position and it is slightly shifted to the left.

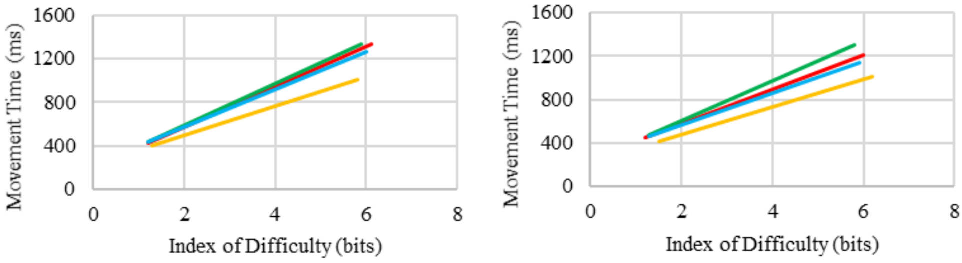
The average slope in fixed placement is 170 ms/bps and in mobile placement 154 ms/bps. The higher slope in fixed setting can be explained with increased fatigue symptoms that might rise during the experiment. Participants rest their arms on their legs in mobile placement. It is also visible that slope values in hover and cruise mode are similar. Previously, it was shown that mean vibration in both modes (depending on environmental factors) are similar. The highest slope value is generated during transition phases where vibrations were at least two times greater. Thus, participants required more time to touch a far and small target.



Vibration	Fixed	Mobile
Transition	$y = 168x + 252, R^2 = 0.30$	$y = 138x + 367, R^2 = 0.16$
Cruise	$y = 144x + 326, R^2 = 0.28$	$y = 110x + 403, R^2 = 0.18$
Hover	$y = 145x + 310, R^2 = 0.32$	$y = 114x + 370, R^2 = 0.22$
Ground	$y = 108x + 308, R^2 = 0.43$	$y = 95x + 320, R^2 = 0.38$

Figure 5.17 All Data for Both Placements

(Graph, Equation and Regression)



Vibration	Fixed	Mobile
Transition	$y = 190x + 211, R^2 = 0.93$	$y = 183x + 236, R^2 = 0.89$
Cruise	$y = 184x + 206, R^2 = 0.94$	$y = 159x + 259, R^2 = 0.95$
Hover	$y = 173x + 222, R^2 = 0.97$	$y = 147x + 276, R^2 = 0.94$
Ground	$y = 136x + 220, R^2 = 0.96$	$y = 126x + 226, R^2 = 0.87$

Figure 5.18 Subgroups for Both Placements

(Graph, Equation and Regression)

5.2.7 IFE - Summary & Research Questions

During the field study the potential impact of vibration, touch target size and placement was evaluated. All factors were found to have a significant impact on error rates. As shown in previous work the target size is the most significant factor, which may be utilized to minimize other degrading factors by selecting an appropriate target size. It was demonstrated that using touch-enabled devices that are fixed in place in vibrating environments produce significantly higher error rates than when the device can be held by the user. Furthermore, it was demonstrated that the location of the interactive element could influence the magnitude of error rates. Throughput values generated in the air were significantly different from ground data. However, level of vibration and placement in the air did not showed a significant difference.

It was demonstrated that binning index of difficulties and taking the average of each group would produce a strong R^2 value. Doing this alleviated individual difference as well as differences in task design. The two constants a and b derived from the regression analyses supported operational observations. The intercept values showed that designers should consider the effect of occlusion. The increased slope in fixed placement showed the effect of fatigue on interaction speed.

There are various opportunities to install touch screen displays in the cockpit. The next study will evaluate the potential impact of display placement more in depth. In this type of aircraft, it was not possible to test touchscreen usability under +Gz conditions. A lab study will try to understand the effect of this phenomenon on touchscreen usability.

The last section of this study will summarise the results and return the first six sub-research questions stated in Chapter 1.

Sub-RQ: *How should be the physical shape of the (fixed) displays, so it supports usability?*

In-flight observations showed that interactions in the fixed placement condition was performed with one hand. Participants always used their preferred hand. They were encouraged to take a break when feeling fatigue in their arms. Eight out of 14 participants were observed to tend to hold on to the device from the side or above. This observation suggests to design displays in such a way that it enables pilots to stabilize their hands from all directions (from behind included) and interactive elements should be placed along the sides.

Sub-RQ: *What is the preferred hold strategy in mobile placement?*

In the mobile placement condition, six participants initially used both of their hands to hold the device, and used their thumbs to tap the task.

Eight participants held the device with their non-dominant hand and performed the experiments with their preferred hand's index finger. In two cases, participants switched from two-handed thumb to one handed index finger grip. The observation suggests that the majority of users would use a mobile device in landscape mode.

It was observed that participants who used both hands had difficulties touching the target at the centre of the tablet. Post experiment interviews revealed that participants prefer to use the tablet device in the mobile condition. In contrast, the fixed placement was described as more fatiguing. In the context of a vibrating environment such as a helicopter cockpit, it is also worth pointing out that by holding the device, the human body is able to absorb vibrations, thereby mitigating for the detrimental effects of vibration on performance, error rates, and throughput.

Sub-RQ: *What is the impact of in-flight vibrations on usability*?

The main finding of this study was that in-flight vibrations have a significant impact on error rates, and that target size can be used to reduce this effect. Average Throughput values on the ground were approximately 18% higher than the average values generated in the air. Average errors generated in the air were 2.8 times higher than the errors on the ground.

The mean Throughput during the flight modes were similar. There was a small (not significant) reduction (3.5%) in Throughput during transition phases. The amount of transitions phase is around 5% of the entire training flight. Average user performance (Throughput) for touch screens during the flight is 4.6 bps.

Sub-RQ: *What are the effects of device placement on usability?*

The effects of holding a device in the hand were significantly different to attaching the device, on ground as well as in the air. Error rates under fixed placement condition were approximately 33% higher than in the mobile placement condition. The difference in Throughput was approximately 6% which was statistically not significant. Results confirmed the hypotheses that participant were likely to make more errors in the fixed condition than in the mobile condition.

Sub-RQ: *Which areas on the display have an increased error rate?*

Targets appeared on a 8 x 10 grid, which enabled further investigation on error rate for specific regions. In the mobile setting, participants had a higher accuracy on the centre of the screen. The error rate gets higher towards the edge of the screen. The error rate at corners for both placements were higher compared to the average error rate.

Sub-RQ: *What is an appropriate target size for touch screens?*

Independent variables were tested systematically, starting broadly at the top level and gradually going into more detail. In the first set of analysis, significant difference for all variables were found. While target sizes between 15 mm and 20 mm were not significantly different, detailed analyses showed that there are few cases where significant difference between 15 and 20 mm exist.

In the second level of analysis, interaction effects between independent variables were examined, which showed that two of three possible combinations have significant interaction effects. The final level of analysis considered each possible case (24) separately and in pairwise comparisons. The provided matrix shows that the effects of placement and vibration disappear with increasing target size. The results recommend to apply 20mm targets for fixed displays and safety critical tasks and to apply 15 mm target for mobile devices and non-safety critical tasks.

5.3 Lab Study - Different Display Positions

5.3.1 Display - Method

Before the field trials, potential factors were considered that might influence the usability of touch screens on the flight deck. Environmental conditions (e.g. in-flight vibration), user interface design and position of display were identified as factors that could have a potential impact on usability. Trials showed that it was not feasible to test all impact factors during the field study. It was decided to limit the levels of display position during the field trial (fixed and mobile placements). After that a lab study was conducted that evaluated the potential impact of various display position on usability of touch screens more in detail.

5.3.2 Display - Participants

10 Participants were recruited from the local university campus. Two were female and two participants were left-handed. The mean age was 27.4 (SD=3.4). All participants had obtained their undergraduate degree and the majority of participants were registered in a post-graduate course. Participants average touch screen usage was 4.75 years. 6 participants reported they frequently played action or strategic games on their smartphones/tablets that require fast and precise interaction. Participants received vouchers for their participation in this research project.

5.3.3 Display - Apparatus

The experiment was conducted on an Acer P3 touch screen tablet running on Windows 10. It has 11.6-inch panel with a resolution of 1366x768 pixels. The tablet was attached to a tripod (Manfrotto 058B). The thread of the tripod mount was changed with a M10 screw with longer thread. The tablet was attached to a rectangular wood sheet, via double sided tape, which is attached to the tripod. This modification was required since the stability of conventional tablet holders did not satisfy the expectations.

5.3.4 Display - Experimental Design

The primary independent variable in this lab study is the display position. This is defined by the angular display position and the displacement in vertical and horizontal direction. Secondary independent variables are controlled through the software, where dependent variables are recorded. Background information, initial design and decisions are described in the following sections. Apart from empirical measurements, participants reflected on their subjective experience by means of a questionnaire. The section closes with a summary of independent and dependent variables.

5.3.4.1 Display - Setting of Experiment

Modern cockpit designs (see Figure 5.19) like Boeing 787, Airbus A380 and Gulfstream G600 were compared. Depending on display size and available area on the dashboard there are 4-5 Head-Down-Displays. There are integrated Electronic Flight Bags (EFB) on the window side of both pilots. Avionics like, Flight Management System (FMS) are located on the pedestal. Depending on the sitting position pilots are likely to operate the aircraft system with their dominant or/and non-dominant hand. After pilot trials, it was decided that a 5x2x2 within-subjects design with repeated measures provided an acceptable compromise between factor levels and demand on participants.

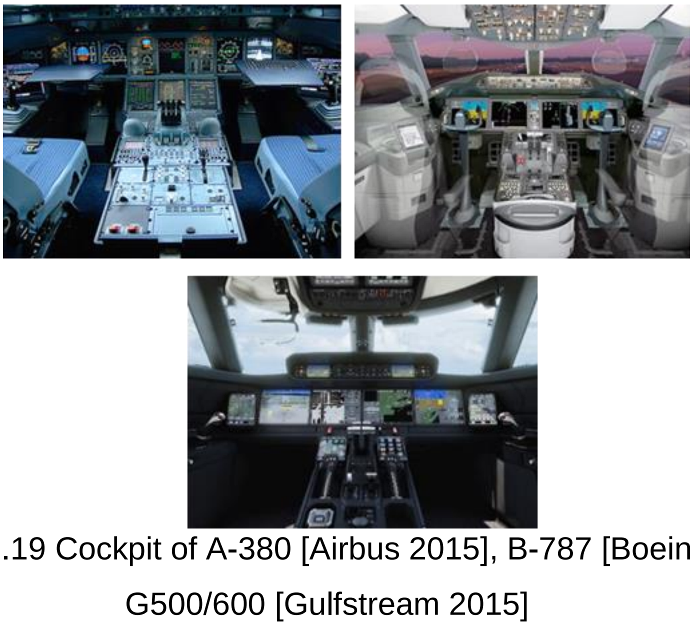


Figure 5.19 Cockpit of A-380 [Airbus 2015], B-787 [Boeing 2015], G500/600 [Gulfstream 2015]

Figure *5.20* illustrates the display positions from above and behind the sitting position. There were 5 display positions (A to E) on angles with 45° increments. Display position C was directly in front of participants and simulated interactions with Head-Down-Displays. Position B and D positioned on the diagonal simulated EFB interaction. Position A and E were placed 90 degrees on either side, simulated systems located on the pedestal. Each position had 2 levels for vertical (near and far) and horizontal (low and high) displacement.

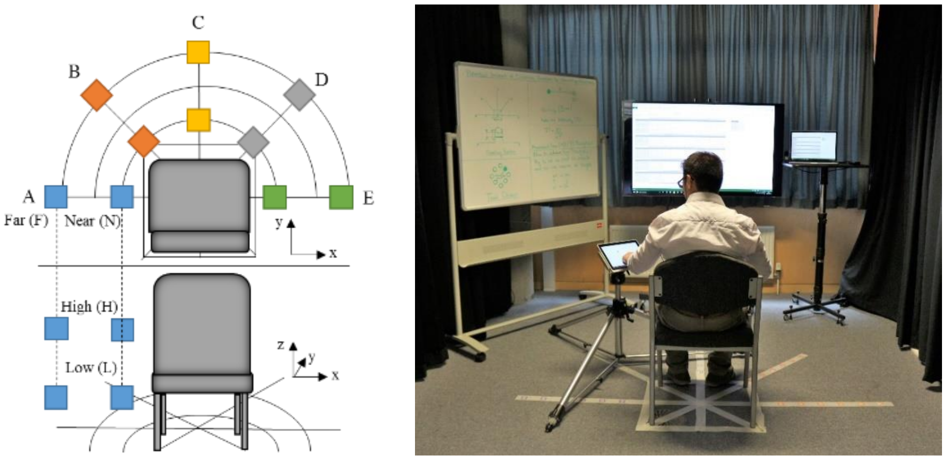


Figure 5.20 Experimental Setting

Near display positions were 40 cm, far display positions were 60 cm from the sitting position. On sides (A and E), low display positions were 60 cm, high display positions 70 cm above ground level. In front (C) and diagonal (B and D), low display positions were 70 cm, high display position were 80 cm above ground level. In position A and E (on sides), the display is parallel to the ground. For position B, C and D the display is tilted toward the participant.

5.3.4.2 Display - Summary of Variables

Table 3 summarizes the independent and dependent variables used in this study. The primary independent variable are levels that defined the screen position. Secondary independent variables are used to gather sufficient quantity of data over a range of task difficulties through measured dependent variables. A subjective rating scale were used to gather general and fatigue indices.

Table 3 Summary of Independent and Dependent Variables.

Variable	Levels	Description
Primary Independent Variables		
Position	5	A, B, C, D and E
Horizontal Displacement	2	Low (L) and High (H)
Vertical Displacement	2	Near (N) and Far (F)
Secondary Independent Variables		
Width	2	50 and 75px
Distance	3	150, 300 and 450px
Targets per Sequence	15	Each 24°
Blocks	5	1, 2, 3, 4 and 5
Dependable Variables for Empirical Measurements		
Movement Time	milliseconds (ms)	
Touch Positions	X and Y Coordinates	
Error Rates	%	
Dependable Variables for Independent Rating Scale		
General Indices	7	7-point scale (higher better)
Fatigue Indices	5	

5.3.4.3 Display - Task Design

The ISO 9241-9 recommended task design and equations for input devices evaluation is shown on Figure 5.1. Applied tapping task software was developed by MacKenzie [2015] using Java SDK 1.6.

Initially, there were 4 levels for distance (75, 150, 300 and 450px) and 3 levels for target width (25, 50 and 75px). Combining all levels would give 12 distinct sequences. Per sequence participants had to hit 20 targets. Sequences with various distance and width levels appeared randomly. After finishing all sequences (240 taps) the block was completed. For each position defined in the previous section, participants repeated the same block 5 times. Thus, participant had to generate 1200 data points per positions. Due to increased fatigue effects and required rest time for recover, completing one position required more than 25 minutes. Discussions with participants, that performed the study during the initial lab trials, showed that target width of 25px (approximately 5 mm) were too small and frustrating to operate. For 75px distance participants said that they do not really move their finger and it gives the impression that they hit the same place.

Therefore, it was decided to remove the first levels for both variables. The ID ranged from 1.58 to 3.32. A wider range of ID is recommended by

Soukoreff and MacKenzie [2004]. Due to nature of the experiment (regarding time), preferences of participants and limited screen area higher ID values could not be included. With the aim to reduce fatigue effects the number of targets per sequence were reduced from 20 to 15. Thus, participant had to tap the screen 90 times per block, 450 times per position and in total (20 position) 9000 times.

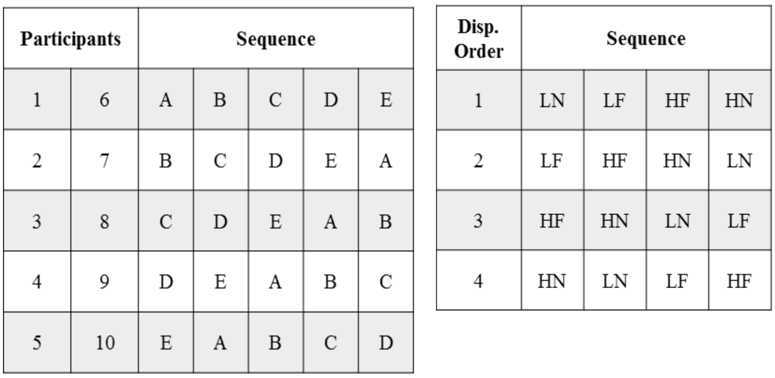
5.3.4.4 Display - Questionnaire

In addition to empirical measurements, an independent rating scale was used to assess impressions of each display position being tested. The independent rating scale taken from ISO-9241 have two group of indices; general and fatigue indices. Questions for general indices are; force required for actuation, smoothness during operation, effort required for operation, accuracy, operation speed, general comfort and overall operation of input device. Questions for fatigue indices are; finger, wrist, arm, shoulder and neck fatigue. On a 7-point scale the questionnaire is formatted in a positive direction, with the highest values being associated with the most positive impressions. As shown on Figure *5.20* participants had a large TV screen in front with questions on a spreadsheet. Between the blocks participant filled out the questionnaire. At the end of each block participants had the possibility to adjust their ratings.

5.3.4.5 Counterbalancing Latin Square

In order to eliminate order effect, the sequence of display positions is counter balanced using 5x5 and 4x4 Latin Square (see Table 4). Both sequences carry on clockwise. For example, participant number 1 starts at Position A with low/near display position. Once Position A is finished position B and displacement order 2 are applied. This carries on in the same way until the participant completes the experiment. The second participant starts at position B with the second displacement order (starting at low/forward position), the rotation continues until all participants finish the experiment.

Table 4 Latin Square for Display Position and Displacement



Participants		Sequence				
1	6	A	B	C	D	E
2	7	B	C	D	E	A
3	8	C	D	E	A	B
4	9	D	E	A	B	C
5	10	E	A	B	C	D

Disp. Order	Sequence			
1	LN	LF	HF	HN
2	LF	HF	HN	LN
3	HF	HN	LN	LF
4	HN	LN	LF	HF

5.3.5 Display - Procedure

The experiment (for one participant) was conducted in three sessions over two days. First, participants filled the pre-experiment questionnaire dealing with demographics and experience and signed the consent form. The investigator explained the aims and objectives of the experiment, before demonstrating how participants could achieve high Throughput values. The investigator asked participants to touch the centre of the target as fast and accurate as possible, but stressed that if becoming fatigued, participants may finish the current sequence and rest until they recover from fatigue symptoms. Since this experiment simulates a flight deck situation where pilots are strapped to the seat, participants were asked not to lean or turn towards the screen as much as possible.

Once participants had familiarised themselves with the procedure the experiment started at the first position. The rule that the investigator applied to decide whether participants were ready to start was if the improvement of Throughput value was below 5% compared to previous block.

With the aim to motivate participants, the overall results (Throughput and Error Rate) of the block were copied on the spreadsheet. After the 3rd block the investigator asked participant to fill the independent rating scale for the current setting. Once the position was finished participants had the opportunity to adjust their ratings. After completing the first position, the first session of experiment was concluded. After a coffee/lunch break (up to 1 hour) participants completed their 2nd and 3rd positions. The final two positions were completed on the following day.

Between the blocks the investigator conducted an informal interview with participants about their experience and observations. After all participants finished the experiment, all mentioned issues were collected and a post-experiment questionnaire was created. On five-point Likert scale participants rated if they would agree with the issues that other participants mentioned.

5.3.6 Display - Results

This section starts with description of raw data pre-processing. Analyses of the distribution characteristics of display positions will be presented. The main results about Throughput, Error Rate and Movement Time analyses will be presented.

5.3.6.1 Display - Data Pre-Processing and Manipulation Checks

99,000 data points (90,000 in lab trials) were imported from the app. Analyses procedure was analogue to the Field Study. Each data point received their key value, which describes the position (A to E), displacement in vertical (N-near, F-far) and horizontal (L-low, H-high) direction. The majority of participants were right-handed. Thus, left-handed participant’s generated data in position A and B were changed with position E and D, respectively. Therefore, position A and B represent experiments conducted with non-dominant hand and position C, D and E represent experiments performed with dominant hand.

The distribution characteristic for Throughput results were assessed. The mean skewness of the distributions, for all conditions, was 0.278. The mean kurtosis was 0.639. Both of these values are low, indicating no overall tendency toward a negative or positive skewness or toward a flat or peaked distribution. A Kolmogorov-Smirnov test with Lilliefors Significance Correction (75% of conditions satisfied this criteria) and a visual inspection of their histograms, normal Q-Q plots, box plots showed that Throughput scores were approximately normally distributed.

For Throughput and Error Rate analyses all independent variables including position, displacement in vertical and horizontal direction were considered separately. In addition to that, the effect of using dominant hand versus non-dominant hand was examined. For the next level of analyses, all independent factors were combined. Results were ranked according to their Throughput results and a matrix show significant differences for all pairwise combinations.

Displayed ID ranged from 1.58 to 3.32. Using the effective index of difficulty, the range increased to 0.47 – 4.56. Creating Fitts' Law prediction models with these data created in some cases some negative off-set values. Previously, ID values below 1.5 were rejected because participants had the feeling there was almost no movement involved. However, to create a realistic Fitts' Law prediction model it is essential to have a wider range of ID values. Thus, an additional experiment (target size and distance levels = 50 and 75 pixel) with an ID range from 0.74 to 1.32 was conducted and added to the results from the previous experiments. The data collection followed the same procedure. The additional participant performed 5 blocks in one condition amounting to 20 blocks per positions. Thus, 9000 data points were collected. This additional experiment was excluded from Throughput and Error Rate analyses. The prediction models were created as described in section 5.2.6.2. First, all data generated from sequences was plotted and then data was binned to subgroups with increments of 0.1.

Statistical results between the 5 blocks that participants had to conduct for a particular position showed no significant difference. This indicated that applied procedure alleviated potential learning and fatigue effects that could manipulate the data set.

5.3.6.2 Display - Throughput Results

The analysis of Throughput revealed that the display position has a large effect on performance. Figure 5.21 shows effect-size for pairwise combinations of different display position. Figure 5.22 shows the mean Throughput values for each participant on different positions. Figure 5.23 has a bar chart showing the mean Throughput values for each particular condition, with 95% confidence intervals. Figure 5.24 presents a matrix that illustrates the significance for pairwise comparisons for each condition. Table 5 presents the mean, standard deviation and the result of statistical tests on Throughput for all conditioning factors. The main results are as follows:

Repeated measures ANOVA tested the effect of various display positions on Throughput. Results indicated a significant large effect of display position on Throughput. Bonferroni post hoc tests showed that the mean score of all display positions were significantly different from each other. All participants had the same trend. Results achieved at non-dominants hands side (worst position) are 26.6% lower than the results achieved at centre position (best position). Smallest Throughput was achieved at the side of the non-dominant hand (A). Throughput was better at the diagonal side of participants' non-dominant hand (B). The lowest Throughput result for the dominant hand was achieved at the side (E). Results were better on the diagonal of the dominant hand (D). Best results were achieved on the centre position where participants could use their dominant hand (C). Cohen's D was used to calculate pairwise effect size. Except two combinations (C&D and B&E) all other combinations showed a large effect. (see Figure 5.21).

	C	D	E	B	Cohen's D Small > 0.20 Medium > 0.50 Large > 0.80
C					
D	0.64				
E	1.38	0.91			
B	2.04	1.65	0.73		
A	3.33	3.04	1.73	0.82	

Figure 5.21 Cohen's D for Angular Display Position

Considering all participants separately, the trend of achieving personal best result on the centre position (C) which is falling continuously to dominant hands diagonal (D) and side (E), to non-dominant hands diagonal (B) and side (A) applied to 8 participants out of 10. The mean Throughput across all participant ranged from 6.26-7.79. The drop in Throughput results ranged from 1.72-2.22 (Figure 5.22). In two cases, it was spotted that participant achieved higher Throughput results at the beginning of the second session compared to the last position at the end of the first session, indicating that fatigue may have impeded the average performance towards the end of prolonged sessions.

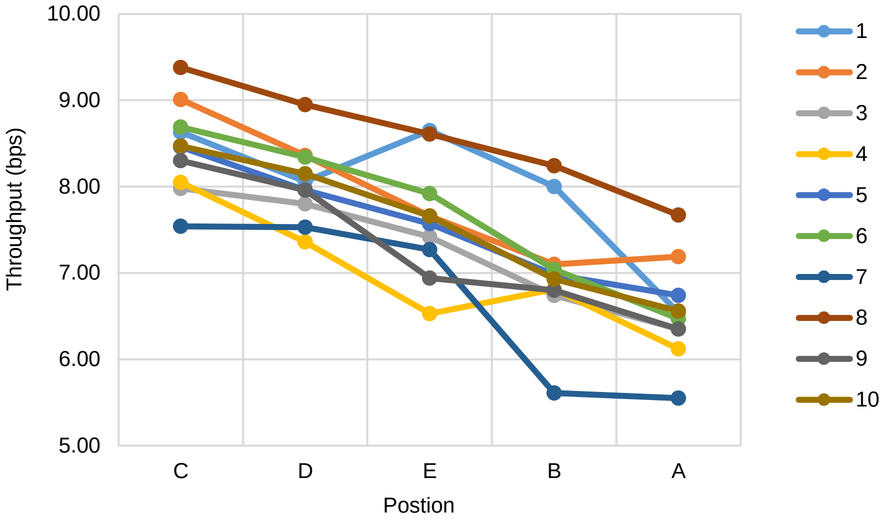


Figure 5.22 Average Throughput Values for all Participants

Figure 5.23 shows that participants achieved a higher Throughput for nearer distances compared to farther distances. Experiments conducted in near distances result in higher Throughput values than for far distances (large effect). The same test was conducted for displacement in vertical direction, which showed no significant difference. There was a significant difference (large effect) in the scores for dominant hand and non-dominant hand conditions.

Bonferroni post-hoc analyses compared all conditions pairwise for significant difference (Figure 5.24). Conditions are ordered according to their mean Throughput value. (The number in brackets are referred to the numbers on Figure 5.24).

- There were five groups where participants achieved similar results (green rectangles). Apart from one pairwise comparison (DLN & CLN), there is no significant effect within the groups (green - 1);
- The Throughput results for comparisons of dominant and non/dominant hand are significant different, with the exception of EFH and BHN pairwise comparison (orange - 2);

- DLN, DLF and ELF (except for ELF and CHF pairwise combination) (from group 1 and 3) did not show a significant effect to positions from group 2. (violet - 3);
- BLN, BLF & BHF and ALN & AHN pairwise combinations are not significantly different. (grey -4)
- All other pairwise combinations, which are not mentioned, are significantly different.

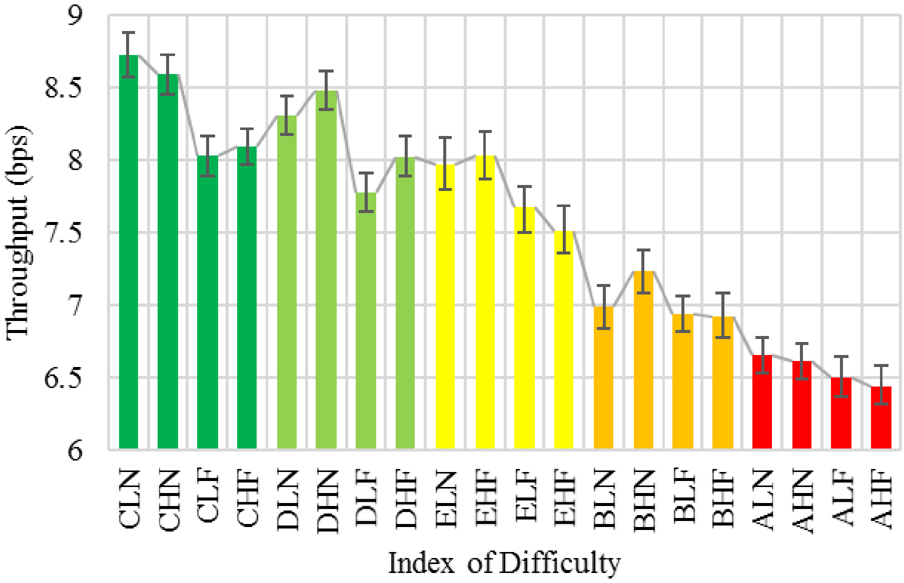
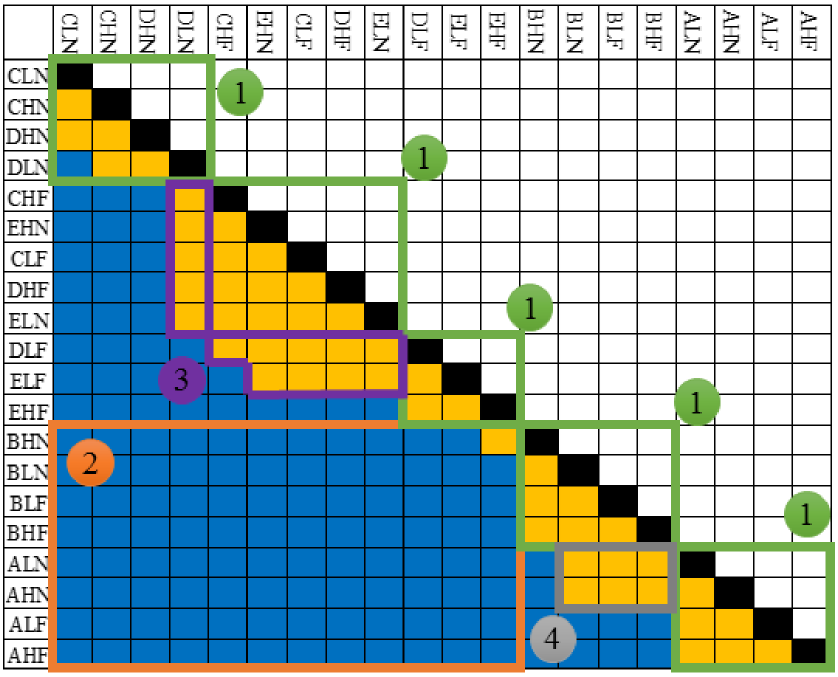


Figure 5.23 Throughput Results for All Conditions



	CLN	CHN	DHN	DLN	CHF	EHN	CLF	DHF	ELN	DLF	ELF	EHF	BHN	BLN	BLF	BHF	ALN	AHN	ALF	AHF
CLN	⬛																			
CHN	🟨	⬛																		
DHN	🟨	🟨	⬛																	
DLN	🟦	🟨	🟨	⬛																
CHF	🟦	🟦	🟦	🟨	⬛															
EHN	🟦	🟦	🟦	🟨	🟨	⬛														
CLF	🟦	🟦	🟦	🟨	🟨	🟨	⬛													
DHF	🟦	🟦	🟦	🟨	🟨	🟨	🟨	⬛												
ELN	🟦	🟦	🟦	🟨	🟨	🟨	🟨	🟨	⬛											
DLF	🟦	🟦	🟦	🟦	🟨	🟨	🟨	🟨	🟨	⬛										
ELF	🟦	🟦	🟦	🟦	🟦	🟨	🟨	🟨	🟨	🟨	⬛									
EHF	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟨	🟨	⬛								
BHN	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟨	⬛							
BLN	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟨	⬛						
BLF	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟨	🟨	⬛					
BHF	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟨	🟨	🟨	⬛				
ALN	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟨	🟨	🟨	⬛			
AHN	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟨	🟨	🟨	🟨	⬛		
ALF	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟨	🟨	⬛	
AHF	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟦	🟨	🟨	🟨	⬛

Figure 5.24 Significant Matrix for All Conditions.

Table 5 Statistical Analyses for Throughput during Lab Experiment

Description	Levels	M	SD	Result
Position	A	6.53	0.58	F (4, 46) =206.7, p <.001, η_p^2 = 0.95
	B	7.09	0.77	
	C	8.50	0.60	
	D	8.15	0.48	
	E	7.62	0.67	
Vertical Displacement	Near	7.76	1.48	F(1,49) = 147.5, p<.001, η_p^2 = 0.75
	Far	7.39	1.37	
Horizontal Displacement	Low	7.57	0.53	F(1,49) = 0.2, p = 0.91
	High	7.58	0.54	
Handiness	Dominant	8.09	0.51	F(1,49) = 452.3 p<.001, η_p^2 =0.90
	Non-Dom.	6.81	0.64	
All Conditions				F (19, 31) = 84.0, p <.001, η_p^2 = 0.98

5.3.6.3 Display - Error Rate Results

The analysis on error rates shows that target size, angular and vertical displacement has a significant impact on error rates. Figure 5.25 shows the different error rates by target size for different positions. Table 6 present the mean, standard deviation and the effects of the independent factors on error rates; the main findings are:

ANOVA compared the effect of various display position on error rate. Results indicated a significant effect of display position on error rate. Bonferroni post hoc tests showed that error rate generated at non-dominant hands diagonal position do not differ significantly from errors generated in both side positions, and error rate at centre position (C) do not differ significantly from errors generated at dominant hands diagonal (D). All other pairwise combinations showed significant difference.

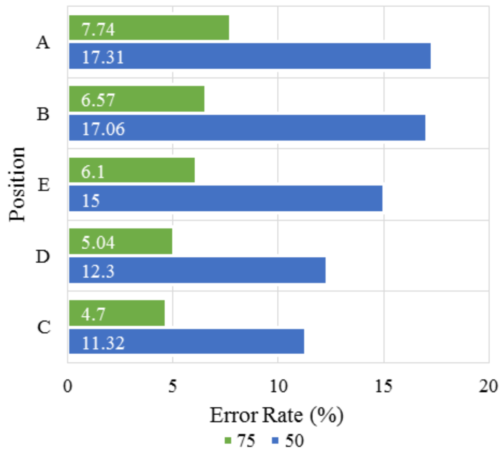


Figure 5.25 Error Rates by Position

Error rates for experiments conducted in near positions were lower than results generated in the farther positions. Displacement in the vertical direction (high and low), showed no significant effect. Participants produced less errors with the dominant hand produced compared to the non-dominant hand. Participants made fewer errors for larger target sizes compared smaller target sizes.

Table 6 Statistical Analyses for Error Rates during Lab Experiment

Description	Levels	M	SD	Result
Position	A	12.41	6.78	F (4, 46) = 3.76, $p<.001$, $\eta_p^2 = 0.25$
	B	11.83	5.66	
	C	8.51	5.08	
	D	8.11	6.81	
	E	10.61	5.20	
Horizontal Displacement	Near	9.86	4.08	F(1,49)= 5.8, p=.020, $\eta_p^2 = 0.11$
	Far	10.73	4.86	
Vertical Displacement	Low	10.45	4.29	F(1,49)=1.7, p=.201
	High	10.13	4.47	
Handiness	Dominant	9.07	4.95	F(1,49)=10.0, p=.002, $\eta_p^2 = 0.19$
	Non-Dominant	12.13	5.90	
Target Size	50	14.60	6.23	F(1,49) = 29.9, $p<.001$
	75	6.03	4.40	
All Conditions				F (19, 31) = 2.73, p=.006

5.3.6.4 Display - Movement Time Results

With regard to Movement Time, the main finding is that the display position has a significant impact to pointing speed. Figure 5.26 shows Fitts' Law Prediction models for all data and mean values for each 0.1 increment group for non-dominants hands side (Position A). *Figure 5.27* shows the prediction models for all positions with their equation and regression. The average R^2 value for equation generated from all data is 49 %. All linear regression models have R^2 value more than 41% (mild correlation). This value is compared to the field trials higher, the reason for that is; the average movement time for a sequence (15 trials per sequence) is plotted. The average R^2 value for subgroups is 95%. The lowest R^2 value was achieved by centre position (C), which is 92%. All regression models showed a strong correlation. This shows that the Fitts Law model is a valid methodology for this setting.

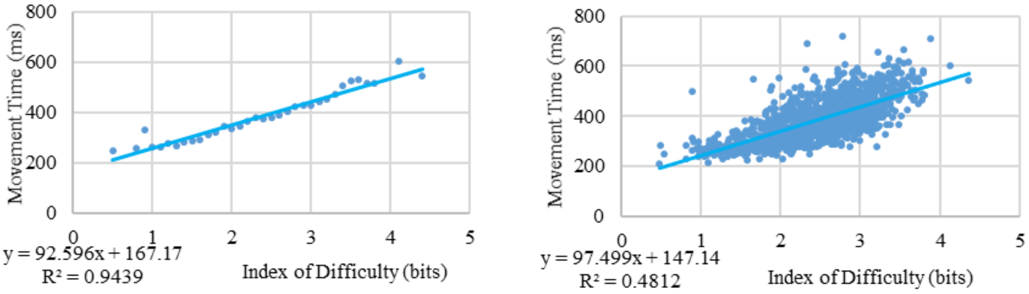


Figure 5.26 Fitts' Law Prediction model for All Data and Subgroups (Graph and Equation for Position A).

It was noticed that experiments conducted with the dominant hand had lower offset value compared to experiments conducted with the non-dominant hand. For equations using average values the average offset value for non-dominant and dominant hand were 171 ms are 149 ms respectively, resulting in a difference of 22 ms. Using all data result an average offset value for non-dominant and dominant hand of 156 ms and 121 ms respectively, resulting to a difference 35 ms. Both approaches yield that participants had a faster reaction time with their dominant hand.

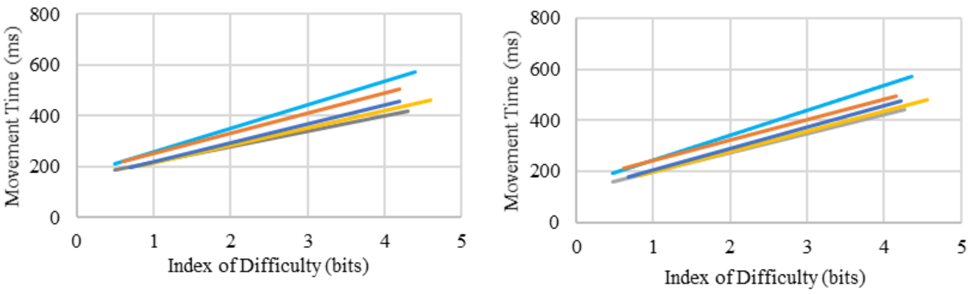
Equivalent, the average slope value for dominant hand was lower than for non-dominant hand for both ways of analyses, showing that participant could move faster to next targets. The average slope for all data at non-dominant hand and dominant hand were 89 ms/bits and 79 ms/bits respectively, resulting in a difference of 10 ms/bits. The average slope for subgroups at non- dominant hand and dominant hand were 86 ms/bits and 68 ms/bits, resulting in a difference of 18 ms/bits.

Due to relatively small screen area it was possible to create a task design which has as ID range between 1.58 and 3.32. Normally it is recommended to apply ID values between 2 and 8 [Soukoreff and MacKenzie 2004]. Using effective Index of difficulty (ID_e), the range increased to 0.5 and 4.6. For the sake of achieving a wider range of ID values, additional data was generated and added to the experimental values.

Binning target into groups gave an average R^2 value of 95 %. A difference was found here by using non-dominant and dominant hand.

Offset and slope values were higher for data generated with non-dominant hand. In respect to averaged values, single values non-dominant hands had 13 % higher slope and 15% higher offset values compared to dominant hands values.

Data generated on ground in fixed placement is comparable with the setting of the lab experiment in centre position. The only difference here was the task design. Compared to the ISO standardize task design, the modified task design (used during the field trials) had 2.2 times higher slope and 62 % higher offset value. A further cause for this could be the instructions during the experiment. In the field trials the investigator requested to favour accuracy than speed. During the lab trials, it was requested to hit target as fast and accurate as possible.



Position	Average	All Data
A	$y = 93\,x + 167, R^2 = 0.94$	$y = 98\,x + 147, R^2 = 0.48$
B	$y = 78\,x + 174, R^2 = 0.95$	$y = 80\,x + 165, R^2 = 0.41$
C	$y = 62\,x + 155, R^2 = 0.92$	$y = 75\,x + 123, R^2 = 0.53$
D	$y = 69\,x + 145, R^2 = 0.97$	$y = 79\,x + 119, R^2 = 0.56$
E	$y = 74x + 147, R^2 = 0.95$	$y = 84\,x + 121, R^2 = 0.47$

Figure 5.27 Fitts' Law Prediction Models for All Data and Subgroups.

5.3.6.5 ISO 9241 – Questionnaire

After completing 3 blocks, participants filled an independent rating scale taken from ISO 9241-411. After the fifth block, when the position was completed, participants had the opportunity to adjust their ratings. The questionnaire includes questions about general as well as fatigue indices. Kruskal Wallis test was applied to levels and positions. Results revealed significant effects for all questionnaire items. Table 7 and Table 8 include the results of the test for general and fatigue indices respectively. In the following, detailed pairwise comparisons will be presented: *(Remember: Position A is non-dominant hands side, Position B is non-dominant hands diagonal, Position C is the centre position, Position D is dominant hands diagonal and Position E is dominant hands side)*. In the following, if results say that particular pairwise comparisons are significant different, all other pairwise comparison which are not mentioned are not significantly different or vice versa.

Actuation force showed significant difference for comparisons for dominant and non-dominant hand. For smoothens during operation, B is not significantly different from both A and E. D is not significantly different from C and E. The effort at E was similar to A and B. In addition, pairwise comparison of B and C was not significantly different. The accuracy and speed at E did not show any significant difference to B and D. Additionally, there was no significant difference for C and D. The comfort and overall operation at C and E was similar to B and D, respectively. Finger fatigue at A was significantly different to C and D. Furthermore, there was a significant difference for B and C. Wrist fatigue at B was not significantly different to A and E. Likewise, D was not significantly different to C and E. Arm fatigue at B was similar to A and E. Pairwise comparison of C and D did not show a significant difference. Shoulder fatigue at C and E was similar to D and B, respectively.

Pairwise comparison was conducted for vertical and horizontal movement. Results are as follows:

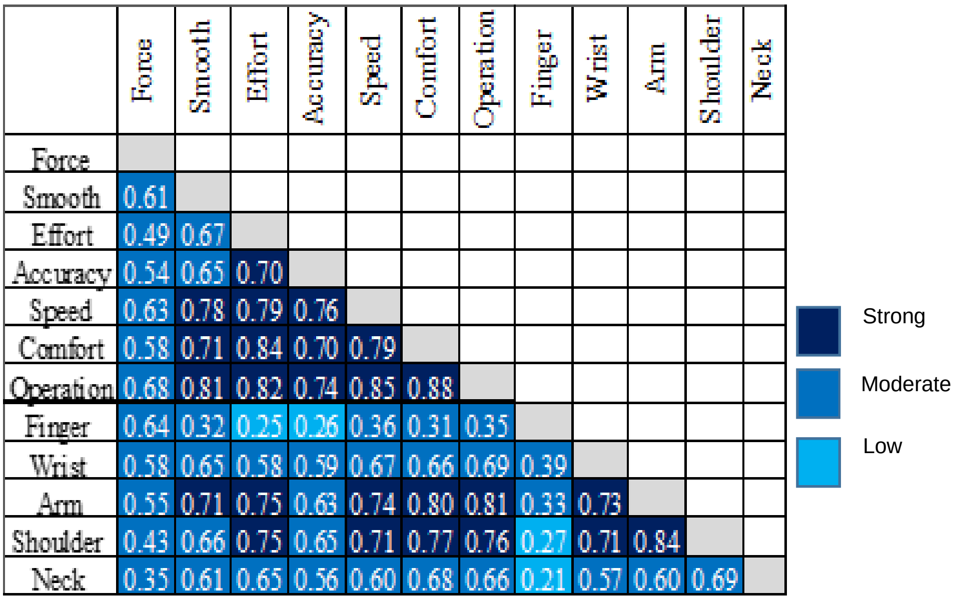
For evaluation in vertical direction (near, far) finger fatigue and smoothness during operation did not show any significant difference. In horizontal direction (high, low) significant differences were only found for shoulder fatigue. Activation force, finger, wrist and neck fatigue showed low or moderate correlation to all other indices. Arm and shoulder fatigue correlate strongly with general indices, except with activation force and accuracy. Within general indices (except activation force), indices had a high correlation with each other (see *Figure 5.28*).

Table 7 General Indices by Position

Desc.	Levels	M. Rank	Result
Force	A	65.86	H(4)=52.1, p<.001
	B	71.75	
	C	138.10	
	D	120.04	
	E	106.75	
Smooth.	A	44.98	H(4)=108.7, p<.001
	B	67.80	
	C	156.93	
	D	133.51	
	E	99.29	
Effort	A	46.04	H(4)=94.8, p<.001
	B	89.20	
	C	154.85	
	D	132.85	
	E	79.56	
Accuracy	A	65.86	H(4)=89.9, p<.001
	B	71.75	
	C	138.10	
	D	120.04	
	E	106.75	
Speed	A	39.28	H(4)=101.1, p<.001
	B	78.99	
	C	153.80	
	D	131.83	
	E	98.61	
Comfort	A	36.01	H(4)=121.3, p<.001
	B	86.85	
	C	158.43	
	D	139.71	
	E	81.50	
Operation	A	36.31	H(4)=119.4, p<.001
	B	82.98	
	C	159.43	
	D	136.73	
	E	87.06	

Table 8 Fatigue Indices by Position

Desc.	Levels	M. Rank	Result
Finger	A	78.41	H(4)=19.3, p=.001
	B	85.84	
	C	118.75	
	D	113.63	
	E	105.88	
Wrist	A	52.43	H(4)=73.0 p<.001
	B	74.14	
	C	141.44	
	D	133.24	
	E	101.26	
Arm	A	46.28	H(4)=95.8, p<.001
	B	73.38	
	C	148.91	
	D	140.91	
	E	93.03	
Shoulder	A	40.06	H(4)=110.1, p<.001
	B	78.45	
	C	160.13	
	D	133.28	
	E	90.84	
Neck	A	51.98	H(4)=99.2, p<.001
	B	95.55	
	C	159.86	
	D	130.28	
	E	64.84	



	Force	Smooth	Effort	Accuracy	Speed	Comfort	Operation	Finger	Wrist	Arm	Shoulder	Neck
Force												
Smooth	0.61											
Effort	0.49	0.67										
Accuracy	0.54	0.65	0.70									
Speed	0.63	0.78	0.79	0.76								
Comfort	0.58	0.71	0.84	0.70	0.79							
Operation	0.68	0.81	0.82	0.74	0.85	0.88						
Finger	0.64	0.32	0.25	0.26	0.36	0.31	0.35					
Wrist	0.58	0.65	0.58	0.59	0.67	0.66	0.69	0.39				
Arm	0.55	0.71	0.75	0.63	0.74	0.80	0.81	0.33	0.73			
Shoulder	0.43	0.66	0.75	0.65	0.71	0.77	0.76	0.27	0.71	0.84		
Neck	0.35	0.61	0.65	0.56	0.60	0.68	0.66	0.21	0.57	0.60	0.69	

Figure 5.28 Correlation Matrix for Subjective Measurements.

During the experiment, the investigator asked participants to reflect on their experience for a particular position and observed participants during the experiments. Participant feedback was recorded. After the experiment, the investigator created a post-experiment questionnaire to test the feedback received from individual participants on all participants. The majority of reported issues were brought up by all participants; however, there were some issues deemed worth asking other participants about. Participants were asked to rate the issues on a 5-point Likert-scale. (1-Strongly Disagree, 2-Disagree, 3-Neither, 4-Agree, 5-Strongly Agree); the results are as follows:

Participants agreed that they performed the experiments better with their dominant hand than non-dominant hand (M=4.4). In addition to that the performance, comfort and effort was better when the display position were closer to participants (M=3.8). The majority agreed that high and far display positions were more fatiguing than lower and near display position (M=4.1). They thought that they would achieve better results in low positions, if their hand would not hide the next target in some cases (Occlusion Problem) (M=3.7). Some participants mentioned that on the sides, the view was limited, which impeded their performance (M=2.8). Everybody agreed that touching smaller targets was frustrating compared to larger targets (M=4.4). Some participants requested to perform the task on the high/near diagonal position of non-dominant hand with their dominant hand. They said that they would use their dominant hand in that particular position if they had to perform a series of interactions (M=3.1). Actuation force and finger fatigue did not change noticeably (M=4.0). The highest mean agreement was achieved by the feedback that not all points from the subjective rating scale have equal contribution to overall performance. E.g. Shoulder and arm fatigue are superior to other fatigue indices (M=4.5);

The questionnaire was applied as it was stated on the ISO 9241-411 standard. The questionnaire is designed in a way that could be applied to a wide range of research areas. For instance, this questionnaire may be applied to compare various input devices in research similar to [Natapov et al. 2009].

In our study, we considered touch screen usage in various display positions. With the exception of activation force, the general indices have shown to be strongly correlated with each other. Actuation force only showed significant differences in subjective ratings for comparisons between dominant hand and non-dominant hand. All participants agreed that the actuation force did not changed noticeably. Thus, for potential studies in which only one particular device is going to be used the question about required actuation force can be excluded.

For smoothness during operation, some participants reported that in various positions their hand obscured the next target. Looking into individual data more closely however showed that the ratings for smoothness were not consistent throughout the conditions; no significant effects could be found.

Participants knew exactly where the next target would appear, (in contrast to the task design applied during the field trials); however, some of them believed that this kind of occlusion would impede their speed. 7 participants agreed that occlusion impacted their speed, while 3 participants neither agreed nor disagreed. Only a few comparisons between the levels of position showed significant difference. There were no significant differences between the levels of vertical and horizontal displacement. This might be explained by the anthropometric differences of participants. Anthropometric measures were not taken from the participants as an analysis of their effects is out of the scope of this work.

Effort was most strongly (inversely) correlated to comfort. As might be expected, there is a strong relation between increasing effort and decreasing comfort.

There is a significant body of research mentioned in Soukoreff and MacKenzie [2004] showing the speed/accuracy trade-off of Fitts' law experiments . In this research, both indices showed the same significant results for all positions. The investigator observed that participants looked to the other general indices and rated the overall operation. This can be seen by the high correlation values with other general indices. From this perspective, the overall impression can be excluded from the questionnaire, for future work similar to this.

Wrist and finger fatigue correlated low and moderately with general indices; the lowest correlation was with effort. Some participants indicated that finger fatigue did not change throughout the study. It was most highly correlated to activation force. Looking at the raw data showed that both indices had the highest average and smallest standard deviation value across all indices.

Arm and shoulder were the indices that affected general indices the most. This was also mentioned by several participants and said that these indices are superior to other indices. This was the post-questionnaire question who had the highest average value. 5 Participants agrees this statement and 5 participants strongly agreed with this statement.

5.3.7 Display - Summary & Research Questions

It was found that the display position has a significant impact on the usability of touchscreen. There was a significant effect between using the dominant hand and the non-dominant hand as well as near display position and far display positions. There was no significant difference between displacement in the horizontal direction. The results of the ISO 9241-9 subjective rating questionnaire were presented and suggestions were made how to customize the questionnaire to similar studies. The obvious limitation of the lab experiment was that the experiment was not conducted in a cockpit setting. There were no simultaneous tasks that participants had to conduct while completing the tapping task.

The next study is related to this and the field trial in the following way, In the +Gz study we will try to understand the potential impact of +Gz on touch screen usability which is a further environmental factor present in the flight of agile aircrafts. Since we will conduct the experiment in fixed placement, it will give us the opportunity to compare the effect of display position with another study that conducted the experiment in mobile placement.

During the last section of this study sub research questions stated in Chapter 1.4 will be addressed.

> ***Sub-RQ:*** *Is there a difference in usability for different display positions?*

The analysis of Throughput revealed that the display position has a large effect on performance. In this experiment the average decrease between the worst position (non-dominant hand side; Position A) and the best position (in front: Position C) is 26.6%.

> ***Sub-RQ:*** *Is there a difference for displacement in vertical and horizontal direction?*

Displacement in both vertical and horizontal direction were tested. Results showed that Throughput for the near placement was significantly better than for the far placements. Results suggest that the Throughput of pilots would be significantly higher if displays were closer. There was no significant difference for horizontal displacement. Error rate results were analogue to Throughput results. There was a significant reduction in error rates for near display position over far display position and there was no significant difference in error rates for low and high display positions.

Sub-RQ: *Does the handedness effect the usability and personal experience?*

Throughput values dropped by moving the screen towards to the side of dominant hands. Conducting experiments with the non-dominant hand produced significantly low Throughput values. Participants made on average 25% less errors with their dominant hand compared to their non-dominant hand. Participants made less than half the amount of errors with 14 mm targets (75px) compared to 9 mm (50px) targets.

5.4 Lab Study - +Gz

Lockheed Martin was one of the early adopters of touch screens that envisioned a panoramic cockpit display (8 by 20-inch panel) in the F-35 Lighting II fighter jet (Figure 5.29). The reduction of switches and mechanical controls on the flight deck, compared to fourth generation jet fighters (e.g. F-16), is noticeable. The aim of touch screen integration was to achieve a user friendly design that reduces pilot workload during combat [Philips 2006].

Figure 5.29 F-35 Cockpit [AHunt 2015] © *Ahunt (Public Domain)*

Pilots flying a fast-jet aircraft are frequently exposed to periods of +Gz during agile flight manoeuvres. Considering the flight deck of the F-35, with its edge to edge display, pilots will have less opportunity to stabilize their hands. Thus, pilots will have less opportunity (especially for interactive areas on the centre of the display) to counterbalance the negative effects of in-flight vibrations and alternating G-forces. Future flight deck concepts incorporate fixed as well as mobile touch screens. For fixed displays, pilots have to extent and raise or lower their arms to interact with the aircraft system; this could be a further degrading factor (assuming no hand support is provided) on usability which needs further investigation. This work presents the results of a lab study that evaluated touch screen performance on fixed displays under simulated +Gz conditions.

5.4.1 +Gz - Method

Figure *5.30* illustrates a person operating a touch screen. Using this figure, a simplified equation (*Equation 5*) can be created that describes the moment (Ma) that applies to the arm of the operator. The two variables which may change by each person is the resulting mass (m) of the arm and the distance (a) to the display. The gravitational force (g) on earth is 9.81 m/s^2.

The gravitational force will be doubled if pilots perform a 60° turn. Thus, the moment (Ma) that applies to pilot's arm will be doubled. Since the gravitational force cannot be increased in the lab, the mass of the arm will be increased to simulate +Gz. There is no study existing that simulated +Gz in a lab environment and this approach was the first method that simulated this factor.

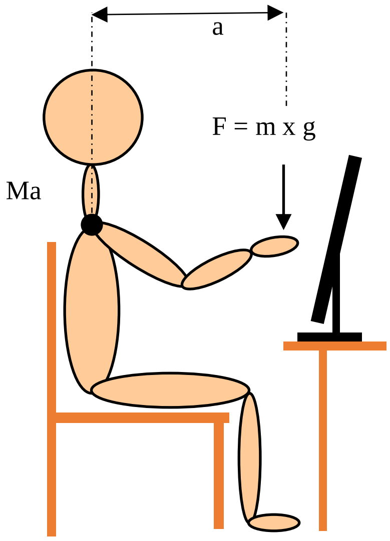


Figure 5.30 Simplified Biomechanics of Touch Screen Users.

$$Ma = m \times g \times a$$

Equation 5

5.4.2 +Gz - Participants

10 male participants were recruited from the local campus. Their age ranged from 23 to 33 years (M=25, SD=2.87). All participants were right handed, owned a touch enabled device (smartphone or/and tabled) and registered in a post graduate course (Master or PhD). The participants'

average touch screen experience was 4.65 years. Six participants frequently played action or strategy games on their devices which requires fast and precise interaction. On a 10-point scale (10 means very good) participants rated their touch screen skills (M=8.40, SD=1.17). Five participants have previously taken part in a Fitts' Law experiment. (Participant information sheet - Appendix IV)

5.4.3 +Gz - Apparatus

Figure 5.31 shows the equipment that was used during the experiment. The task was displayed and executed on a 19-inch resistive touch screen display (Iiyaama Prolite T1932SR) with a resolution of 1280 x 1024 pixels. A portable luggage scale with a graduation of 0.1 kg was used to measure the weight of participant's arm. A weight-adjustable wrist band with 10 pockets (empty weight 0.13 kg) was used to increase the moment that applies to the participant's arm. Required weight were merged with iron bars (0.5 kg) and small iron balls (pellets). A digital weight scale with a graduation of 0.001 kg was used to adjust the total weight that will be added to the wrist band.

Figure 5.31 Equipment used during the Experiment.

5.4.4 +Gz - Experimental Design

A 3x2x3 within-subjects design with repeated measures was used for the experiment. Primary independent variable in this lab experiment was simulated +Gz (3 levels – 1-Gz, 2-Gz and 3-Gz). Secondary independents variables included target width (2 levels – 55 px (15 mm) and 75 px (20mm)) and target distance (3 levels – 100, 300 and 900 px); these were controlled by the software (taken from [MacKenzie 2015]), where dependent variables like movement time, touch position, error rate and throughput were recorded.

5.4.4.1 +Gz - Subjective Questionnaire

In addition to empirical measurements, an independent rating scale based on ISO 9241-9 was used to assess impressions of each simulated +Gz. The independent rating scale is subdivided into two group of indices; general and fatigue indices. Questions for general indices are; Smoothness during operation, effort required for operation, accuracy and operation speed. Questions for fatigue indices are; wrist, arm, shoulder and neck fatigue. On a 7-point scale the questionnaire is formatted in a positive direction, with the highest values being associated with the most positive impressions.

5.4.5 +Gz - Procedure

The investigator explained the aim and objectives of the experiment. After that participants gave their consent by signing a form, and their demographic details were recorded. Participants who had not previously taken part in a Fitts' Law experiment performed a familiarisation task (without weight) before the experiment. Task design and relevant equations were explained. The investigator demonstrated the experiment before participants start with the familiarisation session. Required time and blocks were recorded until participants achieved plateau in TP results and there was no significant improvement. This data set was used to create the power law of practice for this setting and to estimate how long participants needed to practise until they reach their personal maximum performance. The training session terminated, if the investigator or the participant thought they reached their maximum capable TP value, which was important to exclude the learning effect during the experiment.

For participants who have had past experience with this task design the familiarisation session was shortened compared to participants who had no experience. These data set were not used in the power law of experience. After the familiarisation session, there was a break that lasted at least 1 hour for participants who took part in the experiment for the first time and 30 minutes for participants who had prior experience. Breaks between both sessions were set to reduce fatigue effects.

The lab study (Chapter 5.3) investigating the impact of various display positions on touch screen performance found that participants achieved higher TP values and made less errors at display positions which were closer to the participant's body. Compared to far display positions, participants' fatigue indices were also better at near display positions. This information was shared with participants and they were free to adjust their sitting position with respect to the display. Participants used their right hand, which was the dominant hand in all cases. Before the experiment started the investigator asked participants to rest their arm on a portable scale (Figure 5.32). The measurement was repeated a couple of times until similar values were observed. This value was doubled or tripled in 2-Gz and 3-Gz conditions using a weight adjustable wristband.

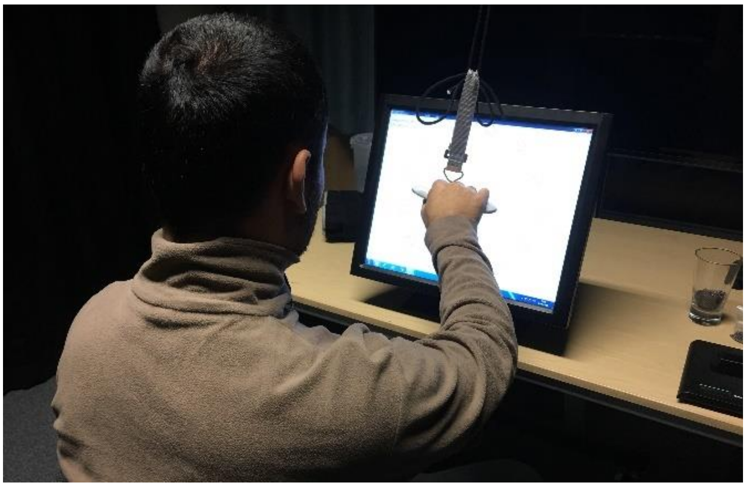


Figure 5.32 Arm Weight Measurement.

Depending on the task order, the investigator prepared the wristband and attached it to the participant's right arm. After attaching the wristband, the weight was checked again with the same method, and then the experiment started. Participants were asked to do the tasks as fast and accurate as possible and to rest if participants felt fatigued.

After the 3rd block the investigator asked participant to fill in the subjective rating scale for the current setting. Once the block was finished participants had the opportunity to adjust their ratings. The other two conditions were repeated in the same manner.

5.4.6 +Gz - Results

Data from 900 sequences was imported. Because of unwanted touches or touching the same target twice, 13 sequences were faulty and excluded from the data set. The distribution characteristic for Throughput (TP) results were assessed. Throughput results were normalized using log transformation. The mean skewness of the distributions, for subgroups defined by level of simulated +Gz, was 0.08. The mean kurtosis was 0.53. Both of these values are low, indicating no overall tendency towards a negative or positive skewness or towards a flat or peaked distribution. A Shapiro-Wilk test and a visual inspection of their histograms, normal Q-Q plots and box plots showed that TP scores were approximately normally distributed.

Statistical results between the 5 blocks that participants had to conduct for a particular setting showed no significant difference. This indicated that applied procedure alleviated potential learning and fatigue effects that could manipulate the data set. Average Throughput and Error Rates values were used to conduct the statistical analyses.

5.4.6.1 Gz+ - Throughput Results

The grand mean values for simulated +Gz are shown in Table 9. As expected participants achieved their best results in the 1-Gz condition without added weight on their wrist. Compared to 1-Gz the decrease in TP values in 2-Gz condition is 6.8% and in 3-Gz condition 20%. With the aim to see the trajectory of TP development one participant was asked to conducted a further condition that simulated a 4-Gz condition. The average TP value across 5 blocks was 50% lower than his TP results for 1-Gz condition. This indicates that the decrease in TP values is exponentially to increase in +Gz. ANOVA showed a significant large effect (η_p^2=0.99) of +Gz to TP results. Bonferroni post-hoc test showed that all levels of simulated +Gz were significantly different from each other. F(2,8)=268, p<.001. Cohens' D was used to compare the effect size pairwise, which showed a large effect in all cases.

Table 9 Throughput for simulated +Gz.

Description	Mean (bps)	SD (bps)
1-Gz	8.32	0.43
2-Gz	7.76	0.59
3-Gz	6.66	0.50

5.4.6.2 Gz+ - Movement Time Results

The grand mean values for simulated +Gz are shown in Table 9. It was observed that participants performing 2-Gz and 3-Gz conditions used more rest time between sequences and blocks, and conducted the experiment in a slower pace. Compared to 1-Gz condition the decrease in movement time in the 2-Gz condition is 10% and in the 3-Gz condition 29%. ANOVA showed a significant medium effect (η_p^2=0.08) of +Gz on movement times. Bonferroni post-hoc test revealed that all levels of simulated +Gz were significantly different from each other. F(2,8)=42.0, p<.001. Cohens' D was used to compare the effect size pairwise which revealed a large effect on all cases.

Table 10 Movement Time for simulated +Gz.

Description	Mean (ms)	SD (ms)
1-Gz	347	14
2-Gz	382	36
3-Gz	449	42

There is a known speed-accuracy trade-off in Fitts' Law experiments [Soukoreff and MacKenzie 2004]. The weight on participant arm decreases the movement time. However, the participants' aiming performance was better. ANOVA proved that +Gz improved the effective width (W_e) significantly, which compensated the difference in TP values. F(2,8)=8.3, p=.004. The total time from beginning of a block to completion provides a more comprehensive view of the impact of +Gz on performance. Participants conducted the 1-Gz condition in 5.30 minutes (SD=1.57) for the 2-Gz and the 3-Gz condition the average time increased by 23% and 38%.

Fitts' Law Prediction Models are shown on Figure 5.34. Equation 6 represent the 1-Gz condition, Equation 7 the 2-Gz and Equation 8 the 3-Gz condition. All equations have a high R^2 value, showing that Fitts' Law is a valid method for this experimental setting. Interceptions should be slightly above 0 ms [Soukoreff and MacKenzie 2004] which is present in all cases. The increase in slope with increasing +Gz shows that participant experiencing high +Gz requires more time to point a target which is small and further away from their current hand position.

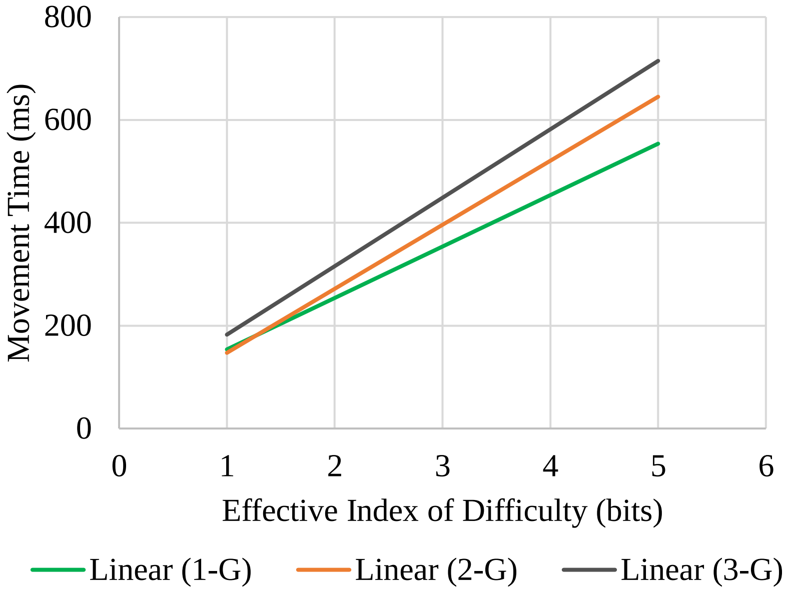


Figure 5.33 . Fitts' Law Prediction Models.

$$MT = 53.8 + 100.0 \times IDe, R^2 = 0.89$$

Equation 6

$$MT = 22.7 + 124.5 \times IDe, R^2 = 0.91$$

Equation 7

$$MT = 49.4 + 133.1 \times IDe, R^2 = 0.93$$

Equation 8

5.4.6.3 Gz+ - Error Rate Results

In this experiment two target sizes were used. 55 px corresponds to 15 mm and 75 px to 20 mm targets. Participants made approximately three times less errors on 20 mm targets (M=1.65%, SD=1.94%) compared to 15 mm targets (M=5.05%, SD=1.99%).

The error rates in different simulated +Gz showed also a significant difference. F(2,8)=4.7, p=.045. Bonferroni post-hoc test revealed that only 3-Gz (M= 2.69%, SD=2.23) and 1-Gz (M=4.04%, SD=1.54 pairwise combinations are significantly different from one other. (2-Gz (M=3.26%, SD=2.41))

5.4.6.4 Learning Curve

Participants performed the Fitts' Law experiment for the first time. During the familiarisation session participants conducted the experiment without any weight on their wrist. TP results for each block were recorded and plotted on Figure 5.34.

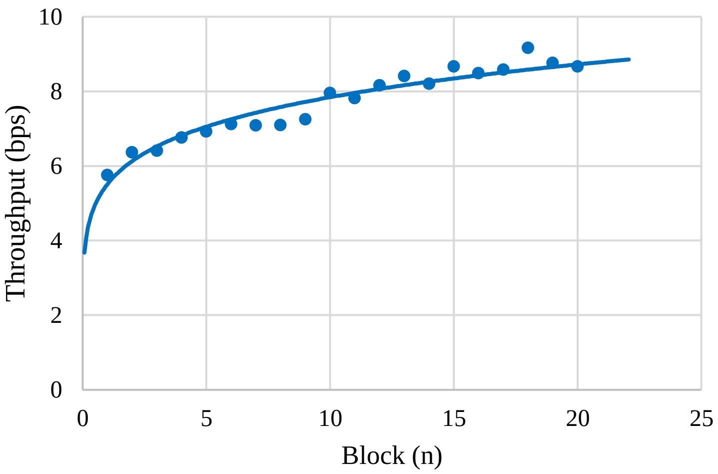


Figure 5.34 Power Law of Practice

$$TP = 5.51 \times n^{0.153}\,, R^2 = 0.93$$

Equation 9

The corresponding Equation 9 gives the power law of practice for this setting. Participants who performed the experiment for the first time have an overall TP of approximately 5.9. Approximately after 20 blocks (1560 taps) participants reach their personal maximum TP values which is around 8.5. A similar mean value was achieved in the previous study investigating the impact of display position (Chapter 5.3). Participants required on average 38 minutes to minimise the effect of learning and to stabilise their TP values. For future projects, it is recommended to offer potential research participants a training that lasts at least 40 minutes. Ideally, the training session should be performed one day before the real experiment to avoid fatigue effects which could be still present from training session.

5.4.6.5 Subjective Ratings

As expected subjective rating scales were not normally distributed, non-parametric tests were applied. Kruskal Wallis H test showed that there was a statistically significant difference in all rating scores between different simulated +Gz. Except for accuracy (p=.032) all other p values were <.001. Table 11 shows mean rank scores and χ^2 results for subjective ratings.

For smoothness, during operation and speed ratings the 1-Gz and the 2-Gz condition did not differ significantly. The other two possible pairwise comparisons differed significantly. For accuracy, the 1-Gz and the 3-Gz condition differed significantly from each other. The other two possible pairwise comparisons did not differ significantly. All other pairwise comparisons which were not mentioned above showed a significant difference.

Table 11 Mean Ranks and χ^2 results for Subjective Ratings.

Description	**1-Gz**	**2-Gz**	**3-Gz**	**χ^2**
Smoothness	23.45	17.25	5.80	21.90
Effort & Comfort	25.20	15.80	5.50	25.96
Accuracy	10.00	17.70	18.80	6.89
Speed	24.75	16.25	5.50	24.87
Wrist	25.45	15.45	5.60	26.71
Arm	25.25	15.75	5.50	25.93
Shoulder	25.30	15.70	5.50	26.28
Neck	25.50	15.50	5.50	27.69

5.4.7 +Gz - Summary & Research Questions

This study investigated the effect of +Gz on touch screen performance. It was confirmed statistically that +Gz has a negative effect on usability. The drop in empirical results as well as subjective ratings is exponential with linear increase in simulated +Gz. There was a small increase in accuracy with increasing +Gz. We seek to transfer this experimental setting to a human centrifuge, where experiments can be conducted under more realistic conditions. Human centrifuges are used to simulate extreme +Gz experienced by fast jet aircraft pilots and astronauts with the aim to train the crew and to develop countermeasures to the impacts of +Gz on the human body.

In the following section sub research questions stated in Chapter 1 will be answered.

Sub-RQ: *What is the impact of increased G-force on error rates and usability?*

Empirical and subjective results largely confirmed the hypotheses of pilots stated that increased Throughput results showed a reduction in mean values with increased +Gz. The trend indicated an exponential fall in TP values. Fitts' Law Prediction Models all yielded high R^2 values showing that this methodology is valid for this research area. The increase in accuracy with increasing simulated +Gz, was the only unanticipated result of the study. Error rates of 20 mm target were approximately three times lower than for 15 mm targets, which suggest to use 20 mm targets on fixed displays on the flight deck.

Sub-RQ: *How are fatigue symptoms affected with increased +Gz?*

Participants subjective ratings supported the overall view. Some participants who performed 3-Gz condition before others changed their ratings after the 1-Gzand the 2-Gz conditions were completed. All participants agreed that compared to the 1-Gz condition the inconvenience in the 2-Gz condition in their arm, shoulder and neck was moderate. However, the 3-Gz condition had a strong effect to these indices compared to the other two conditions. During post-experiment interviews participants said that the 3-Gz condition was painful, and estimated a simulated 4-G condition as their limit where they could finish a sequence (13 taps) before they have to rest their arms.

Sub-RQ: *Can experience and fitness influence overall performance?*

Another limitation worth mentioning are the physical conditions of participants. Pilots flying a fast jet aircraft have to pass medical tests and need to be in a good physical condition. Physical fitness might be a compensating factor that could reduce the effect of +Gz by a certain amount. Previous lab study investigating the potential impact of display position on touch screen usability revealed that personal experience played a significant role in performance rates.

6 Design Study

This chapter presents two studies; the first investigates touch screen based Electronic Flight Bags (EFB) on the specific domain of Search and Rescue (SAR) helicopters. A first set of results aiming to explore and understand potential benefits and challenges of an EFB in a SAR environment will be presented. A review of related work, operational observations and interviews with pilots were conducted previously to understand and specify the use context. A Digital Human Modelling (DHM) software was used to determine physical constraints of an EFB in this type of flight deck. A scenario was developed and distributed to define features, content and functionality that a SAR pilot may wish to see in an EFB. A visual prototype was created and presented alongside the scenario to pilots to support the understanding of the features. Developed initial interface design guidelines and expected features by pilots are presented.

The second research is a user study where a new way of interaction to manipulate radio frequencies of avionics systems is examined. A usability experiment simulating departures and approaches to airports was used to evaluate the interface and compare it with the current system (Flight Management System). In addition, interviews with pilots were conducted to find out their personal impressions and to reveal problem areas of the interface. Potential problem areas were identified and an improved interface is suggested. Key hypotheses driving this work are:

> ***Hypothesis:*** *Participants will be faster and will make less errors on the new developed user interface*
>
> ***Hypothesis:*** *Completion time using the keypad virtual will be similar to physical buttons.*

After this chapter, the framework will be created showing the relation between various variables that could affect touch screen usability on the flight deck. A short summary of all findings will be listed and a preliminary questionnaire will be given that can help avionic designers to evaluate whether a touch screen is an appropriate user interface for their system.

6.1 Electronic Flight Bags in SAR Helicopter Operations

Search and Rescue (SAR) and law enforcement operations requires actively looking outside for targets. Touch screens request users to focus solely on the display which may be acceptable for IFR flights. However, it is likely that this fact will be a significant trade-off against the potential benefits of touch screens.

This study focuses on the specific domain of Search and Rescue (SAR) Helicopters. A scenario was developed (from the interviews described in Chapter 4) which was used to define features, content and functionality that a SAR pilot may wish to see in an EFB. A Digital Human Modelling (DHM) software was used to determine physical constraints of an EFB in this type of flight deck. Developed initial interface design guidelines are presented.

During the second stage of the study a high-fidelity prototype simulating a mobile application customized according to the needs of SAR pilots was created. This was presented alongside with the scenario to pilots. A questionnaire was used to prioritise the features and functionalities of an EFB to be used in this environment.

6.1.1 EFB – Method

Boeing and Airbus have slightly different flight deck design philosophies. However, there is a general agreement that the flight crew is and will remain responsible for the safety of the airplane [K. H. Abbott 2001]. Two-thirds of fatal accidents are caused by human error [Civil Aviation Authority 2008]. Johnstone summarized 11 reports where the use of an EFB has been cited as being a causal or contributing factor for the incidents. These incidents are caused mainly due to human error [Johnstone 2013], which makes designing a usable interface more important.

Potential benefits of applying human centred design philosophy are reduced number of errors, and increased ease of use and learning. ISO 9241-210 [2010] defines human-centred design as *“an approach to systems design and development that aims to make interactive systems* *more usable by focusing on the use of the system and applying human factors/ergonomics and usability knowledge and techniques".*

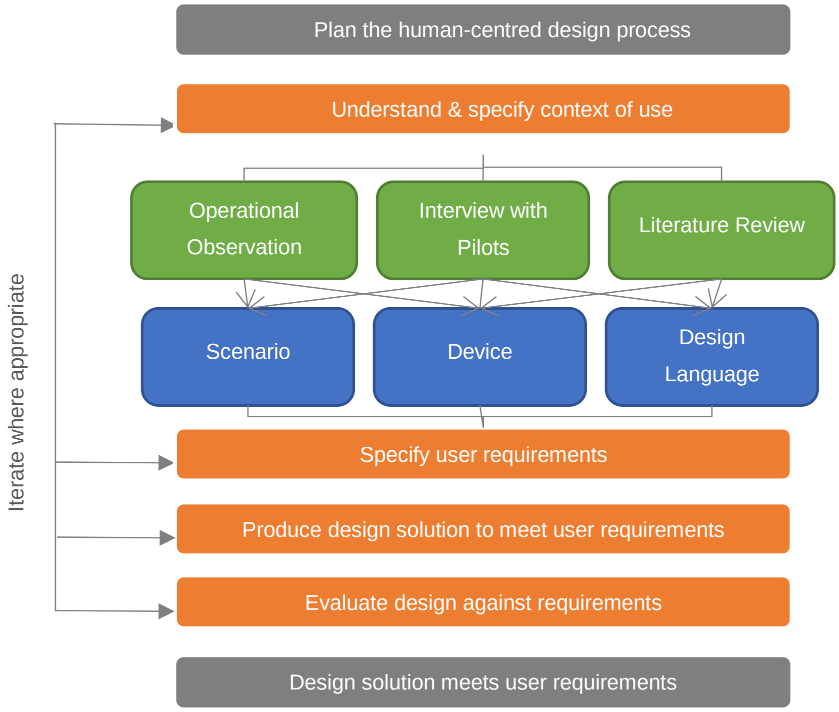


Figure 6.1 . Human-Centred Design Process (based on ISO 9241-210 [International Organization for Standardization 2010])

Figure 6.1 illustrates the human-centred design approach of this research which is based on ISO 9241-210 standards. There are four user centred activities (marked in orange). Spanish Maritime Safety Agency (known as SASEMAR) facilities were visited with the aim to understand the context of use and to define potential application area of an EFB. The investigator was accompanied by pilots and other crew members (rescue swimmer, hoist operator, mechanics and ground operators). The daily routine of pilots was observed on the ground as well as during operations. In order to inform design requirements semi-structured interviews with pilots were conducted to understand their tasks and to define their expectations from an EFB.

As shown on Figure 6.1 interviews and in-flight observations were used to create future scenarios and to define physical measurements of the EFB. Interface design language guidelines were created based on information from the literature review and interviews with pilots. This was done during the first stage of the study. In the second stage, the scenario was presented to focus groups alongside with a visual prototype of the intended EFB application which was designed with the interface design guidelines created during the first stage.

6.1.2 EFB - Device

The first part of the research focused on finding a suitable platform (mobile device) where expected features can be mocked up. A Digital Human Modelling (DHM) software package was used as a supporting tool for hardware selection and design. Project expectations of the DHM package were:

- Integrated anthropometric databases
- Mannequin posture database and modification
- Field of view and reach envelope capability
- Import of Computer Aided Design (CAD) files

A comparative analysis of DHM tools [Poirson et al. 2013] yielded JACK from Siemens [Siemens Industry Software Limited 2013] as a suitable solution for this particular project. CAD files to be imported were generated with SolidWorks.

Interview results showed that physical expectations from a portable EFB are maximised screen real estate, while minimising overall weight. It should fit properly onto the knee and there should be room on the thigh to rest the arms. Strapping the EFB to the knee is likely to have advantages, such as reducing fatigue (pilots could use their legs to support their arms), improving accessibility (the EFB would be within the zone of convenient reach [Pheasant and Haslegrave 2005]), and interacting with one hand, while the other keeps the aircraft under control.

Figure 6.2 shows relaxed seating posture replicated from Rune et al. [2008] (except arm and hand position). The blue rectangle defines the recommended surface area (RSA) for potential EFB’s. The length (L) is defined from the fingertip to the knee and the width (W) is the width of the knee.

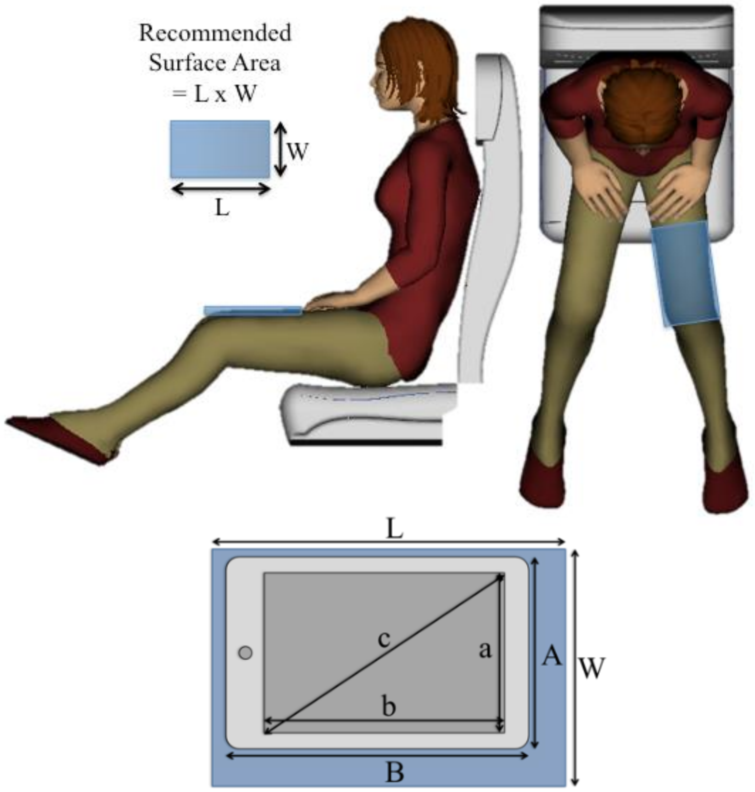


Figure 6.2 Relaxed Seating Posture.

Universal design approach (design for adjustable range) was selected with the aim to achieve minimum fatigue, optimum performance, improved comfort and safety [Happian-Smith 2000]. EFB's are (currently) not safety critical for the operation, so the design limits are established as 5th percentile values for females and 95th percentile values for males. At this point it is worth to repeat that SASEMAR has three female pilots (out of 110). The device would be comfortable to use for the majority (95%) of pilots if it fits to the smallest pilot's knee (5th percentile female).

Integrated anthropometric databases in Jack are: Canadian Land Forces (1997), ANSUR – United States Army Anthropometry Survey (1988), Asian – Indian Database, Ahmedabad, National Institute of Design (1997), German Anthropometric Database, DIN 33402: German Industry Standard (2008), NA_Auto - North American automotive working population, NHANES - National Health and Nutrition Examination Survey (1990) and Chinese adults, report: GB 10000-88 (1989).

The conducted research spans 20 years between the oldest and most recent work. The secular growth in stature per decade for the USA is 10 mm and for Germany is 11.5 mm [Malina 2004; Ahlstrom 2010]. The German database will be used for further analysis because all other sources can be considered as out-of-date. In addition, field trials will be performed with Spanish pilots, and the German data is therefore more likely to represent these more accurately due to closer geographic location.

By accounting for the additive effect of clothing in real world usage [Ahlstrom 2010] RSA values are (L) 223 mm and (W) 142 mm. Suitable devices will be evaluated as followed. All tablet devices which are currently available on the market will be listed, devices that achieve the highest screen area to weight ratio will be selected. The final point is to calculate how well the short-listed devices would fit into the recommended surface area (RSA).

101 tablet devices released since June 2013 were analysed (Information taken from Wikipedia: Comparison of tablet computers). The screen size ranged from 5 inch (127 mm) to 18.4 inch (467 mm). Manufacturers generally supply information about the screen size (see Figure 6.2 – length c), resolution (length a and b in pixel) and weight. These data were used to calculate the screen area/weight ratio (mm^2/g).

The recommended minimum screen size for an EFB is 200 mm (or 7.9 inch measured diagonally) [Civil Aviation Safety Authority Australia 2013], which was considered in the next assessment. 8 Tablet devices that produced the best results in the previous calculation were used for the final evaluation.

The projected surface areas of tablets, were divided by the RSA. The result should be less or in ideal case equal to 1. Results are given in Table 12.

Table 12 Suitable Devices for EFB Application

Model	A (mm)	B (mm)	RSA
ASUS Transformer T90	137	241	1.04
Google HTC Nexus 9	153	228	1.10
Samsung Tab 4 8.0	124	210	0.82
Apple iPad Air 2 9.7	170	240	1.29
Apple iPad Mini 7.9	135	203	0.87
LG G Pad 8.3	127	217	0.87
Samsung TabPro 8.4	128	219	0.89
Samsung TabPro 10.1	171	243	1.31

Samsung GalaxyTabPro 8.4 (Aspect Ratio (AR) 16:10) was the device, which came closest to the ideal value (89%). Predictably, a device with an AR of 16:10 fits better into the RSA since the AR of the RSA is 1.57 (223/142). The next bigger available device is the ASUS transformer T90 Chi with an 8.9-inch display. The length of the device is longer than recommended in RSA. However, the width of the device is more critical because it could collide with the cyclic stick. On the other hand, Samsung GalaxyTabPro 8.4 (290 gram) is 18% lighter than ASUS Transformer. Other devices which seem to be suitable as well are the Apple iPad mini (which is used by some SASEMAR pilots) and the LG G Pad. This simulation confirmed pilots' prediction that the ideal size for a EFB is between 8 and 10 inch.

Another physical consideration is the position of the EFB on the knee. Ideally, the screen surface of the device should be approximately perpendicular to the pilot's line of sight [Pheasant and Haslegrave 2005].

For both extreme cases (95th % male & 5th % female) recommended angle between the thigh-line and EFB is ~ 30° (Figure 6.3). Figure 6.4 shows the improved readability with adjusted EFB angle.

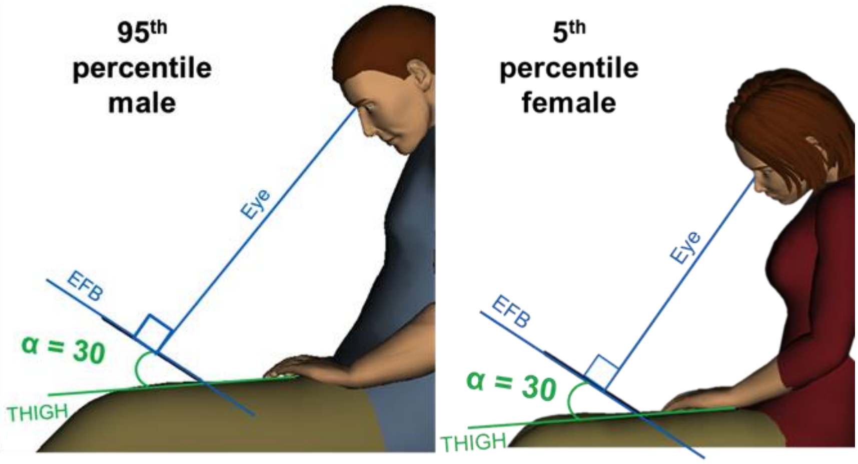


Figure 6.3 Recommended angle between Thigh-Line & EFB.

Figure 6.4 Improved EFB Position on the knee.

6.1.3 EFB - Functional area of the Thumb

Not all of the display surface can be reached with the thumb of the hand that holds the device. Users change or adjust the grip frequently. The functional area of the thumb can be modelled with various approaches [Bergstrom-Lehtovirta et al. 2011].

In this particular case, it is easier to model the functional area of the thumb, since the device is supported by the knee. Pilots could use the edge to stabilize their hand and can move freely alongside the vertical axis. Figure 6.5 shows different hand postures for one handed thumb operation (modelled on an Apple iPad Mini). A 5th percentile female could reach interactive elements up to 51 mm away from the display edge. In addition, it shows the recommended area where the majority of interactive elements should be placed. This will ensure permanent support of the hand, less posture change and enhanced one handed operation. For right hand operation, interactive elements should be placed on the opposite edge.

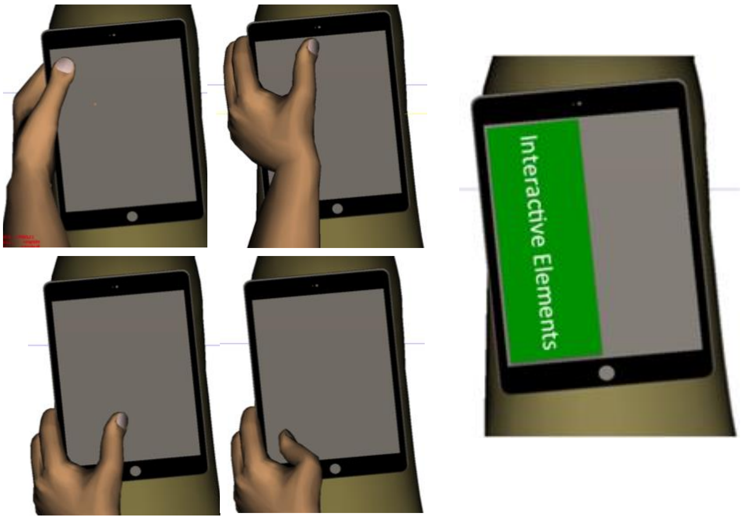


Figure 6.5 Reachable areas for one handed operation.

6.1.4 EFB - Scenario

This scenario was created from the interview results described in Chapter 4. The aim of the scenario is to figure out the features, content and functionality that pilots would like to see in a tablet app. The scenario describes the daily life of SAR pilots in a narrative. The task is to mark the point where pilots think it will improve the overall operation. Features are incorporated in the story are listed below:

Anthony is a SAR pilot based in Valencia. He has an EFB where he can perform various tasks before, during and after the flight.

- Pre-Flight Task

Anthony’s working day starts with checking the state of the aircraft. He has access to aircraft, engine and personal logbooks. The app has also flight rostering capabilities where Anthony can check his upcoming duty times and periods. He checks the NOTAM, TAF, METAR and SIGMET reports and the forecast. Once, he finished his daily routine he receives a mission alert from the responsible MRCC reporting a vessel in distress. He confirms receipt and start with mission preparation.

Anthony tells his crew members that there is a mission briefing in 10 minutes. He downloads the mission file, which includes information about type of mission, target position, number of person, search type and area. The EFB automatically creates a flight plan directly to the target location (including search pattern). He is able to modify the flight plan by adding waypoints. The system calculates and updates Weight & Balance and Performance calculations automatically if a flight plan modification is conducted. The app is set to default (4 crew members and full tank). The pilot adds the weight of SAR equipment and other equipment’s to the weight and balance calculations. The pilot retrieves weather information from target location. The last point is to complete the SAR mission form, which is already partially prefilled by the system using the mission file. The app creates a briefing presentation to all crew members. It is possible to share briefing information or mirror the screen of the EFB to a bigger screen (TV). After the briefing the pilot will tell how much time crew members have to prepare themselves. The device stores all required information and updates it in frequent intervals (e.g. every 30 minutes).

- In-Flight Tasks

Both pilots have access to all types of checklists. The device is communicating with the aircraft system and auto-check it once a task is accomplished. In addition to that he has access to various documents (QRH, POH or IAMSAR Manual). Anthony uploads the flight plan from his tablet to the aircraft system. It shows the own ship position on different maps (aerial, street, VFR and IFR). Anthony uses his tablet as a scratchpad to take note of the clearances received from the ATC. The system has hand writing recognition which offers the possibility to send data (speed, altitude, heading, coordinates and frequencies) to the aircraft system.

During the flight, the pilot can use his tablet as an additional display and is able to mirror PFD, MFD, FLIR and RADAR Displays. Anthony is able communicate, send and receive information from MRCC through his device. He can record specific time stamps (engine start, take off, time on scene, search start and finished, mission completed, landing and engine shut down) which are required afterwards for paperwork. It is also possible to control avionic systems through the device (VOR, NDB, COM, Autopilot). The EFB has the ability to record video footage via FLIR or device camera. The crew found the target and the rescue mission started.

Anthony updates his Weight and Balance calculations after the hoist operation and creates a new flight to the destination airport. The system has also a library with various points of interests (like hospitals or areas with helipads). The system updates the performance data, distance, times and potential fuel usage. Anthony reports the estimated time of arrival to ground units. He has access to approach plates and review the approach plate of the airport before landing.

- Post-Flight - Tasks

The crew enters the room for debriefing. The EFB recorded the path of aircraft for debriefing and for further analyses. It creates a presentation for debriefing where the crew can go through different steps. After the briefing pilots complete the pre-filled paperwork and send it to authorities.

6.1.5 Touch Screen Design Guidelines

In this section, in addition to research conducted within this project, previous studies will be reported that shaped the user interface design, in terms of; layout, button size, font size, colour and symbols.

The most important point might be the need for ease of use during high vibrations. The in-flight experiment was conducted over a duration of one month with 14 crew members, which is already described in Chapter 5.2. The findings from the in-flight study suggested that 15 mm buttons are sufficiently large for non-safety critical Electronic Flight Bag (EFB) applications. For interaction with fixed displays where pilots have to extend their arms, and for safety critical tasks it is recommended to use interactive elements of about 20 mm size. The expected error rate during high vibrations is 3% (likely to occur during transition to hover phases). In the lab study (Chapter 5.4) where we tried to understand the impact of increased G-force on touch screen usability revealed similar results.

These recommendations were based on the results achieved during transition phases, which is the flight mode with highest vibrations and error rates. An avionics engineer stated that not using the interface during transition to hover phases would probably be acceptable for most users. This was also observed during the training flights. Pilots did not interact much with the aircraft system during these phases. Manipulating the frequencies of the avionics system is not safety critical and an error rate below 5% is acceptable. Therefore, interactive elements around 12 mm were used for both studies described in Chapter 6.

Further, the interface should be usable with one hand. From video recordings, it was noticeable that pilots support their hand by grasping the device (fixed displays) and using their index finger or thumb to interact with the screen. The tendency of holding the device was observed in both studies (Chapter 5.2 and 5.3). Interviews with pilots revealed information that was used to determine the physical constraints and user interface layout that meets the pilot's operational requirements. For one hand operation frequently used interactive elements like keypad and switch buttons should be placed alongside the edges. It is recommended to place interactive areas within the recommend area, as shown on Figure 6.5. The majority of pilots could reach interactive elements up to 5 cm away from the display edge.

This should be factored in when designing the hardware as well as interface. For example, the display should be designed in such a way that it enables pilots to stabilize their hands from all directions (from behind included). Pilots identified increased G-force as a potential threat for touch screen usability. The last empirical study, described in Chapter 5.4 and a field study [Le Pape and Vatrapu 2009] revealed that +Gz has a large impact on touch screen usability which increase the importance of design that enables hand stabilisation while interacting with the display.

Worth mentioning is also that this strategy will avoid occlusions which were present in the lab study that evaluated the potential impact of display position on touch screen usability. For differences in handedness pilots should be able to set these interactive elements on the opposite edge. As requested by pilots, the number of interactions to get the desired command should be minimised.

The use of colours and animations on the user interface should be thoroughly investigated. The main reason for using colours is to distinguish and group information on a dense (cluttered) display area [Harris 2004]. To avoid clutter on display area menus, selection and dialogue boxed should be hidden until required. Normal aging of the eye and colour blindness should be considered. Colours should be standardized, consistent in their use and easily distinguishable for all possible flight conditions. Colours should be standardized and consistent with other displays. It is recommended not to use more than 6 colours. **Error! Reference source not found.** shows aviation related colour coding and the functional [Federal Aviation Administration (FAA) 2014]. It is predictable that the EFB will be subordinated in the cockpit. It is expected that pilots will interact with other avionic systems like PFD, MFD and FMS more than with the EFB. Therefore, it is recommended to apply grayscale in a pronounced form and add colour for feedback (or alerting) purposes (EFB applications).

Table 13 Recommended Colours for Features

Feature	**Color**
Warnings	Red
Flight envelope and system limitations exceedances	Red or Yellow
Caution, non-normal sources	Yellow/Amber
Scales, dials, tapes, and associated information elements	White
Earth	Tan/Brown
Sky	Blue/Cyan
Engaged Modes/normal condition	Green
Instrument landing system deviation pointer	Magenta
Divisor lines, units and labels for inactive soft buttons	Light Grey

Today’s operating systems use more symbols/icons in their interface (see iOS and Android OS). Researches showed that symbols can be easily recognized and remembered [Wiedenbeck 1999]. Compared to text (only) there is the possibility that symbols lead to faster recognition [Shepard 1967]. Symbols can reduce the necessity of reading, save space and support the learning of a system [Horton 1994]. Icons may support the learning of a system [Ausubel et al. 1968]. To achieve these benefits icons must be immediately recognisable by the targeted user population [Familant and Detweiler 1993]. Interpreting icons depends on factors like type of software application, text labels and the user's familiarity with the icons [Horton 1994]. Confusion may result if the user is unfamiliar with the icons [Harris 2004]. Labelled icons reduce the risk for wrong interpretations and may significantly increase the usability [Wiedenbeck 1999]. Therefore, it is recommended to label icons

To achieve this benefits symbols must be immediately recognizable to the targeted user population [Familant and Detweiler 1993]. Interpreting a symbol depends on factors like type of software application, text labels and the user's familiarity with the particular symbol [Horton 1994]. Confusion may result if the user is unfamiliar with the symbol [Harris 2004]. Labelled symbols reduce the risk for wrong interpretations and increase the usability significantly [Wiedenbeck 1999]. Symbols which were used in the interface was selected in cooperation with avionic experts and pilots. In addition, each symbol should receive a descriptive text label.

Another study [Kim and Jo 2015] revealed that depending on which finger is used has a significant effect on speed and accuracy. In example, pilots are likely to use their EFBs with their left hand. The majority of the population is right handed. The lab study that evaluated the impact of display position (Chapter 5.3), revealed that there is a significant difference in error rates and interaction speed between dominant and non-dominant hand.

Nowadays, primary usage of EFB is information seeking and processing. Available information are checklists, quick reference handbook (QRH), maps and approach charts. Checklists can be considered as an important interface between pilots and aircraft. The major function of checklists is to provide pilots with a set of sequential tasks in order to configure the aircraft for all imaginable flight modes (e.g. engine start, taxi, take off, cruise and landing) [Federal Aviation Authority (FAA) 1995]. Misusing of checklists were a contributing factor in several aircraft accidents. A review of incident reports, provided by flight crews to the Aviation Safety Reporting System (ASRS), summarized the main issues for checklist related errors, which are [National Transportation Safety Board 2010] (with additional researches which revealed similar findings); failing to use the checklist, skipping items on the checklist, failing to verify settings visually, interruption of checklist flow by outside sources (distraction) and containing error(s) or incompleteness of operator's or aircraft manufacturers checklist. Similar findings were also achieved by Sumwalt [1991] and Ross [2004].

Another research reported that the individual mood (individualism, complacency, humor and frustration) of pilots is an additional factor for deviation from checklists [Degani and Wiener 1994]. It is beneficial to integrate guidelines for checklist design. The following design guidelines are summarized from findings by Degani and Wiener [1992] [1994] and de Ree [1993]; Fonts should be of the sans (without)-serif style, most preferred font is Helvetica, the type size should be 0.10 inch (~8 point) or greater (best readable was 0.11 inch), fonts that have similar looking characters should not be used, long strings of text should be in lower case, when using upper case, the first letter of the word should be larger, font height-to-width ratio should be about 5:3, the vertical spacing between lines should be at least 25-33 percent of the overall font size, the horizontal spacing between characters should be 25 percent of the overall size and at least one stroke width, do not use long strings of words in italics, do not use more than one or two typefaces for emphasis, use black characters on a white or yellow background (best readable is black on yellow), avoid black on dark red, green, or blue.

The average age of SASEMAR pilots, who participated in this study, is above 40 years. Due to old-age-related short-sightedness experienced pilots mentioned that they have difficulties in retrieving information from head down display during high vibration phases. The checklists used in the cockpit are created with a 12 pt font size on prolonged A5 sheets. Therefore, 12 pt font size was used for the user interface in both studies.

Another recommendation was to have pressure activated touch screens to avoid unwanted or accidental touches. Compared to capacitive displays, which are contact activated, on displays with resistive touch technology users have to apply a certain amount of force on interactive elements to activate it. Recently, Apple introduced a new technology called 3D-Touch, which could measure the force applied to the display. Setting a force limit to activate interactive areas could eliminate errors caused by accidental touches.

6.1.6 EFB “Stage 2” - Visual Prototyping Tool

At the end of the first stage of the study, initial design guidelines, possible features and functionalities and the physical size of the tablet device on which EFB applications will run was determined. For the second stage of the study a visual rapid prototyping (RP) tool was required to mock up the interface. 13 different tools were considered. Depending on the level of fidelity RP tools can be categorized in three groups:

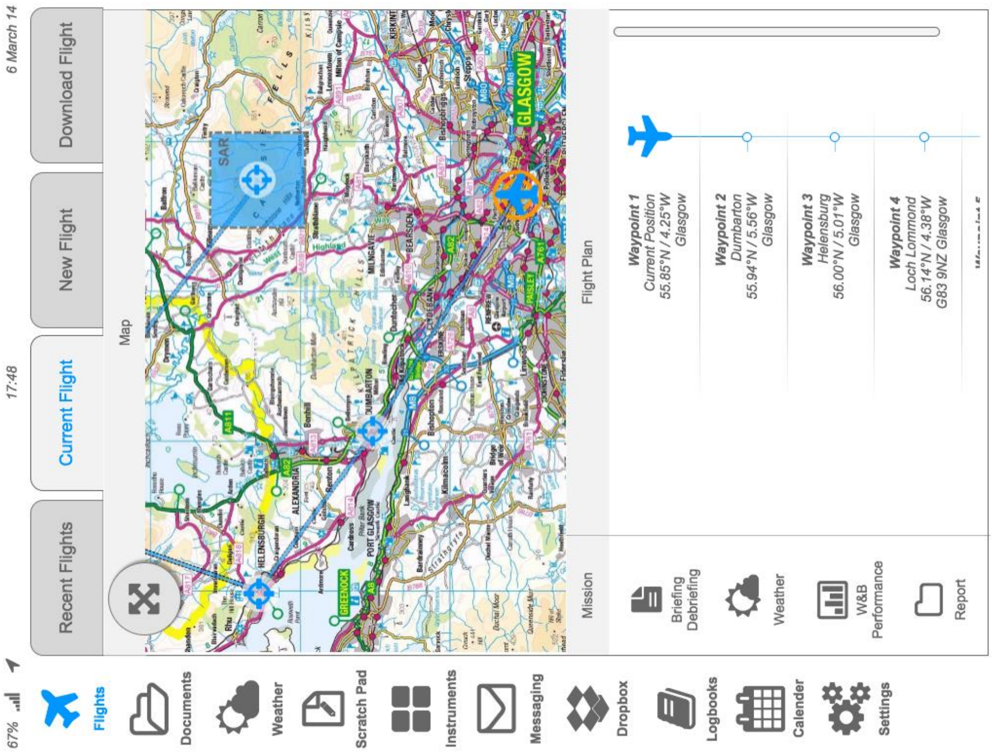
- *Low fidelity* - tools are suitable for describing ideas. It has a “hand drawn” appearance and capability for simple interactions (click operation). (e.g. Balsamiq [2016])
- *Medium fidelity* - tools are able to fully replicate the appearance with limited functionality. (e.g. Fluid UI [2016])
- *High fidelity* - tools are capable to add more features with conditional logic (If-then, Do-Loop operations) or variables. These can be triggered/manipulated by the end-user. (e.g. Axure RP [2016])

The application will be presented directly to potential end-users (pilots). Therefore, Axure RP was selected where we can simulate functionalities as real as possible. Generated prototypes were HTML files, which can be viewed on different web browsers. Possible ways were explored to get the files onto a touch-enabled device. It was decided to use a HTML prototype viewers (Android and iOS devices), which use internet or local storage (offline) to keep and run the RP file (e.g. ProtoSee [2016]). Interaction and performance is not as good as a real application, but it is in an acceptable level. Preliminary designs of the application were shared with avionics experts. Feedback regarding concept of operation, software requirements and design (layout) were received and implemented.

6.1.7 EFB "Stage 2"- Prototype

The following prototype was presented to pilots before the scenario was distributed. This section will describe the functionalities of the proposed EFB application. Each particular step described below is available in Appendix VIII. A few screenshots are presented in this section which should give the reader the idea how the interface looks and operates.

Figure 6.6 shows the main menu of the EFB application. It has a sticky sidebar with buttons (labelled icons) for various functionalities. Through this sidebar pilots have access to flights, documents, weather, scratch pad, instruments, messaging, file sharing, logbooks, calendar and settings. Selected function, in this example "flights", will have a blue symbol and font colour. According to the selection the right side of the display shows the desired information. This is the default position of the sidebar for left-handed operation. For right handed operation users can change the position of the sidebar through settings. The flights section has four tabs; recent, current, new and download flight.



Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow

Waypoint 2
Dumbarton
55.94°N / 5.56°W
Glasgow

Waypoint 3
Helensburg
56.00°N / 5.01°W
Glasgow

Waypoint 4
Loch Lommond
56.14°N / 4.38°W
G83 9NZ Glasgow

Figure 6.6 Main Menu of the EFB App

In recent flights section, pilots can search and review their recent flights. The left-hand side of the page contains a scrollable list box, which shows the recent flights chronologically. Pilots can search recent flights by typing information, like flight date, destination and type of operation, into the search box which is located on top of the list box. Selected recent flight turn to blue and the right side of the display shows flight related information. On top of the page (right side) is the introduction of the incident including information such as; incident number, date, type of operation, location and contact details. The full report and the flight route can be accessed through the labelled icons below the introduction. A brief summary of the incident is given on the bottom of the page.

The current flight shows the active (or most recent created) flight plan. The page is separated in two parts. On top the flight plan is displayed on a dynamic map (moveable by dragging). This section can be enlarged by tapping the expand button which positioned on top left side. The lower part of this section is also divided into two parts. Through the side bar, which is placed on the left-hand side, pilots have access to briefing/debriefing reports and weather information. The right-hand side shows the flight plan. An aircraft (blue symbol) illustrates the current position of the aircraft in this flight plan.

In new flight section, pilots can create a new flight plan by typing a specific incident number and selecting the steps they want to perform for this flight. Available steps are; briefing/debriefing, flight planning, weather information, weight and balance calculation and reports. Since response time is critical in search and rescue operations, responsible MRCC that contacted the flight crew can prepare the mission plan and send pilots a file number. Pilots can use this number to search and download the flight plan, through the “download flight section”. On top of the page is a search box that pilots can use to input the file number. Once the file is found a brief description of the incident will be displayed below. Then the pilots can select which steps they want to perform for this flight. In the following sections, it will be described how pilots can create a new flight plan as shown on Figure 6.6.

If pilots want to create/download a new flight plan the first step is filling the briefing form (which can be skipped). The form is empty if pilots create the flight plan from the beginning or partially (or fully) prefilled if they download a flight plan. The briefing section is divided into two sections; full and short briefing. Both briefings comply with the standards stated in the IAMSAR Manual. After the briefing form is completed pilots can review the information and distribute it to other crew members.

After that the pilots can create the flight plan on the map display. Figure 6.7 shows the map and flight planning page. Similar to main page there is a sticky sidebar on the left-hand side. Labelled icons are; menu, flight plan, synchronisation, ok and undo button, mission, SAR pattern, maps, waypoints and position. Tapping the menu, flight plan, flight information, SAR pattern, maps, way point and position (long tap) will show the functions which are under these buttons. Figure 6.8 shows the interface if all functions are activated. Active buttons have a blue symbol and font colour.

Through the menu button pilots can go back to the main menu. A long tap will put the pilots to the page where they were before they came to the map page or they can directly back flights, documents, messaging, calendar and settings sections. Tapping the flight plan, will display the flight plan window. On this window, which is scrollable, each waypoint of the flight plan is listed.

Waypoints can be selected (the font of selected waypoints will be bold) and may move up or down, edited or deleted or selected as the next destination (through direct to button) via the buttons located on the left-hand side. After pressing the flight info button, flight related information, such as speed, altitude, heading, position, accuracy, distance and time to next waypoint and destination will slide in. With the SAR button pilots, can create a specific search pattern around a selected waypoint. Four search patterns are available; expanding square, sector, ladder, and parallel search patterns.

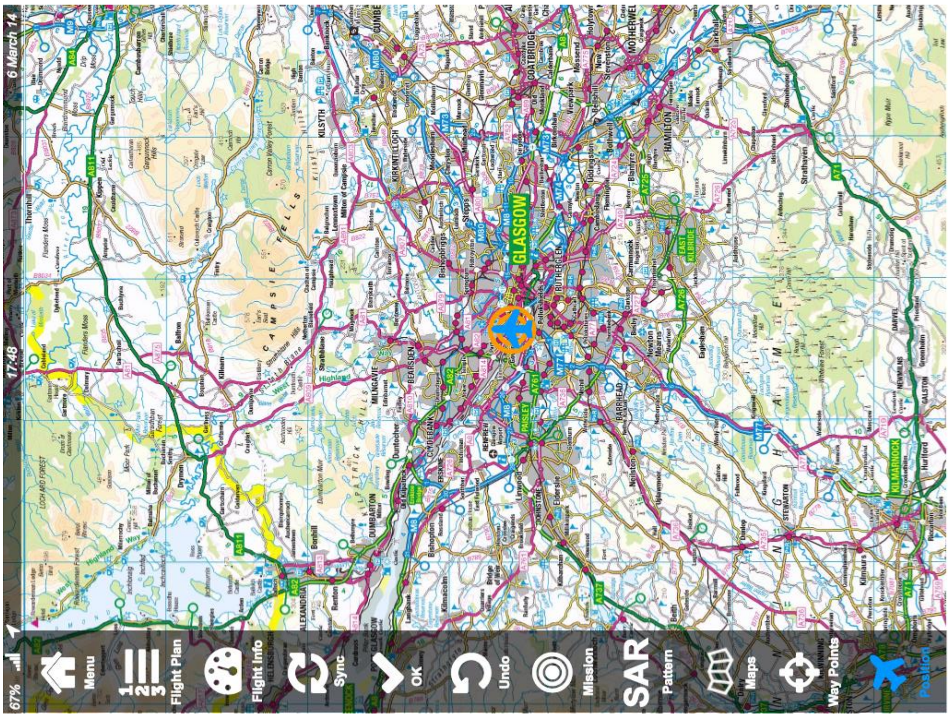
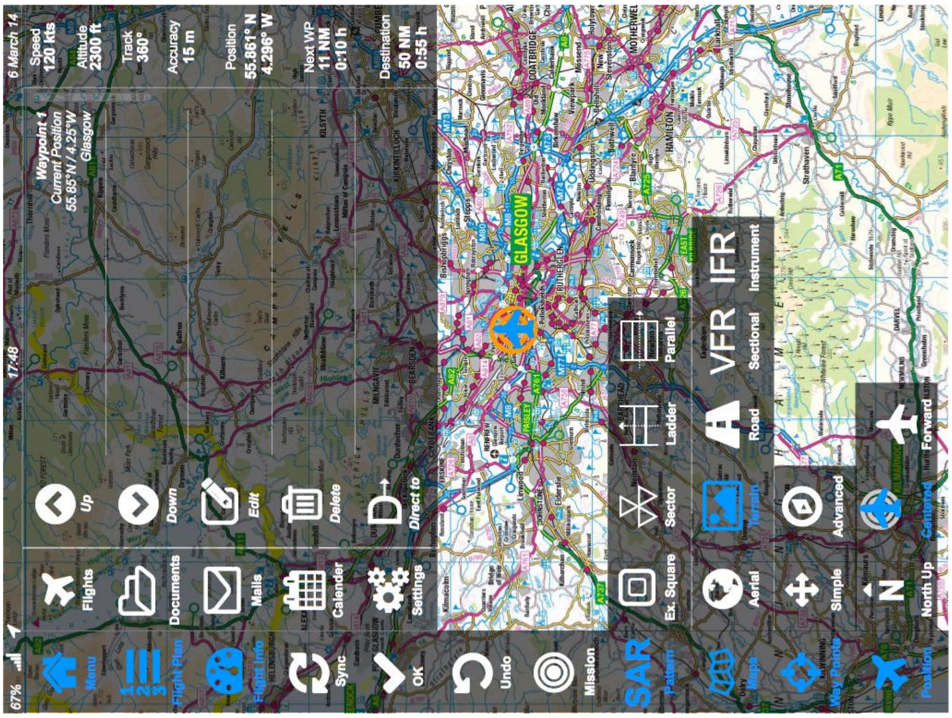


Figure 6.7 Map and Flight Planning Page 1/2



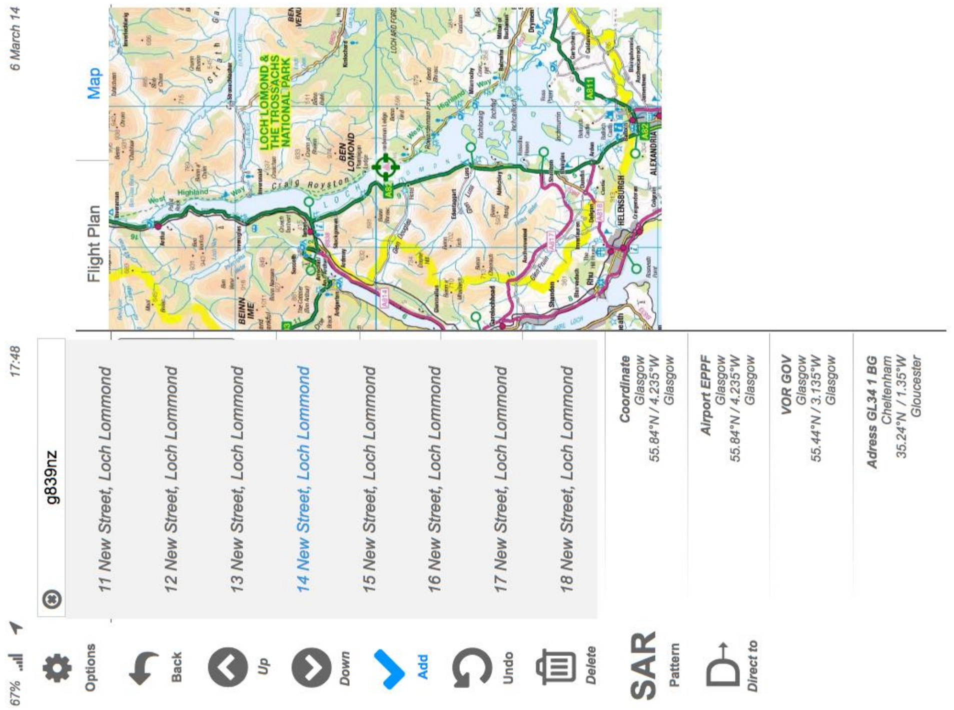
Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow

Speed 120 kts
Altitude 2300 ft
Track 360°
Accuracy 15 m
Position 55.861° N 4.296° W
Next WP 11 NM 0:10 h
Destination 50 NM 0:55 h

Figure 6.8 Map and Flight Planning Page 2/2

When a search pattern is selected the flight plan page will change where pilots can type the parameters (e.g. starting point, turn direction, track, leg space, initial leg length, maximum search radius and speed) of the search pattern. The search pattern will be created if all required fields are filled and the ok button is selected.

Pilots have access to various maps including aerial, terrain, road, sectional and instrumental via the map button. Maps can be zoomed (pinch to zoom). Selected map has a blue symbol and font colour. There are two ways to create waypoints; simple and advanced. Once simple waypoints creation is activated, a crosshair will appear on the centre of the display. Pilots can drag the desired position (below to the crosshair) and tap the ok button to create a waypoint. Another option to create waypoints is through the advanced settings. Selecting advanced waypoints will bring pilots to a new page as shown on Figure 6.9. The design is coherent with the entire application. On the left-hand side, there is a sticky sidebar with interactive buttons. The main page is subdivided into two parts. The left side has a search box where pilots can type information like coordinates, post code, airport (ICAO) and navigation aids codes to find the desired position for the waypoint (or destination). Below is scrollable list box containing recent used waypoints and routes. The right side shows the current flight plan on the list or on the map.

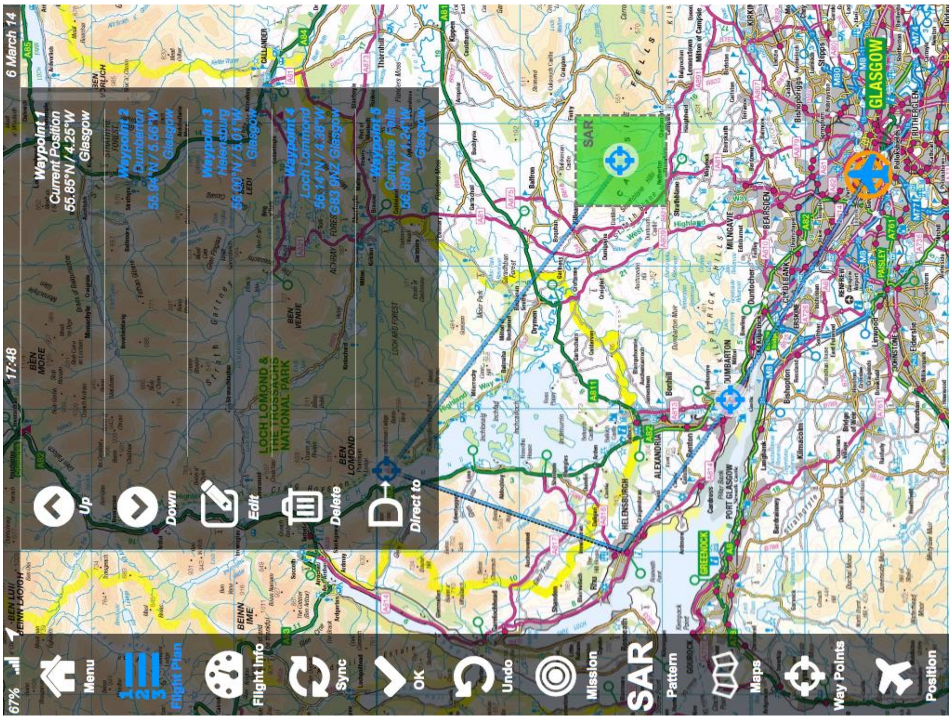

g839nz

11 New Street, Loch Lommond
12 New Street, Loch Lommond
13 New Street, Loch Lommond
14 New Street, Loch Lommond
15 New Street, Loch Lommond
16 New Street, Loch Lommond
17 New Street, Loch Lommond
18 New Street, Loch Lommond

Coordinate Glasgow 55.84°N / 4.235°W Glasgow
Airport EPPF Glasgow 55.84°N / 4.235°W Glasgow
VOR GOV Glasgow 55.44°N / 3.135°W Glasgow
Adress GL34 1 BG Cheltenham 35.24°N / 1.35°W Gloucester

Figure 6.9 Advanced Waypoint Creation

In this example, a post code (g839nz) was used to find the destination. The right display shows the location of this post code and a list box appears where pilots can select the house number for a more accurate positioning of the waypoint. The waypoint will be added to the flight plan by clicking the add button. Assuming that a directly flight to the destination is not possible because there is a temporary restricted area, pilot can use the simple waypoint creation method to add a waypoint in the flight plan. However, creating a waypoint in this method will put this point as the last waypoint in the flight plan. Pilots can change this by tapping the second waypoint on the flight plan and moving it up. The last way of creating waypoint is to press and hold a route which will create another waypoint on this route pilots can drag this waypoint to the desired position. The system will create a new route in green. The former route which is uploaded on the aircraft system is still blue. Tapping the sync button will overwrite the current flight plan with the new one. The modified flight plan will be updated (turn to blue). The last point (Figure 6.10) in this demo was creating a search plan over the last waypoint (e.g. waypoint 5). For this pilots can use the search pattern button on the map page or on the advanced waypoint creation page. After typing the required information, the system will create a search pattern around the selected waypoint.



Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow

Waypoint 2
Dumbarton
55.94°N / 5.56°W
Glasgow

Waypoint 3
Helensburg
56.00°N / 5.01°W
Glasgow

Waypoint 4
Loch Lommond
56.14°N / 4.38°W
G83 9NZ Glasgow

Waypoint 5
Campsie Fells
55.89°N / 1.24°W
Glasgow

Figure 6.10 Simulated Flight Plan

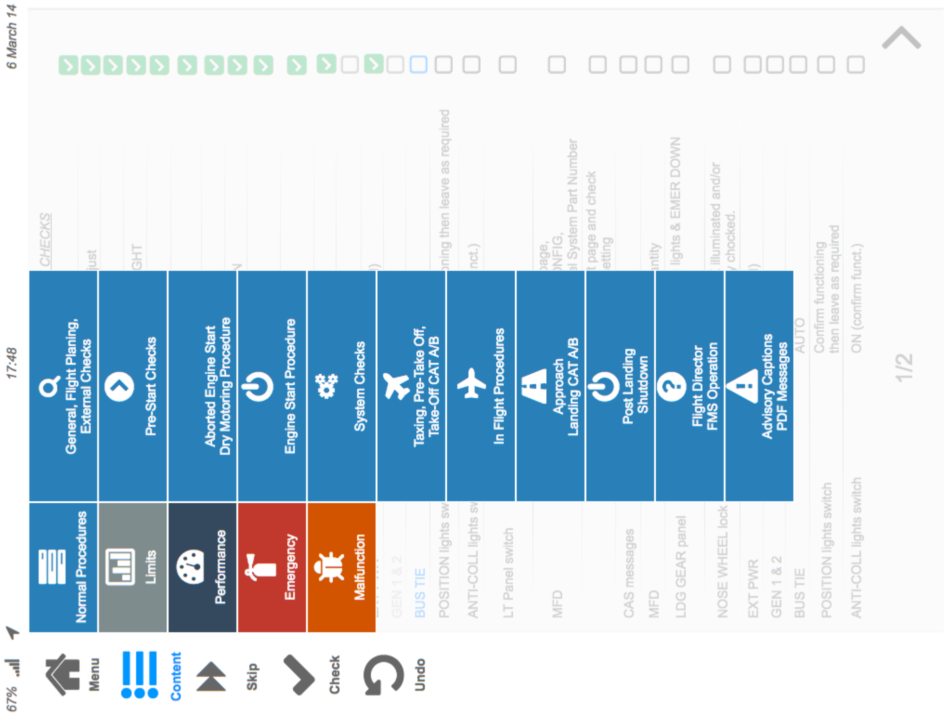
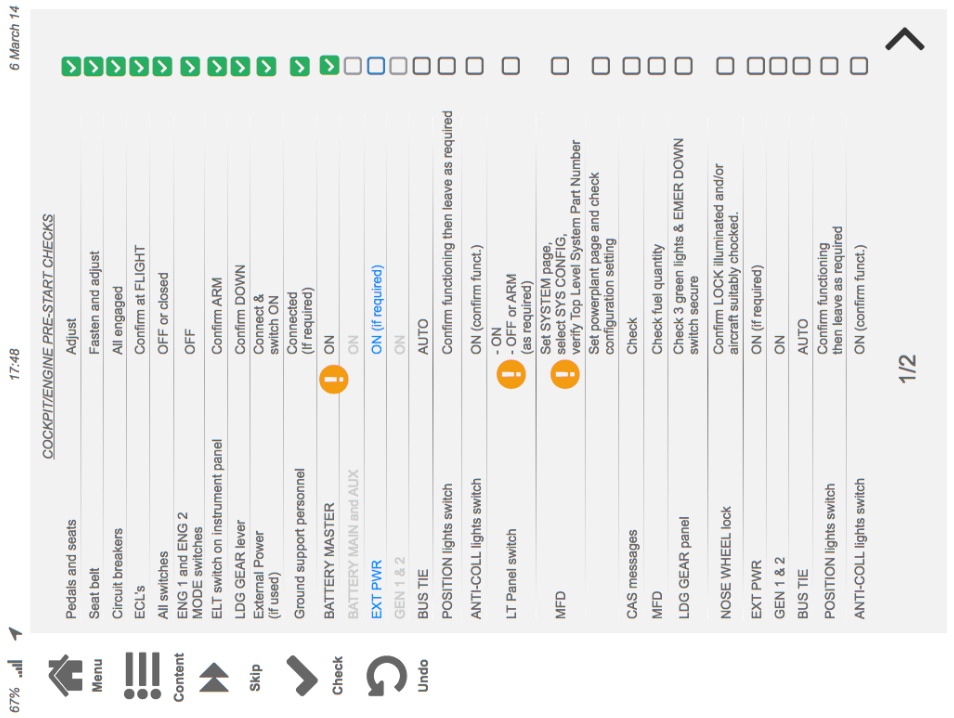


Figure 6.11 Checklist Main Menu

Since there are many waypoints in a search pattern in close proximity the display will show only the search area.

Figure *6.11* shows the main menu for the checklist application which can be accessed through the documents page. There are five types of checklist; Normal, limits, performance, emergency and malfunction checklist. Each checklist group has specific colour coding which is adapted from the original quick reference handbook. Pilots can navigate and find appropriate checklist

The checklist page has the same structure as the rest of the application. On the sticky sidebar positioned on the left are interaction elements like; menu, content, skip, check and undo. The right side of the page is reserved for the content of the checklists. Next task will be highlighted with blue font colour. Checkboxes will be checked once the pilot tap the check button. Then the next task in the checklist will be highlighted with blue. Pilots have the opportunity to skip a particular task and come back later. Pilots can select a task by tapping it. Some checklists may include some message boxes with exclamation marks which pop up and ask a question. In this example the system asks whether external battery is required or not? Yes or No.



COCKPIT/ENGINE PRE-START CHECKS

Pedals and seats	Adjust	☑
Seat belt	Fasten and adjust	☑
Circuit breakers	All engaged	☑
ECL's	Confirm at FLIGHT	☑
All switches	OFF or closed	☑
ENG 1 and ENG 2 MODE switches	OFF	☑
ELT switch on instrument panel	Confirm ARM	☑
LDG GEAR lever	Confirm DOWN	☑
External Power (if used)	Connect & switch ON	☑
Ground support personnel	Connected (If required)	☑
BATTERY MASTER	! ON	☑
BATTERY MAIN and AUX	ON	☐
EXT PWR	ON (if required)	☐
GEN 1 & 2	ON	☐
BUS TIE	AUTO	☐
POSITION lights switch	Confirm functioning then leave as required	☐
ANTI-COLL lights switch	ON (confirm funct.)	☐
LT Panel switch	! - ON - OFF or ARM (as required)	☐
MFD	! Set SYSTEM page, select SYS CONFIG, verify Top Level System Part Number	☐
	Set powerplant page and check configuration setting	☐
CAS messages	Check	☐
MFD	Check fuel quantity	☐
LDG GEAR panel	Check 3 green lights & EMER DOWN switch secure	☐
NOSE WHEEL lock	Confirm LOCK illuminated and/or aircraft suitably chocked.	☐
EXT PWR	ON (if required)	☐
GEN 1 & 2	ON	☐
BUS TIE	AUTO	☐
POSITION lights switch	Confirm functioning then leave as required	☐
ANTI-COLL lights switch	ON (confirm funct.)	☐

Figure 6.12 Checklist Page

According to the answer the checklist will grey out (skip) particular tasks in the checklist. Another function which was presented to pilots was the auto check function. The EFB is able to sense the aircraft system and check tasks in the checklist if they are executed. This was demonstrated with take-off procedure. A take-off was simulated where the speed, altitude and rotor speed changed over time. Once the parameters had reached the desired values on the checklist, the system checked this task automatically.

The last part of the presentation was about demonstrating the weather, scratchpad and messaging functions. The weather section has a similar design as the recent flights section. There is search box on top of a scrollable list box of recent used airports on the left-hand side. The right-hand side shows METAR, TAF, PRIREP, AIRMET, SIGMENT AND NOTAM information. The scratchpad has only two buttons; write and erase, positioned on the bottom of the page. If writing is activated pilots can use their finger to make annotations on the page. If eraser is activated pilots can swipe their finger over the areas to delete the content below. The messaging function is similar to mail application. Pilots can receive mails and flight plans from the intranet, which is another method that the MRCC can use to contact and provide the flight crew with information.

The rest of the features such as instruments, file sharing, logbook and calendar was not integrated at the time where the presentation were conducted. The investigator explained what pilots can expect under these features. In instruments tab pilots, will be able view and control various avionics systems. In logbook section, will include personal logbook as well as the logbook of the aircraft. The calendar feature is a flight rostering program where pilots can check their shift plan.

6.1.8 EFB “Stage 2” - Focus Group

Four focus group sessions with 11 pilots were conducted. Aims and objectives were presented to pilots before the consent forms were signed. First, the prototype was presented as described above in section 6.1.7 then the scenarios were distributed to pilots. Pilots were free to talk, collaborate and decide which features they would prefer in an EFB application.

The majority of pilots would like to have the following features listed in the pre-flight section; logbook, weather, messaging, creating and downloading flight plans, weight & balance and performance calculation and briefing. The only feature where pilots were sceptical about was the information sharing feature. Captains, who normally conduct the briefing, said that briefings would be better if they can mirror the information on a bigger screen (e.g. television) instead of distributing it to other crew members. In-flight features like checklisting, uploading and modification of flight plans, accessing to various maps and approach plates, annotations on scratchpad were the most preferred features. Regarding auto check feature in the checklist pilots said that it would be better if they check all task personally. Pilots were happy with seeing NAV settings, IR Camera and RADAR imaginary, PFD and MFD but were against controlling these avionics devices through the tablet. There were also discussions about whether it make sense to mirror these systems on the device. Some pilots predicted that they would use this function rarely and adding features that users would not use often would increase the complexity of the system. Except flight rostering, all post-flight features listed in the scenario, like debriefing and flight recording was appreciated by the pilots. Regarding flight rostering app pilots said that they shift plan do not change frequently over the week, so this feature is less interesting for them.

After the scenario was completed, pilots were interested about the required time to create an EFB application that they can use during the operations. Some of the requested features that require communication with the aircraft system are subject to approval via type certification (EFB Class 3). In addition, the aircraft system of the AW139 do not enable information exchange with a tablet. Practically, the certification process of a mobile EFB which can communicate with the aircraft system may take long and pilots cannot use a mobile EFB in this type of aircraft. Pilots said having basic functions on a tablet (as described for EFB Class 1 and 2) such as checklisting, access to various charts, creating reports and filling logbooks, which do not need any information exchange with the aircraft system, can be deployed faster and would ease the daily routine tasks.

Pilots recommended to start with a type B software (*include dynamic interactive applications which, could perform various calculations and are able to zoom, pan, and scroll approach charts (to display own-ship position requires further approvals). It has the permission to receive (or update) weather information. An authorised person should validate such applications)* [Federal Aviation Administration 2012] on a EFB Class 1 *(a portable device that is not attached to any aircraft-mounted device. Any data connectivity to the aircraft system is forbidden, and it is not a part of the aircraft configuration. Therefore, a Class 1 device does not require airworthiness approval).* Features that require more time for integration can be considered in future flight decks that enables information transfer between portable devices.

6.2 User Study: Input devices for radio frequencies

A summary of avionics technologies [Blasch et al. 2015] pointed that flight critical systems (FCS- including flight deck displays and controls) and communication, navigation and surveillance (CNS) are important areas essential for maintaining accurate and safe flight. Manipulating radio frequencies of radio communication (COM), very high frequency (VHF) omnidirectional range (VOR), automatic direction finder (ADF) or transponder (XPDR) device are tasks that pilots have to do while flying an aircraft. A new touch screen interface was developed and evaluated in experiments with pilots from the Spanish Maritime Safety Agency (SASEMAR) using a tablet PC and the Flight Management System (FMS) of the Agusta Westland 139 (AW139). The primary aim of this comparative study is to evaluate whether a touch interface developed from the design guidelines created in this thesis is able to cause a significant improvement in usability.

6.2.1 User Study - Definitions of Terminology

This section will define the terminology that is used in this study. The airband, is the name for a group of frequencies in the very high frequency (VHF) radio spectrum allocated for voice communication with other air and ground units. The VHF airband uses the frequencies between 108 and 137 MHz. Each airport has a symbol on a map showing the direction of its runway/s (Figure 6.13a) and the communication frequencies are in near proximity to this symbol. VOR stations (Figure 6.13b) are fixed ground radio beacons that send signals which enable pilots to determine their position through a VOR receiver. Some VOR stations are fitted with distance measuring equipment (DME) which provide the distance between the aircraft and the VOR station (Figure 6.13c). VOR stations use frequencies between 108.00 and 117.95 MHz. A non-directional radio beacon (NDB) (Figure 6.13d) is a radio transmitter that operates in the frequency band of 190 to 535 kHz. Pilots use ADF to determine the direction or bearing to the NDB station relative to their position. A transponder (XPDR) is on board of an aircraft and sends location and altitude information to air traffic controllers. Transponder code (squawk code) is four-digit octal numbers; the dials on a transponder read from zero to seven, inclusive.

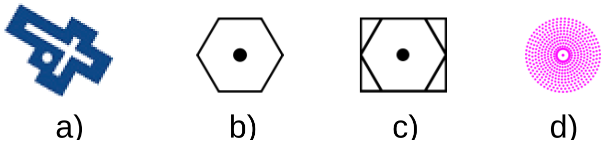


Figure 6.13 Symbology on Maps

6.2.2 User Study - Method

The design rationale was to develop a user interface for radio frequency changes on a touch screen, which is easy to use and learn, error proof and fast to operate. Figure *6.14* shows the “Seven Stages of Action” coined by Norman [1988]. The pilot will define a goal. The “gulf of execution” includes the steps that pilots have to do to achieve this goal. In the “gulf of evaluation” the pilot will check if his actions produced the desired results.

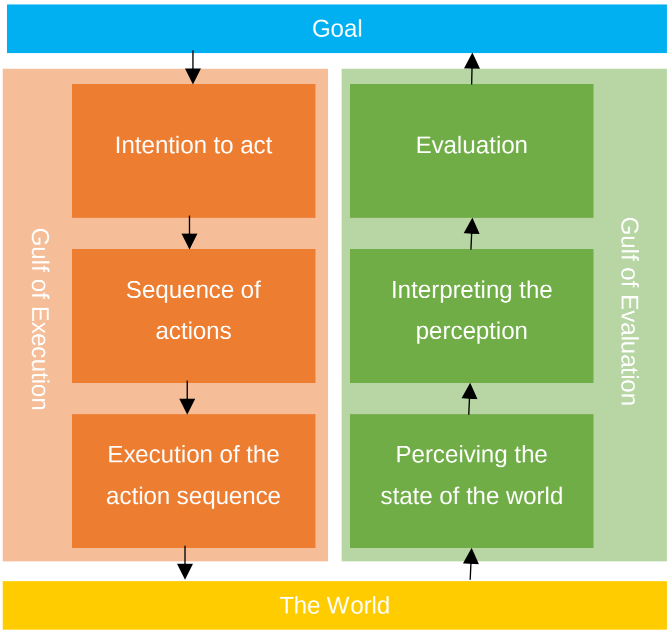


Figure 6.14 Norman’s 7 Stages of Action.

An example that applies to the current study is given below;

Goal – The main objective for the pilot is to operate the aircraft safely. *Forming the Intention* – Navigating from departure airport. *Specifying an Actions Sequence* – Search appropriate VOR frequency. *Execution of Action* – Input frequency into aircraft system.

Perceiving the state of the world – Morse code signal comes up in pilot's headphone. *Interpreting the State of the World* - Pilot listens to the Mors code from the VOR station and takes note. *Evaluation of Outcome* - Pilot is comparing the code with the desired code given on the map.

This example can be repeated for COM, ADF and XPDR devices. The aim of this study was to create an interface that will shorten time between search and execution tasks. The new interface was evaluated and compared with a user study.

6.2.2.1 User Study - Interface

The interface (Figure 6.15) has 2 COMs, 2 VORs, 1 ADF and 1 XPDR devices like in other aircraft that are certified after certification specification 23 (CS 23).

Figure 6.15 User Interface for Avionics Frequency Manipulation.

Figure 6.15a shows the default layout of the interface. It shows the own ship position, the route and waypoints. Users can move the map by dragging it. There are two interactive buttons on the upper left corner. The upper one will trigger the tab that shows the radio frequencies. This is shown on Figure 6.15c left, which will cover half of the page. The right part of the screen, which is not covered by the frequency tab can still be moved. The lower button toggles the visibility of interactive elements. Both buttons are click-activated.

Once interactive elements are activated the symbol of the lower button will change and interactive elements on top of the airports will appear. For demonstration purposes, there is one of each interactive element on the Figure 6.15b. VOR and ADF stations are overlaid with invisible interactive areas. If the pilot wants a particular frequency, he has to drag it towards the “Hot Corner” which slides in after an interactive element is dragged. VOR and ADF stations will turn to transparent white indicating that the pilot is dragging an interactive element (Figure 6.15b).

After dropping the interactive element over the “Hot Corner” the frequency tab and selection tab will slide in (Figure 6.15c). Available frequencies from the airport may be tower, delivery, approach and automatic terminal information service (ATIS). For the experiments the interface was limited to Tower and ATIS frequencies available on the map. The pilot has to select the desired frequency and its destination. The green areas are the active frequencies and the grey areas are pre-set frequencies, which can be switched by tapping the switch button located between the frequencies. The pilot has the option to set (or pre-set) the frequency to a device by clicking the corresponding area. Each manipulation will trigger a visual feedback (flashing). Selecting a VOR station requires only to select its destination (NAV1 or NAV2). Since there is only one ADF device the system will automatically pre-set the frequency once a ADF frequency is selected. The virtual keypad below the radio frequencies can be used for manual input.

6.2.2.2 User Study - Participants

10 male pilots participated in this research project. All participants conducted the user study, however only 8 pilots were available for the post interview. At that time SASEMAR had 3 female pilots (out of 110), which were not on duty. Their age ranged from 32 to 52 (M=42.2, SD=5.6). Logged flight hours ranged from 2500 to 7800 (M=4560, SD=1637). Two of the participants were left handed. All participants are using a touch-enabled device (tablet or smartphone) and rated their touch screens skills on a 10-point sale. (10 means very good) (M=7.8,

SD=0.79). Usage ranged from 1 hours per day to 6 hours per day (M=3.2, SD=1.55). (Participant Information Sheet –Appendix IV)

6.2.3 User Study - Apparatus

Results from interviews and simulations showed that an 8-inch tablet would be sufficiently large to display flight related information. Three pilots already used an iPad Mini as an EFB. Thus, the interface was displayed on an Apple iPad Mini (7.9 inch with capacitive touch screen). In addition, pilots used the FMS of the AW139, which is the current input method for these tasks. Figure 6.16 shows both FMS installed on the pedestal of the flight deck of AW139.

Figure 6.16 Flight Management System of AW139.

6.2.4 User Study - Experimental Design

A 3x3 within-subjects design with repeated measures was used for the user study. Independent variables were 3 scenarios simulating departures and approaches to airports. 3 input methods were compared; physical keypad on the FMS, integrated virtual keypad (Figure 6.15c) and new developed drag and drop strategy. Recorded dependent variables were completion time and error rate.

6.2.4.1 User Study - Task Design

The task is to configure the system for departure (or approach) with a particular input method. Pilots have to manipulate the frequencies of four avionic devices; COM, NAV, ADF and XPDR.

Tasks are given below;

Task 1: Depart from La Guardia

- COM 1 → LGA Control Tower
- COM 2 → LGA ATIS
- NAV 1 → VOR LGA (113.100)
- NAV 2 → VOR SBJ (112.900)
- XPDR → 2466

Task 2: Approach to JFK

- COM 1 → JFK Control Tower
- COM 2 → JFK ATIS
- NAV 1 → VOR JFK (115.900)
- ADF → OGY (414)
- XPDR → 4756

Task 3: Approach to Teterboro

- COM 1 → TEB Control Tower
- COM 2 → TEB ATIS
- NAV 1 → VOR TEB (108.400)
- ADF → TE (214)
- XPDR → 4756

If pilots want to change a particular frequency, they have to look this up on a paper chart, or (if available) on the digital map. The desired frequency then has to be given (copied) into the device. In operational use, usually pilots put the new frequency to pre-set before they make the change. Once they intend to make the change, they will press the switch button to set the frequency. To achieve consistency throughout the experiment, it was requested to put the frequency first to pre-set position and then set it.

Pilots setting a COM or NAV device via FMS have to make at least 5 inputs (without zeros at the end) to get the frequency on the scratchpad. Then they will pre-set and set frequencies. In total, they have to conduct at least 7 key strokes. These are 5 for ADF and 6 for XPDR. Virtual keypad does not require the separating dot (.) the system will automatically put the dot at the desired position once a destination is selected. This means pilots were able to make one keystroke less compared to FMS input.

Touch interaction requires dragging and dropping the interactive element over the “Hot Corner”. Possibly if there is only one frequency (like in VOR and NDB stations) than it is preselected, if not the user has to select the desired one and select its destination. For COM, NAV, ADF devices the number of interaction is 4, 3 and 2 respectively. Since the squawk code (XPDR) is not fixed and usually given by the air traffic control. This input was performed via the virtual keypad.

The number of interactions required for task 2 and 3 are same. Input via FMS require for task 1 and 2&3 34 and 32, via virtual keypad 30 and 28 and for touch interaction users have to make 20 and 19 interactions respectively.

6.2.4.2 Counter Balance (Latin Square)

In order to eliminate order effect, the sequence of task and input method is counter balanced using 3x3 Latin Square. Participants were assigned sequentially to one of the three groups. Table 14 shows the tasks order of the groups. Table 15 shows the sequence of input device.

Table 14 Order of Tasks.

Group			
1	Task 1	Task 2	Task 3
2	Task 2	Task 3	Task 1
3	Task 3	Task 1	Task 2

Table 15 Order of Input Devices.

Seq.			
1	FMS	Keypad	Touch
2	Keypad	Touch	FMS
3	Touch	FMS	Keypad

Participants assigned to first group performed the tasks in the following order; Task 1 (Sequence 1), Task 2 (Sequence 2) and Task 3 (Sequence 3). Participants assigned to second group performed the tasks in the following order; Task 2 (Sequence 2), Task 3 (Sequence 3) and Task 1 (Sequence 1). Participants assigned to second group performed the tasks in the following order; Task 3 (Sequence 3), Task 1 (Sequence 1) and Task 2 (Sequence 2). All group settings were repeated 3 times with 9 participants. Participant number 10 conducted the experiments as described for group 1.

6.2.5 User Study - Procedure

The investigator explained the aim and objectives of the experiment. It was clarified that the aim was not to test the abilities of participants. The main objective is to find out how the current status of the new interface is and to detect problem areas. After that participants gave their consent. The investigator demonstrated the user interface, then pilots had a familiarization session for 5-10 minutes. The investigator gave instructions like “set COM1 to La Guardia ATIS” or “NAV1 to JFK”.

Once the familiarization session finished participants opened the route for their first task. The investigator provided the task written on a paper (as stated in section 6.2.4.1). Pilots searched the frequencies they need to use in the current task. Once ready participants used the desired interaction method to manipulate radio frequencies. To achieve consistency in data input it was requested to put the frequency to pre-set and press the switch button to set it. In addition, it was requested to perform the tasks in the pace as they would do in a real operation. Participants held the tablet device during all input methods. Input errors were recorded and participants were requested to repeat the task.

Additionally, participants could repeat the task if they thought they could improve the completion time.

There were always two pilots on duty. One pilot performed the experiments while the other rested. The entire experiment lasted on average 30 minutes. The completion time and error rates were recorded. After the experiment, there was an informal interview with pilots about their experiences and impressions.

6.2.6 User Study - Results

Completion time results from 90 measurements were imported to SPSS. The distribution characteristic for completion results were assessed. The mean skewness of the distributions, for input methods was 0.85, for tasks was 0.57. The mean kurtosis was 1.31 and 0.66 respectively. Both of these values are low, indicating no overall tendency towards a negative or positive skewness or towards a flat or peaked distribution. A Shapiro-Wilk test and a visual inspection of their histograms, normal Q-Q plots and box plots showed that completion time scores for keypad and tasks were approximately normally distributed. The p-value for FMS (p=.047 for input device) was slightly below the cut-off value of 0.05. Therefore, parametric tests were applied. All mean (M) and standard deviation (SD) values are in seconds. Few input errors were made by the participants using the physical and virtual keypad. These were excluded from the analyses and pilots repeated the task. ANOVA could not detect a significant difference between tasks ($F_{2,8}$=2.60 p=.080). Therefore, average completion time per participant was used for statistical analysis.Figure 6.17 shows the mean completion time and standard deviation for all input devices. ANOVA revealed a large effect (η_p^2=0.85) in input methods ($F_{2,8}$=22.8 p<.001). Touch interaction (drag drop) was the fastest interaction method (M=33.0, SD=6.3). Bonferroni post hoc test revealed that there was no significant difference between FMS (M=39.8, SD=8.2) and virtual keypad (M=40.2, SD=8.6). Other pairwise comparisons showed significant differences.

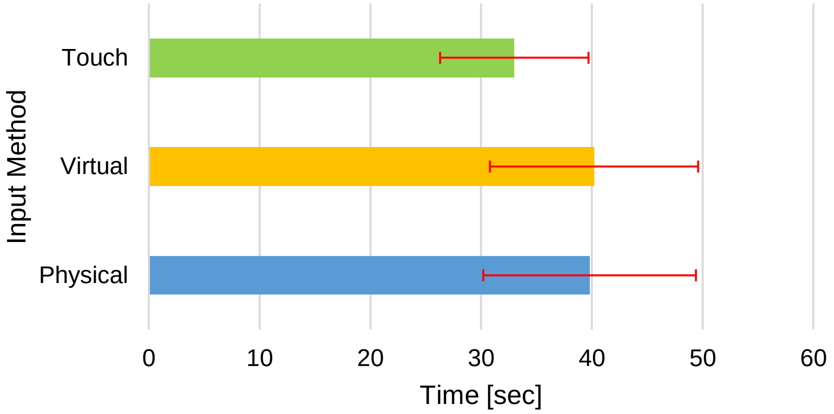


Figure 6.17 Mean Completion Time and SD for All Device

6.2.7 User Study - Post Experiment Interview

After the experiments the investigator performed an unstructured interview with pilots. Identified flaws were used to improve the current design. In the following section interview results will be presented.

> *Pilot 1: "It (touch screen interface) was very easy to use and I learned immediately how I should use the interface…"*

> *Pilot 4: "I like that I was able to use it only with one hand. …I think as improvement you can consider a design where I can put my hand… that will compensate (vibrations during the flight)…"*

Overall all participants had a positive impression from the new developed way of interaction. They found the key idea design for "one hand operation" (placing interactive elements alongside the edges) a good countermeasure for in-flight vibrations. Pilots confirmed that this interaction strategy is easier to learn and to use than the current system. In another study [Riley et al. 1993] pilots often comment that the interface design of FMS appears to have been done from the perspective of the engineer, rather than the pilot. Riley [1996] stated that avionics systems would be much easier to learn and use if their underlying logic would match the task demands of the pilots.

Another point which is not directly related to interface design was the request for arm support if the display would be fixed on the dashboard. This was also requested in a different study where pilots tested a new interface on a laptop with touch panel [Ragland 1987]. The size of the font and interactive elements was mentioned previously during the interviews. This was considered in this design. Pilots were asked whether the size of both are sufficiently large.

> *Pilot 3: "Yes, the size of the text and buttons are large enough. I think that would not cause any problem in the air…"*
>
> *Pilot 2: "Yes… but I think the device was a bit too small… I would prefer a larger (touch) screen, because the map area was too small and the frequency (radio tab) covered too much place…"*
>
> *Pilot 1: "I agree (with Pilot 2). You should look to the displays of the new Agusta Westland 189. I think they large enough for this type of interaction"*

Each pilot agreed that the size of the font and interactive elements were large enough for operational use. Pilots said that the 8-inch display is too small for this type of interaction. Some pilots mentioned that they had difficulties with moving the map while the radio tab was retracted, because the draggable area was too small. This was also found by Hamblin [2003]. Their recommendation was to display this system on a larger display. Some pilots estimated the size of displays like in the Agusta Westland 189 (AW189, with four 13-inch head down displays) may be large enough to perform this task easily.

> *Pilot 5: "…it is nice to see the name of the station, but we usually know which frequency belongs to which station… so, you could delete that and the interface would be "cleaner"."*

In addition, pilots said that it was nice to see the name of the station above the radio frequencies. However, if that could save space and provide more area for the map, it should be avoided. Pilots would prefer to fix the radio frequencies to its place (rather than making it retractable).

A previous research conducted in military vehicles [Hong et al. 2011] suggested not to perform drag operations with touch screen on a moving vehicle. This was reminded to pilots and asked if they would think that might be an issue for their domain.

> *Pilot 4: “Personally I did not have any problems with activation (dropping the interactive area over the “Hot Corner”)”*

> *Pilot 3: “No this was not a problem… I think it would be a problem if you had to drop it precisely over a point. In this case, I was able to swipe the button (interactive element) over the edge (“Hot Corner) and it was activated. If this is a problem to other pilots maybe you can create a design with only click operations”*

> *Pilot 2: I did not had problem with dragging the item, but sometimes I had the problem to find right button. … New York is a very dense airspace with lots of stations and airports. Interactive elements overlapped and it was difficult to point the right interactive element…”*

The way of drag and drop interaction was found to be easy and intuitive. Pilots opinion was that it would not cause a problem since there is no precision drag required to select the frequency. The current way of interaction requires click and drag operations. The invisible interactive area over navigation aids caused mapping problems. Some pilots suggested to use only click interactions. Pilots stated that they had sometimes difficulties finding the location of the invisible interactive element especially if interactive elements overlapped. The most difficult part of this interaction method was to identify and point the interactive element, the rest seemed to be easy and straight forward.

> *Pilot 5: “… You can try to make all interactive elements visible with an icon. Maybe it would make easier to spot the right interactive element”*

Their suggestion was to put visible interactive elements over VOR and NDB stations like on airports. So, clicking a navigation aid will open a message box asking for its destination. Pilots predicted that using solely click operations would make the process easier. A common request was to have a button that centers the own ship position (north up and track up).

> *Pilots 2: "If I have the possibility to set my frequency in this way (touch interaction), I would rarely use the (virtual) keypad. Maybe only to set transponder code… "*
>
> *Pilot 1: "Yes (agree with Pilot 2), maybe you can make this extra (separate it from radio tab). Thus, you have more area on the screen for the map*

Another suggestion was to separate the virtual keypad for manual input. Pilots assumed that they would use rarely the virtual keypad if they would have the option to tune radio frequencies that way. One of the most frequently requested feature was to integrate the ability to create and modify flight plans and to display air traffic on this interface. Pilots stated that flight planning is performed through the alphanumeric keypad on the FMS. Since, the input is done manually there is room for human error. Pilots reported some incidents where pilots input wrong coordinates into the aircraft system. Some scenarios were discussed how an interface can be created which is more error robust. Last but not least, a further request was to design the interface for portrait as well as landscape mode (adaptive view).

Data saturation was achieved after the interviews with the 5th and 6th participant. Last two pilots did not produce any new information.

6.2.8 User Study - Improved Interface

Feedback from pilots and observations were integrated into the new design. Figure 6.18 shows the new design which is designed for a 13.3-inch display. Figure 6.18a shows the default view of the improved interface. The frequencies are now fixed alongside the edge, which can be mirrored to the opposite side. In the previous design, there were 3 buttons for each frequency (pre-set, active and switch). For the sake of saving space this was reduced to one button with description, active frequency (large font) and pre-set frequency on it. This button will be used to switch frequencies. Near bottom edge there are 3 buttons; activation switch for interactive elements, centering own ship position and keypad.

In Figure 6.18b interactive elements and the keypad are activated. This will be visualized with a light blue background color. The key pad and interactive elements over airports and navigations aids are displayed on Figure 6.19 a, b and c show the interactive elements over VOR, NDB and Airports respectively. Some airports incorporate navigation aids. Rather than placing two interactive areas in close proximity a new icon Figure 6.19d) was designed showing that both frequencies. Both frequencies can be found by clicking this interactive element. Clicking on an interactive element will open a new window with available frequencies (up to 15 per page). On the example shown on Figure 6.18c the interactive element over John F. Kennedy Airport is selected. On the page, there are interactive elements describing the frequency, description and the destination device. Once the desired frequency is selected, possible destinations will turn to light blue (in this example Com1 and Com 2). Pilots selecting the destination will receive a visual feedback (flashing). The system will put the frequency to pre-set first, another click is required to activate it.

As it was present in the first version of the interface selecting a VOR station requires only to select the destination and another click to activate it. Selecting a NDB station requires only an activation click. As requested the entire operation is executed with clicks. A comparison study [MacKenzie et al. 1991] revealed that pointing at targets is significantly faster than dragging them. The weakest part of the design may be still the size (8mm) of interactive elements over navigation aids and airports. This design was tried out in a static environment and users found the size sufficiently large. An in-flight experiment could show whether the size is large enough. Three participants recruited from the local university campus conducted a pilot study. The task was displayed on a 27-inch touch screen monitor (Iiyama Projective Capacitive Touch Screen VESA 27" Monitor), however the interface size was as 13.3 inch.

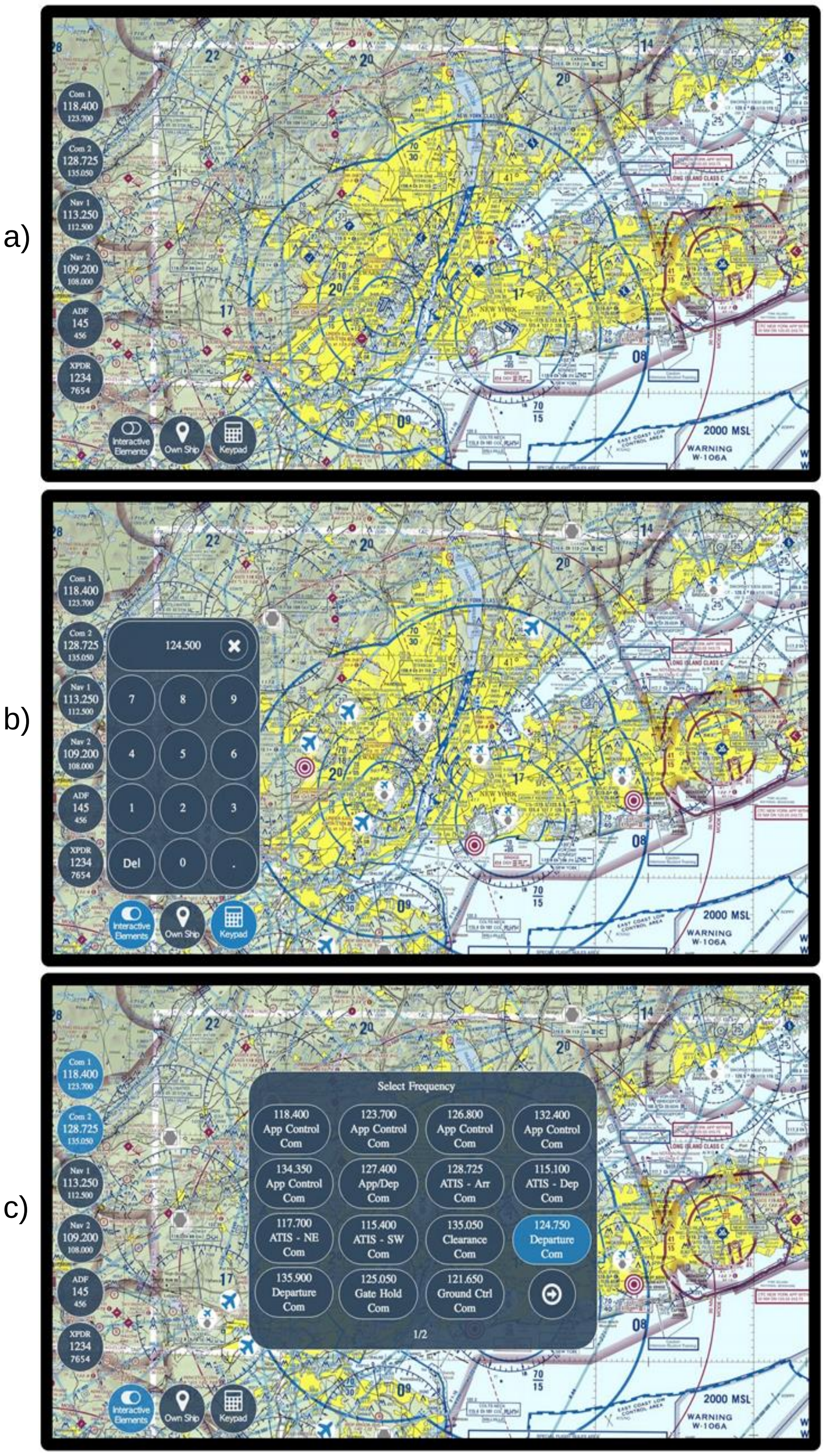


Figure 6.18 Improved User Interface.

Figure 6.19 Interactive Elements

The distance from the seating position was approximately the same distance as that between pilots and the main instrument panel. After a brief introduction and familiarization session participants simulated the same take-off and approaches as described in the main study. Figure 6.20 shows the results of the improved interface compared to the previous results recorded during the main trial. The main completion time of the improved interface (Touch 2) was 26.5 seconds with a standard deviation of 3.5 seconds which is significantly shorter than the previous interaction strategies.

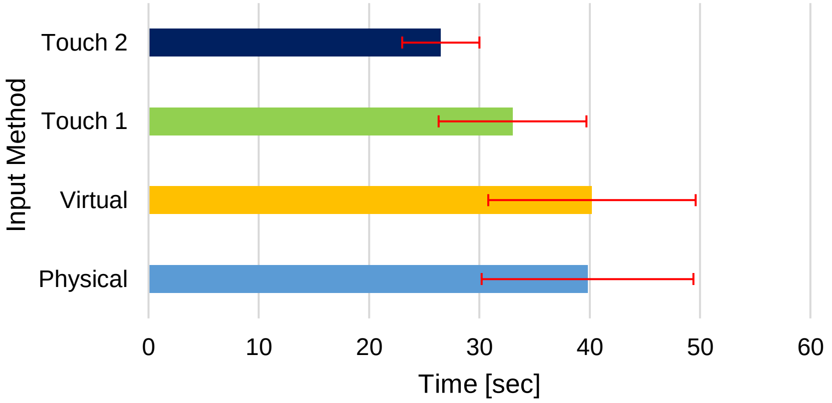


Figure 6.20 Initial Results of the Improved Interface Compared to Previous Results

However, this reduction in completion time is not a result of interface improvement. The experimental setting in the pilot study is not comparable with the main study. Pilots conducted the experiment in mobile placement where participant in the pilot study conducted the experiments in fixed placement. In addition, the interface size in the pilot study was significantly larger than the interface on the tablet device which improved the interaction speed significantly. Participants in the pilot study did not had to move the map to select the interactive element because all required interactive elements were visible. In the main study, sometimes pilots had to adjust the map which caused a higher variability in the mean completion times. Another reason for a reduced variability in completion time in the pilot study can be explained with the number of participants. Compared to the pilot study, which was conducted with 3 participants, the main study was conducted with 10 pilots. Thus, it is predictable that the improvement in completion time is caused by change in interface design and display size.

6.2.9 User Study - Summary & Research Questions

A new way of interaction to manipulate frequencies of the avionics system was presented. Analyses of task completion time showed that the touch interface is significantly faster and less prone to user input errors than the conventional input method (via physical and virtual keypad). Results revealed that designing user interfaces that represent their real-word counterparts (skeuomorphism) will not improve the usability and the design of user interface plays a key role in performance. An improved interface is proposed that was shaped by interviews with pilots and personal observations. In the following section the last sub research question which was raised in Chapter 1.4 will be addressed.

> ***Sub-RQ:*** *Which input method provides the best and safest interaction method for radio frequency changes?*

There were only 2 (out of 30) task sequences where the input with FMS was faster than touch (drag drop) interaction. Comparison of physical and virtual keypad showed no significant difference. Results revealed that designing user interfaces that represent their real-word counterparts will not affect the completion time significantly.

Pilots opinion that the hardest part to localize the target interactive element and to point it was coherent with the investigators observation. After the familiarization session pilots swiped the interactive element over (sometimes slide over the edge) the ”Hot Corner” without paying attention to its location.

Two pilots performed the experiments at the same time. The majority of pilots were right handed. Pilots sitting on the right-hand side had to use their non-dominant hand to make inputs via FMS. In touch interaction participants always used their preferred hand. This could be another factor that increased the difference between the input methods.

7 Framework

Figure 7.1 shows the framework, which was developed from research presented in this thesis and other relevant studies. The framework sets out relationships between four key kinds of factors: environmental, user, physical, and virtual factors. The direction of arrows visualizes which aspect(s) influence another aspect(s).

Solid lines are quantitative findings, derived from empirical measurements and statistical analyses. Dotted lines are qualitative findings from interviews, questionnaires and informal conversations with experts and participants. In the following sections studies, will be introduced briefly and findings will be summarised to provide the rationale for the framework. ***Superscriptions (numbers) at the end of each finding are provided in Figure 6.1***

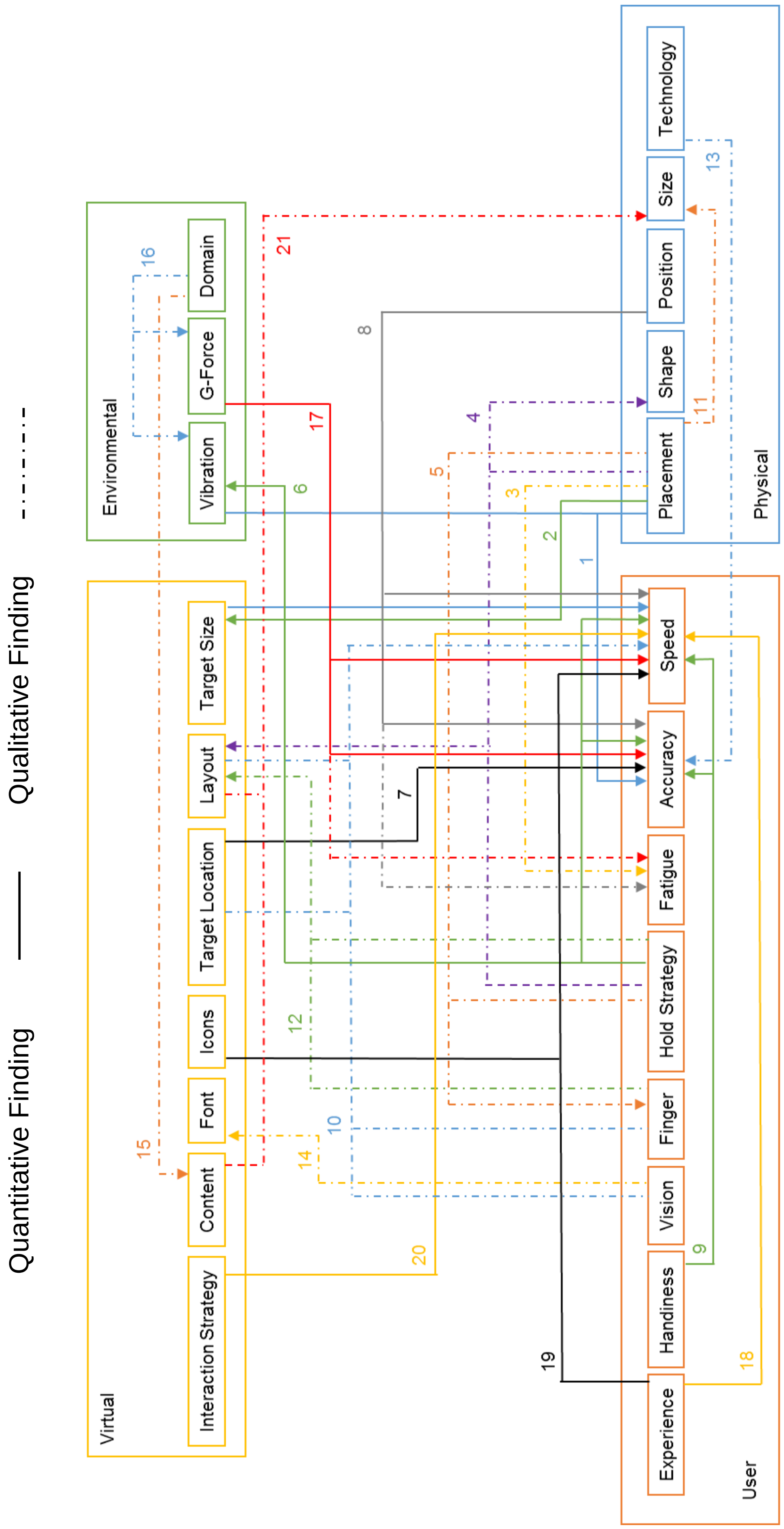


Figure 7.1 Framework of the Research for Touch Screen Integration on Flight Decks

7.1 Inflight Vibrations (Chapter 5.2)

In this study, the impact of inflight vibrations on touch screen usability was investigated. A 2x3x4 within-subjects design with repeated measures was used for the experiment. Independent variables in this experiment were device placemen, vibration and target size. For safety reasons pilots, did not participate in this study. Participants were hoist operators and rescue swimmers on board of the helicopter. On a tablet device participants performed a modified Fitts' Law Experiment. Tasks were performed with two different device placements; mobile and fixed.

Main implications for the framework are: device placement, vibration and target size have significant effects on targeting accuracy and performance (1). However, increasing target size eliminates the negative effects of placement and vibration in most cases. The findings suggest that 15 mm targets are sufficiently large for non-safety critical Electronic Flight Bag (EFB) applications. For interaction with fixed displays where pilots have to extend their arms, and for safety critical tasks it is recommended to use interactive elements of about 20 mm (2).

It was observable and it was reported by participants that conducting experiments in fixed setting was more fatiguing than performing the experiments in mobile placement (3). Participants tried to stabilize (hold) their hands while interacting with the device in fixed placement. This phenomenon was also observed by pilots interacting with the aircraft system installed on the pedestal (centre console). Fixed displays should be designed such a way that it enables pilots to stabilize their hands from all directions and interactive elements should be placed along the sides (4). In mobile placement participants held the device always in landscape mode. The majority of participants held the device with their non-dominant hand and performed the experiments with their dominant hand's index finger. In few cases participants hold the device with both hands and used their thumbs to conduct the experiments (5). Vibration measurements revealed that the human body is able to absorb a certain amount of vibration. In mobile placement participants were able to use the device inside the "zone of convenient reach [Pheasant and

Haslegrave 2005]" causing the device to vibrate similarly to their body. Results revealed that participants were significantly faster and more accurate in mobile placement (6). Participants had a higher accuracy on tapping targets displayed on the centre of the display. The error rate increased for target displayed near the edge of the screen (7).

7.2 Display Positions (Chapter 5.3)

The display position within the cockpit was identified as a potential factor that could affect touch screen usability, which was confirmed by a lab study. A 5x2x2 within-subjects design with repeated measures was used for the experiment. The primary independent variable in this study was display position, displacement in vertical and horizontal direction. Participants performed the tapping task on a 10-inch tablet attached to a tripod.

Results revealed that display position has a large impact on touch screen usability. As expected best results were achieved when the display was directly in front of participants, worst results were achieved on side position where participants used their non-preferred hand. Participants performed better and were more accurate at near display positions than far display positions. There was no significant difference found for vertical displacement. Subjective experience for general and fatigue indices were analogue to empirical results (8). There was a significant difference for experiments in performance and accuracy conducted with dominant and non-dominant hand (9). Participants mentioned that in some display positions their hand occluded the next target and they mentioned that this slowed down their movement. Placing interactive elements along the edges (except top edge) and preserving the centre of the display to display information, as suggested in the field trials, would prevent occlusions (10).

7.3 Content, Features and Functionality (Chapter 6.1)

Many air carriers have recognized the potential benefits of paperless cockpit and adopted (or are in transition phase) tablets to replace conventional flight bags. A study was conducted with the aim to explore and understand potential benefits and challenges of an Electronic Flight Bag (mobile device) in a search and rescue (SAR) environment. The primary aim of this research was to define features and functionalities of a mobile device within a flight deck environment. A review of related work, operational observations and interviews with SAR pilots were conducted to understand and specify the use of context within this particular area.

Physical expectations from a portable EFB are maximised screen real estate, while minimising overall weight. It should fit properly onto the knee and there should be room on the thigh to rest the arms. A Digital Human Modelling Software was used to determine physical constraints of the device. Results revealed that 8.5 inch tablets attached to a kneeboard would meet these requirements (11). For flight decks with dedicated mounting device it is recommended to have bigger tablets. In the field studies, it was suggested to use 20 mm targets for fixed devices, this is approximately 33% larger than recommended target size for mobile devices. This will decrease the area on the display which can be used to display information. Another request was that the device should be usable with one hand (thumb), because pilots would use the other hand to hold the control stick. The majority of pilots could reach up to 5 cm away from the display edge. Placing interactive elements within these limits would enhance supported one hand operation (12). Pilots suggested to have a kneeboard that can be tilt up to adjust viewing angle and a design that prevent heat transformation from the tablet onto the knee. Pilots mentioned that in addition to in-flight vibrations, increased G-Force might have a decremented effect on touch screen usability. To avoid accidental touches pilots suggested to use a pressure activated touch screen technology. (14).

A scenario was generated with the aim to figure out features, content, and functionality that pilots would like to see in their EFB, which was distributed to other pilots. It is predictable that each domain (military, commercial or parapublic operations) will have their own specific requirements and expectations (15). It is intended to be a future work to investigate other domains to see differences in expectations. For new applications system designers, should involve pilots from the beginning of planning and development phase. Each stage of the development should be evaluated with user studies. An example for user studies is given below in the following section.

7.4 Increased G-Force (Chapter 4.4)

In the previous study pilots stated that increased G-Force might have an impeding factor on touch screen usability. A lab study was conducted to understand the potential impact of increased G-Force on touch screen usability (fixed display position). The magnitude of in-flight vibration and alternating G-Force depends on the domain, operational conditions, weather and size/type of the aircraft (16). Primary independent variable in this lab experiment was simulated G-force. A weight adjustable wristband was used to mimic increased G-force. On a 17-inch resistive touch screen display participants performed a two-dimensional tapping task (designed after ISO 9241-9).

The key finding is that increased G-force has a large effect on performance and fatigue indices. While the simulated G-force increased linearly, performance decreased exponentially, and movement time increased exponentially. This was also reflected by subjective ratings across all conditions. Controversially the error rate was better with increasing G-force, due to the unusual condition that slowed participant's movement speed down (17). Personal fitness and experience with touch screen usage was found to be a compensating factor (18). Since the lab study did not simulate increased G-force in a realistic way it was recommended to transfer this setting to a human centrifuge where ecological valid results can be achieved.

7.5 Comparative User Study (Chapter 6.2)

A usability experiment simulated departures and approaches to airports evaluated a new developed touch interface and compared it with the current system. Three scenarios and three input methods were compared. These were the physical keypad on the FMS, the integrated virtual keypad and, the new developed drag and drop strategy on the tablet device. An 8-inch tablet was used for input via virtual keyboard and drag & drop strategy. The FMS was used for input via physical keyboard.

The interface was constructed from findings mentioned in previous sections. Interface elements which were out of scope of the research area were colour and icon (symbology) usage. Advisory circular 25-11B explain colour coding in aviation and the functional meaning related with each colour [Federal Aviation Administration (FAA) 2014]. To avoid distraction grayscale was used in a pronounced form and other used colours comply with this standard. Using symbols have potential benefits like fast recognition [Shepard 1967], reduction of the necessity to read, saving space and supporting learning of a system. To achieve these benefits symbols must be immediately recognizable to the targeted user population [Horton 1994]. Therefore, the experience of pilots plays a key role in selecting appropriate icons. Some icons were used in the interface which were selected with pilots and avionics experts (19).

Analyses of task completion time showed that touch interface is significantly faster and error proof than conventional input methods (via physical and virtual keypad). Results revealed that designing user interfaces that represent their real-word counterparts (skeuomorphism) will not improve the usability and that the design of user interface plays a key role in performance (20). Post interviews with pilots revealed that an 8-inch tablet is not sufficiently large for this task and interface. Pilots said that searching on a small area was difficult (21).

7.6 Questionnaire for Touch Screen Integration

This section will list a series of questions that designers can take into account to evaluate whether touch screen technology is a suitable input device for their system.

Does the task require pilots to focus solely on the screen? Touch screen technology requires users to look always at the screen while interacting with it. For operations conducted under instrument flight rules (IFR), this might be not an issue. Except at take-off and landing pilots are not relying on looking outside. This could raise a bigger problem for operations (e.g. SAR and military) where pilots have to look outside frequently. Generally, helicopter operations require looking outside. An analogue system is a better solution if pilots are likely to use the system while they are looking outside.

Is the magnitude of vibration/turbulence acceptable? In-flight vibration and turbulence degrade the speed of interaction and more important the accuracy. For future designs, it is recommended to explore the environment in which pilots will interact with touch screens. The type and weight of an aircraft, operation altitude, speed and weather are major factors that will determine the magnitude of movements (e.g. vibrations) within the flight deck. Preferable, evaluation experiments should be conducted under worst case (turbulent, vibrating) conditions.

Don't pilots wear gloves? The majority of commercial and general aviation pilots do not wear gloves. Other domains like military or SAR operations require pilots to wear heat resistant gloves. Current, capacitive touch screen technology should be avoided if pilots are likely to use gloves during operation. It is predictable that wearing gloves will increase errors which is asked in the following question.

Are accidental touches acceptable? Previous studies showed that the biggest drawback of using touch screens are unwanted and accidental touches. Therefore, safety critical tasks should receive a safety layer in form of a confirmation box or replaced with traditional physical switches.

Will the device be large enough for interactive elements and information? The recommended size for interactive elements for interactive displays are significantly larger than interfaces designed for mouse or trackpad usage. This will consequently decrease the space for displaying information. As a result, designers will require a larger space (display).

Will the position of the screen provide adequate ergonomics? The position of the display has a significant impact on performance and fatigue. The number and frequency of interaction will play a significant role in addressing this problem. Since the flight deck has a limited space an interface which will be used rarely can be positioned at a place which is uncomfortable to view and use.

Can pilots stabilize their hands while interacting? Pilots are likely to hold the device to stabilize their hands while interacting with the system. Another solution is to design a padding underneath the arms. Providing a design that enables hand stabilization would improve the accuracy. It would be beneficial if the touch screen technology can perform palm rejection as then pilot could stabilize their hands against the screen. This would be an advantage for larger screens where not all areas of the screen can be operated whilst stabilizing against the bezel.

Answering “Yes” to many of the questions above suggest that a touch screen interface is a suitable solution for the intended device. Answering “No” to a given question does not mean that touch screen technology is not a suitable solution. It should be considered how the associated factor might affect the device usability and safety. Potential countermeasures to mitigate degrading factors are given in the previous sections. These questions should provide avionics designers with an initial idea whether a touch screen interface is worth considering.

8 Discussion

This chapter presents a general discussion of the main contributions that emerge from this research. In previous chapters, discussion and analysis of results to individual studies was given. Therefore, this chapter will not discuss specific results or data at a detailed level. Instead, there is a synthesis of results which would lead to conclusions presented in the last chapter.

The discussion of this thesis will begin with revisiting the research problem and the concerns that motivated the research. An investigation about the applied methodology will be conducted with a main focus on the experienced benefits and challenges. A broad analysis of the main research questions (environmental, physical, virtual and user) will be performed. Related sub research questions will be used to address the main four research questions. Each section will include discussion about the primary (identifying potential benefits and challenges of touch screen technology on flight decks) and secondary (design implications for touch screen interfaces) contributions. Thereafter, the results from this research would be discussed in relation to existing knowledge in the field. Limitations of each study, particularly those that restrict the generalisability of the results will be presented. Generalisable results will be examined, as well. Finally, there will be ideas of opportunities for future work.

8.1 Revisiting Problem Definition and Motivation

Interviews with avionics experts and a review of statements of avionics company representatives regarding touch screen integration on flight decks revealed that leading avionics manufacturer want to integrate touch screens because they think that touch screens offer a better user experience/performance than current input devices. However, the HCI community demonstrated that potential benefits, which are stated by the manufacturer, can only be achieved if designers understand the flight deck environment and develop design solutions that supports touch screen usability.

At the beginning of the project there were only few research that studied touch screens on flight decks. The scope of these research were limited and nobody made a broad approach to identify and understand the relation of various factors that could affect touch screen usability. Research that were conducted in similar environments (e.g. vehicles) showed that this area has many open research questions and opportunities for explorations. Therefore, “Exploratory Design” which is a particular Mixed Methods Approach was adopted.

8.2 Applied Methodology

One of the biggest drawbacks of applying “Exploratory Design” is that the sequential process requires considerable time to implement. However, the approach of collecting qualitative data, and then quantitative data is a logical and intuitive approach [Onwuegbuzie and Leech 2006]. This is especially true for research areas where important variables and relationships are unknown. Findings from qualitative research have been validated through quantitative research which provided a better understanding of the topic. All findings, mentioned in the previous chapters, could not be achieved with only quantitative or qualitative methods alone. Experienced benefits and drawbacks are coincident with the literature. Researchers who are working on projects in the size of this work and have similar conditions at the beginning could apply this research strategy. In the following sections experienced advantages and challenges of qualitative and quantitative methods will be presented.

8.2.1 Qualitative Methods

The initial qualitative research was done with semi-structured interviews with avionics experts and pilots. It was possible to ask for clarification and to add questions which enabled the investigator going deeper into the topic and to receive valuable information. Interviewees shared their ideas, expectations and insight views. Since these interviews were done with multiple participants, more information was gathered from discussions between participants. Such information could not be captured in a survey. Analysing open-ended questions, and discussions made the interviews the most challenging part to analyse. It was even harder and more time consuming than the field study where the investigator had a limited control over the experiment.

Observations were conducted during the field study in a natural environment to see how crew members are using mobile and fixed devices during the operation and to understand the process of operations. This was an essential task to understand the way how crew members operating the aircraft system. Observations were used to predict the way how pilots would use touch screen interfaces in the future. During the first set of trainings flights it was difficult to follow the operations. It was easier to follow the structure of SAR operations after a few flights and post flight interviews with pilots.

International Standard Organisation (ISO) questionnaires dealing with general and fatigue indices supported the understanding and interpretation of quantitative data in lab-based studies. Especially, questionnaires that were generated with participant statements and distributed to participants once the empirical work was finished provided a more comprehensive understanding of the overall outcome. On the other hand, the EFB scenario was initially a questionnaire that was distributed to pilots. A low response rate in this type of data collection is known problem. Therefore, the method was altered and data collection was conducted with semi structured interviews.

8.2.2 Quantitative Methods

Pilot studies played a key role in evaluating experimental settings. Problem areas that were identified saved significant time. Problems in lab experiments may cause a moderate setback. However, in the field studies (e.g. in-flight and human-centrifuge study) we had limited access and time, so an issue in experimental design could have caused a significant problem. Another advantage was understanding potential benefits and challenges of a setting in a real-world usage. At the beginning the in-air interaction solution (Chapter 10) seemed to be a good countermeasure for the effects of display position. However, the initial results of the pilot study showed that there will be more problems than benefits. So, the decision was not to conduct the main study which saved a lot of time and effort.

Participants might behave differently in a lab experiments due to the fact of being observed and in a different environment. Being observed can cause participants to make short-term improvements which would not be the case in a real-world situation [Landsberger 1958]. Therefore, results achieved in a field study have a higher ecological validity. The biggest limitation is that the investigator has less control over the experiment, which makes it difficult for another researcher to replicate the study.

Lab based experiments have the advantage of conducting the experiment in a controlled environment. Compared to field trials the investigator has the freedom to decide where and when the experiment will be conducted. Since a standardized procedure is used it is easier for another researcher to replicate a laboratory experiment. As mentioned before, the majority of touch screen evaluation experiments is conducted in a lab environment. Therefore, it is easier to compare the results with other studies and to position the work in the literature.

The findings from all the research conducted within this research project and other relevant studies were used to create the interface which was used in the comparative user study. Creating a prototype of the intended interface is a cost and time effective way to evaluate high level design choices. It is possible to optimise the design through fast design cycles. In the experiment touch screen technology proofed to have the potential to be a good input device, if certain aspects are considered in the design process. The user study showed that touch screen interface (even if it had room for improvement) compared to conventional input methods is a better solution for frequency manipulation tasks.

8.3 Environmental

Pilots are operating in a non-stationary environment. Various factors were stated by avionics experts and pilots that can cause movements within the aircraft. These were; domain, in-flight vibration and G-force. These factors formed the group “environmental factors” in the framework. Two sub-research questions were used to address the main research question about environmental factors. These were;

> ***Main RQ:*** *What are the environmental factors which can cause movements in the flight deck and how much will these factors affect touch screen usability?*
>
> ***Sub RQ:*** *What is the impact of in-flight vibrations on usability?*
>
> ***Sub RQ:*** *What is the impact of +Gz on touch screen usability?*

8.3.1 In-flight Vibration

This section will address the following sub research question: *What is the impact of in-flight vibrations on usability?*

The main finding of the field trials was that in-flight vibrations have a significant impact on touch screen usability. Degrading effect on touch screen performance in non-stationary environments were also detected in other studies; walking [Conradi et al. 2015] motion platform [Lin et al. 2010], tractor [Baldus and Patterson 2008], car simulator [H. Kim et al. 2014], car [Ahmad et al. 2015] and flight simulator [Dodd et al. 2014].

Average Throughput values on the ground were approximately 18% higher than the average values generated in the air. Error rate were approximately 3 times higher in the air than results achieved on the ground. The obvious reason for this difference are the vibrations during the flight, which were found to have a significant effect. The mean Throughput during hover and cruise were similar. There was a small reduction (3%) in Throughput during transition phases. The amount of transitions phase is around 5% of the entire training flight.

Further, the demand on the participants’ attention is substantial whilst in the air. During the flight, performing the experiment had a secondary order. For example, participants had to listen and communicate with voice and hand gestures, and look out for target. They frequently also had to hand over the tablet to their fellow crew member to concentrate on a task, or to take a break due to fatigue. In addition to in-flight vibrations, these types of activities increased the movement time between targets and consequently reduced the Throughput. Divided attention was investigated by several researchers (e.g. [Schildbach and Rukzio 2010]. [Bergstrom-Lehtovirta et al. 2011], [Conradi et al. 2015], [Hayes et. al 2014], [Mizobuchi et. al 2005] and [Lin et. al 2007]). All studies revealed a negative impact of divided attention situation on touch screen usability. Therefore, we can say that this variable has a confounding effect on vibration results. On the other hand, the current data set can be considered as a more ecological valid data.

There were various limitations in the field trial. The major limitation in the field study was that pilots could not participate in the experiment. Crew members who performed the experiments were not strapped to the seat all the time and had compared to pilots more space. Rescue swimmer’s tasks is completely different (except looking out for targets) to pilots and these require a higher physical effort. In addition to fatigue symptoms mentioned in the field study, the fatigue caused by the simulated rescue mission may have impacted the results. Another factor worth to mention is the weather. All flights were performed between May and June 2015. In all flights, there were no clouds below 5000 feet and the visibility were at least 10 kilometres. There were no thunderstorms which could increase the vibrations/turbulences felt by the participants. Challenging weather conditions are likely in the winter months.

8.3.2 Domain

The amount of movements depends on the domain. In comparison to commercial aircraft, general aviation aircraft and helicopters are smaller, lighter and operating at lower altitudes. A commercial pilot who flies a modern passenger aircraft at an altitude of 40000 ft feels less movements in the cockpit than a SAR pilot who operates a helicopter at sea level. This was the starting point of the research, where the hypothesis was that results achieved in a commercial aircraft setting is not transferrable to other domains.

Dodd et. al [2014] conducted a simulator study focusing on commercial jets. Similar to our study a baseline (without vibrations) determination was conducted. The reduction in accuracy compared to our study is significantly lower. The increase in error rates with increasing vibration was also visible during the field study. Vibration were significantly higher in transition phases than during cruise/hover. Statistical results revealed a significant difference between these two conditions. The task designs were different therefore, the speed of interaction is not comparable. During the field study we applied a modified Fitts' Law experiment. Dodd used a data entry task. From both studies and other relevant studies, we can see that there is an increase in standard deviation for interaction speed recorded under vibration. Error Rate analysis suggests that results achieved in a domain are not easily transferrable to other domains.

The HCI literature showed already that using touch screen devices in non-stationary environments results in higher error rates. Therefore, this significant difference was expected. In-flight vibrations have a larger effect on accuracy than interaction speed. The more important finding gathered from this research is that the magnitude of vibration influences the amount of error rates.

Average user performance (Throughput) for touch screen during the flight is4.6 bps. Soukoreff and MacKenzie [2004] reviewed studies that applied ISO 9241-9 standard. Throughput values for the mouse ranged from 3.7 bps to 4.9 bps. The field trial described may be considered as a semi-controlled field experiment. Keeping in mind that the task design applied during the field study required additional search time for the next target, what our findings show is that touch input even in the air is better (in terms of interaction speed) for pointing tasks than a mouse in an office environment.

The primary contribution of this work is: the in-flight vibration has a significant effect on touch screen usability (interaction speed, error rates and fatigue). As shown in previous work the size of interactive elements, can be utilized to minimize this effect. The secondary contribution that should minimise this degrading effect will be discussed in the virtual factors section under the heading target size.

8.3.3 Task Design for Touch Enabled Devices

In this place it is worth to say that this data in this form could not be collected without the new developed Fitts' Law Experiment. A series of pilot studies were undertaken in a lab setting prior to moving to more open-ended field trials in a real-world setting. Pilots studies demonstrated that the tapping task design as described in ISO 9241-9 is not suitable for devices with multi-touch capability. Participants tended to hover their finger over the next target before clicking the current target with the other hand. This kind of predictability would lead to contrived movement time measurements compared to realistic operational use. This could cause a problem because one of our objectives were to observe how potential users are going to use the device in a real-world situation.

Therefore, a task design was created in which the size and the distance of each target varied dynamically from the previous one. An advantage of applying this task design was that it was possible to record results from a large ID range (1.2 – 6.2), which would be not possible if following the ISO standards that recommend targets appear around a circle. In this case the width of the device is the limiting factor. For tested target sizes (5 mm – 20 mm) the maximum achievable ID value on a 7.9" tablet would be 4.5.

The main contribution of this modified task design is that it enables researchers to observe how potential users would use touch screen devices in particular environment. In addition, it also shows that the interface design will influence how users would hold and interact with touch screen display.

8.3.4 +Gz

This section will address the following sub research question: *What is the impact of +Gz on usability?*

Empirical and subjective results of the +Gz study, largely confirmed the hypotheses of pilots stated that increased. Throughput results showed a reduction in mean values with increased +Gz. The trend indicated an exponential fall in TP values. Rest time to recover from fatigue were not reflected in the TP values. Therefore, it was important to consider the movement time analyses. Analysing movement time and the overall time needed to complete a condition provided a more comprehensive view of the potential impact of +Gz on touch screen usability. Fitts' Law Prediction Models all yielded high R^2 values showing that this methodology is valid for this research area.

Comparing movement time results with the latency time results from La Pape and Vatrapu [2009] shows that placement of the device (fixed or mobile) plays a significant role in overall performance. A similar finding was also achieved in the previous study investigating the effects of in-flight vibrations. Average latency results from La Pape and Vatrapu showed also an exponential increase with linear increase in +Gz. This suggests that the experimental setting mimics increased +Gz with a weight adjustable wristband in a way that ecological validity is achieved to some extent. This study also investigated negative Gz (-1-Gz and -2-Gz). -1-Gz condition showed an increase and -2-Gz showed a decrease in latency time compared to +1-Gz. Authors did not discuss the potential reason why participants were faster in pointing the target in -2-Gz condition. A possible explanation could be carry on and learning effects because -2-Gz condition was always the last condition in the sequence.

The increase in accuracy with increasing simulated +Gz, was the only unanticipated result of the study. It was assumed that participants would not decelerate properly and overshoot targets due to the additional weight on their wrist, which was in fact the case. It was observable that participants who made a movement from the top of the screen towards the bottom overshoot targets and had to adjust. However, participants were able to increase their accuracy, due to the unusual condition that slowed their movement speed down. This can be explained with speed-accuracy trade off stated by Soukoreff and MacKenzie [2004], which basically says a reduction in interaction speed would increase the accuracy. The increase in accuracy compensated for differences in TP values, which were smaller compared to the mean movement time. Error rates of 20 mm target were approximately three times lower than for 15 mm targets, which suggest to use 20 mm targets on fixed displays on the flight deck.

The primary contribution of this work was that the device placement has an additional negative effect on +Gz factor. The secondary contribution recommends to provide hand and arm support for stabilisation and support. This should mitigate the detrimental effects of fatigue and error rates. This can be considered as a generalisable recommendation for all type of operations and aircraft. How the shape of displays should look like will be discussed in the next section under the heading “Shape”

8.4 Physical Factors

Several physical factors were frequently stated during the initial interviews. Investigating these variables revealed further variables that might affect touch screen usability on the flight deck. Following factors were identified and investigated during this research: placement, shape, position, size and technology. These factors formed the group “physical factors” in the frame work. Five sub research questions were used to address the main research questions about physical factors. These were;

> ***Main RQ:*** *What physical/hardware factors are existing that can influence touch screen usability on a flight deck situation?*
>
> ***Sub RQ:*** Is there a difference in performance for device placement?
>
> ***Sub RQ:*** *Is there a difference in usability for different display positions?*

Sub RQ: *Is there a difference for display displacement in vertical and horizontal direction?*

Sub RQ: *How should be the physical shape of the (fixed) display, so it supports usability?*

Sub RQ: *What are physical expectations from a mobile device?*

8.4.1 Placement (Mobile and Fixed)

This section will address the following sub research question: *Is there a difference in performance for device placement?*

There are two types of displays envisioned in future flight deck concepts; mobile and fixed. The position of the display in mobile placement is similar for all users. However, there are various opportunities on the flight deck to install a touch screen display. This section will focus solely on mobile and fixed placements. Schedlbauer [2007], Tsang [2013], Colle and Hiszem [2004] and Parhi and Karlson [2006] performed keypad input experiments in different display placements. It was noticeable that studies conducted in fixed placement had a higher error rate compared to experiment conducted in mobile placement. This motivated us to investigate this variable in a flight deck situation. This factor was investigated during the field study, which produced one of the primary contributions.

The in-flight study confirms that without support this increases the likelihood to make more errors in a vibrating environment in fixed placement. The effects of holding a device in the hand were significantly different to attaching the device, on ground as well as in the air. Error rates in the fixed placement condition were approximately 33 % higher than in the mobile placement condition. The difference in Throughput was approximately 6% which was statistically not significant. The difference in error rates may be explained by increased fatigue during the fixed placement condition where participant had to extent their arms to reach the screen, and by the bodily absorption of vibration when holding the device (mobile placement condition).

New cockpit designs have fixed as well as mobile touch screens integrated. Pilots have to extend their arms towards the dashboard to interact with the aircraft systems. The in-flight study confirms that without support this increases the likelihood to make errors in a vibrating environment. In the mobile setting the user was able to pull the device inside his “zone of convenient reach” [Pheasant and Haslegrave 2005], causing the device to vibrate similarly to the human body, ‘absorbing’ a certain amount of vibration, which is not the case in the fixed condition. Results confirmed the hypotheses that participant were likely to make more errors in the fixed condition than in the mobile condition. This variable showed the importance of designs where pilots can stabilise and rest their hand and arms.

8.4.2 Position (Fixed Display)

This section will address the following sub research question: *Is there a difference in usability for different display positions? and Is there a difference for display displacement in vertical and horizontal direction?*

Due to experimental design and other limitations it was not possible to analyse this variable during the field trials. Therefore, a separate lab study was conducted, ehich revealed that display position has a large effect on touch screen usability. This was the first study that evaluated the potential impact of various display positions on usability, using a Fitts Law design.

In everyday stationary screen usage, such as when using ATM’s or public terminals people can adjust their position relative to the screen. The only research that evaluated the effect of sitting orientation on touch screen performance was conducted by Chourasia et al. [2013]; however, only two device positions were tested, and the study did not follow a Fitts’ Law design. Depending on the physical design of the ATM or terminal, wheelchair users have to position themselves parallel to the screen. They found a decrement of 36-48%. Future flight deck environment is another domain, in which screen position has a potential impact on touch screen performance. As mentioned in Chapter 5.3, Gulfstream makes frequent use of touch screens in their Symmetry Flight Deck. This design incorporates 10 touch screens (2 overhead, 4 head down, 2 EFB and 2 on the pedestal), which are be operated by two pilots. Keeping in mind that pilots are usually strapped to the seat, the freedom of movement is limited. In this experiment the average decrease between the worst position (non-dominant hand side; Position A) and the best (in front: Position C) is 26.6%. Best results were achieved when the screen was directly in front of participants (Position C).

Displacement in both vertical and horizontal direction were tested. Results showed that Throughput for the near placement was significantly better than for the far placements. The far position was 60 cm from the sitting position, approximately the same distance that pilots are from their control panels in the AW-139 Helicopter, which was used during the field trials. Results suggest that the Throughput of pilots should be significantly higher if displays were closer. However, it should not be assumed that getting display position as close as possible to the body of users would automatically produce higher Throughput results. Throughput results may get better only up to a point at which the performance is likely to diminish; however, where this point lies were not subject of this study and may need to be investigated in future work.

There was no significant difference for horizontal displacement (low vs. high). The height of the low position is approximately similar to the pedestal on AW-139. The increase in height did not lead to a significant difference in throughput. The reason for this can be explained by the relatively small displacement. The difference was only 10 cm between the two levels. During experiments, 7 out of 10 participants mentioned that conducting the experiment in the higher position is more fatiguing. It could be the case that if the difference between the two levels were larger, fatigue effects may play a role and have a significant effect on Throughput values. In the lab study, error rates results were analogue to Throughput results. There was a significant reduction in error rates for near display position over far display position and there was no significant difference in error rates for low and high display positions.

The primary contribution of the lab study was that the display position has a large effect on touch screen usability. However, this variable can be used to minimise the detrimental effects of display position. The results of the lab study can be used to optimise the display position within the flight deck for touch interaction.

8.4.3 Shape

This section will address the following sub research question: *How should be the physical shape of the (fixed) display, so it supports usability? and What are physical expectations from a mobile device?*

During the in-flight study interactions in the fixed placement condition was performed with one hand. Participants always used their preferred hand. They were encouraged to take a break when feeling fatigue in their arms. Eight out of 14 participants were observed to tend to hold on to the device from the side or above. To avoid bias participants were asked not to hold on to the device. However, the observation suggests that people tend to hold on to the screen to stabilize their hands. Video recordings revealed that pilots stabilize their hand while interacting with aircraft system. This could be factored in when designing the hardware as well as the user interface. For example, the display could be designed in such a way that it enables pilots to stabilize their hands from all directions (from behind included) and interactive elements should be placed along the sides. This will enable interaction with aircraft system while maintaining hand stabilisation.

For mobile devices without any dedicated mounting device pilots pointed two important factors; weight and screen area. Basically, the screen area should be maximised while the overall weight is minimised. Usually, if the screen area increases the overall weight increases, as well. So, an acceptable trade-off between screen size on weight need to be found. Additional information about how pilots are using mobile devices currently was used to simulate and define appropriate EFB devices. Pilots, using a mobile device, stated that a mobile device fit properly onto the knee, while there should be room on the thigh to rest the arms

Simulations were performed so that the recommendations applied (can be used by) to the majority of pilots. Results revealed that mobile device between 8.5 and 9 inch provided the best results for these expectations.

For applications in other domains designers should determine whether the mobile device would have a dedicated mounting device in the flight deck or not. This will influence the size of the mobile device. In the next chapter target size results will be discussed and will reveal that fixed display should have larger target size. This will consequently increase the minimum display size. The recommendation that displays should be designed in such a way that it enables hand stabilisation is a generalisable recommendation.

8.5 Virtual Factors

A significant part of the initial interviews was focused on the interface design. Variables investigated in this research formed this group in the framework. Following factors were identified and investigated during this research: target size, layout, target location, icons, fonts, content and interaction strategy. Five sub-research questions were used to address the main research questions about virtual factors. These were;

> ***Main RQ:*** *How should be the interface design so it is ultimately usable by pilots in a flight deck environment?*
>
> ***Sub RQ:*** *What is an appropriate size for interactive elements on a touch screen installed on a flight deck?*
>
> ***Sub RQ:*** *Which areas on the display have an increased error rate?*
>
> ***Sub RQ:*** *What are interface design guidelines for one handed thumb operation?*
>
> ***Sub RQ:*** *What features, functionality and content are pilots expecting from a mobile device?*
>
> ***Sub RQ:*** *Which input method provides the best and safest interaction method for radio frequency changes?*

8.5.1 Target Size

One of the main independent variables that was investigated during the field study was the target size. This section will address the following sub research question: *How should be the interface design so it is ultimately usable by pilots in a flight deck environment?*

Independent variables were tested systematically, starting broadly at the top level and gradually going into more detail. In the first set of analysis, significant difference for all variables were found. While target sizes between 15 mm and 20 mm were not significantly different, detailed analyses showed that there are few cases where significant difference between 15 and 20 mm exist. In the second level of analysis, interaction effects between independent variables were examined, which showed that two of three possible combinations have significant interaction effects. The final level of analysis considered each possible case (24) separately and in pairwise comparisons. The provided matrix shows that the effects of placement and vibration disappear with increasing target size.

Target sizes beyond 20 mm were not tested; however, helicopters are able to absorb higher vibrations. Keeping previous works in mind it is unlikely that targets bigger than 20 mm would lead to significant improvement. Therefore, it is recommended to use 20 mm targets for fixed devices for which pilots have to extend their arms to reach, and for safety critical tasks. In the worst case, the expected error rate for 20 mm targets during the transition phase (strongest vibrations) with a fixed placement is 3 %. These results were presented to avionics experts during a conference. An engineer said that not using the interface during transition to hover would probably be acceptable to most users. This was also observed during the training flights and were confirmed with video recordings. Pairwise comparison revealed that errors caused in fixed placement during transition phases produce a significant difference between 20 mm targets and 15 mm targets. For such applications where it is acceptable to not use the interface during transition phases, it is recommended to use 15 mm targets.

Airlines are increasingly interested in the integration of portable touch screen devices into the cockpit. In 2011, the FAA has authorized the use of the Apple iPad as EFB [Murphy 2011]. Currently, many Airlines are in the transition phase to a paperless cockpit. American Airlines (AA) was the first major commercial carrier that completed their EFB program. The software, used by AA, has the following features [Pschierer et al. 2012]: Enroute charts and airport diagrams (displays own-ship position), arrival, departure and approach procedures, change notifications (terminal and enroute).

Mobile devices are (currently) not used for safety critical task. Thus, 15 mm targets for mobile devices may be sufficiently large for non-safety critical tasks, such as in an EFB. The expected error rate for 15 mm targets during transition (strongest vibration) when the device is held rather than fixed is 3%. During cruise and hover which covers the majority of the flight 10 mm targets would produce 7-8% error which might be acceptable for such applications.

In Chapter 2.3.1, recommendation and design guidelines from mobile device suppliers (Apple [2014], Microsoft [2014] Google [2014]) were presented. These recommendation are acceptable for daily usage however in a safety critical environment a higher accuracy is required. There recommendations from Ubuntu [2008] the American National Standard Institute / Human Factors and Ergonomics society ANSI/HFES 100– [2007] are more suitable for this application area.

8.5.2 Target Location

Another sub research question was: *Which areas on the display have an increased error rate?*

In the field study, targets appeared on a 8 x 10 grid, which enabled further investigation on error rate for specific regions. The results were consistent with previous findings mentioned in Henze et al. [2011], Park and Han [2010] and Avrahami [2015]. In the mobile setting, participants had a higher accuracy on the centre of the screen. The error rate gets higher towards the edge of the screen. The error rate at corners for both placements were higher compared to the average error rate. The findings of this work were consistent with the literature. An inspection of Barstow's [2012] summary of widely used EFB applications largely shows that the interface designs make use of the centre of the screen to display information (e.g., charts or checklists), and the edges are designated for interactive elements. Due to the likelihood of occlusion, the top of the screen is not recommended to place interactive elements. However, it is still recommended to place interactive elements along the edges (left and right). This will enable hand stabilisation while holding the device. Confirmation boxes can appear on the centre of the screen.

8.5.3 Layout (One Handed Operation)

This section will address the following sub research question: *What are interface design guidelines for one handed thumb operation?*

For several tasks during the flight requires the interface to be usable with one hand. From video recordings, it was noticeable that pilots support their hand by grasping the device (fixed displays) and using their index finger or thumb to interact with the screen. The tendency of holding the device was observed in both studies. Interviews with pilots revealed information that was used to determine the physical constraints and user interface layout that meets the pilot's operational requirements. For one hand operation frequently used interactive elements like keypad and switch buttons should be placed alongside the edges. It is recommended to place interactive areas within the recommend area, as shown on Figure 6.5. The majority of pilots could reach interactive elements up to 5 cm away from the display edge.

This should be factored in when designing the hardware as well as interface. For example, the display should be designed in such a way that it enables pilots to stabilize their hands from all directions (from behind included). Pilots identified increased G-force as a potential threat for touch screen usability. The last empirical study, described in Chapter 5.4 and a field study [Le Pape and Vatrapu 2009] revealed that +Gz has a large impact on touch screen usability which increase the importance of design that enables hand stabilisation while interacting with the display.

As mentioned in the literature review, an acceptable error rate for this application area has not been established. However, it is expected that authorities will establish guidance for acceptable error rates for different tasks (safety critical and non-safety critical tasks). This research seeks to inform such decision-making. If designers require a higher accuracy, it is not recommended to increase the target size beyond the recommended values. Instead, adding an additional safety layer with message box saying: “Do you want to proceed?” would make the interface more error proof.

To give another example, “shutting down engines” may be classified as a safety critical task, accidental shutting down must be avoided. The interaction may be designed to minimize the error probability in the following way. To shut the engines off, the pilot would need to navigate to a menu item, select and touch the ‘off’ button, upon which the system would prompt the pilot to confirm if they want to shut down the engines. In total, the pilot would have to take three steps within the system to shut down the engine. If we assume all interactive elements have the recommended size (transition), the error rate is at worst 3% per layer. Adding three layers will reduce the probability of shutting down the engines by accident to 0.0027% (0.03x0.03x0.03=0.000027). However, alternatively, certain safety-critical actions may only be supported by traditional physical switches.

8.5.4 Content

This section will address the following sub-research question*: What features, functionality and content are pilots expecting from a mobile device?*

Expected features, functionality and content of EFBs were defined with interviews and surveys. A scenario created from the interviews was distributed to other pilots. The outcome of both approaches was coherent. Pilots want to have a tablet application where they can access to all required documents (e.g. checklist or maps), perform calculations, fill reports, create/manipulate/upload flight plans. Pilots were against controlling any kind of avionics system through the app. Automation like the auto check function in the checklist was found to be not suitable. Pilots thought that such automation would take them out of loop and said that self-checking is better.

Pilots would appreciate to have basic functions like chelating, carry maps and other documents, filling reports which does not require communication with the aircraft system and consequently no certification as soon as possible to use the benefits of an EFB.

Requested features were compared with requirements from other domains. The primary contribution of this work is that each domain has its own specific requirements and the results achieved in this study are not transferable onto another domain.

8.5.5 Interaction Strategy

This section will address the following sub research question: *Which input method provides the best and safest interaction method for radio frequency changes?*

Keyboard studies (e.g. [Kim et al. 2012], Sears et. al [1993] and [J. H. Kim et al. 2014]) comparing physical and virtual (touch) keyboards) showed that user interfaces representing their real-word counterparts (skeuomorphism) will worsen the usability (speed and accuracy). This indicated that the interaction design of the user interface should be optimised for touch interaction. In the user study an interface was created which was optimised for touch interaction.

Hong et. al [2011] recommend not to perform drag operations in a moving military vehicle. A new way of interaction was proposed in the last experiment where pilots could manipulate frequencies by dragging and dropping targets over a “Hot Corner”. This design revealed that drag operations are acceptable if there is less precision required. While the results on throughput are encouraging for in-flight use of touch screens, further in-air investigation is required for interaction methods like drag and drop, pinch to zoom or swipe operations.

There were only 2 (out of 30) task sequences where the input with FMS was faster than touch (drag drop) interaction. These measurements were taken at task 1 where some movements to the left were required to get VOR SBJ. Comparison of physical and virtual keypad showed no significant difference. However, Lee and Zhai [2009] found that input via virtual keypad is significantly faster than its physical counterpart. A reason for that could be the experience of using the FMS on a daily basis and the virtual keypad was used for the very first time in this setting. Results revealed that designing user interfaces that represent their real-word counterparts will not affect the completion time significantly. Advantage of skeuomorphism is that users understand the purpose of the system immediately and there is no additional training required. However, considered in the long term, such novel designs as shown in this study are more efficient in terms of completion time and error robustness.

The New York airspace is one of the densest airspaces in the world. Consequently, there were interactive areas that overlapped. This caused the following problems; pilots could not detect immediately where they have to put their finger first or they dropped the wrong interactive area over the “Hot Corner”. This would likely be less a problem in areas not as densely covered by airports and navigation aids. Pilots suggestions to perform the entire interaction by clicking interactive areas is integrated in the new design. This has the advantage that pilots will immediately spot the interactive element and click it, which will produce consequently its disadvantage by adding more clutter onto the map. Another requested feature is displaying traffic information. A study [Endsley et al. 1999] found that pilots’ traffic situation awareness improved when traffic information is displayed on the map.

Pilots opinion that the hardest part to localize the target interactive element and to point it was coherent with the investigators observation.

After the familiarization session pilots swiped the interactive element over (sometimes slide over the edge) the ”Hot Corner” without paying attention to its location. This interaction method seemed intuitive and fluent.

In addition, the size of interactive elements was 8 mm, which is optimal for usage in a static environment but not for dynamic environments. Making the size of icons bigger could cover important information. So, using this strategy has the trade-off between acceptable error rate/speed of interaction.

Two pilots performed the experiments at the same time. The majority of pilots were right handed. Pilots sitting on the right-hand side had to use their non-dominant hand to make inputs via FMS. In touch interaction participants always used their preferred hand. Results from the lab study that explored the potential impact of display position on usability revealed that handedness plays a significant role in touch screen performance. This could be another factor that increased the difference between the input methods. The primary contribution of this work is that interfaces representing their real-world counterpart will not improve the usability of touch screen devices.

8.6 User Factors

Avionics experts were largely concerned about which variables (environmental, physical and virtual) could affect the usability (user - speed, accuracy and fatigue). In addition, there were also other factors that influenced variables in the framework. Following factors formed the user factors in the framework: Hold Strategy, handedness, experience, vision and finger. Four sub-research questions were used to address the main research questions about user factors. These were;

Main RQ: *What are the personal factors between users that can cause a difference in performance?*

Sub RQ*: Does handedness effect the usability?*

Sub RQ: *Can experience and fitness influence overall performance?*

Sub RQ: *How will pilots use mobile devices on the flight deck?*

***Sub RQ**: How are fatigue symptoms affected with +Gz?*

8.6.1 Handedness

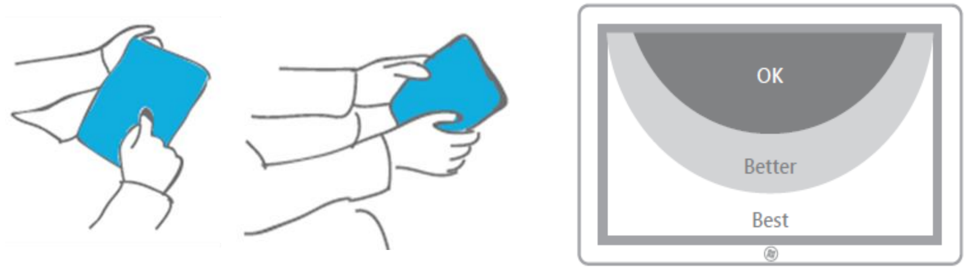
Trudeau et al. [2016], Perry and Hourcade [2008] and Kim and Jo [2015] focused on grip and how user use touch screen in different conditions. There is no existing study that investigated the effects of handedness and finger use in a flight deck environment. One of the objectives during the lab study investigating the impact of display position was to examine the effect of handedness on touch screen usability.

The main finding was coherent with the literature. Results revealed that throughput values dropped by moving the screen towards to the side of dominant hands. Conducting experiments with the non-dominant hand produced significantly low Throughput results. Participants made on average 25% less errors with their dominant hand compared to their non-dominant hand. Participants made less than half the amount of errors with 14 mm targets (75px) compared to 9 mm (50px) targets. Ina flight deck environment it is likely that a pilot would use his non-dominant hand to interact with the aircraft system. A generalisable recommendation for cockpit designers is to create interfaces to be usable with non-dominant hand.

The best Throughput values were achieved by participant number 8. In his pre-experiment questionnaire, he mentioned that he is able to use both hands well and that he uses touch enabled devices several times a day. Participant number 7 produced the lowest average Throughput across all blocks. This participant stated that he only had limited touch screen experience. He also mentioned he rarely uses his smartphone, he does not play games which require fast and precise interaction, and he does not use any other touch-enabled devices. The average drop in Throughput between the best position and the worst position is between 2.5-3.5. These findings suggest that experience may have a non-negligible effect on Throughput.

8.6.2 Hold strategy

In the mobile placement condition, six participants initially used both of their hands to hold the device, and used their thumb to tap the task (see Figure 8.1b). Eight participants held the device with their non-dominant hand and performed the experiments with their preferred hand’s index finger (see Figure 8.1a). In two cases, participants switched from two-handed thumb to one handed index finger grip. We could say that the majority of users would use a mobile device in landscape mode. However, it is recommended to apply adaptive views to user interfaces.



a) b) c)

Figure 8.1 Tablet Holding Strategies used in the Experiment & Recommended Interactions Areas for Two Hands Holding, Thumbs Interaction [Microsoft 2014]

It was observed that participants who used both hands had difficulties touching the target at the centre of the tablet. Participants had to readjust their grip frequently. This is a known drawback of this holding strategy. Figure 8.1c shows recommended interaction areas for two-handed holding. Holding the device with the non-dominant hand and using the dominant hand’s index finger has the advantage that users can reach any location of the screen without readjusting the grip.

However, there is the risk of occlusion. Participants pointed out that sometimes the next target was covered by their hands. This was also a factor that was mentioned in the next lab study investigating the impact of various display positions. Another point that might lead participants to use their index finger and hold the device with their non-dominant hand is the fact that the width of the thumb is usually wider than the index finger which can cause a significant difference in accuracy [MacKenzie 2015].

This was also found by Kim and Jo [2015] that compared finger and thumb input.

Post experiment interviews revealed that participants prefer to use the tablet device in the mobile condition. In contrast, the fixed placement was described as more fatiguing. Hong et al. [2011] also found that participants preferred to use a handheld device which they can hold in their hands. In the context of a vibrating environment such as a helicopter cockpit, it is also worth pointing out that by holding the device, the human body is able to absorb vibrations (as shown in Chapter 5.2.6.1), thereby mitigating for the detrimental effects of vibration on performance, error rates, and throughput.

8.6.3 Experience

Another limitation worth mentioning are the physical conditions of participants. Pilots flying a fast jet aircraft have to pass medical tests and need to be in a good physical condition. Physical fitness might be a compensating factor that could reduce the effect of +Gz by a certain amount. Previous lab study investigating the potential impact of display position on touch screen usability revealed that personal experience played a significant role in performance rates. Aside from these limitations this experiment provides evidence that +Gz is a potential impeding factor on touch screen usability. It is recommended to transfer this setting to a human centrifuge, where the effect of +Gz can be studied in a more realistic way.

The main question is about whether touch displays are suitable for such challenging environments? This study is part of a research project that investigates potential benefits and challenges of touch screens on flight decks. The framework showed that there are many factors (e.g. inflight vibration, location of the display, interface design and interaction strategy) that affect performance. Overall, all impeding factors should be considered before making a decision whether touch screen technology is a suitable interface for the desired aircraft system. However, based on current findings, we can say that there is a break-even point between 2-

Gz and 3-Gz; below this point pilots can benefit from touch screen technology. Towards 3-Gz and beyond it will be a challenging task to interact with fixed displays. Therefore, for tasks that are likely to be beyond this point, it is recommended to use hard controls which are in close proximity (on control stick or throttle) to pilots.

8.6.4 Fatigue

Participants subjective ratings supported the overall view. Some participants who performed 3-Gz condition before others changed their ratings after the 1-Gz and the 2-Gz conditions were completed. The reason for this was to highlight the effect of +Gz to fatigue indices. All participants agreed that compared to the 1-Gz condition the inconvenience in the 2-Gz condition in their arm, shoulder and neck was moderate. However, the 3-Gz condition had a strong effect to these indices compared to the other two conditions. Figure 8.2 shows a participant who conducted the experiment in 3-Gz condition. Their discomfort was visible in that participant tried to counterbalance the effect of the weight adjustable wristband by leaning to the left. During post-experiment interviews participants said that the 3-Gz condition was painful, and estimated a simulated 4-G condition as their limit where they could finish a sequence (13 taps) before they have to rest their arms.

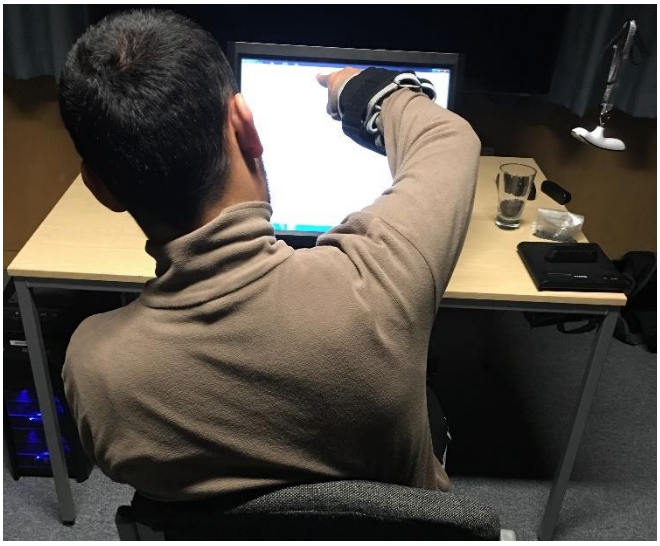


Figure 8.2 Participant during 3-Gz Condition.

In comparison, Pape and Vatrapu study showed no significant difference in subjective satisfaction and wellness across all Gz conditions. The reason could be because the mobile device was on the thigh of participants (smaller moment on the arm) and there was less arm, shoulder and neck movement required.

The realism of the current study's simulation of increased +Gz is limited. Experienced weight increase in this setting was created by adding additional weight to a certain point (wrist) which is not the case in a real flight. During a steep turn the increase of G-Force is experienced by the whole body, equally. +Gz can cause a reduction in the pilot's brain blood pressure, and it takes a certain amount of time until the body can compensate for this change. A study investigated the effects of ±Gz acceleration on cognitive performance revealed performance degradation in tracing, system monitoring and a strategic resource management task [Morrison et al. 1994].

9 Conclusion

The main research question was: “What are the potential benefits and challenges of touch screens on the flight deck?”. Therefore, interviews with avionics experts and pilots were conducted to figure out potential variables that could affect touch screen usability.

Identified variables were used to construct the foundations of the framework. Within this research project 18 research questions are addressed, which have been iteratively developed from the literature review and empirical findings. To address all research questions one field study, two lab studies, one observational study, one simulation study and one comparative user study were conducted. All findings contributed to form the big picture that showed potential benefits and challenges of touch screens on the flight deck.

Field study results revealed that all tested factors (in-flight vibration, placement and target size) have a significant impact on error rates. The target size is the most significant factor, which may be utilized to minimize other degrading factors by selecting an appropriate target size. It was demonstrated that using touch-enabled devices that are fixed in place in vibrating environments produce significantly higher error rates than when the device can be held by the user. Target size recommendation for mobile and fixed displays are given. The analyses of throughput were not consistent with the error rates. The Throughput during cruise and hover, which covers the majority of the flight, were similar. As expected, vibrations during transition phases result in lower throughput values. It was demonstrated for both experiments, binning index of difficulties and taking the average of each group would produce a strong R^2 value. Doing this would alleviate individual differences as well as differences in task design.

A modified Fitts’ Law task was applied to see how users would operate a multitouch enabled device in a real-world environment. The modified task design enabled further investigation on error rate for specific regions. In the mobile setting, participants had a higher accuracy on the centre of the screen. The error rate gets higher towards the edge of the screen. The error rate at corners for both placements were higher compared to the average error rate. This factor should be considered in the design process.

As stated before, trials showed that it was not feasible to test all impact factors during the field study. A lab study was conducted that evaluated the potential impact of various display position on usability of touch screens more in detail.

Statistically it was found that the display position has a significant impact on the usability of touch screens. There was a significant effect between using the dominant hand and the non-dominant hand as well as near display position and far display positions. There was no significant difference between displacement in the horizontal direction. The results of the ISO 9241-9 subjective rating questionnaire were presented and suggestions were made on how to customize the questionnaire to similar studies. Participants stated occlusion problems in some display positions. This effect was also observed and mentioned by the crewmembers during the field trials.

A lab study was conducted to understand the potential impact of +Gz on fixed touch screen displays. It was confirmed statistically that +Gz has a negative effect on usability. The drop in empirical results as well as subjective ratings is exponential with linear increase in simulated +Gz. There was a small increase in accuracy with increasing +Gz. Comparison with another study showed that using a weight adjustable wristband to simulate +Gz produced ecological valid results in some extent. Personal fitness and experience with touch screen usage was found to be a compensating factor.

A study was conducted with the aim to explore and understand potential benefits and challenges of an Electronic Flight Bag (mobile device) in a search and rescue (SAR) environment. Operational observations and interviews with SAR pilots were conducted to understand and specify the use of context within this particular area.

Based on requirements physical (device size) and virtual (interface design) factors were defined using a Digital Human Modelling (DHM) software. Developed initial interface design guidelines and expected features by pilots were presented. A scenario and an EFB prototype was developed and presented to pilots during the second stage of the study. Features, content and functionality that SAR pilots would like to see in a tablet app was presented.

Based on findings in this work and other related work a new touch screen interface was developed and evaluated in experiments with pilots from the Spanish Maritime Safety Agency (SASEMAR) using a tablet PC and the Flight Management System (FMS) of the Agusta Westland 139 (AW139). Results revealed that touch interface is significantly faster and error proof than conventional input method. That showed that designing user interfaces that represent their real-word counterparts (skeuomorphism) will not improve the usability and the design of user interface plays a key role in performance. An improved interface is proposed that was shaped by interviews with pilots and personal observations.

Findings from these studies were used to construct a framework that shows the relations between the four key factors (environment, physical, virtual and user). A preliminary questionnaire that avionics designer can use to determine whether touch screen technology is a suitable interface for their system was presented.

The overall conclusion from this thesis is that touch screen devices has the potential to be a good alternative input device for the flight deck if certain aspects are considered during the design process. Flight deck designers should understand the flight deck environment and create design solutions that meet the requirements of pilots. Touch screen interfaces would be not suitable if pilots have to interact with the system without looking on it. For this type of tasks and safety critical tasks it is recommended to use hard control.

10 Future Work

This chapter presents two potential future works resulting from the current studies. The first one aims to evaluate the potential benefits of Free-Air Interaction on flight decks. One of the primary finding in the lab study investigating the potential impact of display position on usability was that participants were significantly faster and more convenient in near display positions. The control stick or the joke (except side stick configuration) in front of pilots could limit designers to create a flight deck with displays that are close to pilots. Therefore, the idea came up to separate touch from screen where pilots can make mid-air gestures to interact with displays without touching it.

Since the lab study did not simulate +Gz in a realistic way it was recommended to transfer this setting to a human centrifuge where ecological valid results can be achieved. The second potential future works give a brief description of the proposed human centrifuge project.

10.1 Future Work: Separating Touch from Screen

There is no definition for “Free-Air Interaction”. In this context, we defined Free-Air Interaction as; “Human-Computer-Interaction where users do not touch a physical device to make an input”.

Free-Air-Interaction (finger and hand tracking) is a new way of interaction. Camera based devices that meet this definition are for examples; LEAP Motion [2015] and Microsoft Kinect [2016]. At the beginning of the project, camera based optical systems were able to distinguish between 3-4 fingers. This could be a limiting factor in multi crew cockpits where both pilots want to use the system at the same time. Therefore, a different technology was used during the preliminary study.

ZeroTouch (ZT) is a multi-touch sensing technology, which is based on detecting visual-hulls in an interactive area, created with daisy chained modules fitted with infrared (IR) sensors and light emitting diodes (LED). The shape of the interactive area can be customized according to special needs and requirements. A ZT frame attached onto the screen of a display, will transform it into a touch screen. Compared to other technologies this method of retrofitting is relatively simple. An interaction in ZT can be initiated with any physical object. Display and light quality is exactly the same, because there is layer between the user and display. Current ZT frames can track over 20 objects at a time. [J Moeller et al. 2011]

ZT offers the opportunity to separate “touch” from “screen”. Free-Air-Interaction with ZT was tested with a digital projected finger painting application. A frame, equipped with ZT modules, was placed in direct line-of-sight between participants and a projected canvas. Participants were able to paint on the canvas by putting their hands, fingers and other objects inside the frame. Participants found this kind of interaction engaging. However, the lack of tactile feedback lead to problems in distinguishing the activation threshold for the system. Another difficulty was precision in targeting a specific location. For an effective user experience, pre-activation feedback is essential. Authors suggested to use an extra layer of sensors, which can be used to create pre-activation feedback on the screen. [Jon Moeller et al. 2011].

10.1.1 Effect of Display Size & Aspect-Ratio

Until 2003 most computers had a display with 4:3 aspect ratio. In 2008, the computer industry started to move from 4:3 to 16:10 (or 16:9) (wide-screen) as the standard aspect ratio for monitors and laptops. Since 2012, displays with an aspect-ratio of 21:9 (ultra-wide screen) are available [Wikipedia 2015]. Display evolution shows that displays are getting wider and wider. Since the majority of interactive elements (e.g. buttons) are placed alongside the edges the distance between interactive areas will be bigger. Figure 10.1 shows the extreme cases for all mentioned aspect ratios. The flight deck of the Lockheed Martin Lightning II F-35 incorporates an ultra-wide touch screen [LockheedMartin 2014].

In this scenario, the user will close an app (blue target) and move to start button (red target) to open a new app. This operation is a frequent interaction for computer users (especially for Windows OS). The way of operation for a mouse is described in the first picture (blue line). The distance between two interactive areas will increase with increasing aspect ratio and screen size. Touch screen users may split the display and use the first half with the one hand and the other part with the other hand, which would result in shorter movement distance.

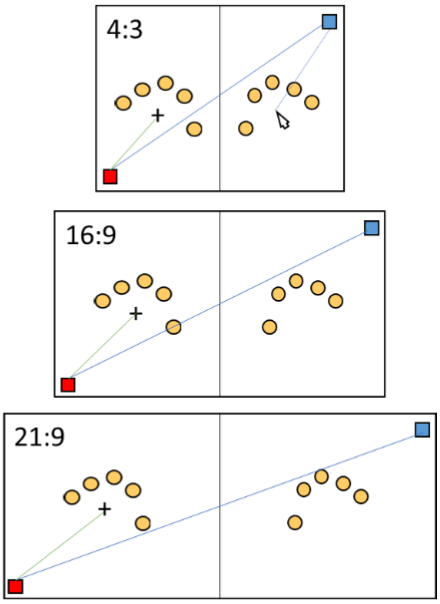


Figure 10.1 Maximum distance between two interactive areas on different displays

10.1.2 Conditions and Configurations

Displacement between display and ZT frame are conditions that might impede the usability. The aim of preliminary study is to investigate the effect of different displacement distances. Configurations are feedback methods that should compensate conditional drawbacks and improve user interaction. The effect of audial, and visual feedback method was tested separately or in combination.

10.1.2.1 Perspective in Free-Air Interaction

It is predictable that ZT attached directly onto the screen would produce similar performance to other touch screen technologies. The perception of letters, buttons and symbols will decrease by increased displacement in vertical direction (y direction). Pointing small targets on a touch screen is a known problem in the HCI Literature. As seen on Figure 10.2 increased displacement would generate an offset problem.

From user perspective, there are two different locations for the interactive area. However, the system is calibrated to one (green circle).

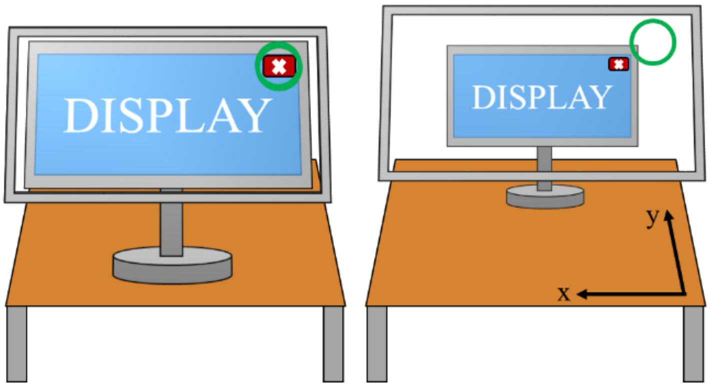


Figure 10.2 Perspective Issue in Free-Air Interaction.

10.1.2.2 ZT Study - Calibration

For an easy operation and to minimize the perspective effects following requirements are necessary:

- Line-of-sight, centre of ZT frame and screen are coincident.
- Display surface is approximately perpendicular to the line-of-sight.
- Both surfaces (display and ZT) are parallel.

A spirit level attached on the top of the ZT frame will be used for fine adjustments. A digital laser measure will be used to measure the displacement distance at various points. A height adjustable desk will be used to line up the centre of the ZT frame and screen with the line-of-sight of participants.

10.1.3 ZT study - Apparatus

Figure 10.3 shows the experimental setup of the ZT study. The interactive area of the ZT frame is 750 x 350 mm and is capable to track up to 20 objects. Two Line Lasers (50mW 405nm Blue-violet) with a fan angle of 110° will create a visual pre-activation feedback. Two ordinary speakers will provide audial feedback. With aid of brackets, the ZT frame will be attached to a height and angle adjustable fixture. Housing for laser modules are integrated in the brackets.

Figure 10.3 Experimental Setup for ZT study.

10.1.4 Possible Touch Strategies

There are three touch interaction strategies that can be used; First contact touch, last contact touch and hover mode

In “first contact touch” the interaction occurs once the finger touches the display (in this case if an object is within the interactive area). This kind of interaction is susceptible against accidental touches. In Free-Air-Interaction the user does not see exactly where he/she is pointing and this will lead to further problems. (Windows Touch is operating with this method)

The risk of accidental touches and unwanted selections is reduced with “last contact touch”. In this method, the interaction occurs after the finger leaves the interactive area. This kind of interaction could solve the pre-activation feedback problem. The user could put his/her finger into the interactive area, the cursor will move to the specific point, the user will drag the cursor to the target and lift his/her finger, which will be recognized as a click.

The second approach seems to increase the accuracy. However, it is predictable that overall operation (movement and selection time) will be longer. A cost/benefit analysis should show which approach is more appropriate.

The last touch strategy is the hover mode. In this method, the user has to hover over the interactive element for a certain amount of time. The interaction will be initiated once the time over target is exceeded.

10.1.5 Pilot Study on for Multi Directional Taping Task

A pilot study was conducted with 3 participants recruited from the local university campus. Multi directional taping task as described in ISO-9241-9, with first contact touch method, was conducted. Displacement distances were 200 mm and 400 mm. A baseline was created with 0 displacement.

The main finding was that the accuracy and interaction speed decreases with increasing displacement distance. Average throughput values were in 200 mm displacement condition 19% and in 400 mm displacement 55% lower than the baseline. As stated by Moeller [2011]; without any feedback, interacting through ZT was found to be difficult (compared to mouse). Finding (distinguishing) the interactive area increases the cognitive effort and frustration. For a click operation, the finger should not move within the interactive area otherwise it will be recognized as a drag operation. This was another disadvantage stated by participants.

Based on these findings it is recommended to add an extra layer of ZT sensors for positional and pre-activation feedback for first contact touch. The user could see the position of the cursor by going into the first layer of sensors. Users can move the cursor to the desired location and activate the interaction (click) by pushing through the second layer.

However, this can be solved easily by adapting the “last contact touch” strategy (and a second layer of ZT would be obsolete). In addition, other interaction strategies like drag, swipe, pinch and pan can be tested. Visual feedback was tested. Since the human eye can only focus to one location, users tended to focus to the display. So, visual feedback on the screen would be more beneficial than visual feedback before the ZT frame.

Another finding was that participants, after gaining experience, tended to hover their finger over the next target before pointing the current target with the other hand. This was also observed during the pilot trials for the field study. Following results (Figure 10.4 & Figure 10.5) are created using the tool from Wallner [2010]. The orange line represents the results generated with a mouse. The blue line is created with ZT (displacement distance 450 mm). Hovering over the next target manipulated movement times. The standard deviation in movement time increased with increasing ID values. Participants stated that the main reasons for this variation was the smaller targets (10 mm). The effective throughput for the mouse indicates a normal distribution. ZT results looks like that two processes are overlapping. Thus, this task design is not suitable for multi touch input devices.

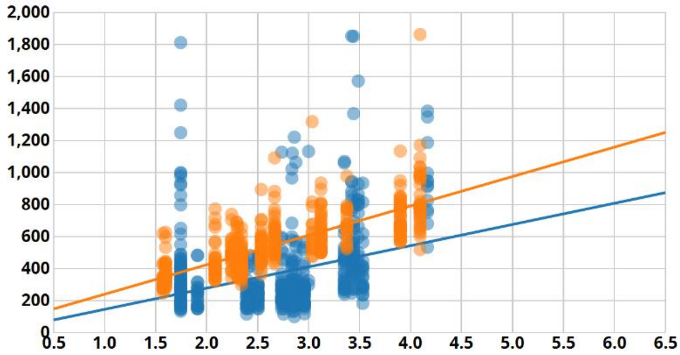


Figure 10.4 Effective IDe over movement time (ms) (Displacement 400 mm)

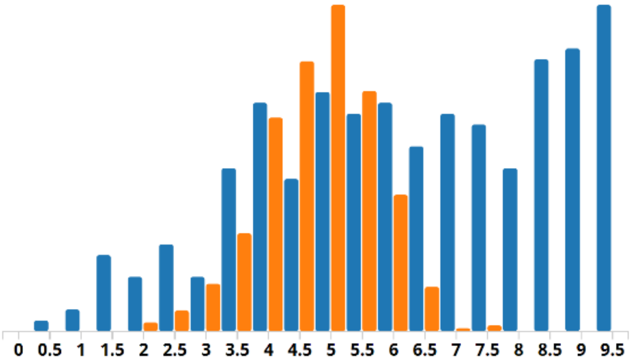


Figure 10.5 Effective Throughput Histogram (Displacement 400)

10.1.6 Additional Task Designs for Input Device Evaluation

In addition to two-dimensional tapping task there are additional task designs, stated in ISO-9241-9 [2007], that can be considered during the evaluation process. These are;

One directional tapping and dragging task (Figure 10.6). This is the original input device evaluation method proposed by Paul Fitts in 1954 [Fitts 1954]. This task can be performed as a tapping or dragging task where two rectangles will be presented to the participants. The aim in the tapping task is to click back-and-forth between the two rectangles. The aim in the dragging task is to drag a rectangle from one rectangle to another. Clicking or dropping the square outside (completely) of the rectangle will be recorded as error. The distance and width will change dynamically after each tap/drag.

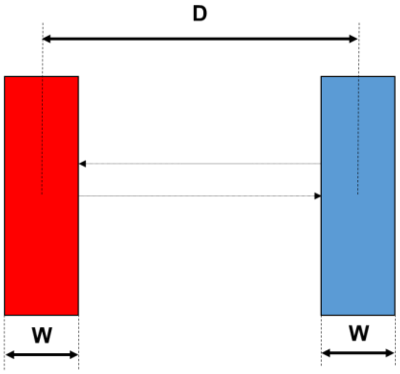


Figure 10.6 One-Directional Tapping and Dragging Task

Path following task (Figure 10.7). The task is to drag a circle through a “channel”. An error will be recorded if the circle touches the border. The path can be shaped to a multidirectional design/

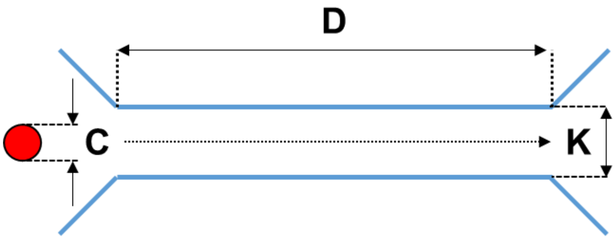


Figure 10.7 Path Following Task

Tracing task (Figure 10.8). The task is to trace a moving object with constant speed. The time where the cursor (or finger) is outside of the target will be recorded as error.

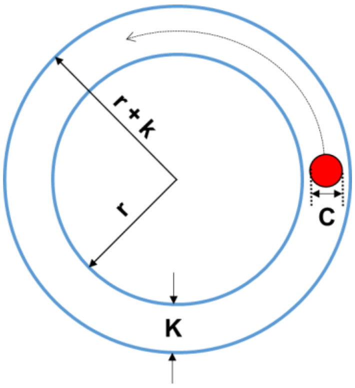


Figure 10.8 Tracing Task

10.1.7 Outcome and Decision of Zero Touch Study

A video footage of the trials was shared with pilots. Pilots said that they would prefer the inconvenience of bending towards the screen but having a haptic feedback than not having it. In addition, empirical results revealed that “Free-Air Interaction” as described and executed in the pilot study will have reduced accuracy and movement speed. This kind of interaction may be suitable for stationary usage where the accuracy has a lower priority than user satisfaction. The project was not continued after it was clear that “Free-Air Interaction” is not suitable for flight decks.

10.1.8 Future Work – Human Centrifuge

We seek to transfer this experimental setting to a human centrifuge, where experiments can be conducted under more realistic conditions, such as QinetiQ’s human centrifuge [QinetiQ 2016] (Figure 10.9), which is one of 20 centrifuges available worldwide. It has the added advantage of more closely replicating the ergonomics of a fast-jet cockpit, and can include pilot worn equipment, ejection seat and harness. It is used to simulate extreme +Gz experienced by fast jet aircraft pilots and astronauts with the aim to train the crew and to develop countermeasures to the impacts of +Gz on the human body. It is capable of simulating 9-Gz turns for manned experiments and 30-Gz for equipment testing. The following proposal is a brief version of the original proposal which was created in cooperation with QinetiQ engineers.

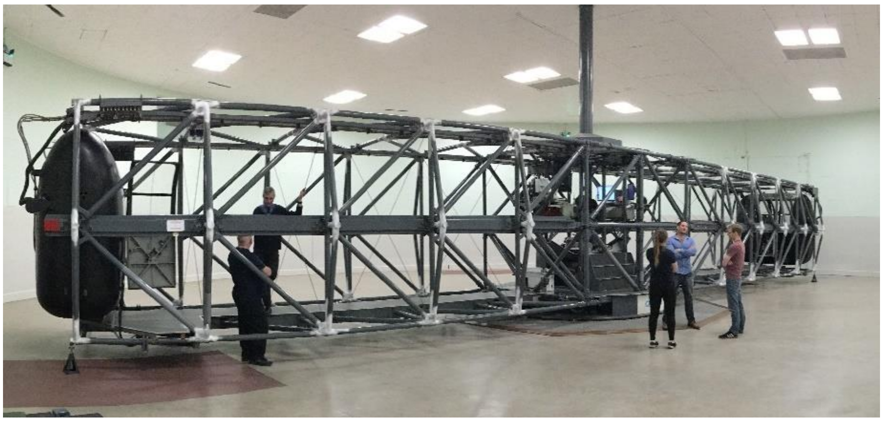


Figure 10.9 QinetiQ Human Centrifuge

10.2 Aim of Human Centrifuge Project

The aim of the proposed study is to determine whether (and to what extent) +Gz acceleration affects the performance on a touch screen. The modifying effects of task parameters (target size, distance of movement) will also be investigated. A further aim is to establish whether prior +Gz exposure proves detrimental effects to the performance of a touch screen task at +1Gz. In addition, due to subjective reports of perceived fatigue performing the task in the laboratory (Chapter 5.4) at simulated +3Gz, measures of arm, neck and upper back muscle activity will be acquired. Qualitative data (subjective questionnaires, as used in the initial study) will also be collected. Both results will be used to derive design recommendations and guidelines for touch screen interfaces on the flight deck. The aim is to identify ways in which these human-machine interfaces can be better designed (physical and software) to improve effectiveness and ease of use in both civilian and military applications.

10.2.1 Method of the Human Centrifuge Project

The study is separated into 3 phases:

- **Phase 1: Pilot study**. This phase will aim to prove that the task and protocol of testing to be applied in Phase 3 is feasible. It is possible that task difficulty, duration, number of repeats and duration of exposure will require slight modifications to the experimental design. If changes are required these will only reduce the risks associated with the trial (i.e. reduced task difficulty, fewer repeats, lower +Gz level, shorter duration of exposure). Only one volunteer is required. It is suggested that this volunteer will be an experienced centrifuge user, who has freely expressed his willingness to volunteer for the pilot study.

- **Phase 2: Task familiarisation.** With the aim to reduce learning effects during the main trial, participants will be familiarised with the aims and objectives of the study and task design one day before the main trial. The familiarisation session will be conducted at +1Gz with the participant seated in an office chair (i.e. not in the centrifuge gondola). Familiarisation session will take no longer than 1 hour. During this session, the aims and objectives of the experiment will be explained. An experimenter will then demonstrate the task which the participants will then practice. They will be instructed to conduct the task as quickly and as accurately as possible. They will also be informed that they may stop between blocks and rest to recover from fatigue if it should develop. Familiarisation will be completed at the point where improvement in throughput values with each successive performance reach a plateau.
- **Phase 3: Main trial.** The main trial will be composed of two centrifuge sessions; the first will involve 9 centrifuge exposures and the second 6 centrifuge exposures. A total of 15 separate exposures, each maintained for no longer than 90 seconds, will be performed. The maximum +Gz level used will not exceed +4Gz.

10.2.2 Apparatus

It is envisioned to conduct the experiment in the QinetiQ man-rated centrifuge at Farnborough, UK. The tapping task will be performed using a large touch screen (27 inch) fixed in the centrifuge gondola. The screen will be mounted on an adjustable bracket in order to accommodate the arm length of different subjects. The location and angle of the screen will replicate, as far as practicable, the position of the touch screen display in the Lightning II cockpit. Prior to any manned runs the safety and functioning of the screen will be confirmed via unmanned testing to +9Gz.

During all exposures participants, will be harnessed in an ejection seat and will wear the following aircrew equipment assemblies: aircrew coverall, inflatable socks, anti-G trousers (AGT), Mk10b helmet and type P/Q oronasal mask. The AGT will be pressurised using the Typhoon aircrew systems package, which commences inflation at +2Gz at 10kPa.G-1. Positive pressure breathing will not be used during this study. Instead the output from the regulator will be capped and the participants will breathe ambient air.

10.2.3 Experimental Design

The main trial is split into two parts. The first will investigate the effect of different levels of +Gz acceleration on task performance and the second the influence of prior +Gz exposure on subsequent performance of the task at +1Gz (i.e. examine whether there is any carry-over effect from the preceding +Gz exposure).

In the first phase participants, will be exposed to +Gz accelerations of +1Gz (i.e. centrifuge static), 2, 3 and 4Gz, repeated three times (i.e. 9 discrete +Gz acceleration exposures). During each the multi-directional tapping task, will be performed. A period of rest will be given between successive exposures, the duration of which will be dependent on the +Gz level at which it was undertaken, with one minute rest between repeats at +1Gz increasing in a linear fashion with +Gz level to 4 minutes at +4Gz. These rest intervals are based on previous performance in the lab study. A rest interval of 10 minutes will follow before the second (and final) part of testing commences.

In the second part the participants will be exposed to +2Gz and +4Gz, each repeated 3 times. The runs will be of similar duration and format to those used in the first phase (see Procedure) except that the participant will not perform the task, instead maintaining their hands in the stick and throttle position as if flying the aircraft. Upon return to +1Gz, and following a 20 second period to allow for the disorientating effects of centrifuge motion to subside, the participant will execute the tapping task.

The duration of +Gz will be sufficient to ensure that the participant can finish the task. Initial lab study showed that this can be completed in less than 50 seconds. To accommodate some scope for increased response times and to provide a period at the start of the +Gz profile for the participant to ready themselves, the centrifuge will be operated in manual mode, whereby the controlling engineer manually (via joystick input) controls onset and offset from the plateau. Once the task is complete the experimenter will call the controlling engineer to terminate the run. Notwithstanding the above, a maximum of 90 seconds will be pre-set for the +Gz exposures. Thus, regardless of the actions of the controlling engineer or experimenter, +Gz exposure will not exceed 90 seconds. The order of acceleration exposures will be randomly determined.

10.2.3.1 Muscle Activity

During the main trial muscle activity, will be recorded from the deltoid (shoulder), trapezius (shoulder/neck) and extensor digitorum (forearm) muscles. This will allow assessment of the levels of muscle activity required during the task and determination of the extent of any fatigue that has developed. Muscle activity will be recorded from small self-adhesive electrodes attached to the skin overlying the muscle of interest. These will be connected to wireless transmitters which will be located in the pockets of the coveralls that the subject is wearing.

10.2.3.2 Post-run Questionnaire

An independent rating scale based on ISO 9241-9, but modified to ensure its relevance to the user interface, will be employed to record subjective impressions of the ease of performing the task under each experimental condition (see Appendix D). The questionnaire is subdivided into two groups of indices; general usage and fatigue. Questions for general usage are; Smoothness during operation, effort required for operation, accuracy and operation speed. Fatigue questions are directed at identifying the regions (wrist, arm, shoulder and neck) and extent of fatigue. On a 7-point scale the questionnaire is formatted in a positive direction, with the highest values being associated with the most positive impressions. The experimenter or supervising medical officer will verbally administer the questionnaire after the three repeats for each condition are complete.

10.2.3.3 Post Experiment Questionnaire

After completing the test session an experimenter will conduct an informal debrief with participants about their experience and observations. After all participants, have finished the experiment, all mentioned issues will be collated and a post-experiment questionnaire created (as described in Chapter 5.3.6.5), summarising common issues. On a five-point Likert scale participants will rate if they would agree with the issues raised. This questionnaire will be sent to the participants via email.

10.2.4 Main Trial Procedure

Each participant will attend for a half-day. On arrival at the centrifuge facility the participant's fitness to undergo centrifuge exposure will be confirmed by the medical officer. The experiment will last for ~1 hour, with a prior ~30 minutes preparation required for attaching medical monitoring, donning aircrew clothing and installation in the gondola. In total the session should not exceed two hours.

Before entering the gondola maximal voluntary contractions (MVC) of the deltoid (shoulder), trapezius (shoulder/neck) and extensor digitorum (forearm) muscles will be performed to identify the maximal amount of muscle activity that can be recorded from each muscle. Static MVCs will be performed with the subject either trying to extend their fingers, raise their arm to the side or raise their shoulder for each muscle, respectively, with the movement being manually resisted by the experimenter. Three MVC will be performed for each muscle with a minimum of 1 minute rest between contractions. The data obtained from these will be used to normalise the data acquired under +Gz.

Once harnessed in the ejection seat in the centrifuge gondola an experimenter will give the same instructions as during the familiarisation session. The participant will be told to perform the task as quickly and accurate as possible, to rest if they feel fatigued and to use their left hand. Their right hand will be placed on the arm rest and positioned so that they can easily activate the centrifuge stopping mechanism, if required.

The participant will first perform the task at +1Gz (i.e. centrifuge static) a total of five times. The first two repeats will be used to re-familiarise the participant with the task and the data will be disregarded while the remaining three repeats will be used to provide baseline data. Once complete and a suitable rest period has been taken, the medical officer will inform the participant of the +Gz level of the upcoming exposure and will ask them to confirm they are ready to proceed. The centrifuge will then be accelerated to the desired G Level at an onset rate of 0.3 G.s-1. Once the plateau acceleration level is reached the medical officer will inform the participant that they can commence the tapping task. After finishing a sequence participants will continue onto the next. Once a block is finished the experimenter will inform the centrifuge controlling engineers to terminate the run. A period of rest will be taken (1-4 minutes, dependent on +Gz level) before the next run is commenced. Once three repeats at the same +Gz level have been completed an experimenter or medical officer will administer the questionnaire marking participant’s responses using the 7-point scale.

The second part of the main trial is performed almost identically except that the participants will be asked to place their left hand on the throttle during the centrifuge exposure (the right hand will remain on the arm rest and in a position to easily activate the centrifuge stop button). The duration of the run will be 60 seconds. Post exposure, and after a period of 20 seconds rest the medical officer will inform the participant to commence the task.

10.2.5 Participants

Centrifuge trained participants who have consented to being contacted about future centrifuge studies will be approached. This will be done via an e-mail to each participant drawing their attention to the fact that the current study is being conducted. Attached to this e-mail will be the participant information sheet which potential volunteers will be asked to read, if they are interested in taking part in the study. It will be explicitly stated in the e-mail that potential participants are under no obligation to volunteer for the study.

In the event that insufficient participants are recruited from the existing participant pool an advert will be placed on the QinetiQ intranet requesting participants for a centrifuge study. If interested in volunteering, individuals will be asked to contact the principal investigator who will provide a copy of the participant information sheet.

In all cases, once an individual has been given the participant information sheet they will have a minimum of 24 hours to read it. They will then be provided the opportunity to discuss the study and any questions they may have with the principal investigator.

10.2.6 Statistical analysis

Throughput is the principle dependent variable for the study calculated automatically by the software task following completion of the task. A Shapiro-Wilk test will be used to examine the data and if not normally distributed an appropriate transformation will be performed. If assumptions of normality are still not met following transformation a non-parametric equivalent to the statistical tests described below will be used.

The effects of +Gz acceleration on task performance under +Gz and post +Gz will be assessed separately. For the performance of the task *under* +Gz a one-way repeated measure analysis of variance (RMANOVA) will be used with +Gz level as the factor (4 levels: 1, 2, 3 and 4Gz). The initial study determined an effect size (f) of 0.547 (partial η2 = 0.23). Allowing for a reduction in effect size (30%) due to 'field' conditions, correction for non-sphericity (we would expect between subject variation in task performance to increase with higher +Gz levels) and a correlation between repeated measurements of 0.6 recorded previously, a sample size of 10 is required to find a statistically significant difference at an alpha of 0.05 with a power of 80%. The effect of +Gz on task performance at +1Gz post exposure will be analysed with one-way RMANOVA with preceding +Gz level as the factor (3 levels: 1, 2, and 4Gz). A medium-large effect size is of interest (partial η^2 = 0.11), as smaller changes in task performance are unlikely to significantly influence operational output. Using the assumptions described earlier, to find a statistically significant difference at an alpha of 0.05 with a power of 80% a study sample size of 15 is required. Considering the above, and to provide some scope for participant dropout 16 subjects will be recruited. Statistical power calculations were performed using G*power (v3.1.9.2).

11 References

K. Abbott et al. 1996. *The interfaces between flightcrews and modern flight deck systems*, Washington.

T.S. Abbott and W.H. Rogers. 1993. Functional categories for human-centered flight deck design. *Proc. 12th Am. Inst. Aeronaut. Astronaut. Electr. Electron. Eng. Digit. Avion. Syst. Conf.* (1993), 66–74. DOI:http://dx.doi.org/10.1109/DASC.1993.283569

V. Ahlstrom. 2010. *HFDS 2003: chapter 14 anthropometry and biomechanics*, Washington.

Bashar I. Ahmad, Patrick M. Langdon, Simon J. Godsill, Robert Hardy, Lee Skrypchuk, and Richard Donkor. 2015. Touchscreen usability and input performance in vehicles under different road conditions. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '15*. New York, New York, USA: ACM Press, 47–54. DOI:http://dx.doi.org/10.1145/2799250.2799284

AHunt. 2015. Lockheed Martin F-35 Lightning II. (2015). Retrieved May 20, 2016 from https://en.wikipedia.org/wiki/Lockheed_Martin_F-35_Lightning_II

Airbus. 2015. A380. (2015). Retrieved June 6, 2015 from http://www.airbus.com/aircraftfamilies/passengeraircraft/a380family/commonality/

Jetstar Airways. 2011. Boeing 787 Dreamliner flight deck. (2011). Retrieved July 7, 2014 from https://www.flickr.com/photos/jetstarairways/6769120147/

Pär-Anders Albinsson and Shumin Zhai. 2003. High precision touch screen interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '03*. New York: ACM Press, 105. DOI:http://dx.doi.org/10.1145/642611.642631

DM Anderson. 2014. *Design for manufacturability: how to use concurrent engineering to rapidly develop low-cost, high-quality products for lean production*,

Damon P. Andrew, Paul M. Pedersen, and Chad D. McEvoy. 2011. Research methods and design in sport management. (2011).

DH Andrews, J. Brown, J. Byrnes, J. Chang, and R. Hartman. 1998. *Enabling technology: analysis of categories with potential to support the use of modeling and simulation in the united states air force.*, Mesa, AZ.

Michael Angrosino. 2007. *Doing ethnographic and observational research*, 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom: SAGE Publications Ltd.

Apple. 2014. *iOS human interaface guidlines (iOS 8)*,

Apple. 2016. Take advantage of 3D touch. (2016). Retrieved January 8, 2016 from https://developer.apple.com/ios/3d-touch/

D.P. Ausubel, J.D. Novak, and H. Hanesian. 1968. *Educational psychology: A cognitive view. New York, NY: Holt, Rinehart and Winston*, Morocco: Culture, Cognition, and Literacy.

Daniel Avrahami. 2015. The effect of edge targets on touch performance. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. New York, New York, USA: ACM Press, 1837–1846. DOI:http://dx.doi.org/10.1145/2702123.2702439

Huseyin Avsar, Joel Fischer, and Tom Rodden. 2016a. Designing touch-enabled electronic flight bags in sar helicopter operations. In *International Conference on Human Computer Interaction in Aerospace: HCI-Aero*. Paris: ACM Press. DOI:http://dx.doi.org/http://dx.doi.org/10.1145/2950112.2964591

Huseyin Avsar, Joel Fischer, and Tom Rodden. 2016b. Designing touch screen user interfaces for future flight deck operations. In *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. Sacramento: IEEE. DOI:http://dx.doi.org/10.1109/DASC.2016.7777976

Huseyin Avsar, Joel Fischer, and Tom Rodden. 2016c. Future flight decks: impact of +Gz on touchscreen usability. In *International Conference on Human Computer Interaction in Aerospace: HCI-Aero*. Paris: ACM Press. DOI:http://dx.doi.org/http://dx.doi.org/10.1145/2950112.2964592

Huseyin Avsar, Joel Fischer, and Tom Rodden. 2016d. Mixed method approach in designing flight decks with touchscreens: A framework. In *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. Sacramento. DOI:http://dx.doi.org/10.1109/DASC.2016.7778066

Huseyin Avsar, Joel Fischer, and Tom Rodden. 2016e. Physical and environmental considerations for touchscreen integration on the flight deck. *Unpublished* (2016).

Huseyin Avsar, Joel Fischer, and Tom Rodden. 2015. Target size guidelines for interactive displays on the flight deck. In *2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC)*. Prague: IEEE, 3C4-1-3C4-15. DOI:http://dx.doi.org/10.1109/DASC.2015.7311400

Axure Software Solutions. 2016. Axure. (2016). Retrieved November 26, 2016 from axure.com

K. Bailey. 2008. *Methods of social research*, New York: The Free Press.

Maxim Bakaev. 2008. Fitts' law for older adults: considering a factor of age. *Proc. VIII Brazilian Symp. Hum. Factors Comput. Syst.* , Id (2008), 260–263.

Thorsten Baldus and Patrick Patterson. 2008. Usability of pointing devices for office applications in a moving off-road environment. *Appl. Ergon.* 39, 6 (2008), 671–677. DOI:http://dx.doi.org/10.1016/j.apergo.2008.01.004

Mark. Balnaves and Peter. Caputi. 2001. *Introduction to quantitative research methods : an investigative approach*,

Balsamiq Studios. 2016. Balsamiq. (2016). Retrieved November 26, 2016 from balsamiq.com

J. Banks, J.S. Carson, B. Nelson, and D.M. Nicol. 2001. Verification and validation of simulation models. In *Discrete-event system simulation*. Prentice-Hall, Upper Saddle River (NJ, 367–397.

Dave Barstow. 2012. The aviation iPad revolution. *J. Air Traffic Control* 54, 8 (2012), 4–8.

B. Baumgarten and C. Lahusen. 2006. Politiknetzwerke—Vorteile und Grundzüge einer qualitativen Analysestrategie. *Qual. Netzwerkanalyse* (2006), 177–197.

D. Beerbaum. 2016. *Towards an XBRL-enabled corporate governance reporting taxonomy.: An empirical study of NYSE-listed Financial Institutions*, BoD- Books on Demand.

Woodrow Bellamy. 2013. NBAA 2013: Touchscreen cockpit systems and next generation business aircraft. (2013). Retrieved December 31, 2016 from http://www.aviationtoday.com/av/topstories/NBAA-2013-Touchscreen-Cockpit-Systems-and-Next-Generation-Business-Aircraft_80495.html

Herbert D. Benington. 1983. Production of large computer programs. *IEEE Ann. Hist. Comput.* 5, 4 (October 1983), 350–361. DOI:http://dx.doi.org/10.1109/MAHC.1983.10102

Joanna Bergstrom-Lehtovirta, Antti Oulasvirta, and Stephen Brewster. 2011. The effects of walking speed on target acquisition on a touchscreen interface. *Proc. 13th Int. Conf. Hum. Comput. Interact. with Mob. Devices - MobileHCI '11* (2011), 143–146. DOI:http://dx.doi.org/10.1145/2037373.2037396

M. Bertucco and T.D. Sanger. 2014. Speed-Accuracy Testing on the Apple iPad(R) Provides a Quantitative Test of Upper Extremity Motor Performance in Children with Dystonia. *J. Child Neurol.* 29, 11 (November 2014), 1460–1466. DOI:http://dx.doi.org/10.1177/0883073813494265

Erik Blasch, Paul Kostek, Pavel Pačes, and Kathleen Kramer. 2015. Summary of avionics technologies. *IEEE Aerosp. Electron. Syst. Mag.* 30, 9 (2015), 6–11. DOI:http://dx.doi.org/10.1109/MAES.2015.150012

Claire Bless, Craig Higson-Smith, and Ashraf Kagee. 2006. *Fundamentals of social research methods: an african perspective*,

LTM Blessing and A. Chakrabarti. 2009. *DRM: A Design Reseach Methodology*,

Barry W. Boehm. 1988. A spiral model of software development and enhancement. *Computer (Long. Beach. Calif).* 21, 5 (1988), 61–72. DOI:http://dx.doi.org/10.1109/2.59

Boeing. 2015. Boeing 787. (2015). Retrieved May 11, 2015 from http://www.boeing.com/commercial/aeromagazine/articles/2012_q1/3/

Boeing. 2007. MEDA Investigation Process. *Aeromagazine* (2007), 15–16.

Boeing. 2012. Statistical summary of commercial jet airplanes accident. (2012). Retrieved October 11, 2012 from www.boeing.com/news/techissues/pdf/statsum.pdf

Boeing. 2016. Touchscreens come to 777X flight deck. (2016). Retrieved October 18, 2016 from http://www.boeing.com/features/2016/07/777x-touchscreen-07-16.page

R. Bogdan and M. Ksander. 1980. Policy data as a social process: A qualitative approach to quantitative data. *Hum. Organ.* (1980).

Stefano Bonelli et al. 2013. The usability evaluation of a touch screen in the flight deck. In *Cases on Usability Engineering*. IGI Global, 270–297. DOI:http://dx.doi.org/10.4018/978-1-4666-4046-7.ch012

Stefano Bonelli and Valentina Barsotti. 2014. ALICIA Project. (2014). Retrieved October 17, 2016 from http://www.alicia-project.eu/CMS/

Stefano Bonelli and Linda Napoletona. 2013. *Flying with complexity ; bringing touch screens into the cockpit*,

I. Brace. 2008. *Questionnaire design: How to plan, structure and write survey material for effective market research*, Kogan Page Publishers.

N. Britten. 1999. Qualitative interviews in healthcare. In C. Pope & N. Mays, eds. *Qualitative research in health care*. London: BMJ Books, 11–19.

A. Bryman. 2006. Integrating quantitative and qualitative research: how is it done? *Qual. Res.* 6, 1 (February 2006), 97–113. DOI:http://dx.doi.org/10.1177/1468794106058877

Califlier001. 2014. Interflug Airbus A310. (2014). Retrieved July 7, 2016 from https://www.flickr.com/photos/calflier001/13155123603

DT Campbell and DW Fiske. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* 59, 2 (1959), 81–105. DOI:http://dx.doi.org/http://dx.doi.org/10.1037/h0046016

DC Chandra, M. Yeh, V. Riley, and SJ Mangold. 2003. Human factors considerations in the design and evaluation of Electronic Flight Bags (EFBs), Version 2. *http//ntl. bts. gov/lib/34000/ …* (2003).

Karen B. Chen, Anne B. Savage, Amrish O. Chourasia, Douglas A. Wiegmann, and Mary E. Sesto. 2013. Touch screen performance by individuals with and without motor control disabilities. *Appl. Ergon.* 44, 2 (2013), 297–302. DOI:http://dx.doi.org/10.1016/j.apergo.2012.08.004

Amrish Chourasia, Douglas Wiegmann, Karen Chen, and Mary Sesto. 2013. Effect of sitting orientation on touchscreen performance, touch characteristics, user preference, and workload. *IIE Trans. Occup. Ergon. Hum. Factors* 1, 4 (October 2013), 235–245. DOI:http://dx.doi.org/10.1080/21577323.2013.847873

Civil Aviation Authority. 2008. *CAP 780 - Aviation safety review*,

Civil Aviation Safety Authority Australia. 2013. *Electronic Flight Bags*, Canberra.

Nicola Clark. 2013. Touch screens are tested for piloting passenger jets. (2013). Retrieved December 31, 2016 from http://www.nytimes.com/2013/07/06/technology/passenger-jets-testing-touch-screen-technology.html

Herbert Colle and Keith Hiszem. 2004. Standing at a kiosk: Effects of key size and spacing on touch screen numeric keypad performance and user preference. *Ergonomics* 47, 13 (October 2004), 1406–1423. DOI:http://dx.doi.org/10.1080/00140130410001724228

Jessica Conradi, Olivia Busch, and Thomas Alexander. 2015. Optimal touch button size for the use of mobile devices while walking. *Procedia Manuf.* 3 (2015), 387–394. DOI:http://dx.doi.org/10.1016/j.promfg.2015.07.182

Antonia S. Conti, Florian Kremser, Michael Krause, Dae-jin An, and Klaus Bengler. 2015. The effect of varying target sizes and spaces between target and non-target elements on goal-directed hand movement times while driving. *Procedia Manuf.* 3 (2015), 3168–3175. DOI:http://dx.doi.org/10.1016/j.promfg.2015.07.866

Alan Cooper, Robert. Reimann, Dave Cronin, and Alan Cooper. 2007. *About face 3 : the essentials of interaction design*, Wiley Pub.

Coordination Committee. 2014. *State of harmonisation document*,

JW Creswell. 2013. *Research design: Qualitative, quantitative, and mixed methods approaches*, Sage Publications.

JW Creswell and VLP Clark. 2007. Designing and conducting mixed methods research. (2007).

Curimedia. 2011. Airbus A320-214. (2011). Retrieved July 7, 2016 from https://www.flickr.com/photos/calflier001/13155123603

John. Daintith, Edmund (Thomas Edmund Farnsworth) Wright, and Oxford University Press. 2008. *A dictionary of computing.*, Oxford University Press.

Kiran Dandekar, Balasundar I. Raju, and Mandayam A. Srinivasan. 2003. 3-D finite-element models of human and monkey fingertips to investigate the mechanics of tactile sense. *J. Biomech. Eng.* 125, 5 (2003), 682–691. DOI:http://dx.doi.org/10.1115/1.1613673

Matthew David and Carole D. Sutton. 2004. *Social research: the basics*,

A. Degani and EL Wiener. 1994. On the design of flight-deck procedures. (1994).

Asaf Degani. 1992. On the typography of flight-deck documentation. (1992), 40.

Asaf Degani, Everett A. Palmer, and Kristin G. Bauersfeld. 1992. “Soft” Controls for hard displays: still a challenge. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Human Factors Society, 52–56. DOI:http://dx.doi.org/10.1177/154193129203600114

Department of Defense. 2012. *Design Criteria Standard (MIL-STD-1472G)*,

Department of Defense. 1994. *Directive 5000.59 Modeling and Simulation (M&S) Management*,

Department of Defense. 1988. *DoD STD 2167 A, “Defense system software development,”*

K. Musante DeWalt and BR DeWalt. 2002. *Participant observation: A guide for fieldworkers. Walnut Creek*, Alta Mira Press.

Amit Dhir. 2004. *The digital consumer technology handbook: a comprehensive guide to devices, standards, future directions, and programmable logic solutions*, Elsevier.

DA Dillman. 2000. *Mail and internet surveys: The tailored design method*, New York, USA: Wiley.

DA Dillman, JD Smyth, and L. Melani. 2011. *Internet, mail, and mixed-mode surveys: the tailored design method*, Toronto: Wiley & Sons.

Alen Dix, Janet Finlay, Gregoru Abowd, and Russell Beale. 2004. Human-computer interaction. In Pearson Education Limited 5, 194–196.

Sonia Dodd, Jeff Lancaster, Andrew Miranda, Steve Grothe, B. DeMers, and Bill Rogers. 2014. Touch screens on the flight deck: The impact of touch target size, spacing, touch technology and turbulence on pilot performance. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* 58, 1 (September 2014), 6–10. DOI:http://dx.doi.org/10.1177/1541931214581002

L. Doyle, A.M. Brady, and G. Byrne. 2009. An overview of mixed methods research. *J. Res. Nurs.* 14, 2 (March 2009), 175–185. DOI:http://dx.doi.org/10.1177/1744987108093962

JS Dumas and JE Fox. 2009. Usability testing: Current practice and future directions. *Human-Computer Interact. Dev.* (2009).

MR Endsley. 2016. *Designing for situation awareness: An approach to user-centered design*, CRC press.

David A. Erlandson, Edward L. Harris, Barbara L. Skipper, and Steven Allen. 1993. *Doing naturalistic inquiry: A guide to methods.*, SAGE Publications.

M. Elliot. Familant and Mark C. Detweiler. 1993. Iconic reference: evolving perspectives and an organizing framework. *Int. J. Man. Mach. Stud.* 39, 5 (November 1993), 705–728. DOI:http://dx.doi.org/10.1006/imms.1993.1080

Federal Aviation Administration. 2007. *Fact sheet: NextGen*,

Federal Aviation Administration. 2012. *Guidelines for the certification, airworthiness, and operational use of electronic flight bags*, Washington.

Federal Aviation Administration (FAA). 2014. *AC 25-11B - Electronic flight displays*,

Federal Aviation Administration (FAA). 2015. *Guide for aviation medical examiners*,

Federal Aviation Authority (FAA). 1995. *Human performance considerations in the use and design of aircraft checklists*, Washington DC.

Leah Findlater, Jon E. Froehlich, Kays Fattal, Jacob O. Wobbrock, and Tanya Dastyar. 2013. Age-related differences in performance with touchscreens compared to traditional mouse input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’13*. New York, New York, USA: ACM Press, 343–347. DOI:http://dx.doi.org/10.1145/2470654.2470703

A. Fink. 2003. How to ask survey questions. *Surv. kit* (2003), 79–82. DOI:http://dx.doi.org/10.4135/9781412984393

Paul M. Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* 47, 6 (1954), 381–391. DOI:http://dx.doi.org/10.1037/h0055392

Fluid. 2016. Fluid UI. (2016). Retrieved November 26, 2016 from fluidui.com

W. Foddy. 1994. *Constructing questions for interviews and questionnaires: Theory and practice in social research*, Cambridge university press.

Kevin Forsberg and Harold Mooz. 1994. The relationship of system engineering to the project cycle. *12th INTERNET World Congr. Proj. Manag.* , June 1994 (1994), 12. DOI:http://dx.doi.org/10.1002/j.2334-

5837.1991.tb01484.x

Kyle Frost. 2013. Report: iPad lowering number of pilot injuries. (2013). Retrieved October 18, 2016 from http://www.todaysiphone.com/2013/05/report-ipad-lowering-number-of-pilot-injuries/

Q. Gao and Q. Sun. 2015. Examining the usability of touch screen gestures for older and younger adults. *Hum. Factors J. Hum. Factors Ergon. Soc.* 57, 5 (August 2015), 835–863. DOI:http://dx.doi.org/10.1177/0018720815581293

Anna Arasa Gaspar. 2011. *Design and evaluation of haptic feedback for in-vehicle touch screens*.

Jason Gauci, Nathalie Cauchi, Kevin Theuma, David Zammit-Mangion, and Alan Muscat. 2015. Design and evaluation of a touch screen concept for pilot interaction with avionic systems. In *2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC)*. Prague: IEEE, 3C2-1-3C2-19. DOI:http://dx.doi.org/10.1109/DASC.2015.7311398

DT Gilbert, ST Fiske, and G. Lindzey. 1998. *The handbook of social psychology*, McGraw-Hill.

P. Gill, K. Stewart, E. Treasure, and B. Chadwick. 2008. Methods of data collection in qualitative research: interviews and focus groups. *Br. Dent. J.* 204, 6 (March 2008), 291–295. DOI:http://dx.doi.org/10.1038/bdj.2008.192

B. Gillham. 2008. *Developing a questionnaire*, A&C Black.

Lisa M. Given. 2008. *The Sage encyclopedia of qualitative research methods*, Sage Publications.

Google. 2014. *Android interface design guidlines (Android 5.0)*,

DE Gray. 2004. *Doing research in the real world*,

Samuel B. Green, Neil J. Salkind, and Theresa M. Akey. 1997. *Using SPSS for windows : analyzing and understanding data*, Prentice Hall.

J.C. Greene, V.J. Caracelli, and W.F. Graham. 1989. Toward a conceptual framework for mixed-method evaluation designs. *Educ. Eval. Policy Anal.* 11, 3 (January 1989), 255–274. DOI:http://dx.doi.org/10.3102/01623737011003255

E.R.F.W. Grossman. 1960. The information-capacity of the human motor-system in pursuit tracking. *Q. J. Exp. Psychol.* 12, 1 (January 1960), 01–16. DOI:http://dx.doi.org/10.1080/17470216008416694

SK Grove and N. Burns. 2005. *The practice of nursing research: Conduct, critique & utilization*, St.Louis: Elsevier Sanders.

Tiago Guerreiro, Hugo Nicolau, Joaquim Jorge, and Daniel Gonçalves. 2010. Towards accessible touch interfaces. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '10*. New York, New York, USA: ACM Press, 19. DOI:http://dx.doi.org/10.1145/1878803.1878809

Gulfstream. 2015. Gulfstream symmetry flight deck. (2015). Retrieved May 1, 2015 from http://www.gulfstream.com/technology/symmetry-flight-deck

Hamblin C. 2003. Electronic Flight Bags (EFBs) with small screens significantly increase information retrieval times. In *Proceedings of 12th International Symposium on Aviation Psychology*. Dayton OH, 463–468.

Martyn Hammersley and Paul Atkinson. 1995. *Ethnography: principles in practice*, Routledge.

PA Hancock, DA Vincenzi, JA Wise, and M. Mouloua. 2008. *Human factors in simulation and training*,

Julian Happian-Smith. 2000. *An introduction to modern vehicle design*,

C. Hardyck and L.F. Petrinovich. 1977. Left-handedness. *Psychol. Bull.* 87, 3 (1977), 385–404.

Don Harris. 2004. *Human factors for civil flight deck design*, Gower Publishing, Ltd.

MR Harwell. 2011. Research design in qualitative/quantitative/mixed methods. In *The SAGE Handbook for Research in Education: Pursuing ideas as the keystone of exemplary inquiry*. Sage Publications, 147–163.

S.T. Hayes, E.R. Hooten, and J.A. Adams. 2014. Tablet interaction accuracy and precision: seated vs. walking. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* 58, 1 (September 2014), 432–436. DOI:http://dx.doi.org/10.1177/1541931214581090

Niels Henze, Enrico Rukzio, and Susanne Boll. 2011. 100,000,000 taps: analysis and improvement of touch performance in the large. *Proc. 13th Int. Conf. Hum. Comput. Interact. with Mob. Devices Serv. - MobileHCI '11* (2011), 133. DOI:http://dx.doi.org/10.1145/2037373.2037395

SN Hesse-Biber. 2011. *The handbook of emergent technologies in social research*, Oxford University Press.

SN Hesse-Biber and P. Leavy. 2010. *Handbook of emergent methods*, Guilford Press.

Honeywell. 2015. Honeywell brings modern touch to gulfstream cockpit. (2015). Retrieved June 1, 2015 from https://aerospace.honeywell.com/news/honeywell-brings-modern-touch-to-gulfstream-cockpit.

Honeywell. 2014. New Honeywell technologies including touchscreens in cockpit help gulfstream's newest aircraft take flight. (2014). Retrieved December 29, 2016 from https://www.honeywell.com/newsroom/pressreleases/2014/10/new-honeywell-technologies-including-touchscreens-in-cockpit-help-gulfstream-s-newest-aircraft-take-flight

T.C. Hong, N.A. Kuan, T.K. Kiang, and S.K.T. John. 2011. Evaluation of Input Devices for Pointing, Dragging and Text Entry Tasks On A Tracked Vehicle. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* 55, 1 (September 2011), 2078–2082. DOI:http://dx.doi.org/10.1177/1071181311551433

William Horton. 1994. The icon book: visual symbols for computer systems and documentation. (1994).

Andrea Huguely. 2013. *American airlines completes electronic flight bag implementation*,

SB Hulley. 2007. *Designing clinical research*, Lippincott Williams & Wilkins.

Human Factors and Ergonomics Society. 2007. *ANSI/HFES 100-2007 human factored engineering of computer workstations*,

International Organization for Standardization. 2010. *ISO 9241-210: Ergonomics of human-centred system interaction - part 210: Human-centred design for interactive systems*, Geneva.

International Standard Organisation. 2015. *ISO/DIS 9241-11 Ergonomics of human-system interaction -- Part 11: Usability: Definitions and concepts*,

International Standard Organisation. 2007. *ISO 9241-9 Ergonomic requirements for office work with visual display terminals*,

Allen Ivey, Mary Ivey, Carlos Zalaquett, and Kathryn Quirk. 2011. *Essentials of intentional interviewing: Counseling in a multicultural world*, Cengage Learning.

R.B. Johnson, A.J. Onwuegbuzie, and L.A. Turner. 2007. Toward a definition of mixed methods research. *J. Mix. Methods Res.* 1, 2 (April 2007), 112–133. DOI:http://dx.doi.org/10.1177/1558689806298224

N. Johnstone. 2013. *The electronic flight bag friend or foe? Report Nr 104*,

D. Jones. 1990. *Three input concepts crew interaction presented electronic for flight with information on a large-screen cockpit display*,

E.R. Jones. 1967. *Simulation applied to education*, McDonnell Co.

Patrick W. Jordan. 1998. *An introduction to usability*, Taylor & Francis.

N. Jørgensen. 2006. The Boeing 777: Development Life Cycle Follows Artifact. *World Conf. Integr. Des. Process* (2006).

K. H. Abbott. 2001. *The Avionics Handbook, Chapter 9: Human factors engineering and flight deck design*, Boca Raton: CRC press LLC.

Sridher Kaminani. 2011. Human computer interaction issues with touch screen interfaces in the flight deck. In *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*. IEEE, 6B4-1-6B4-7. DOI:http://dx.doi.org/10.1109/DASC.2011.6096098

Waldemar Karwowski. 2006. International Encyclopedia of Ergonomics and Human Factors, Second Edition. (2006), 3728. DOI:http://dx.doi.org/10.1201/9780849375477

Christian Kath. 2006. Concorde flight deck layout. (2006). Retrieved May 28, 2016 from www.wikipedia.org/wiki/Concorde

Panagiotis Kenterlis. 2012. ODICIS Website. (2012). Retrieved October 17, 2016 from http://www.odicis.teipir.gr/index.php

Heejin Kim, Sunghyuk Kwon, Jiyoon Heo, Hojin Lee, and Min K. Chung. 2014. The effect of touch-key size on the usability of in-vehicle information systems and driving safety during simulated driving. *Appl. Ergon.* 45, 3 (2014), 379–388. DOI:http://dx.doi.org/10.1016/j.apergo.2013.05.006

Inki Kim and Jang Hyeon Jo. 2015. Performance comparisons between thumb-based and finger-based input on a small touch-screen under realistic variability. *Int. J. Hum. Comput. Interact.* 31, 11 (November 2015), 746–760. DOI:http://dx.doi.org/10.1080/10447318.2015.1045241

J.H. Kim, L. Aulck, M.C. Bartha, C.A. Harper, and P.W. Johnson. 2012. Are there differences in force exposures and typing productivity between touchscreen and conventional keyboard? *Proc. Hum.*

Factors Ergon. Soc. Annu. Meet. 56, 1 (September 2012), 1104–1108. DOI:http://dx.doi.org/10.1177/1071181312561240

J.H. Kim, L. Aulck, O. Thamsuwan, M.C. Bartha, and P.W. Johnson. 2014. The effect of key size of touch screen virtual keyboards on productivity, usability, and typing biomechanics. *Hum. Factors J. Hum. Factors Ergon. Soc.* 56, 7 (November 2014), 1235–1248. DOI:http://dx.doi.org/10.1177/0018720814531784

J.C. Knight. 2002. Safety critical systems: challenges and directions. *Proc. 24th Int. Conf. Softw. Eng. (ICSE, 2002)* (2002), 547–550. DOI:http://dx.doi.org/10.1109/ICSE.2002.1007998

CR Kothari. 2004. *Research methodology: Methods and techniques*,

Masaaki Kurosu. 2013. Human-Computer Interaction. Human-centred Design Approaches, Methods, Tools, and Environments. In *HCI International 2013*. Las Vegas. DOI:http://dx.doi.org/10.1007/978-3-642-39232-0_10

Steinar Kvale. 1996. *Interviews : an introduction to qualitative research interviewing*, Sage Publications.

HA Landsberger. 1958. Hawthorne revisited: management and the worker, its critics, and developments in human relations in industry. (1958).

Jonathan. Lazar, Jinjuan Heidi. Feng, and Harry. Hochheiser. 2010. *Research methods in human-computer interaction*, Wiley.

Leap Motion. 2015. Leap Motion. (2015). Retrieved November 8, 2016 from https://www.leapmotion.com/

Seungyon Lee and Shumin Zhai. 2009. The performance of touch screen soft buttons. In *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. New York, New York, USA: ACM Press, 309. DOI:http://dx.doi.org/10.1145/1518701.1518750

PN Lehner. 1998. *Handbook of ethological methods*, Cambridge University Press.

R. Leitão and PA Silva. 2012. Target and spacing sizes for smartphone user interfaces for older adults: design patterns based on an evaluation with users. *Conf. Pattern Lang. Programs* (2012), 19–21.

James R. Lewis. 2006. Sample sizes for usability tests. *interactions* 13, 6 (November 2006), 29. DOI:http://dx.doi.org/10.1145/1167948.1167973

R. Likert. 1932. A technique for the measurement of attitudes. *Arch. Psychol.* 22, 140 (1932), 55.

Chiuhsiang Joe Lin, Chi No Liu, Chin Jung Chao, and Hung Jen Chen. 2010. The performance of computer input devices in a vibration environment. *Ergonomics* 53, 4 (April 2010), 478–490. DOI:http://dx.doi.org/10.1080/00140130903528186

Min Lin, Rich Goldman, Kathleen J. Price, Andrew Sears, and Julie Jacko. 2007. How do people tap when walking? An empirical investigation of nomadic data entry. *Int. J. Hum. Comput. Stud.* 65, 9 (September 2007), 759–769. DOI:http://dx.doi.org/10.1016/j.ijhcs.2007.04.001

Norman K. Lincoln and Yvonna S. Denzin. 2000. *The SAGE handbook of qualitative research* 3rd ed., Thousand Oaks: SAGE.

LockheedMartin. 2013. Cretaing the blackbrid. (2013). Retrieved May 28, 2016 from http://www.lockheedmartin.co.uk/us/100years/stories/blackbird.html

LockheedMartin. 2014. F-35 simulator comes to Marana. (2014). Retrieved February 15, 2016 from https://www.f35.com/news/detail/f-35-simulator-comes-to-marana

S. Love. 2005. *Understanding mobile human-computer interaction*, Butterworth-Heinemann.

I. Mackay and K. Weinstein. 1998. *Asking questions*, CIPD Publishing.

I. Scott MacKenzie, Abigail Sellen, and William a S. Buxton. 1991. A comparison of input devices in element pointing and dragging tasks. *Proc. SIGCHI Conf. Hum. factors Comput. Syst. Reach. through Technol.* , 1978 (1991), 161–166. DOI:http://dx.doi.org/10.1145/108844.108868

Scott I. MacKenzie. 2015. Fitts' throughput and the remarkable case of touch-based target selection. In *Proceedings of the 16th International Conference on Human-Computer Interaction - HCII 2015 (LNCS 9170)*. Switzerland: Springer, 238–249. DOI:http://dx.doi.org/10.1007/978-3-319-20916-6_23

Robert M. Malina. 2004. Secular trends in growth, maturation and physical performance : A review. *Anthropol. Rev.* 67 (2004), 3–31. DOI:http://dx.doi.org/10.1159/0000101076

C. Marshall and G. Rossman. 1989. *Designing qualitative research*, Sage Publications.

Catherine Marshall and Gretchen B. Rossman. 2011. *Designing qualitative research*, Sage.

Martin N. Marshall. 1996. Sampling for qualitative research - Sample size. *Fam. Pract.* 13, 6 (1996), 522–525. DOI:http://dx.doi.org/10.1093/fampra/13.6.522

Nigel Mathers, Nick Fox, and Amanda Hunn. 2009. *Surveys and questionnaires.*, Nottingham.

S.A. McLeod. 2015. Observation methods. (2015). Retrieved July 23, 2016 from www.simplypsychology.org/observation.html

Microsoft. 2016. Microsoft Kinect. (2016). Retrieved August 11, 2016 from https://developer.microsoft.com/en-us/windows/kinect/develop

Microsoft. 2014. *Touch interfaces for windows*,

Sachi Mizobuchi, Mark Chignell, and David Newton. 2005. Mobile text entry - relationship between walking speed and text input task difficulty. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services (MobileHCI'05)*. New York, New York, USA: ACM Press, 122–128. DOI:http://dx.doi.org/10.1145/1085777.1085798

J. Moeller, A. Kerne, and S. Damaraju. 2011. ZeroTouch: A zero-thickness optical multi-touch force field. In *Conference on Human Factors in Computing Systems - Proceedings*. DOI:http://dx.doi.org/10.1145/1979742.1979710

Jon Moeller, N. Lupfer, B. Hamilton, and H. Lin. 2011. intangibleCanvas: free-air finger painting on a projected canvas. *CHI EA '11 Proc. 2011 Annu. Conf. Ext. Abstr. Hum. factors Comput. Syst.* (2011), 1615–1620. DOI:http://dx.doi.org/10.1145/1979742.1979817

David L. Morgan. 1998. Practical strategies for combining qualitative and quantitative methods: applications to health research. *Qual. Health Res.* 8, 3 (May 1998), 362–376. DOI:http://dx.doi.org/10.1177/104973239800800307

J.G. Morrison, E. Forster, E.M. Hitchcock, C.A. Barba, T.P. Santarelli, and M.W. Scerbo. 1994. Cumulative effects of +Gz on cognitive performance. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* 38, 1 (October 1994), 46–50. DOI:http://dx.doi.org/10.1177/154193129403800108

J.M. Morse. 1991. Approaches to qualitative-quantitative methodological triangulation. *Nurs. Res.* 40, 2 (1991), 120–123. DOI:http://dx.doi.org/10.1097/00006199-199103000-00014

Victor C. de Munck and Elisa J. Sobo. 1998. *Using methods in the field : a practical introduction and casebook*, Rowman Altamira.

Kate Murphy. 2011. The paperless cockpit. *New York Times* (July 2011), B6.

Daniel Natapov, Steven J. Castellucci, and I. Scott MacKenzie. 2009. *ISO 9241-9 evaluation of video game controllers*, Canadian Information Processing Society.

National Transportation Safety Board. 2010. *A review of flightcrew-involved major accicdents of U.Si air carriers, 1978 through 1990*,

Hugo Nicolau and Joaquim Jorge. 2012. Touch typing using thumbs. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*. New York, New York, USA: ACM Press, 2683. DOI:http://dx.doi.org/10.1145/2207676.2208661

J. Nielsen. 1994. *Usability engineering*, Taylor & Francis.

Donald Norman. 1988. *The Design of Everyday Things*, New York: Basic Books.

Jan M. Noyes and Alison F. Starr. 2007. A comparison of speech input and touch screen for executing checklists in an avionics application. *Int. J. Aviat. Psychol.* 17, 3 (June 2007), 299–315. DOI:http://dx.doi.org/10.1080/10508410701462761

Anthony Onwuegbuzie and Nancy Leech. 2006. Linking research questions to mixed methods data analysis procedures 1. *Qual. Rep.* 11, 3 (2006).

J. Orlansky and J. String. 1977. *Cost-Effectiveness of flight simulators for military training. Volume 1. use and effectiveness of flight simulators*,

Tony Osborne. 2013. Thales eyes touch-screen avionics for commercial jets by 2020. (2013). Retrieved December 31, 2016 from http://aviationweek.com/awin/thales-eyes-touch-screen-avionics-commercial-jets-2020

Arbuckle P. Douglas, Abbott Kathy H., Abbott Terence S., and Schutte Paul C. 1998. Future flight decks. In *21st Congress International Council of the Aeronautical Sciences*. Melbourne, 98–6.9.3.

Michael T. Palmer, William H. Rogers, Hayes N. Press, Kara A. Latorella, and Terence S. Abbott. 1995. *A* *crew-centered flight deck design philosophy for High-Speed Civil Transport (HSCT) aircraft*, NASA Langley Technical Report Server.

Marc A. Le Pape and Ravi K. Vatrapu. 2009. An experimental study of field dependency in altered Gz environments. In *Proceedings of the 27th international conference on Human Factors in Computing Systems - CHI 09*. New York: ACM Press, 1255. DOI:http://dx.doi.org/10.1145/1518701.1518890

Pekka Parhi, Amy K. Karlson, and Benjamin B. Bederson. 2006. Target size study for one-handed thumb use on small touchscreen devices. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services - MobileHCI '06*. New York, New York, USA: ACM Press, 203. DOI:http://dx.doi.org/10.1145/1152215.1152260

Yong S. Park and Sung H. Han. 2010. Touch key design for one-handed thumb interaction with a mobile phone: Effects of touch key size and touch key location. *Int. J. Ind. Ergon.* 40, 1 (2010), 68–76. DOI:http://dx.doi.org/10.1016/j.ergon.2009.08.002

HR Parsaei and WG Sullivan. 2012. *Concurrent engineering: contemporary issues and modern design tools*,

A. Patenaude. 1996. *Study on the effectiveness of modeling and simulation in the weapon system acquisition process.*, Washington DC.

Jason Paur. 2011. FAA approves iPads for pilots' electronic charts. (2011). Retrieved October 17, 2016 from http://edition.cnn.com/2011/TECH/innovation/02/28/FAA.approves.iPads/index.html

Kb Perry and Jp Hourcade. 2008. Evaluating one handed thumb tapping on mobile touchscreen devices. *Proc. Graph. Interface 2008* (2008), 57–64.

S. Pheasant and CM Haslegrave. 2005. *Bodyspace: Anthropometry, ergonomics and the design of work*,

Edwards H. Philips. 2006. F-35 cockpit tech. *Aviat. Week Space Technol.* 164, 6 (2006), 59.

Emilie Poirson, Matthieu Delangle, and Ecole Centrale De Nantes. 2013. Comparative analysis of human modeling tools DHM tools comparison : methodology. *2nd Int. Digit. Hum. Model. Symp.* (2013), 1–7.

Giaime Porcu. 2013. Thales unveils avionics 2020, the cockpit of the future. (2013). Retrieved October 17, 2016 from https://www.thalesgroup.com/en/worldwide/aerospace/press-release/thales-unveils-avionics-2020-cockpit-future

Jenny Preece, Yvonne Rogers, and Helen Sharp. 2002. *Interaction design: beyond human-computer interaction*, J. Wiley & Sons.

ProtoSee. 2016. Protosee - an offline HTML prototypes viewer for iOS. (2016). Retrieved November 26, 2016 from protosee.com

C. Pschierer, T. Thompson, R. Ellerbrock, and S. Haffner. 2012. From captain Jeppesen's little black book to the iPad and beyond. In *2012 IEEE/AIAA 31st Digital Avionics Systems Conference (DASC)*. IEEE, 1A2-1-1A2-11. DOI:http://dx.doi.org/10.1109/DASC.2012.6382277

QinetiQ. 2016. Centrifuge. (2016). Retrieved February 2, 2016 from http://www.qinetiq.com/about-us/our-history/our-facilities/Pages/centrifuge.aspx

Michael A. Ragland. 1987. Application of ground/air data link to general aviation operations. *IEEE Aerosp. Electron. Syst. Mag.* 2, 6 (June 1987), 15–19. DOI:http://dx.doi.org/10.1109/MAES.1987.5005414

B.C. Read. 1996. Developing the next generation cockpit display system. *IEEE Aerosp. Electron. Syst. Mag.* 11, 10 (1996), 25–28. DOI:http://dx.doi.org/10.1109/62.538798

H. De Ree and Gerco Blok. 1993. Font to use in cockpit documentation. In *7th International symposium --- 1993 Apr: Columbus; OH*. 115–118.

Florence Reuzeau and Rene Nibbelke. 2012. Flight deck design process. In Don Harris, ed. *Human factors for civil flight deck design*. Farnham: Ashgate, 33–54.

V. Riley, E. Lyall, and E. Wiener. 1993. *Analytic methods for flight-deck automation design and evaluation, phase two report: Pilot use of automation*, Washington DC.

Victor Riley. 1996. What avionics engineers should know about pilots and automation. *IEEE Aerosp. Electron. Syst. Mag.* 11, 5 (1996), 3–8. DOI:http://dx.doi.org/10.1109/62.494182

Colin. Robson. 2002. *Real world research: a resource for users of social research methods in applied settings.*, Wiley.

Rockwell Collins. 2016. Boeing 777X to feature touchscreen flight displays from Rockwell Collins. (2016). Retrieved December 29, 2016 from https://www.rockwellcollins.com/Data/News/2016-Cal-

Yr/CS/FY16CSNR51-777x.aspx

Rockwell Collins. 2012. Flight testing for aviation's first touch-screen primary flight displays takes off at Rockwell Collins. (2012). Retrieved January 20, 2015 from http://www.rockwellcollins.com/Data/News/2012

Rockwell Collins. 2015. Pro Line Fusion. (2015). Retrieved December 12, 2015 from https://www.rockwellcollins.com/Data/Products/Integrated_Systems/Flight_Deck/Pro_Line_Fusion.aspx

William H. Rogers and Paul C. Schutte. 1997. Clean slate flight deck design- Theory based solutions provide useful design goals. In *International Symposium on Aviation Psychology 9th*. Columbus, 859–864.

Patrick Ross. 2004. Human factors Issues of the Aircraft Checklist. *J. Aviat. Aerosp. Educ. Res.* , Winter (2004), 9–14.

Su Rune, Xue Hongjun, and Song Bifeng. 2008. Ergonomic assessment method for cockpit layout of civil aircraft x based on virtual design. In *Proceedings of 26th International Congress of the Aeronautical Sciences*.

Karl. Sabbagh. 1996. *21st century jet: the making of the Boeing 777*, Pan Books.

Maggi Savin-Baden and Claire Howell. Major. 2012. *Qualitative research : the essential guide to theory and practice*, Routledge.

M. Schedlbauer. 2007. Effects of key size and spacing on the completion time and accuracy of input tasks on soft keypads using trackball and touch input. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* 51, 5 (October 2007), 429–433. DOI:http://dx.doi.org/10.1177/154193120705100501

Bastian Schildbach and Enrico Rukzio. 2010. Investigating selection and reading performance on a mobile phone while walking. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI '10*. New York, New York, USA: ACM Press, 93. DOI:http://dx.doi.org/10.1145/1851600.1851619

Rudy Schusteritsch, Carolyn Y. Wei, and Mark LaRosa. 2007. Towards the perfect infrastructure for usability testing on mobile devices. In *CHI '07 extended abstracts on Human factors in computing systems - CHI '07*. New York, New York, USA: ACM Press, 1839. DOI:http://dx.doi.org/10.1145/1240866.1240908

Andrew Sears. 1991. Improving touchscreen keyboards: design issues and a comparison with other devices. *Interact. Comput.* 3, 3 (December 1991), 253–269. DOI:http://dx.doi.org/10.1016/0953-5438(91)90016-U

Andrew Sears and Julie A. Jacko. 2012. *Human computer interaction handbook: Fundamentals, evolving technologies, and emerging applications* Andrew Sears, ed., New York: Taylor & Francis.

Andrew Sears, Doreen Revis, Janet Swatski, Rob Crittenden, and Ben Shneiderman. 1993. Investigating touchscreen typing - The effect of keyboard size on typing speed. *Behav. Inf. Technol.* 12, 1 (January 1993), 17–22. DOI:http://dx.doi.org/10.1080/01449299308924362

Andrew Sears and Ying Zha. 2003. Data entry for mobile devices using soft keyboards - understanding the effects of keyboard size and user tasks. *Int. J. Hum. Comput. Interact.* 16, 2 (October 2003), 163–184. DOI:http://dx.doi.org/10.1207/S15327590IJHC1602_03

SESARJU. 2016. SESAR Joint undertaking. (2016). Retrieved October 17, 2016 from http://www.sesarju.eu/

M.K. Shamo, R. Dror, and A. Degani. 1998. Evaluation of a new cockpit device: The integrated electronic information system. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* 42, 1 (October 1998), 138–142. DOI:http://dx.doi.org/10.1177/154193129804200131

Marcia Kuskin Shamo, Ravit Dror Avionitek, Ltd Haifa, and Israel Asaf Degani. 1999. A multi-dimensional evaluation methodology for new cockpit systems. In *Proceedings of the 10th International Aviation Psychology Symposium*. Columbus.

KJ Sharma and B. Bowonder. 2004. The making of Boeing 777: a case study in concurrent engineering. *Int. J.* (2004).

Roger N. Shepard. 1967. Recognition memory for words, sentences, and pictures. *J. Verbal Learning Verbal Behav.* 6, 1 (1967), 156–163. DOI:http://dx.doi.org/10.1016/S0022-5371(67)80067-7

Ben Shneiderman. 1997. Direct manipulation for comprehensible, predictable and controllable user interfaces.

In *Proceedings of the 2nd international conference on Intelligent user interfaces - IUI '97*. New York, New York, USA: ACM Press, 33–39. DOI:http://dx.doi.org/10.1145/238218.238281

Siemens Industry Software Limited. 2013. Jack DHM. (2013). Retrieved January 2, 2016 from www.plm.automation.siemens.com/en_gb/products/

David Silverman. 2009. *Doing qualitative research*, Sage Publications.

R.D. Smith. 1998. Designing a model of a real or imagined system and conducting experiments with that model. In *Encyclopedia of Computer Science*.

C. Snyder. 2003. *Paper prototyping: The fast and easy way to design and refine user interfaces*,

JA Sokolowski and CM Banks. 2011. *Principles of modeling and simulation: a multidisciplinary approach*,

R. William Soukoreff and I. Scott MacKenzie. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *Int. J. Hum. Comput. Stud.* 61, 6 (2004), 751–789. DOI:http://dx.doi.org/10.1016/j.ijhcs.2004.09.001

Carry Spitzer, Uma Ferell, and Thomas Ferrell. 2000. *Digital Avionics Handbook*, CRC.

Neville A. Stanton, Catherine Harvey, Katherine L. Plant, and Luke Bolton. 2013. To twist, roll, stroke or poke? A study of input devices for menu navigation in the cockpit. *Ergonomics* 56, 4 (2013), 590–611. DOI:http://dx.doi.org/10.1080/00140139.2012.751458

Jane E. Stentz, Vicki L. Plano Clark, and Gina S. Matkin. 2012. Applying mixed methods to leadership research: A review of current practices. *Leadersh. Q.* 23, 6 (2012), 1173–1183. DOI:http://dx.doi.org/10.1016/j.leaqua.2012.10.001

R.L. Sumwalt. 1991. Checking the checklist. *Prof. Pilot* (1991), 62–64.

Xu Sun and Andrew May. 2013. A comparison of field-based and lab-based experiments to evaluate user experience of personalised mobile devices. *Adv. Human-Computer Interact.* 2013 (2013), 1–9. DOI:http://dx.doi.org/10.1155/2013/619767

Jenny E. Symonds and Stephen Gorard. 2010. The death of mixed methods: research labels and their casualties. *Br. Educ. Res. Assoc.* (2010), 1–19. DOI:http://dx.doi.org/10.1080/09500790.2010.483514

Abbas. Tashakkori and Charles. Teddlie. 1998. *Mixed methodology : combining qualitative and quantitative approaches*, Sage.

Edwin van Teijlingen and Vanora Hundley. 1998. The importance of pilot studies. *Nurs. Stand.* 16, 40 (1998), 33–36. DOI:http://dx.doi.org/10.7748/ns2002.06.16.40.33.c3214

Thales. 2014. Thales unveils avionics 2020 for helicopters. (2014). Retrieved June 6, 2014 from https://www.thalesgroup.com/en/worldwide/aerospace/press-release/thales-unveils-avionics-2020-helicopters

Thales. 2015. Touchscreens in the cockpit: pure hype or genuinely helpful? (2015). Retrieved December 31, 2016 from http://onboard.thalesgroup.com/2015/05/22/touchscreens-cockpit-pure-hype-genuinely-helpful/

Matt Thurber. 2015. Touchscreens clean up Gulfstream symmetry flight deck. (2015). Retrieved December 6, 2015 from http://www.ainonline.com/aviation-news/business-aviation/2015-01-02/touchscreens-clean-gulfstream-symmetry-flight-deck

MA Tinker. 1963. Legibility of print. (1963).

Stephen Trimble. 2016. Boeing selects iPad-like touch screens for 777X flightdeck. (2016). Retrieved December 29, 2016 from https://www.flightglobal.com/news/articles/boeing-selects-ipad-like-touch-screens-for-777x-flig-426890/

Matthieu B. Trudeau, Deanna S. Asakawa, Devin L. Jindrich, and Jack T. Dennerlein. 2016. Two-handed grip on a mobile phone affords greater thumb motor performance, decreased variability, and a more extended thumb posture than a one-handed grip. *Appl. Ergon.* 52 (January 2016), 24–28. DOI:http://dx.doi.org/10.1016/j.apergo.2015.06.025

Steve N.H. Tsang, H.S. Chan, and K. Chen. 2013. A study on touch screen numeric keypads: effects of key size and key layout. *Proc. Int. MultiConference Eng. Comput. Sci.* II (2013), 13–16.

Ubuntu. 2008. *UME guide / designing for finger UIs*,

C. Vieyra and R. Vieyra. 2015. Physics Toolbox Apps. (2015). Retrieved June 17, 2015 from www.vieyrasoftware.net

LE Wallace. 1994. Airborne Trailblazer: Two decades with NASA Langley's 737 flying laboratory. (1994).

Simon Wallner. 2010. An interactive visualisation of fitts's law with JavaScript and D3. (2010). Retrieved November 6, 2016 from http://simonwallner.at/ext/fitts/

L. Wang, Q. Cao, J. Chang, and C. Zhao. 2015. The effect of touch-key size and shape on the usability of flight deck MCDU. In *ACHI 2015: The Eight International Conference on Advances in Computer-Human Interactiions*. IARIA, 234–238.

D. Watkins, D. Gioia, and DC Watkins. 2015. *Mixed methods research*, Oxford University Press.

Susan Wiedenbeck. 1999. The use of icons and labels in an end user application program: An empirical study of learning and retention. *Behav. Inf. Technol.* 18, 2 (January 1999), 68–82. DOI:http://dx.doi.org/10.1080/014492999119129

Wikipedia. 2015. Display aspect ratio. (2015). Retrieved June 6, 2015 from ttp://en.wikipedia.org/wiki/Display_aspect_ratio.

Miles D. Woken. 2013. *Advantages of a pilot study*, Springfield.

Wright Brothers Aeroplane Company. 2010. Wright Flyer. (2010). Retrieved May 5, 2016 from http://www.wright-brothers.org/

Linda Wulf, Markus Garschall, Michael Klein, and Manfred Tscheligi. 2015. Young vs old – landscape vs portrait: a comparative study of touch gesture performance. *J. Assist. Technol.* 9, 3 (September 2015), 136–146. DOI:http://dx.doi.org/10.1108/JAT-10-2014-0029

J. Xiong, S. Muraki, and K. Fukumoto. 2014. The effects of touch button size on touchscreen operability. *J. Mech. Eng. Autom.* 4 (2014), 667–672.

Appendix I

Participant Information Sheet

Target Sizes for Interactive Displays in Vibrating Environments

Aims

The purpose of this research is to establish design guidelines and recommendations for target sizes on fixed and mobile touchscreens on the flight deck. The study will investigate the impact of vibration and turbulence to targeting accuracy and movement time on touch enabled devices.

Experiment

A tablet application has been created for this experiment. Participants task is tap or drag targets of different sizes to a specific location. The experiment will start with base line determination which will be conducted on ground. The second stage is conducting the experiment while flying (in training flights). This will highlight the negative effect of vibration and turbulence to the overall performance in two modes of use. The experiment will be conducted with the rear personal. They have the freedom to do the experiments in time frames where the rear personal has no task to do and this will not cause any safety issue. In the mobile mode participants, will hold the device while he/she conducts the experiment. In the fixed mode, the device will be attached to a fixture.

Data Generation & Collection

Investigator will define the sequence of the experiments. As mentioned before participants will decide when they will conduct the experiment. Accuracy and movement time will be recorded on the device. Another device will be used to measure the vibration. A camera will record the experiment for post-hoc analyses. However, participants have the option to reject that and withdraw from the experiment at any time. Participants should know that the investigator is not testing the performance of the participants. Investigator is interested to see which target sizes are easier to tap or drag and what happens if vibration changes during the flight. Participating in this research is voluntary and participants may withdraw from the research project at any stage without prejudice or negative consequences. This will include the deletion of any data that they have generated up until that point, even if it is after the experiment has finished.

Confidentially of Personal Details

Personal details of participant will be not shared with anybody else, except members of the University of Nottingham. Participant will not be identifiable in any published material. Data will be kept in accordance with the Data Protection Act. Participants could contact the investigator or supervisor if they require further information about the research, and they may contact the Research Ethics Coordinator of the School of Education, University of Nottingham, if they wish to make a complaint relating to their involvement in this research.

Appendix II

Participant Information Sheet

Pilot Interaction with Aircraft System

Aims

The aim of this study is to understand how pilots currently using the devices, located on the pedestal. Another aim is to prioritize the devices according to the frequency and duration of interaction. In addition to that the investigator is interested in impact of mission type, environment and time of operation to the frequency of usage.

Experiment

Leading suppliers for cockpit equipment like Honeywell, Thales and Rockwell Colllins are currently performing research about the integration of touch screens in and around the cockpit. GE Aviation is working on a design specialized for para public operations. In this design, a single touchscreen control and display unit (TCDU) should be used for complex and strategic system interaction. Basically, the touchscreen device should replace (or compress) all components, which are fitted in the center console.

A questionnaire dealing with the demographics will be filled by the participants. The research will start with a virtual flight. In this stage both pilots will sit in the cockpit and asked to perform a virtual flight (on ground). It is requested to think load and explain each step they are conducting. Investigator will take notes about the procedure and way of interaction. The second stage will be conducted during the training flight. Investigator sitting in the back will record the interactions. There is no additional task for pilots during the flight. The last stage of the experiment is a post interview where the investigator will discuss (summarize) the flight with pilots. Pilots will be asked what they would change in the cockpit and describe the problems they are facing with current interface.

Data Generation & Collection

The investigator will use an app to record the interactions. The flight will be video recorded for post analyses. Recordings will be not used to monitor the performance of pilots and it is entirely for research purposes. Gained information will be used to improve technologies. Pilots could finish and withdraw from the experiments at any point. This will include the deletion of any data that they have generated up until that point, even if it is after the experiment has finished.

Confidentially of Personal Details

Personal details of participant will be not shared with anybody else, except members of the University of Nottingham. Participant will not be identifiable in any published material. Data will be kept in accordance with the Data Protection Act. Participants could contact the investigator or supervisor if they require further information about the research, and they may contact the Research Ethics Coordinator of the School of Education, University of Nottingham, if they wish to make a complaint relating to their involvement in this research.

Appendix III

Participant Information Sheet

Features, content and functionality requirements for EFB’s

Aims

The aim of the research is to figure out the features, content and functionality that pilots would like to see in an electronic flight bag.

Experiment

The investigator will perform an interview with pilots. He will ask what kind of information they require in daily basis and how this information is gathered. Once, the interview is finished, participant will receive a scenario describing the daily routine of a search and rescue pilot, who use his tablet pc to perform various tasks. Pilots are asked to tick the points what they would prefer to see in an EFB app. After the experiment the investigator will create a “card sorting task”. Pilots are asked to group and label the features how they would like.

Data Collection

The investigator will collect the worksheets to analyze the data. The experiment will be video recorded for post-hoc analyses. Participants will be not identifiable on the recordings. However, pilots could finish and withdraw from the experiments at any point. This will include the deletion of any data that they have generated up until that point, even if it is after the experiment has finished.

Confidentially of Personal Details

Personal details of participant will be not shared with anybody else, except members of the University of Nottingham. Participant will not be identifiable in any published material. Data will be kept in accordance with the Data Protection Act. Participants could contact the investigator or supervisor if they require further information about the research, and they may contact the Research Ethics Coordinator of the School of Education, University of Nottingham, if they wish to make a complaint relating to their involvement in this research.

Appendix IV

Participant Information Sheet

Impact of Display Position and G-Force to the Usability of Touchscreens

Aims

The aim of this research is to explore the potential impact of different display positions and the increase in G-Force to the usability of touchscreens.

Experiment

First Experiment

At the beginning, participants will fill a questionnaire dealing with the demographics. A tablet, which is fixed on a tripod, will be used for the experiment. The experiment will be conducted in the MRL lab area. The task (ISO 9241) is to tap targets, which are displayed in sequential order, while sitting on a chair. After each sequence the investigator will record the results, which will give participants time to recover. Once a session is completed, the investigator will change the position of the display and request to repeat the task. It is requested to take a break if participants feel fatigue in their arms.

There are 20 different positions. Completing tasks for one particular position takes in average two minutes. The experiments will be performed in two sessions (10 position per session). After each experiment participant, will fill a questionnaire (taken from ISO 9241) regarding the physical and cognitive effort. A semi-structured post hoc interview will be performed to gather feedback. These interviews will be recorded (audio) for further analyses.

Second Experiment

The same experimental setup and task will be used during the second experiment. Increased G-Force which is likely to occur on a Fighter Jet will be emulated by adding weights to a wrist band that participant will wear. Depending on the weight of participants arms additional weight will be added so it will simulate 2G and 3G turns.

Data Generation & Collection

The investigator will record overall results simultaneously. Raw data like, movement time, error rate, touch position, target position, target size and distance between target will be stored locally on the tablet for further analyses. Recordings will be used entirely for research purposes. Participants could finish and withdraw from the experiments at any point. This will include the deletion of any data that they have generated up until that point, even if it is after the experiment has finished.

Confidentially of Personal Details

Personal details of participant will be not shared with anybody else, except members of the University of Nottingham who are involved in this study. Participant will be not identifiable in any published material. Data will be kept in accordance with the Data Protection Act. Participants could contact the investigator or supervisor if they require further information about the research, and they may contact the Research Ethics Coordinator of the School of Education, University of Nottingham, if they wish to make a complaint relating to their involvement in this research.

Appendix V

Participant Information Sheet

Input Devices for Future Flight Decks

Aims

The research will focus on change, set and manipulate radio frequencies of COM, NAV and XPDR devices. The aim is to understand how information is received and processed by the pilot's currently and how new technologies and interaction strategies could support this process. This experiment will evaluate touchscreen technology for data input and compare it with the current system.

Experiment

The experiment will start with a questionnaire dealing with demographics and personal experience with smart devices (tablet pc and smartphones). A short interview will be used to review human-human (how is information received) and human-computer (how is information processed) interaction in this particular topic. From this interview, possible scenarios will be developed which occur in this specific area. Taking this scenarios as a base, the main task is to set, manipulate and change frequencies. Investigator will introduce participant into the new way of interaction with touch screens. After familiarization, the investigator will give participants the task written on a paper (e.g. set COM1 to 121.900). Participant will perform the tasks on the current system (via FMS) as well as on the newly developed system. After the experiments pilots will be asked to fill another questionnaire describing their experience with new interaction method. In the last part of the experiment, a discussion will be performed about the pros and cons of the interaction strategy in respect to the scenarios.

Data Generation & Collection

Investigator will define the sequence of the experiments. The investigator will record time on task and error rate. The experiments will be video recorded for post-hoc analyses. However, participants have the option to reject that. Participants should know that the investigator is not testing the performance of the participants. Investigator interest is on the usability of the device and interaction strategy. Participating in this research is voluntary and participants may withdraw from the research project at any stage without prejudice or negative consequences. This will include the deletion of any data that they have generated up until that point, even if it is after the experiment has finished.

Confidentially of Personal Details

Personal details of participant will be not shared with anybody else, except members of the University of Nottingham. Participant will not be identifiable in any published material. Data will be kept in accordance with the Data Protection Act. Participants could contact the investigator or supervisor if they require further information about the research, and they may contact the Research Ethics Coordinator of the School of Education, University of Nottingham, if they wish to make a complaint relating to their involvement in this research.

Appendix VI

Participant Consent Form

Researcher's name: Huseyin Avsar

Supervisor's name: Prof. Thomas Anthony Rodden, Joel Fischer

- I am over 18 years old
- I have read the Participant Information Sheet and the nature and purpose of the research project has been explained to me. I understand and agree to take part.
- I understand the purpose of the research project and my involvement in it.
- I understand that I may withdraw from the research project at any stage and that this will not affect my status now or in the future.
- I understand that I will be videotaped/audiotaped during the experiment.
- I understand that while information gained during the study may be published, I will not be identified and my personal results will remain confidential.
- I understand that data will be stored by the University of Nottingham where only members of the University will have access to it.
- I can finish and withdraw from the experiments at any point. This will include the deletion of any data that they have generated up until that point, even if it is after the experiment has finished
- I understand that I may contact the researcher or supervisor if I require further information about the research, and that I may contact the Research Ethics Coordinator of the School of Education, University of Nottingham, if I wish to make a complaint relating to my involvement in the research.

Signed ………………………………………………… (research participant)

Print name ……………………………… Date ………………………………

Contact details

Researcher: Huseyin Avsar
psxha6@nottingham.ac.uk
Tel: 00447453176918
Mixed Reality Lab
School of Computer Science
University of Nottingham

Supervisor: Thomas Anthony Rodden
psztar@nottingham.ac.uk

School of Education Research Ethics Coordinator: educationresearchethics@nottingham.ac.uk

Appendix VII

Pre-Start Participant Questionnaire

Name: ……………………………………………………………………………….

Gender (male/female): ………………………………………..………………..

Age: ……………………………………………………………………………….

Nationality: ……………………………………………………………………...

Numberof Flight Hours: ……………………………………………………………..

Do you use a smartphone? Yes No

Do you use a table pc? Yes No

Rate your touchscreen skill on a scale 1-10 (10 best) : ………………………….

Average usage per day for smartphone & tablet:………………………………….

(for example 3 hours/day)

Please describe situations that triggers you to change, set or manipulate radio frequencies?

…………………………………………………………………………………………
…………………………………………………………………………………………

Please write your most used applications (max. 5)

Note: This could be any app like; social media, game, mobile banking, messaging etc.

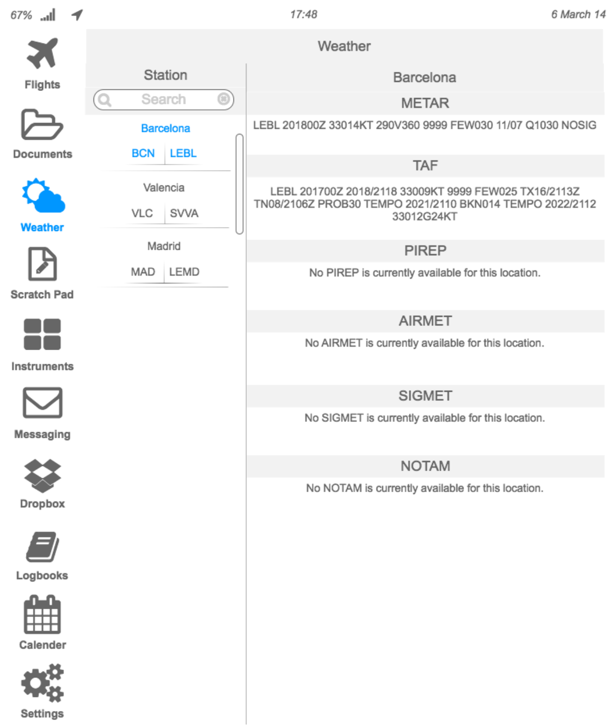
…………………………………………………………………………………………
…..
....

Do you use aviation related applications? If yes, which one (max.5)

Note: e.g. apps for checking the weather, checklist, flight planning, time table etc.

…………………………………………………………………………….....................
....

Appendix VIII



METAR

LEBL 201800Z 33014KT 290V360 9999 FEW030 11/07 Q1030 NOSIG

TAF

LEBL 201700Z 2018/2118 33009KT 9999 FEW025 TX16/2113Z TN08/2106Z PROB30 TEMPO 2021/2110 BKN014 TEMPO 2022/2112 33012G24KT

PIREP

No PIREP is currently available for this location.

AIRMET

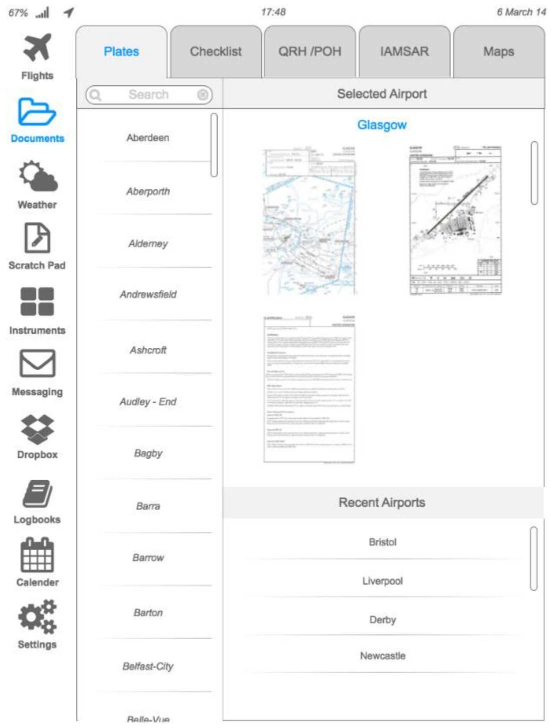
No AIRMET is currently available for this location.

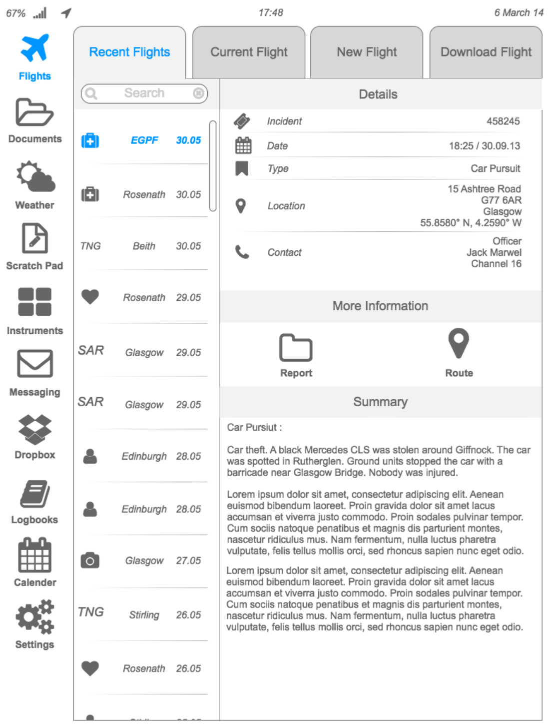
SIGMET

No SIGMET is currently available for this location.

NOTAM

No NOTAM is currently available for this location.





	EGPF	30.05
	Rosenath	30.05
TNG	Beith	30.05
	Rosenath	29.05
SAR	Glasgow	29.05
SAR	Glasgow	29.05
	Edinburgh	28.05
	Edinburgh	28.05
	Glasgow	27.05
TNG	Stirling	26.05
	Rosenath	26.05

Details

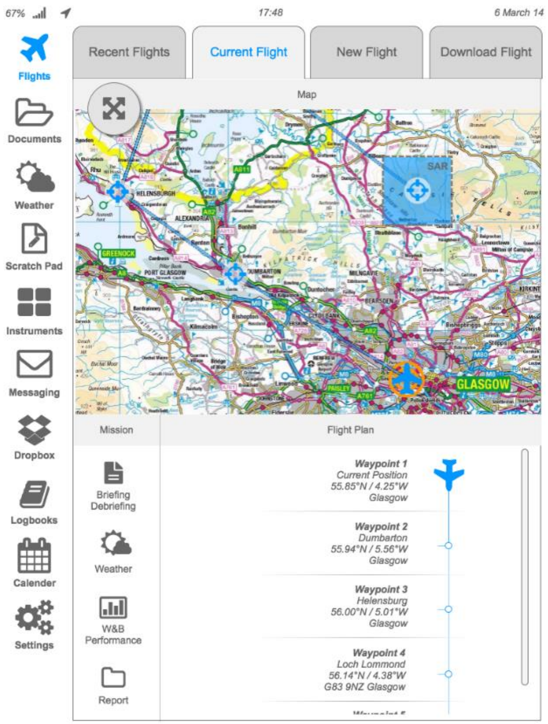
Incident	458245
Date	18:25 / 30.09.13
Type	Car Pursuit
Location	15 Ashtree Road G77 6AR Glasgow 55.8580° N, 4.2590° W
Contact	Officer Jack Marwel Channel 16

Car Pursiut :

Car theft. A black Mercedes CLS was stolen around Giffnock. The car was spotted in Rutherglen. Ground units stopped the car with a barricade near Glasgow Bridge. Nobody was injured.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean euismod bibendum laoreet. Proin gravida dolor sit amet lacus accumsan et viverra justo commodo. Proin sodales pulvinar tempor. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Nam fermentum, nulla luctus pharetra vulputate, felis tellus mollis orci, sed rhoncus sapien nunc eget odio.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean euismod bibendum laoreet. Proin gravida dolor sit amet lacus accumsan et viverra justo commodo. Proin sodales pulvinar tempor. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Nam fermentum, nulla luctus pharetra vulputate, felis tellus mollis orci, sed rhoncus sapien nunc eget odio.

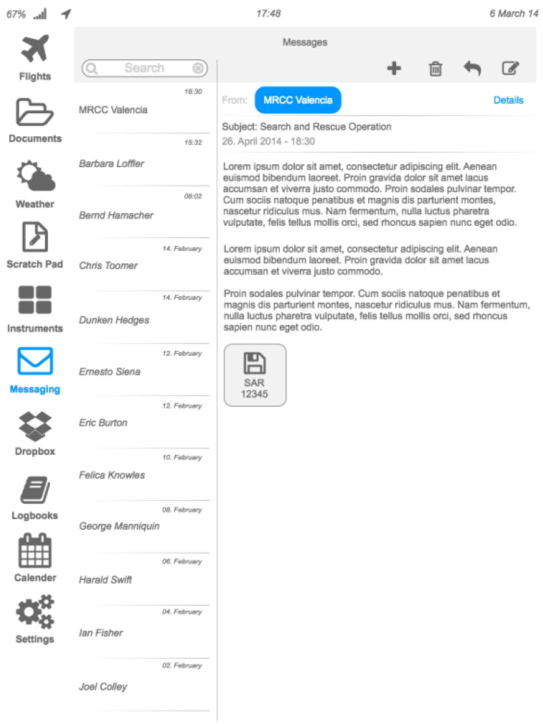


Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow

Waypoint 2
Dumbarton
55.94°N / 5.56°W
Glasgow

Waypoint 3
Helensburg
56.00°N / 5.01°W
Glasgow

Waypoint 4
Loch Lommond
56.14°N / 4.38°W
G83 9NZ Glasgow



MRCC Valencia	18:30
Barbara Loffler	15:32
Bernd Hamacher	08:02
Chris Toomer	14. February
Dunken Hedges	14. February
Ernesto Siena	12. February
Eric Burton	12. February
Felica Knowles	10. February
George Manniquin	08. February
Harald Swift	06. February
Ian Fisher	04. February
Joel Colley	02. February

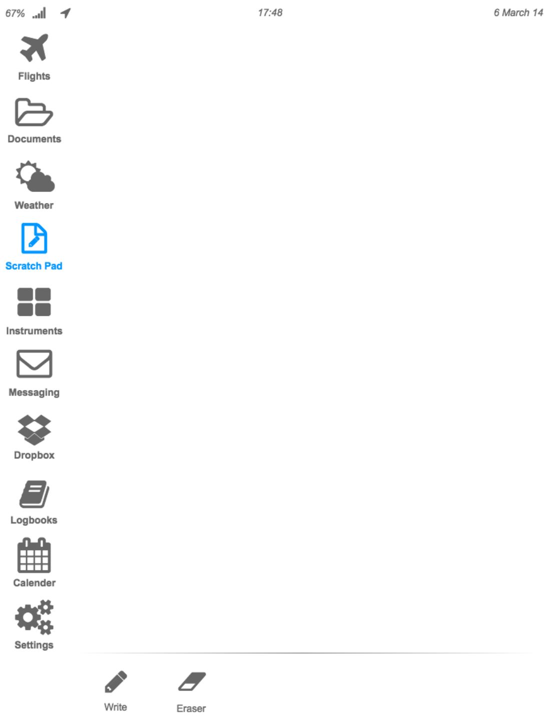
From: MRCC Valencia

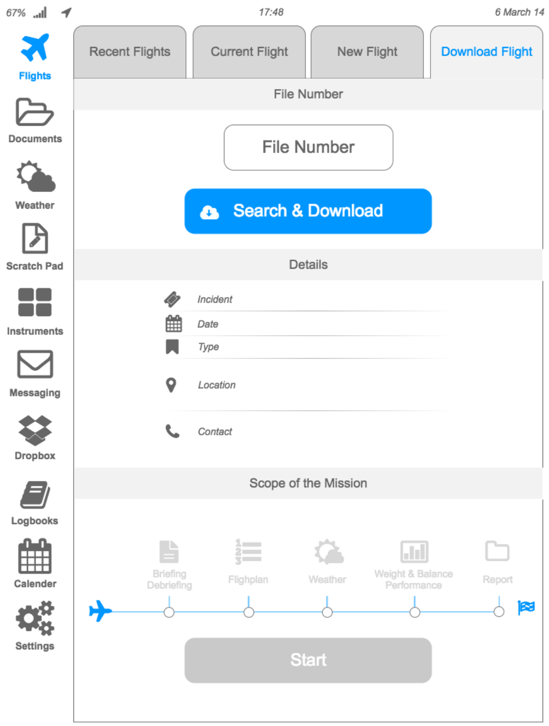
Subject: Search and Rescue Operation
26. April 2014 - 18:30

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean euismod bibendum laoreet. Proin gravida dolor sit amet lacus accumsan et viverra justo commodo. Proin sodales pulvinar tempor. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Nam fermentum, nulla luctus pharetra vulputate, felis tellus mollis orci, sed rhoncus sapien nunc eget odio.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean euismod bibendum laoreet. Proin gravida dolor sit amet lacus accumsan et viverra justo commodo.

Proin sodales pulvinar tempor. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Nam fermentum, nulla luctus pharetra vulputate, felis tellus mollis orci, sed rhoncus sapien nunc eget odio.





File Number

File Number

Details

Incident

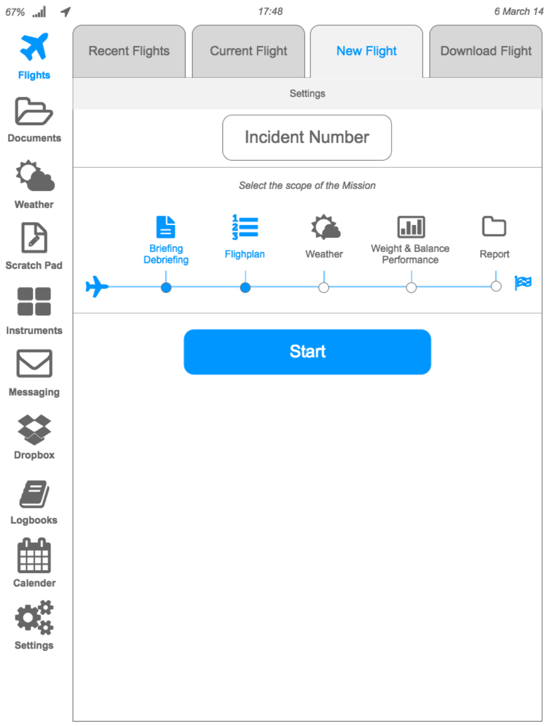
Date

Type

Location

Contact

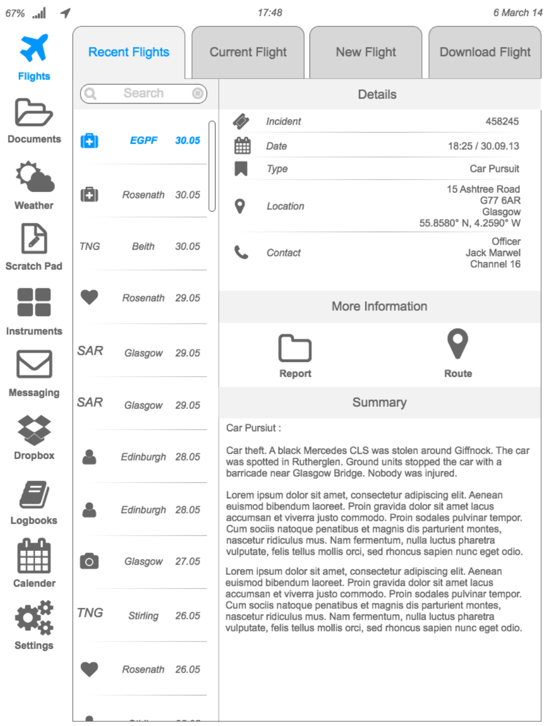
Scope of the Mission



Settings

Incident Number

Select the scope of the Mission



	EGPF	30.05
	Rosenath	30.05
TNG	Beith	30.05
	Rosenath	29.05
SAR	Glasgow	29.05
SAR	Glasgow	29.05
	Edinburgh	28.05
	Edinburgh	28.05
	Glasgow	27.05
TNG	Stirling	26.05
	Rosenath	26.05

Details

Incident	458245
Date	18:25 / 30.09.13
Type	Car Pursuit
Location	15 Ashtree Road G77 6AR Glasgow 55.8580° N, 4.2590° W
Contact	Officer Jack Marwel Channel 16

More Information

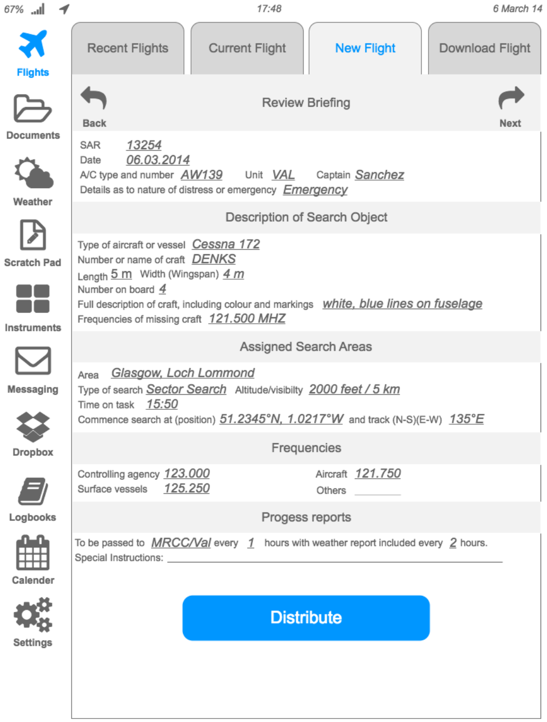
Summary

Car Pursiut :

Car theft. A black Mercedes CLS was stolen around Giffnock. The car was spotted in Rutherglen. Ground units stopped the car with a barricade near Glasgow Bridge. Nobody was injured.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean euismod bibendum laoreet. Proin gravida dolor sit amet lacus accumsan et viverra justo commodo. Proin sodales pulvinar tempor. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Nam fermentum, nulla luctus pharetra vulputate, felis tellus mollis orci, sed rhoncus sapien nunc eget odio.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean euismod bibendum laoreet. Proin gravida dolor sit amet lacus accumsan et viverra justo commodo. Proin sodales pulvinar tempor. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Nam fermentum, nulla luctus pharetra vulputate, felis tellus mollis orci, sed rhoncus sapien nunc eget odio.



Review Briefing

SAR 13254
Date 06.03.2014
A/C type and number AW139 Unit VAL Captain Sanchez
Details as to nature of distress or emergency Emergency

Description of Search Object

Type of aircraft or vessel Cessna 172
Number or name of craft DENKS
Length 5 m Width (Wingspan) 4 m
Number on board 4
Full description of craft, including colour and markings white, blue lines on fuselage
Frequencies of missing craft 121.500 MHZ

Assigned Search Areas

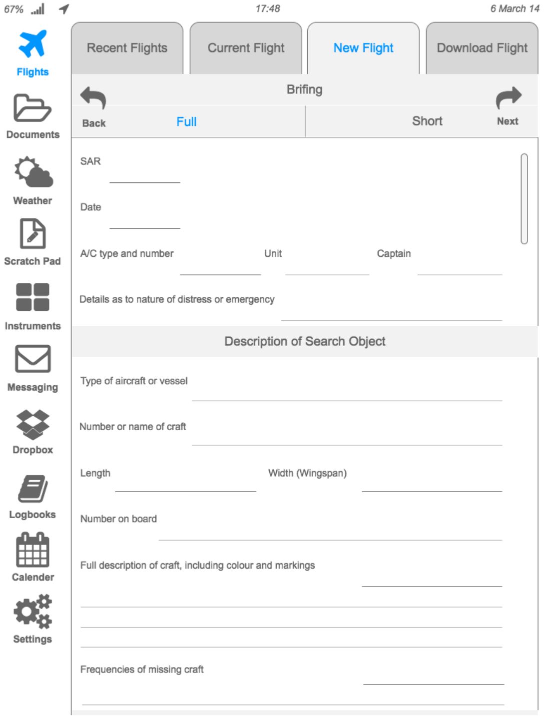
Area Glasgow, Loch Lommond
Type of search Sector Search Altitude/visibilty 2000 feet / 5 km
Time on task 15:50
Commence search at (position) 51.2345°N, 1.0217°W and track (N-S)(E-W) 135°E

Frequencies

Controlling agency 123.000 Aircraft 121.750
Surface vessels 125.250 Others ______

Progess reports

To be passed to MRCC/Val every 1 hours with weather report included every 2 hours.
Special Instructions: ______

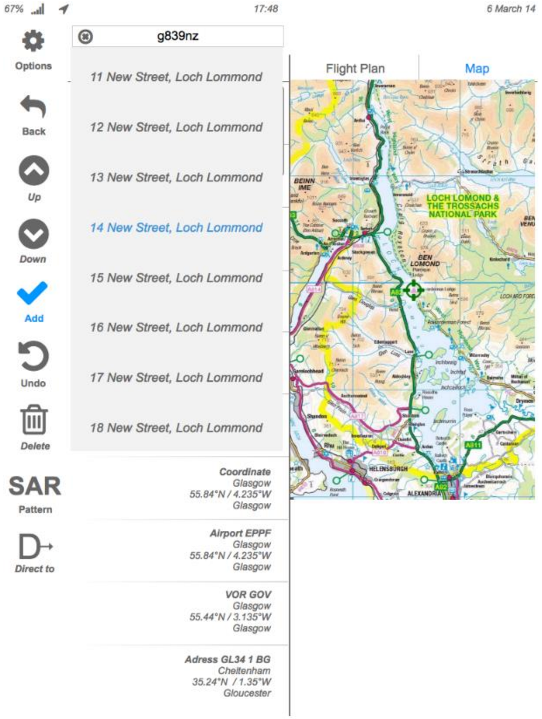


Brifing

SAR ______
Date ______
A/C type and number ______ Unit ______ Captain ______
Details as to nature of distress or emergency ______

Description of Search Object

Type of aircraft or vessel ______
Number or name of craft ______
Length ______ Width (Wingspan) ______
Number on board ______
Full description of craft, including colour and markings ______
Frequencies of missing craft ______



g839nz

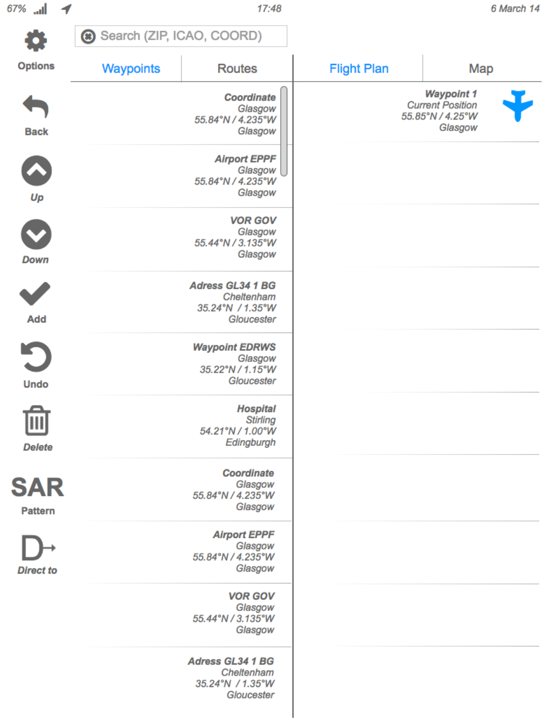
11 New Street, Loch Lommond
12 New Street, Loch Lommond
13 New Street, Loch Lommond
14 New Street, Loch Lommond
15 New Street, Loch Lommond
16 New Street, Loch Lommond
17 New Street, Loch Lommond
18 New Street, Loch Lommond

Coordinate
Glasgow
55.84°N / 4.235°W
Glasgow

Airport EPPF
Glasgow
55.84°N / 4.235°W
Glasgow

VOR GOV
Glasgow
55.44°N / 3.135°W
Glasgow

Adress GL34 1 BG
Cheltenham
35.24°N / 1.35°W
Gloucester



Search (ZIP, ICAO, COORD)

Coordinate
Glasgow
55.84°N / 4.235°W
Glasgow

Airport EPPF
Glasgow
55.84°N / 4.235°W
Glasgow

VOR GOV
Glasgow
55.44°N / 3.135°W
Glasgow

Adress GL34 1 BG
Cheltenham
35.24°N / 1.35°W
Gloucester

Waypoint EDRWS
Glasgow
35.22°N / 1.15°W
Gloucester

Hospital
Stirling
54.21°N / 1.00°W
Edingburgh

Coordinate
Glasgow
55.84°N / 4.235°W
Glasgow

Airport EPPF
Glasgow
55.84°N / 4.235°W
Glasgow

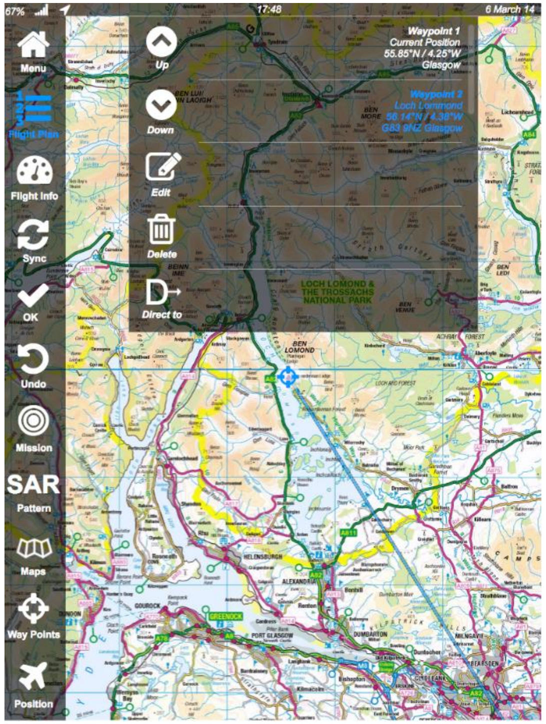
VOR GOV
Glasgow
55.44°N / 3.135°W
Glasgow

Adress GL34 1 BG
Cheltenham
35.24°N / 1.35°W
Gloucester

Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow

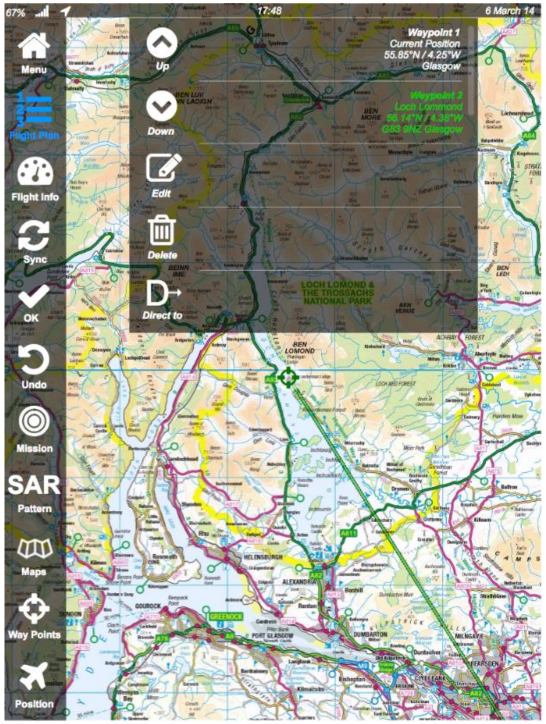
Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow

Speed 120 kts
Altitude 2300 ft
Track 360°
Accuracy 15 m
Position 55.861° N 4.296° W
Next WP 11 NM 0:10 h
Destination 50 NM 0:55 h



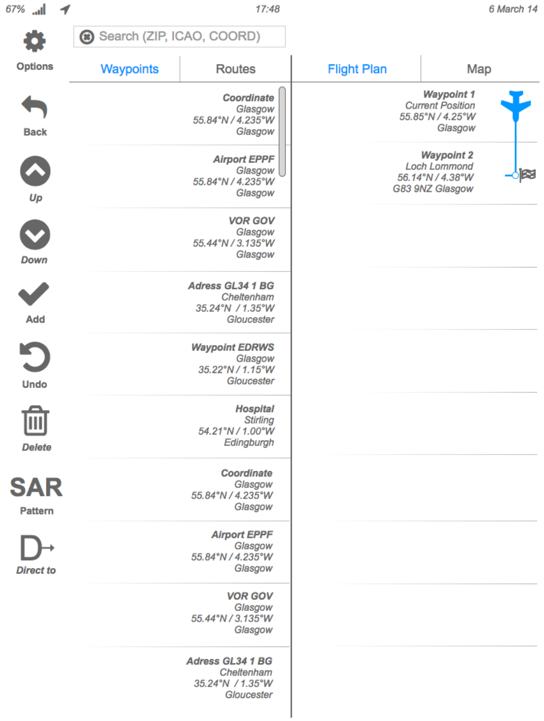
Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow

Waypoint 2
Loch Lommond
56.14°N / 4.38°W
G83 9NZ Glasgow



Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow

Waypoint 2
Loch Lommond
56.14°N / 4.38°W
G83 9NZ Glasgow



Coordinate
Glasgow
55.84°N / 4.235°W
Glasgow

Airport EPPF
Glasgow
55.84°N / 4.235°W
Glasgow

VOR GOV
Glasgow
55.44°N / 3.135°W
Glasgow

Adress GL34 1 BG
Cheltenham
35.24°N / 1.35°W
Gloucester

Waypoint EDRWS
Glasgow
35.22°N / 1.15°W
Gloucester

Hospital
Stirling
54.21°N / 1.00°W
Edingburgh

Coordinate
Glasgow
55.84°N / 4.235°W
Glasgow

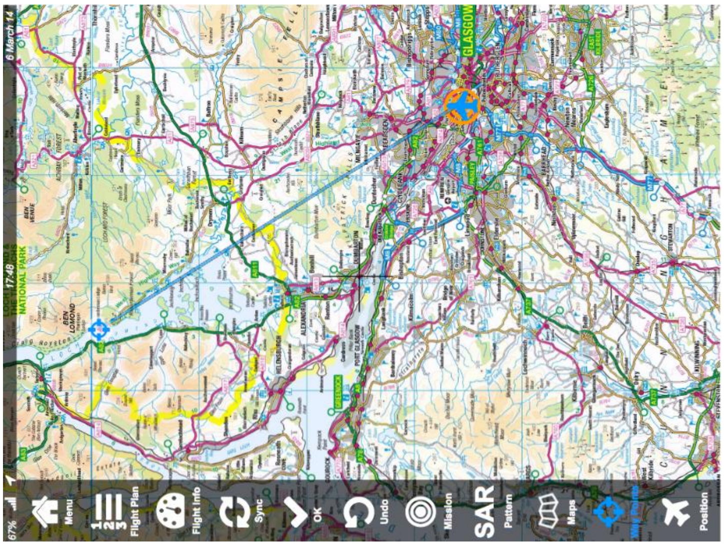
Airport EPPF
Glasgow
55.84°N / 4.235°W
Glasgow

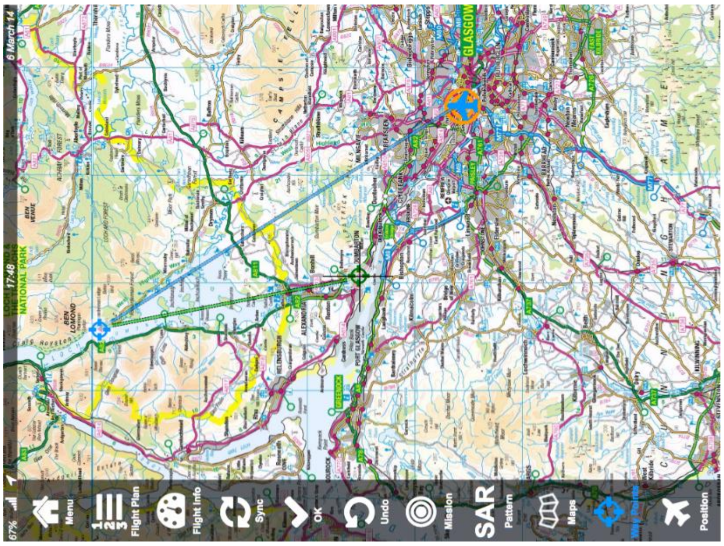
VOR GOV
Glasgow
55.44°N / 3.135°W
Glasgow

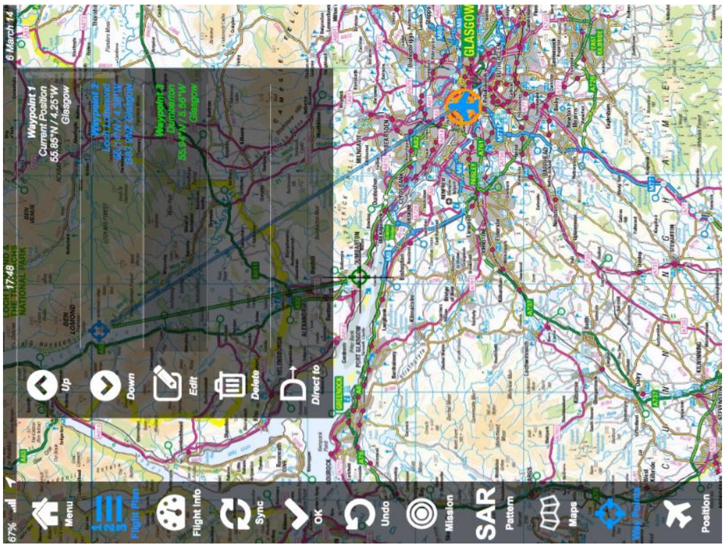
Adress GL34 1 BG
Cheltenham
35.24°N / 1.35°W
Gloucester

Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow

Waypoint 2
Loch Lommond
56.14°N / 4.38°W
G83 9NZ Glasgow

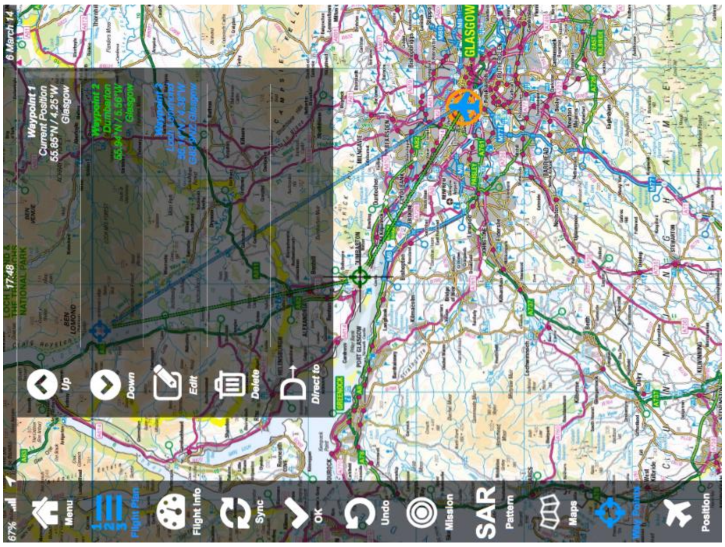


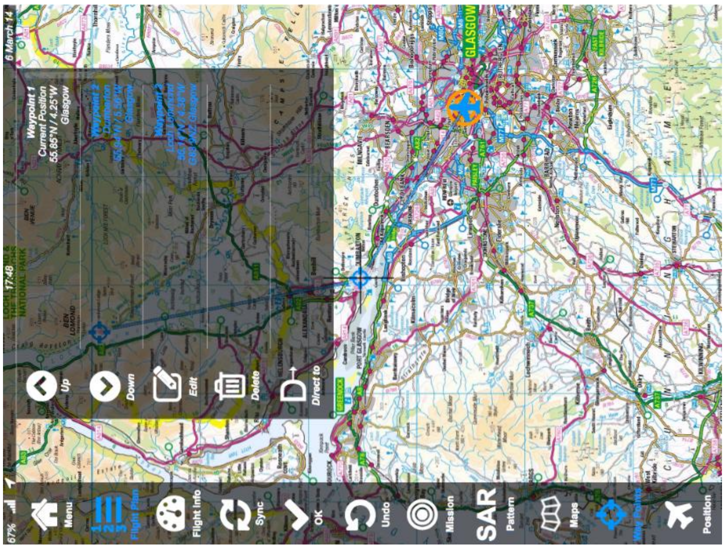


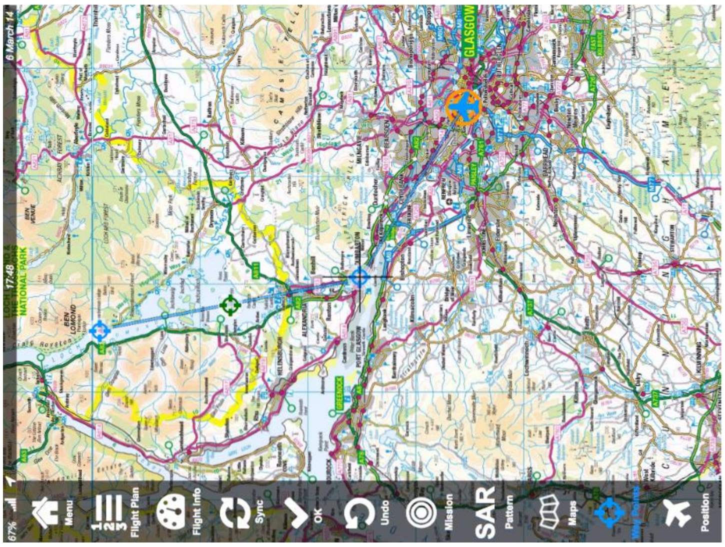

Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow

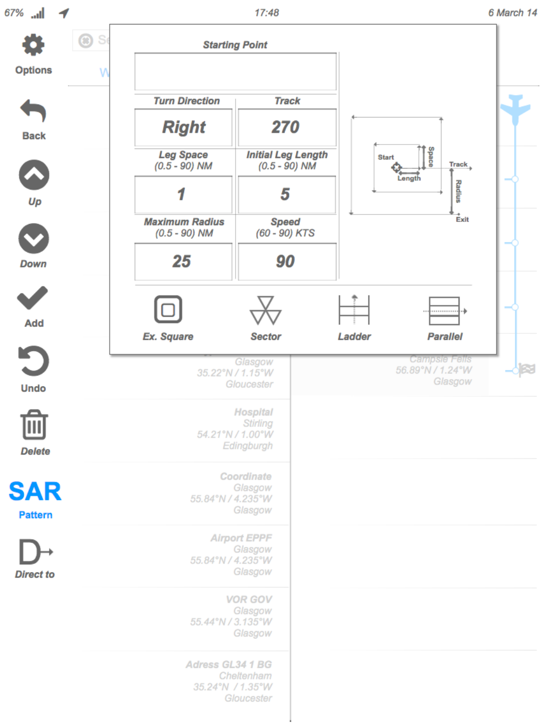
Waypoint 2
Loch Lomond
56.14°N / 4.39°W
G83 9NZ Glasgow

Waypoint 3
Dumbarton
55.94°N / 5.56°W
Glasgow


67%
6 March 14
17:48
Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow
Waypoint 2
Dumbarton
Glasgow
Up
Down
Edit
Delete
Direct to
Menu
Flight Plan
Flight Info
Sync
OK
Undo
Mission
SAR
Pattern
Maps
Way Points
Position
GLASGOW


67%
6 March 14
17:48
Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow
Up
Down
Edit
Delete
Direct to
Menu
Flight Plan
Flight Info
Sync
OK
Undo
Mission
SAR
Pattern
Maps
Way Points
Position
GLASGOW


67%
6 March 14
17:48
Menu
Flight Plan
Flight Info
Sync
OK
Undo
Mission
SAR
Pattern
Maps
Way Points
Position
GLASGOW



Starting Point

Turn Direction	Track
Right	270
Leg Space (0.5 - 90) NM	Initial Leg Length (0.5 - 90) NM
1	5
Maximum Radius (0.5 - 90) NM	Speed (60 - 90) KTS
25	90

Glasgow
35.22°N / 1.15°W
Gloucester

Hospital
Stirling
54.21°N / 1.00°W
Edingburgh

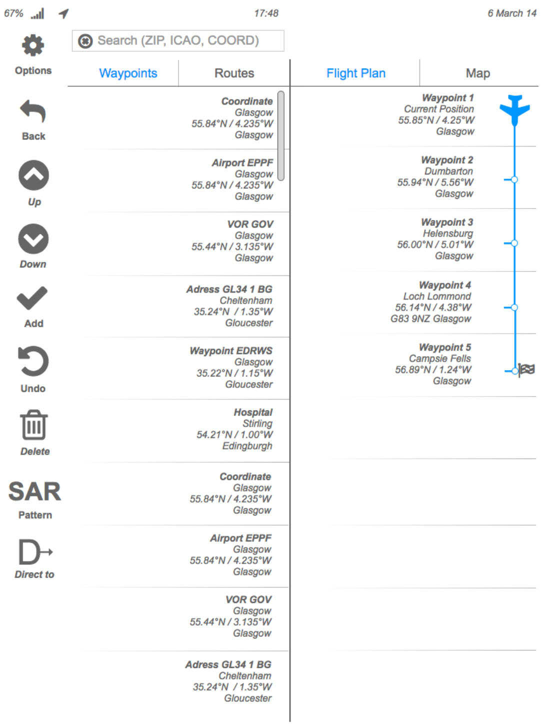
Coordinate
Glasgow
55.84°N / 4.235°W
Glasgow

Airport EPPF
Glasgow
55.84°N / 4.235°W
Glasgow

VOR GOV
Glasgow
55.44°N / 3.135°W
Glasgow

Adress GL34 1 BG
Cheltenham
35.24°N / 1.35°W
Gloucester

Campsie Fells
56.89°N / 1.24°W
Glasgow



Coordinate
Glasgow
55.84°N / 4.235°W
Glasgow

Airport EPPF
Glasgow
55.84°N / 4.235°W
Glasgow

VOR GOV
Glasgow
55.44°N / 3.135°W
Glasgow

Adress GL34 1 BG
Cheltenham
35.24°N / 1.35°W
Gloucester

Waypoint EDRWS
Glasgow
35.22°N / 1.15°W
Gloucester

Hospital
Stirling
54.21°N / 1.00°W
Edingburgh

Coordinate
Glasgow
55.84°N / 4.235°W
Glasgow

Airport EPPF
Glasgow
55.84°N / 4.235°W
Glasgow

VOR GOV
Glasgow
55.44°N / 3.135°W
Glasgow

Adress GL34 1 BG
Cheltenham
35.24°N / 1.35°W
Gloucester

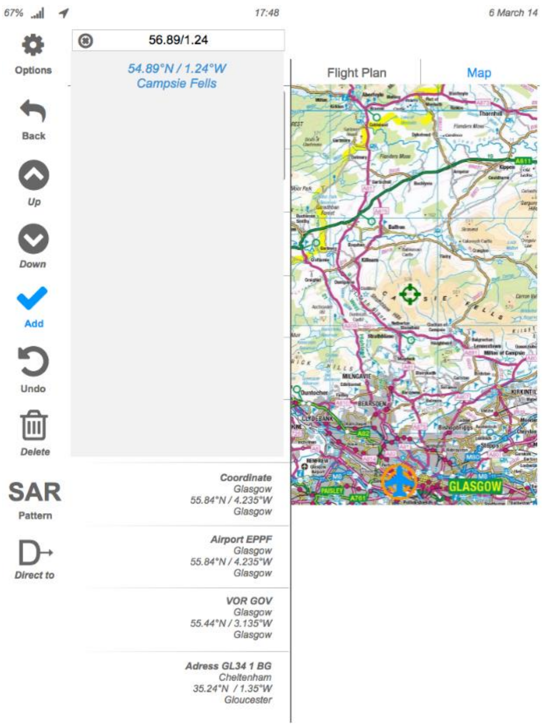
Waypoint 1
Current Position
55.85°N / 4.25°W
Glasgow

Waypoint 2
Dumbarton
55.94°N / 5.56°W
Glasgow

Waypoint 3
Helensburg
56.00°N / 5.01°W
Glasgow

Waypoint 4
Loch Lommond
56.14°N / 4.38°W
G83 9NZ Glasgow

Waypoint 5
Campsie Fells
56.89°N / 1.24°W
Glasgow



56.89/1.24

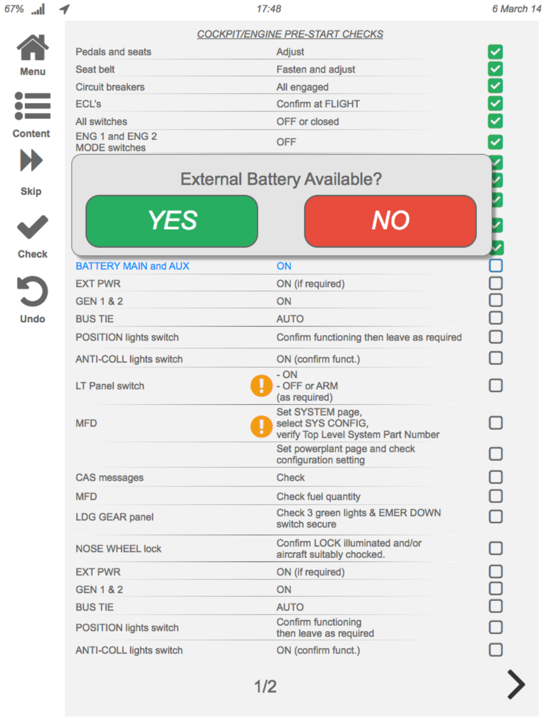
54.89°N / 1.24°W
Campsie Fells

Coordinate
Glasgow
55.84°N / 4.235°W
Glasgow

Airport EPPF
Glasgow
55.84°N / 4.235°W
Glasgow

VOR GOV
Glasgow
55.44°N / 3.135°W
Glasgow

Adress GL34 1 BG
Cheltenham
35.24°N / 1.35°W
Gloucester



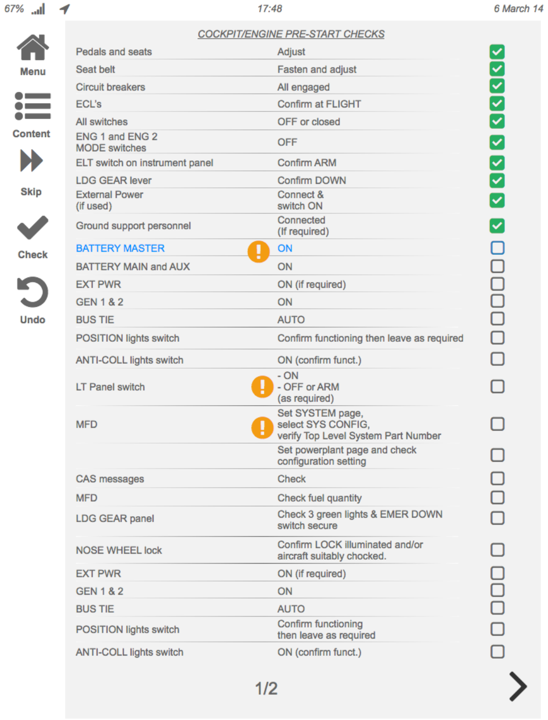
COCKPIT/ENGINE PRE-START CHECKS

Pedals and seats	Adjust	☑
Seat belt	Fasten and adjust	☑
Circuit breakers	All engaged	☑
ECL's	Confirm at FLIGHT	☑
All switches	OFF or closed	☑
ENG 1 and ENG 2 MODE switches	OFF	☑
BATTERY MAIN and AUX	ON	☐
EXT PWR	ON (if required)	☐
GEN 1 & 2	ON	☐
BUS TIE	AUTO	☐
POSITION lights switch	Confirm functioning then leave as required	☐
ANTI-COLL lights switch	ON (confirm funct.)	☐
LT Panel switch	! - ON - OFF or ARM (as required)	☐
MFD	! Set SYSTEM page, select SYS CONFIG, verify Top Level System Part Number	☐
	Set powerplant page and check configuration setting	☐
CAS messages	Check	☐
MFD	Check fuel quantity	☐
LDG GEAR panel	Check 3 green lights & EMER DOWN switch secure	☐
NOSE WHEEL lock	Confirm LOCK illuminated and/or aircraft suitably chocked.	☐
EXT PWR	ON (if required)	☐
GEN 1 & 2	ON	☐
BUS TIE	AUTO	☐
POSITION lights switch	Confirm functioning then leave as required	☐
ANTI-COLL lights switch	ON (confirm funct.)	☐

External Battery Available?

YES NO

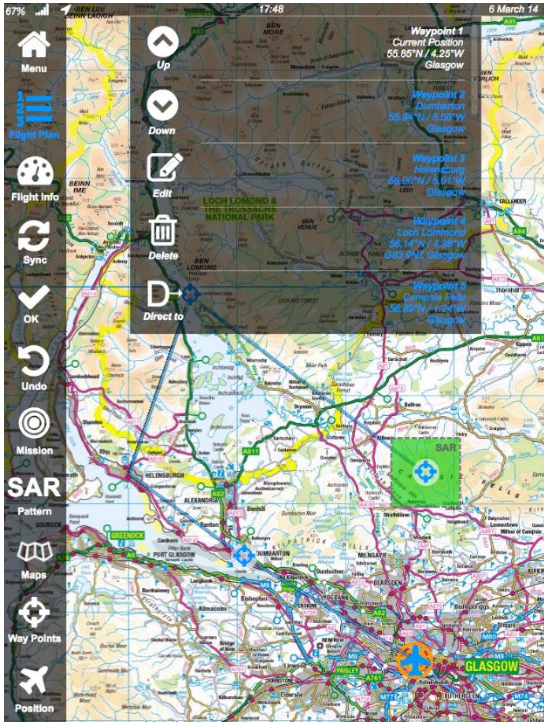
1/2

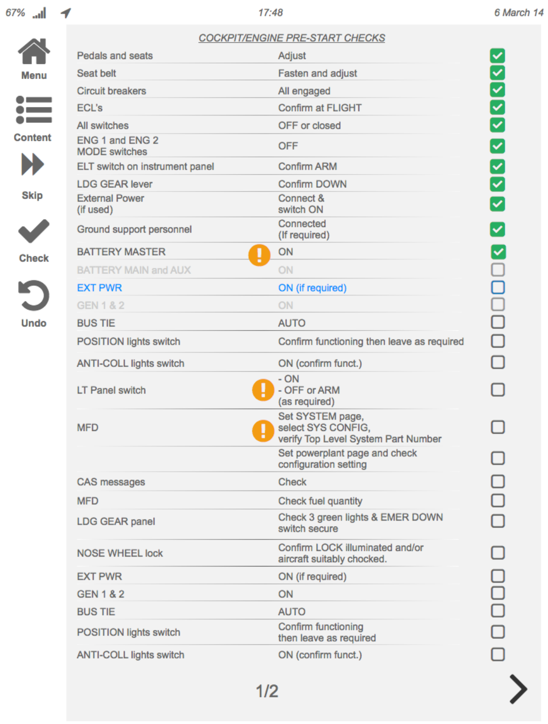


COCKPIT/ENGINE PRE-START CHECKS

Pedals and seats	Adjust	☑
Seat belt	Fasten and adjust	☑
Circuit breakers	All engaged	☑
ECL's	Confirm at FLIGHT	☑
All switches	OFF or closed	☑
ENG 1 and ENG 2 MODE switches	OFF	☑
ELT switch on instrument panel	Confirm ARM	☑
LDG GEAR lever	Confirm DOWN	☑
External Power (if used)	Connect & switch ON	☑
Ground support personnel	Connected (if required)	☑
BATTERY MASTER	! ON	☐
BATTERY MAIN and AUX	ON	☐
EXT PWR	ON (if required)	☐
GEN 1 & 2	ON	☐
BUS TIE	AUTO	☐
POSITION lights switch	Confirm functioning then leave as required	☐
ANTI-COLL lights switch	ON (confirm funct.)	☐
LT Panel switch	! - ON - OFF or ARM (as required)	☐
MFD	! Set SYSTEM page, select SYS CONFIG, verify Top Level System Part Number	☐
	Set powerplant page and check configuration setting	☐
CAS messages	Check	☐
MFD	Check fuel quantity	☐
LDG GEAR panel	Check 3 green lights & EMER DOWN switch secure	☐
NOSE WHEEL lock	Confirm LOCK illuminated and/or aircraft suitably chocked.	☐
EXT PWR	ON (if required)	☐
GEN 1 & 2	ON	☐
BUS TIE	AUTO	☐
POSITION lights switch	Confirm functioning then leave as required	☐
ANTI-COLL lights switch	ON (confirm funct.)	☐

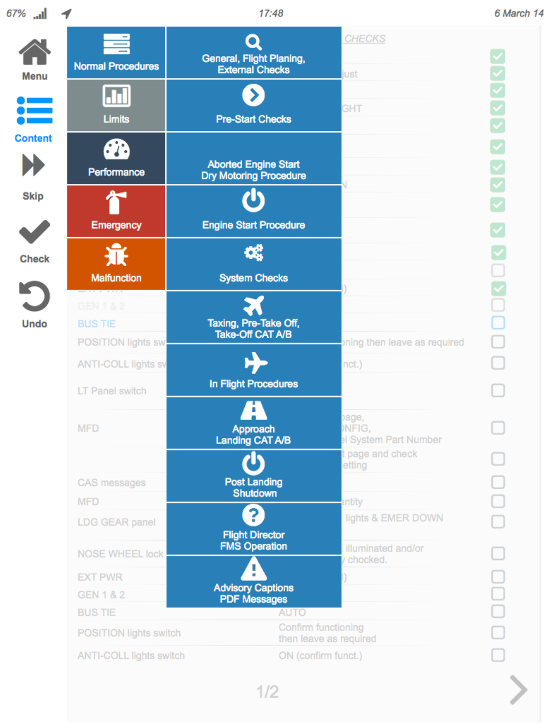
1/2

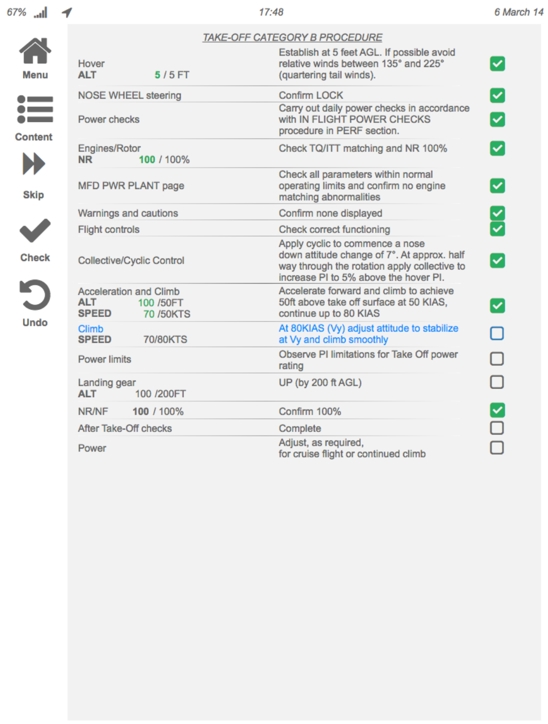




COCKPIT/ENGINE PRE-START CHECKS

Pedals and seats	Adjust	☑
Seat belt	Fasten and adjust	☑
Circuit breakers	All engaged	☑
ECL's	Confirm at FLIGHT	☑
All switches	OFF or closed	☑
ENG 1 and ENG 2 MODE switches	OFF	☑
ELT switch on instrument panel	Confirm ARM	☑
LDG GEAR lever	Confirm DOWN	☑
External Power (if used)	Connect & switch ON	☑
Ground support personnel	Connected (If required)	☑
BATTERY MASTER	! ON	☑
BATTERY MAIN and AUX	ON	☐
EXT PWR	ON (if required)	☐
GEN 1 & 2	ON	☐
BUS TIE	AUTO	☐
POSITION lights switch	Confirm functioning then leave as required	☐
ANTI-COLL lights switch	ON (confirm funct.)	☐
LT Panel switch	! - ON - OFF or ARM (as required)	☐
MFD	! Set SYSTEM page, select SYS CONFIG, verify Top Level System Part Number	☐
	Set powerplant page and check configuration setting	☐
CAS messages	Check	☐
MFD	Check fuel quantity	☐
LDG GEAR panel	Check 3 green lights & EMER DOWN switch secure	☐
NOSE WHEEL lock	Confirm LOCK illuminated and/or aircraft suitably chocked.	☐
EXT PWR	ON (if required)	☐
GEN 1 & 2	ON	☐
BUS TIE	AUTO	☐
POSITION lights switch	Confirm functioning then leave as required	☐
ANTI-COLL lights switch	ON (confirm funct.)	☐





TAKE-OFF CATEGORY B PROCEDURE

Hover ALT 5 / 5 FT	Establish at 5 feet AGL. If possible avoid relative winds between 135° and 225° (quartering tail winds).	☑
NOSE WHEEL steering	Confirm LOCK	☑
Power checks	Carry out daily power checks in accordance with IN FLIGHT POWER CHECKS procedure in PERF section.	☑
Engines/Rotor NR 100 / 100%	Check TQ/ITT matching and NR 100%	☑
MFD PWR PLANT page	Check all parameters within normal operating limits and confirm no engine matching abnormalities	☑
Warnings and cautions	Confirm none displayed	☑
Flight controls	Check correct functioning	☑
Collective/Cyclic Control	Apply cyclic to commence a nose down attitude change of 7°. At approx. half way through the rotation apply collective to increase PI to 5% above the hover PI.	☑
Acceleration and Climb ALT 100 /50FT SPEED 70 /50KTS	Accelerate forward and climb to achieve 50ft above take off surface at 50 KIAS, continue up to 80 KIAS	☑
Climb SPEED 70/80KTS	At 80KIAS (Vy) adjust attitude to stabilize at Vy and climb smoothly	☐
Power limits	Observe PI limitations for Take Off power rating	☐
Landing gear ALT 100 /200FT	UP (by 200 ft AGL)	☐
NR/NF 100 / 100%	Confirm 100%	☑
After Take-Off checks	Complete	☐
Power	Adjust, as required, for cruise flight or continued climb	☐