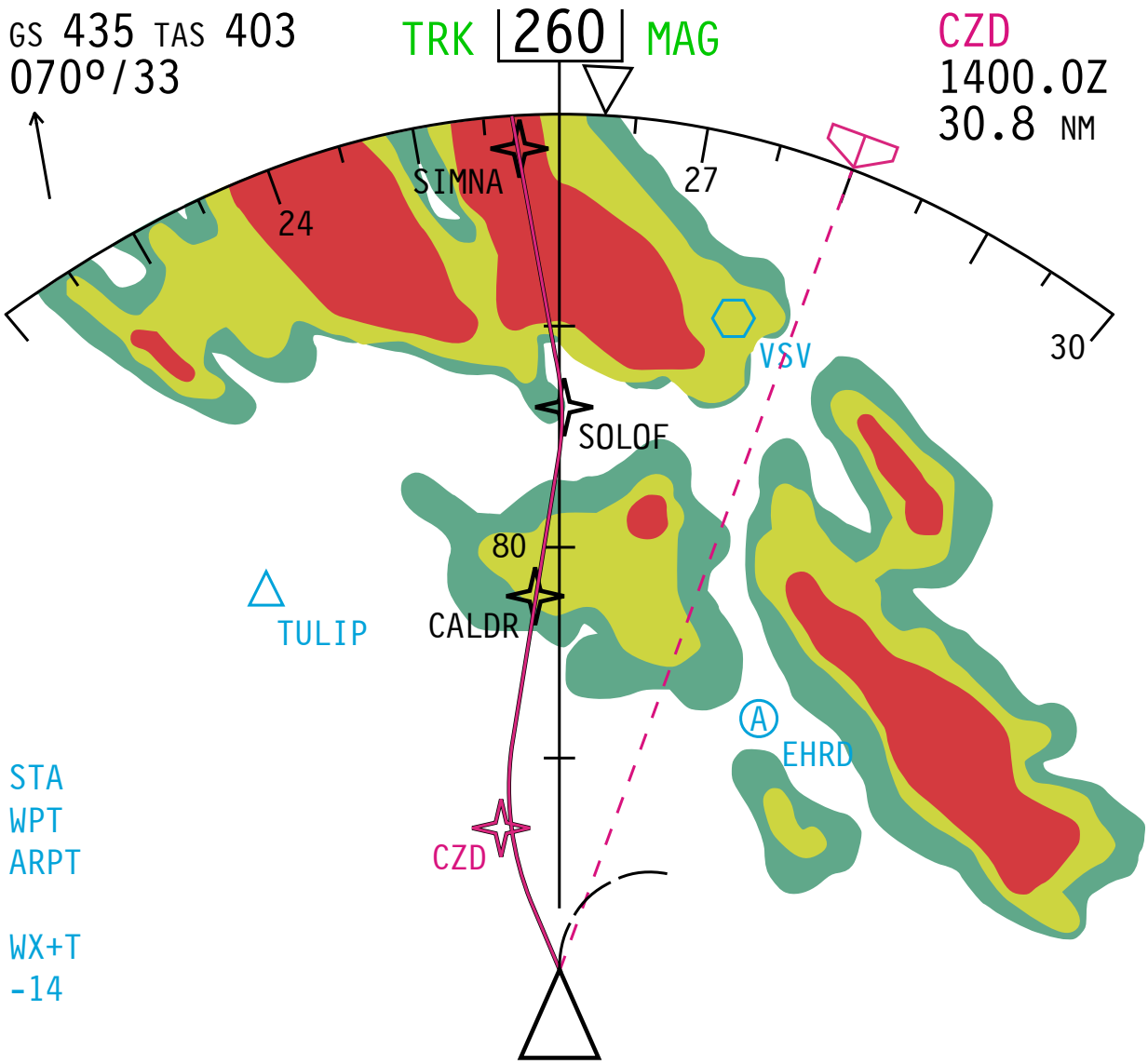
Evaluation of a touch-based navigation display for lateral weather avoidance

N.C.M. van Zon

August 30, 2017



TU Delft

Evaluation of a touch-based navigation display for lateral weather avoidance

Master of Science Thesis

For obtaining the degree of Master of Science in Aerospace Engineering
at Delft University of Technology

N.C.M. van Zon

August 30, 2017

Faculty of Aerospace Engineering · Delft University of Technology

Delft University of Technology

Delft University Of Technology
Department Of
Control and Simulation

The undersigned hereby certify that they have read and recommend to the Faculty of Aerospace Engineering for acceptance a thesis entitled **"Evaluation of a touch-based navigation display for lateral weather avoidance"** by **N.C.M. van Zon** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: August 30, 2017

Readers:

prof.dr.ir. M. Mulder

dr.ir. M. M. van Paassen

dr.ir. C. Borst

dr.ir. D. M. Pool

ir. T. J. Mulder

dr.ir. W. J. C. Verhagen

Acknowledgements

Max, I remember our discussion about internships back in October 2015, when you mentioned that the master thesis should become the most fun part of my studies. The enthusiasm, professional attitude and fun of the Control & Simulation department have undoubtfully lived up to, and surpassed, this expectation. Not only have I learned a lot, I have really enjoyed the entire curriculum and working with everyone I came across. To complete the picture, I had the chance to work with the SIMONA, a definite upgrade from the simple, wooden basement flight simulator I built as a high school project (see image).

Basement flight simulator

Max, I have a lot of respect for how much time you and your department invest in your students. Thank you for all your input, motivating me to reach the highest level possible. Clark, I want to thank you specifically for our various, and often extensive, brainstorm sessions about parts of my thesis. These were filled with interesting discussions and a good portion of humor. The main idea of this thesis topic came from you, where you were able to devise a subject that fit my rather vague request *"to do something operational."* Daan, I want to thank you especially for your help with my Fitts' law experiment. I hope I was able to shed some light onto this supplementary way of looking at a manual control task. Rene, I want to thank you specifically for your help with getting me started with DUECA and C++. No matter how elementary the question was that I posed, you were always a helping hand. Olaf, I want to thank you notably for your help with getting both of my DUECA projects running on the SRS. No C++, DUECA, OpenGL or hardware roadblock was too much a challenge for you. Last but not least Alwin, I want to thank you for helping me to get the SRS configured just in time for my final experiment. To all of you, your almost immediate availability for any question or request to discuss or brainstorm about a part of my thesis was remarkable. Thank you for this very enjoyable period of my studies.

Nout

Delft, August 30, 2017

Contents

List of Figures

List of Tables

Nomenclature

List of Abbreviations

A	Amplitude
A/THR	Auto-throttle
ADIRS	Air Data Inertial Reference System
AFDS	Autopilot Flight Director System
AH	Abstraction Hierarchy
ANOVA	Analysis of Variance
ATC	Air Traffic Control
ATM	Air Traffic Management
B777	Boeing 777
CCD	Cursor Control Device
CDU	Control Display Unit
CPDLC	Controller-Pilot Data Link Communications
CRT	Choice Reaction Task
CV	Control Variables
DL	Decision Ladder
DLP	Digital Light Processing
DMI	Direct Manipulation Interface
DUECA	Delft University Environment for Communication and Activation

DV	Dependent Variables
EFIS	Electronic Flight Instrument System
EID	Ecological Interface Design
ESVD	Ecological Synthetic Vision Display
FAA	Federal Aviation Administration
FDSU	Flight Data Storage Unit
FL	Flight Level
FMC	Flight Management Computer
FMS	Flight Management System
HDG SEL	Heading Select Mode
HMI	Human-Macine Interface
IAS	Indicated Airspeed
IEEE	Institute of Electrical and Electronics Engineers
ID	Index of Difficulty
ISO	International Organization for Standardization
KBB	Knowledge-Based Behaviour
LCD	Liquid Crystal Display
LNAV	Lateral Navigation
LSK	Line Select Key
MCP	Mode Control Panel
MT	Movement Time
ND	Navigation Display
PFD	Primary Flight Display
PTT	Push-To-Talk (Switch)
RBB	Rule-Based Behaviour
RMS	Root Mean Square
RSME	Rating Scale Mental Effort
RTA	Required Time of Arrival
SBB	Skill-Based Behaviour
SESAR	Single European Sky ATM Research
SRK	Skill-, Rule-, and Knowledge-Based Behaviour

SRS	SIMONA Research Simulator
SVGCAS	Synthetic Vision Ground Collision Avoidance System
TBO	(4D) Trajectory Based Operations
TND	Touch-Based Navigation Display
TP	Throughput
VNAV	Vertical Navigation
W	Width

List of Symbols

A	Amplitude or distance to target	[–]
A_{ij}	Amplitude for movement from key i to key j	[–]
α	Angular amplitude or distance to target	[–]
H_j	Height of key j	[–]
ID_e	Effective index of difficulty	[–]
$ID_{e_{ij}}$	ID, adjusted for W_e, for a specific experiment condition and participant	[–]
MT_{ij}	Movement time for a specific experiment condition and participant	[–]
ω	Angular width or tolerance of target	[–]
ω_e	Effective angular width of target	[–]
R^2	Coefficient of determination	[–]
σ_ψ	Standard deviation in angular movement endpoints	[–]
σ_a	Bivariate standard deviation of finger accuracy	[–]
σ_{xy}	Bivariate standard deviation of movement endpoints	[–]
W	Width or tolerance of target	[–]
W_e	Effective width	[–]
W_{ij_e}	Effective width for movement from key i to key j	[–]
W_j	Width of key j	[–]

Chapter 1

Introduction

The modern-day flight management system (FMS) was introduced on the Boeing 767 in 1982 [1] to assist pilots in both lateral and vertical navigation (LNAV, VNAV). As an interface to the FMS, the control display unit (CDU) was introduced and remains the industry standard to date. For example, whilst Boeing replaced the CDU on the Boeing 787 with a digital copy, the look and feel remains the same, see Figure 1-1. According to two KLM captains, interviewed during this research, the CDU remains effective in handling present day LNAV task. Once the necessary commands are learned, executing various procedures can be done quickly and error-free.

Figure 1-1: Pilot interacting with control display unit (CDU), entering en-route wind information (inset: Boeing-style CDU interface showing the conventional alphabetic keyboard and button interface)[1]

However, looking ahead at future developments in LNAV procedures, the necessity to modernize the FMS interface becomes evident. The SESAR Joint Undertaking expects the number of flights in European airspace to have increased by 52% in 2035 compared to 2012 [2]. As a response, 4D trajectory based operations (TBO) is being developed amongst other mechanisms to cope with increased congestion. The basic principle of 4D TBO is to impose time constraints at which aircraft are required to arrive at specific waypoints in their flight plan [3], [4]. As a result, Huisman et al. [5] expects an increased frequency of en-route route adjustments. Hence, Van Marwijk et al. [6] call for *"a redesign of the navigation planning interface [due to] increasing punctuality in, [amongst others,] European SESAR concepts, [which will] make airborne flight plan amendment increasingly complex"*.

Touchscreens have the possibility to reduce workload whilst increasing situation awareness due to their highly intuitive and direct interface [7]–[11]. Furthermore, by directly adjusting the flight plan route using a touch-based navigation display, the LNAV functionalities of the FMS can remain engaged and work to ensure compliance with 4D TBO. Therefore, aircraft and equipment manufacturers have been proposing touchscreens on their newest flight decks to cope with increased complexity in navigation tasks. Boeing will introduce touchscreen displays on the 777X [7], Gulfstream ships their G650 with touch-based interfaces [12] and the Garmin G5000 is now also touch-enabled [13]. However, concerns have been voiced about the loss of tactile feedback, usability in dynamic environments and physical fatigue of operation [8], [14]–[16].

Research has been done evaluating touchscreen interfaces in general and comparing them to less direct interfaces such as trackballs, trackpads and rotary controllers [14], [17]–[20]. Furthermore, Dodd et al. [16] found increased task execution time, error rates and subjective workload for touchscreen usage in turbulence and at specific cockpit positions. In addition, Fitts' law, first published in 1954, has been used by human-machine interface (HMI) researchers as a predictive model of movement time (MT), and a means to compare interfaces using the *throughput* measure [21]–[23]. The latter describes how many *bits* of task difficulty, defined by an index of difficulty (ID), an interface can handle per second. However, Fitts' law has not yet been applied to interfaces on the flight deck. Finally, Damveld et al. [15], [24] propose interesting concepts for novel touch-based FMS interfaces. However, they would benefit from a comprehensive evaluation based on operational and future LNAV demands.

Objective Hence, the main objective of this thesis was two-fold. First, the FMS interfaces were evaluated on a low technical level by analysing the individual components. Fitts' law was used to develop an accuracy and throughput model of the mode control panel (MCP), CDU and a novel touch-based navigation display (TND). Furthermore, an evaluation was performed on the applicability of Fitts' law on flight deck interfaces. Second, a high-level approach was taken to evaluate a TND within an operational context. The SIMONA Research Simulator was used, with 12 commercial pilots participating in a within-subjects experiment to evaluate a touch-based navigation display, compared to a conventional interface. Participants were faced with a lateral weather avoidance re-planning task at varying levels of difficulty. Results of this experiment were compared with the Fitts' law models in order to present a hollistic evaluation of a TND. Recommendations are given concerning the design of the interface between the pilot and the flight management system, in order to meet future lateral navigation demands. Seperate objectives are given below, for each of the constituent parts.

IEEE Article — The objective of the IEEE article is to expose commercial pilots, engaged in a lateral weather avoidance task, to both a touch-based navigation display and a conventional interface to investigate the effectiveness of the touchscreen interface.

IEEE Companion Article — The objective of the IEEE companion article is to engage experiment participants in rapidly aimed movement tasks, in order to develop an accuracy and throughput model of a CDU, MCP and TND using Fitts' law. In addition, the effectiveness of Fitts' law in describing flight deck interfaces is evaluated. The accuracy and throughput models are referred to in the IEEE article.

Literature Review — The objective of the literature review is two-fold. The first is to gain a comprehensive understanding of contemporary lateral navigation systems and procedures, complemented with a detailed review of possible next generation interfaces, such as direct manipulation interfaces and touchscreens. The second is to evaluate a touchscreen flight deck interface, more specifically a touch-based lateral navigation display, both from a manual and cognitive perspective.

Appendices — The objective of the appendices is to provide, where necessary, additional insight for the preceeding parts of this master thesis.

Scope In order to accomplish the previously determined objectives, a scope has been defined to frame the research conducted during this master thesis. A specific look at airborne route adjustment is chosen, because given high airspeeds a larger amount of time pressure exists on the completion of these tasks when compared to those executed on the ground. Subsequently, the lateral weather avoidance task was chosen based on interviews conducted with commercial pilots. Each of them mentioned weather avoidance to be the most often occuring en-route re-planning task. Due to the opportunity to further build on previous research and a prototype developed by Damveld et al. [15], [24], the Boeing-style interface and operating procedures will be used as a baseline. Environmental effects, such as vibration, glare or turbulence, on the use of touchscreens fall beyond the scope of this thesis.

Research Question The following research question will be addressed in this master thesis.

> How can a **touch-enabled navigation display** decrease **task execution time and error rate** whilst supporting the **decision making process** during a **lateral weather avoidance task** with **minimal re-route delay?**

Thesis Outline The first part of this thesis discusses the IEEE article written to describe an experiment conducted in the SIMONA Research Simulator, where twelve commercial pilots were tasked to execute a lateral weather avoidance task using either a touchscreen or conventional interface. The second part discusses an IEEE companion article written to describe a preparatory experiment that used Fitts' law to develop accuracy and throughput models of the touchscreen and conventional interfaces used in the experiment described in part one. This third part of this thesis will enclose a literature review, which has been subdivided into three topics. Finally, the fourth part of this thesis are the appendices containing, amongst others, additional results for both experiments and briefing documents used.

Part I

IEEE Article

Evaluation of a Touch-based Navigation Display for Lateral Weather Avoidance

N.C.M. van Zon
Control & Simulation, Department Control & Operations
Delft University of Technology, Delft, The Netherlands
Supervisors: C. Borst, D.M. Pool, M.M. van Paassen and M. Mulder

***Abstract*—Industry is proposing touch-based solutions to modernize the flight management system, a necessary step to cope with increasing demands on lateral navigation as airspace congestion grows. However, research evaluating touchscreen effectiveness for navigation tasks on the flight deck is lacking. The SIMONA Research Simulator was used, with 12 commercial pilots participating in a within-subjects experiment to evaluate a touch-based navigation display, compared to a conventional interface. Participants were faced with a lateral weather avoidance re-planning task at varying levels of difficulty. In addition, Fitts' law and the Decision Ladder were used to provide a theoretical basis. Results show that, where ample time is present to execute route adjustments, a touch-based navigation display is a promising replacement of the control display unit. However, for time-critical tasks participants failed at avoiding weather, providing more inputs and scoring worse on secondary task and mental effort rating. A more direct control of the aircraft was necessary, such as the heading select mode. Touchscreen interfaces are very promising, however industry is urged to carefully determine which tasks are suitable for touch input.**

I. Introduction

The modern-day flight management system (FMS) was introduced on the Boeing 767 in 1982 [1] to assist pilots in both lateral and vertical navigation (LNAV, VNAV). As an interface to the FMS, the control display unit (CDU) was introduced and remains the industry standard to date. For example, whilst Boeing replaced the CDU on the Boeing 787 with a digital copy, the look and feel remains the same. According to two KLM captains, interviewed during this research, the CDU remains effective in handling present day LNAV tasks. Once the necessary commands are learned, executing various procedures can be done quickly and error-free.

However, looking ahead at future developments in LNAV procedures, the necessity to modernize the FMS interface becomes evident. The SESAR Joint Undertaking expects the number of flights in European airspace to have increased by 52% in 2035 compared to 2012 [2]. As a response, 4D trajectory based operations (TBO) is being developed amongst other mechanisms to cope with increased congestion. The basic principle of 4D TBO is to impose time constraints at which aircraft are required to arrive at specific waypoints in their flight plan [3], [4]. As a result, Huisman et al. [5] expects an increased frequency of en-route route adjustments. Hence, Van Marwijk et al. [6] call for *"a redesign of the navigation planning interface [due to] increasing punctuality in, [amongst others,] European SESAR concepts, [which will] make airborne flight plan amendment increasingly complex"*.

Touchscreens have the possibility to reduce workload whilst increasing situation awareness due to their highly intuitive and direct interface [7]–[11]. Furthermore, by directly adjusting the flight plan route using a touch-based navigation display, the LNAV mode of the FMS can remain engaged and work to ensure compliance with 4D TBO. Therefore, aircraft and equipment manufacturers have been proposing touchscreens on their newest flight decks to cope with increased complexity in navigation tasks. For example, Boeing will introduce touchscreen displays on the 777X [7], Gulfstream ships their G650 with touch-based interfaces [12] and the Garmin G5000 is now also touch-enabled [13]. However, concerns have been voiced about the loss of tactile feedback, usability in dynamic environments and physical fatigue of operation [8], [14]–[16].

Research has been done evaluating touchscreen interfaces in general and comparing them to less direct interfaces such as trackballs, trackpads and rotary controllers [14], [17]–[20]. Furthermore, Dodd et al. [16] found increased task execution time, error rates and subjective workload for touchscreen usage in turbulence and at specific cockpit positions. Damveld et al. [15], [21] proposes an interesting concept for a novel touch-based FMS interface. However, they would benefit from a comprehensive evaluation based on operational and future LNAV demands. Hence, previous research by the author developed an accuracy and throughput model, using Fitts' law, of a mode control panel (MCP), CDU and a touch-based navigation display (TND) [22]. These models suggest that the CDU scores highest in both throughput and accuracy. However, the research lacked a thorough evaluation in the context of a complex lateral navigation task, which remains a relatively untouched subject in flight deck design research.

The goal of this paper is to evaluate a TND for a lateral weather avoidance re-planning task. A specific look at airborne route adjustment is chosen, because given high airspeeds a larger amount of time pressure exists on the completion of these tasks when compared to those executed on the ground. Subsequently, lateral weather avoidance was chosen based on interviews conducted with commercial pilots. Each of them mentioned weather avoidance to be the most often occuring en-route re-planning task.

A TND, based on work in [15], [21], is proposed as a replacement of the conventional interface used in aircraft

today. The Boeing implementation of the latter was used. The TND is analyzed from two different angles, to provide a comprehensive evaluation. First, the FMS interfaces are scrutinized on a low technical level by analyzing the individual components. In addition to the Fitts' law models developed in [22], the Decision Ladder designed by Rasmussen [23] was used to be able to describe the individual controls in both their usage and decision making support. Second, a high-level approach was taken to evaluate a TND within an operational context. The SIMONA Research Simulator was used, with 12 commercial pilots participating in a within-subjects experiment to evaluate a TND, compared to a conventional interface. Participants were faced with a lateral weather avoidance re-planning task at varying levels of difficulty. Results of this experiment were compared with Fitts' law models and the Decision Ladder analysis in order to present a hollistic evaluation of a TND. Recommendations are given concerning the design of the interface between the pilot and the FMS, in order to meet future lateral navigation demands.

The control task analysis is discussed in Section II. Section III introduces the experiment design followed by an analysis of the results in Section IV. Next, these results are discussed in Section V. The paper closes with a conclusion.

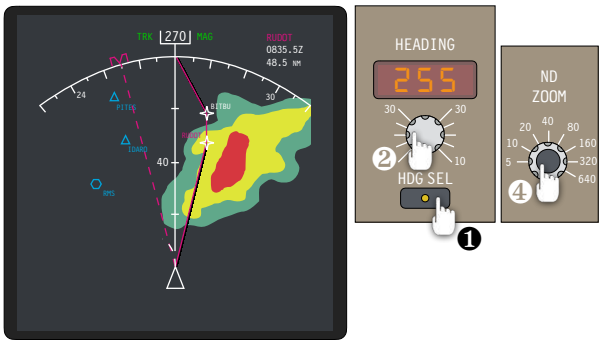
II. Control Task Analysis

A. Lateral Weather Avoidance

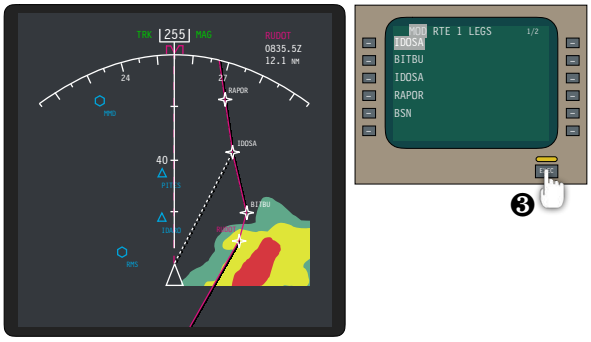
Five commercial pilots were interviewed, in addition to a review of Boeing operating manuals [24], to determine the exact nature of a lateral weather avoidance task. Weather avoidance is always done using the heading select mode. In this mode the pilot selects a heading, instead of requesting the FMS to follow a lateral flight plan, and as a result the autoflight system will roll the aircraft towards the commanded heading. The controls for the heading select mode are found on the MCP. This method is used because the weather radar onboard of commercial aircraft scans weather by looking ahead, and as a result will only show the initial set of weather returns, hiding potential weather behind it. This makes it very difficult for airline crews to perform long-term weather re-routing. Furthermore, weather can be very dynamic. According to the interviewed pilots, weather often intensifies, with gaps in between cells dissapearing, when in closer proximity. The heading select mode allows pilots to respond quickly to new and updated weather radar returns, by immediately commanding a new heading when necessary. This makes it an effective interface for the task. Following safe passage, pilots will return to the original route by adjusting it using the CDU. Throughout a weather avoidance procedure the pilots mentioned their goal was to avoid weather safely and comfortably, whilst minimizing total re-route delay.

B. Human-Machine Interface

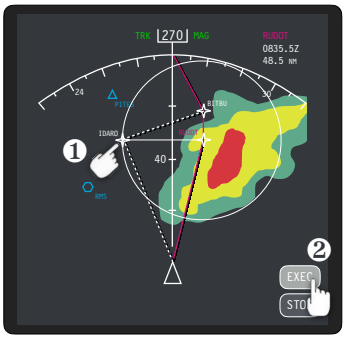
1) Conventional Interface: The operating procedure using the conventional interface for a lateral weather avoidance task can be illustrated. First, the heading select mode is activated on the MCP (❶ in Figure 1a). Second, the necessary heading



(a) Conventional interface showing the heading select mode (using the MCP) and ND range adjustment (using the EFIS panel)



(b) Conventional interface showing a route adjustment (using the CDU) to fly direct-to a waypoint further along the route

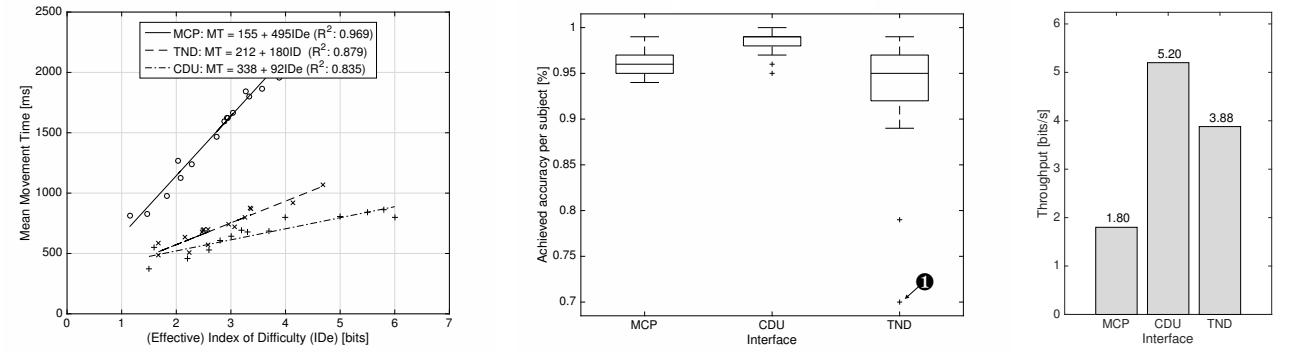


(c) Touchscreen interface including steps necessary to execute control task

Fig. 1. Conventional and touchscreen lateral weather avoidance interfaces including steps necessary to execute control task

can be commanded using a rotary controller on the MCP (❷ in Figure 1a). Following safe passage, the CDU is used to adjust the flight plan route to fly direct to a suitable waypoint (see ❸ in Figure 1b) in compliance with ATC. Next, the LNAV mode is re-activated and the autopilot returns to following the flight plan route. Throughout this process, a zoom knob (❹ in Figure 1a) on the electronic flight instrumentation system (EFIS) control panel can be used to adjust the range of the navigation display (ND).

2) Touchscreen Interface: The information presented on a TND is identical to the ND in the conventional interface. However, pilots can now directly interact with the flight plan



(a) Final Fitts' law models of each individual interface plotted on the same graph for comparative purposes

(b) Observed accuracy scores per participant per experiment.

(c) Measured throughput of each interface.

Fig. 2. Results of previous research aimed at developing an accuracy and throughput model of the MCP, CDU and TND [22]

route and its surroundings. A two-finger pinch zoom gesture can be used to adjust the range of the display. A two-finger swipe gesture allows the pilot to pan across the display, for example, to scroll forward towards a waypoint further along the route. With a single finger, route waypoints can be selected and moved towards a new desired location (❶ in Figure 1c). A preview of the modified route is drawn separately, allowing the pilot to confirm that it matches the desired flightpath, prior to executing the new route (❷ in Figure 1c).

C. Fitts' Law Accuracy and Throughput Models

Fitts' law, first published in 1954, has been used by human-machine interface (HMI) researchers as a predictive model of movement time (MT), and a means to compare interfaces using the *throughput* measure [25]. The latter describes how many *bits* of task difficulty, defined by an index of difficulty (ID), an interface can handle per second. The law can be used to describe the speed-accuracy tradeoff inherent for humans engaged in rapid aimed movement tasks [26], [27]. The complete Fitts' law model to describe MT as a function of ID in a high-accuracy pointing task is presented in Eq. (1). Here, a and b are empirical linear regression constants, A is the amplitude (distance to be traversed) and W_e is the effective width of the target. The latter is empirically calculated using the standard deviation of measured endpoint coordinates.

$$MT = a + b(ID) = a + b\log_2\left(\frac{A}{W_e} + 1\right) \text{ [seconds]} \quad (1)$$

Literature finds that using W_e reflects what participants actually did, rather than what they were asked to do [28]. The spread in movement endpoints will not always align with the specified target width, resulting in inconsistent error rates across various conditions and making comparisons difficult. Participants are found to under- or overperform due to their tendency to *"to cheat on easier ID conditions"* [25] by not moving fast enough or by not covering the entire distance. Given that Fitts' law is designed for rapid aimed movements, *"participants who take their time compromise Fitts' law"* [25], hence, the W_e correction can be useful. The full background and derivation of Fitts' law has been reviewed, see [26], [27].

The usefulness of Fitts' law is twofold. First, it can help build models of task execution time using a particular interface. Second, it can provide a quantitative description of the FMS interface by comparing the throughput of individual interfaces. Eq. (2) defines the throughput in bits per second, which is calculated by dividing the ID by the measured MT for each participant and experimental condition. The total number of conditions and participants is defined by x and y respectively. $ID_{e_{ij}}$ defines the index of difficulty, adjusted using the effective width W_e, and MT_{ij} the movement time, both for a specific experimental condition and participant.

$$TP \text{ (Throughput)} = \frac{1}{y}\sum_{i=1}^{y}\left(\frac{1}{x}\sum_{j=1}^{x}\frac{ID_{e_{ij}}}{MT_{ij}}\right) \text{ [bits/s]} \quad (2)$$

The basic law applies to one-dimensional situations, however, multiple variations of Fitts' law have been consequently researched, published and verified (see [25], [29]–[32]). In previous research an appropriate variation was found, and used for the MCP, CDU and TND, respectively [22].

Main results are reproduced in Figure 2. The W_e correction was not used for the TND, given that without the correction produced a better model fit, see Figure 2a. The CDU was found to score highest in both accuracy and throughput (see Figure 2b and Figure 2c, respectively), of which the latter implies that it can handle more *bits* of difficulty in the same time compared to the other two interfaces. Higher throughput is illustrated graphically in Figure 2a by a smaller slope. Nonetheless, high throughput does not imply a lower total MT if the number of individual inputs is substantial. The outlier (❶ in Figure 2b) in accuracy measurements using the TND was produced by the only left-handed participant. This participant had difficulty using the TND with his right hand, however, had no difficulty using the MCP or CDU. This suggests that the conventional interface is less sensitive to handedness.

D. Decision Ladder Analysis

Fitts' law provides information on the low-level effectiveness of the individual elements of an interface. The Decision Ladder (DL), developed by Rasmussen [23], can be used to analyze the control task from a high-level operational point-of-view. The DL was used to provide a qualitative description of different steps and strategies used to complete the lateral weather avoidance task, see Figure 3. Subsequently, based on Rasmussen's skill-, rule- and knowledge-based taxonomy, they can be compared on cognitive demand.

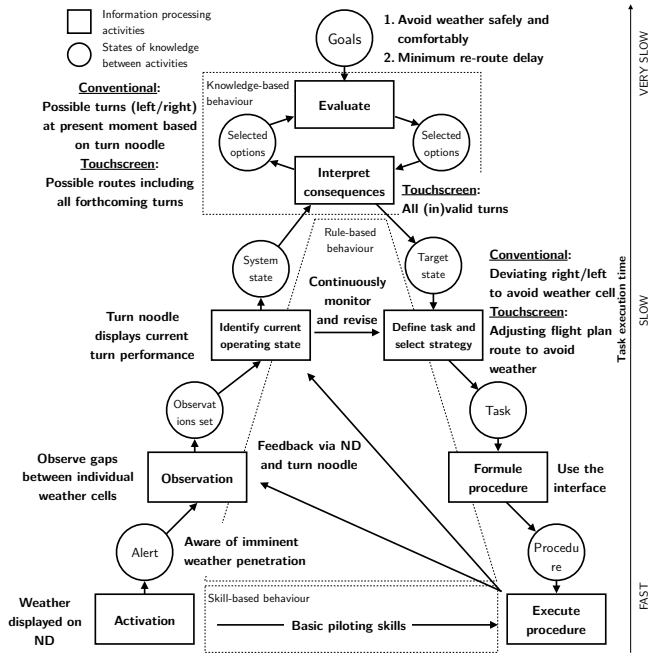
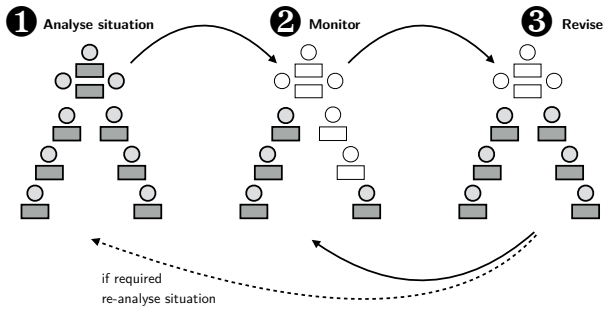


Fig. 3. Decision ladder applied to the lateral weather avoidance task [23]

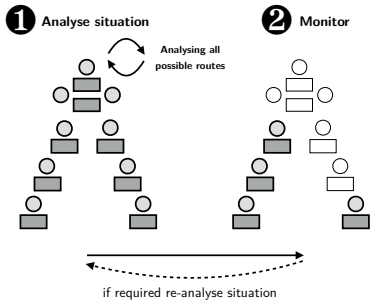
The lateral weather avoidance task can be analyzed using the Decision Ladder, and subsequent differences between both interfaces can be identified. In the bottom-left of the ladder a pilot is activated by the weather displayed on his or her navigation display (ND). Safe gaps are observed between weather cells, leading to the next decision making stage. The current operating state is represented on the ND by the current heading or track angle and a turn noodle which illustrates the instantaneous turn performance of the aircraft. The key difference between either interface is found in the knowledge-based behaviour domain at the top of the decision ladder. Here the various options are evaluated, and tested to see if they meet the set goals.

With the conventional interface a heading is commanded, and therefore the pilot has only to check if the commanded heading, with the resulting turn performance, is sufficient to avoid the closest weather cell. This is a straightforward task, allowing the pilot to quickly exit the cognitively intense knowledge-based decision making region. As illustrated by Figure 4a, following a heading command, the pilot reverts to an easier and faster decision making phase that involves monitoring aircraft trajectory and safe avoidance of weather. Subsequently, as the pilot circumnavigates the weather, minor adjustments are necessary, which involve moving the heading bug as required. Hence, the cognitively intense situation analysis phase (❶ in Figure 4a) is quickly abandoned to progress to easier monitoring and revision phases (❷, ❸ in Figure 4a).

The touchscreen interface allows the pilot to adjust the entire flight plan with the goal to execute a trajectory that avoids all visible and forthcoming weather. As a consequence the pilot will remain for a longer period of time in the upper, knowledge-based decision making situation analysis process (❶ in Figure 4b). Each possible route will need to be evaluated on its ability to meet the goals within given turn performance constraints. However, if successful the pilot can then retract into a monitoring phase (❷ in Figure 4b) and will, given an unchanged obstacle environment, remain in this phase. Only when new obstacles appear does that pilot have to return to the analysis phase. Hence, the touchscreen interface has the ability of easing the entire decision making process, if the time spent in the knowledge-based decision making process at the top of the ladder remains limited and manageable. This suggests that easy and comprehensible tasks, with only a few isolated weather cells, are more manageable using a touchscreen. However, difficult, dynamic and unpredictable tasks seem less workable using a touchscreen. For these tasks, a conventional interface would seem to be less cognitively demanding.



(a) Expected decision making process for conventional interface. Active areas are coloured grey.



(b) Expected decision making process for touchscreen interface

Fig. 4. Decision ladder applied to conventional and touchscreen interfaces. Active areas are coloured grey.

III. Experiment Design

The findings of the theoretical control task analysis, discussed in Section II, are tested using a within-subjects and balanced experiment. The goal of the experiment was to evaluate the effectiveness of a touchscreen interface, in comparison to the conventional one, for a lateral weather avoidance task. Participants were faced with an aircraft cruising on autopilot in differing weather scenarios. Using either the conventional or touchscreen interface they were asked to safely and comfortably avoid weather with minimum re-route delay, restricted to movement only in the lateral plane. Results of this experiment were compared with Fitts' law models of each interface and interfaces were compared on their ability to decrease cognitive workload.

A. Primary Task

The primary task was to avoid weather encountered enroute using the given interface. Re-route time delay had to be kept at a minimum, while considering the safety and comfort of their passengers. When using the conventional interface, participants used the heading select mode on the mode control panel to direct the aircraft around weather. Following safe passage, the control display unit was used to adjust the flight plan route and proceed direct to a suitable waypoint on the original route. This procedure was taken, given that it was found to be most frequently applied. With the touchscreen interface, the participants were able to directly manipulate and adjust the flight plan route so as to circumnavigate weather.

B. Secondary Task

In order to determine how each interface influenced pilot workload, and their ability to simultaneously complete other tasks, a realistic secondary task was introduced. Throughout each measurement run continuous Air Traffic Control (ATC) radio chatter was played in the headset participants were given. Various messages for different aircraft were played interchangeably, and the participant was asked to recognize ATC messages designated for them, recognisable by their callsign `KLM9TU` (pronounced *"KLM niner tango uniform"*). Participants were asked to acknowledge these messages by depressing a push-to-talk (PTT) switch (see Section III-I).

Messages were played at pre-defined positions along each scenario, characterized by a specified latitude (or longitude when necessary). This had two inherent advantages, first it allowed for comparison between measurements given that each message was played at the exact same geographic position. Furthermore, it allowed for defining strategic measurement positions during the most cognitive-intense parts of each scenario. As a result, the secondary task could describe a participants' response time, during key positions along each scenario when the primary task was most demanding. Designated messages, with the `KLM9TU` callsign, occurred every one to five messages, in between messages with other callsigns. Sufficient different messages were recorded such that each ATC tape sounded differently during measurement runs. The ATC messages were written and recorded by the author. An illustration is given in Figure 5, where it can be seen that `KLM9TU` messages, combined with other messages were played a specified geographic positions.

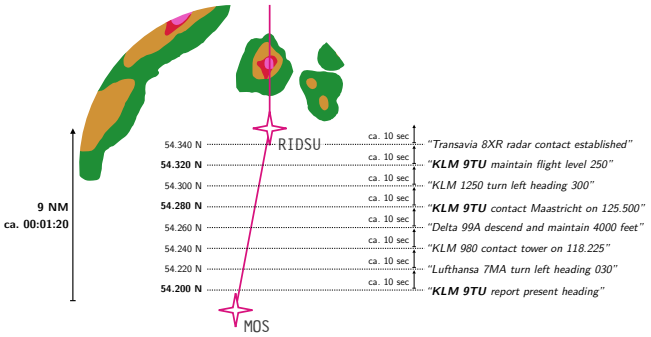


Fig. 5. Illustration of aural radio chatter used for the secondary task.

C. Participants and Instructions

A total of twelve commercial pilots were selected as participants for this experiment. The motivation to only select commercial pilots was twofold. First, due to the complexity of, amongst others, the CDU and the LNAV algorithm, selecting pilots whose day-to-day job is using these systems alleviates the requirement for extensive training. Second, it gave the opportunity to receive useful subjective feedback based on years of operational experience. A short profile of each of the participants is presented in Table I. In addition to the profiles mentioned below, all participants were male and right-handed.

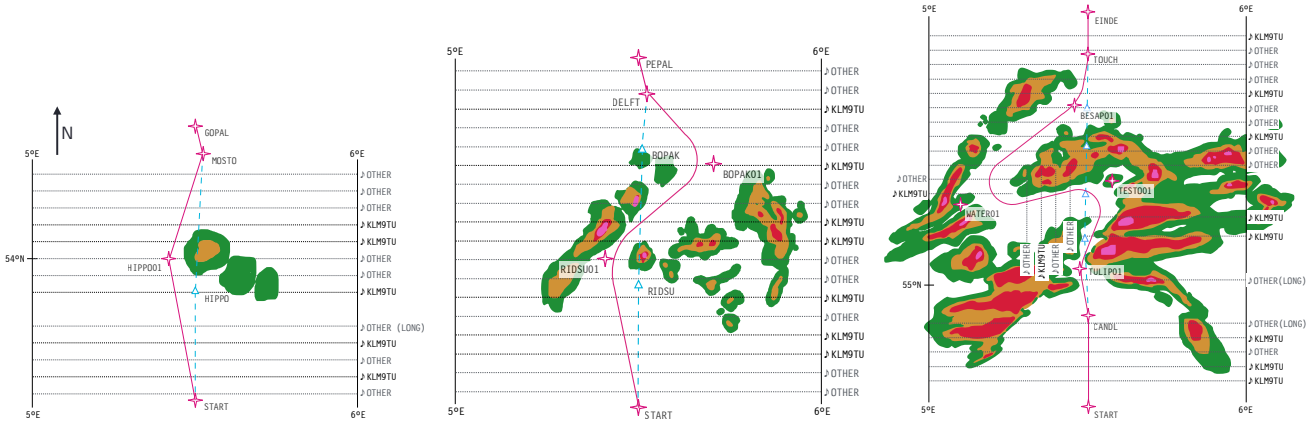
TABLE I
Participant Profiles

#	Age	Current Position	Last Typerating	Total Hours
1	58	KLM captain (retired)	Boeing 777	15,070
2	56	KLM captain	Boeing 777/787	20,000
3	59	KLM captain (retired)	Airbus A330	18,500
4	25	KLM second officer	Boeing 777	1,500
5	56	KLM captain	Boeing 747	19,500
6	40	KLM first officer	Boeing 777/787	9,800
7	30	KLM first officer	Boeing 737	2,000
8	55	KLM captain + FI	Boeing 777/787	16,000
9	47	Research pilot	Cessna Citation	1,900
10	56	KLM captain + FI	Boeing 777/787	16,500
11	41	Research pilot	Cessna Citation	3,500
12	26	KLM second officer	Boeing 777/787	2,000

Participants received a short briefing document a few days prior to the experiment, describing both the primary and secondary task to be conducted and the expected time schedule. In addition, a verbal briefing was conducted directly prior to the experiment. A verbal briefing checklist was used to ensure that each participant received the same information.

D. Independent Variables

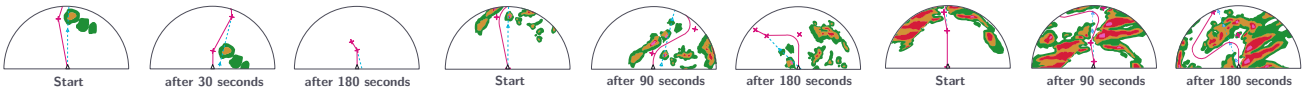
A within-subjects design was chosen as it allowed participants to experience both interfaces and provide relevant



(a) Easy scenario includes one isolated weather cell

(b) Medium scenario includes several scattered weather cells, requiring an S-shaped trajectory

(c) Difficult scenario includes a large squall line with a narrow passage and large turns demanding high speed and accuracy from the interface and participants



(d) Visibility of weather radar, easy scenario

(e) Visibility of weather radar, medium scenario

(f) Visibility of weather radar, difficult scenario

Fig. 6. The three weather scenarios used during the experiment, including the visibility of the weather radar throughout each scenario (at start and after 90 and 180 seconds). The original flight plan route is presented as a stippled cyan line and a possible solution trajectory is illustrated by the solid magenta line.

feedback on how they compare. The combination of two independent variables resulted in six unique experimental conditions.

1) Interface: The first variable was the interface participants were given to complete the task. The conventional interface included a primary flight display (PFD), ND, MCP, EFIS control panel and CDU. The touchscreen interface included a PFD, ND and pedestal-mounted TND. The set-up of these is discussed in Section III-I.

2) Task difficulty: Literature suggests that the effectiveness of a touchscreen interface is highly task dependent [17], [18], [20] and therefore the difficulty of the task was the second variable. Varying task difficulty has the additional advantage of providing information on how the interfaces compare under different levels of pilot effort and stress. Task difficulty was adjusted by defining three weather scenarios.

E. Weather Scenarios

Three weather scenarios, of increasing level of difficulty, were carefully designed for this experiment, each of which is presented in Figure 6. The original flight plan route is illustrated by a magenta line from south to north. Waypoints are designated by a four-point diamond with its name adjacent to it. Weather is drawn in green, yellow, red or magenta based on respective intensity.

Each weather scenario was designed such that only one logical solution existed. This was done to ensure that all participants flew similar, and hence comparable, trajectories. The easy scenario (see Figure 6a) included one isolated weather cell, with a simple solution of passing it towards the west. The medium scenario (see Figure 6b) included several scattered weather cells, with a S-shaped weather avoidance trajectory. Finally, the difficult scenario (see Figure 6c) included a large squall line, including embedded weather cells with a narrow passage and large turns demanding high speed and accuracy from both the interface and participants. The difficult scenario was designed specifically to challenge and stress participants, such that the interface effectiveness under a stressful situation could be evaluted. Furthermore, given the necessary speed and accuracy of the interface the results could be compared with Fitts' law accuracy and throughput models developed for each of the interfaces, see Section II-C [22].

Finally, scenarios were initiated approximately 10 nm prior to encountering the first weather cell. Although this led to a substantially shorter planning horizon than normal for commercial flights this was chosen deliberately to allow for shorter and less boring scenarios, and hence allowing for more measurements per participant. However, this had two implications. First, more waypoints needed to be included such that, using the touchscreen interface, sufficient waypoints were present to allow participants to succesfully adjust the route to avoid weather. Second, the control law was adjusted to allow up to 45 degrees of roll angle in order to make the tight turns in the difficult scenario feasible.

Nonetheless, each of the three weather scenarios were thoroughly tested with an active KLM captain (subject matter expert with 20,000 flying hours), outside of the participant group, and found to hold sufficient operational realism.

F. Weather Radar

The weather radar was designed such that it presented a top-down view of weather, differing in intensity from green to magenta. The weather radar was refreshed every second, and its range was limited to 20 nautical miles (nm). Given scenario lengths of 30 to 50 nm, the range was chosen such that the entire weather picture was not immediately visible, as illustrated by Figure 6(d), (e) and (f). This increased the operational realism of the task, and imposed time pressure on the participants. For example, pilots entered the weather within the first 90 seconds of the difficult scenario, hence, needed to successfully execute the first route adjustments within that time period, as illustrated by Figure 6(f).

Nonetheless, the weather itself remained static, hence, there was no wind and individual cells did not change or move whilst flying. This also implied that gaps observed between or behind weather cells, remained so throughout each scenario. Given the experiment goal to evaluate the effectiveness of a touch-based navigation display in directly adjusting the flightplan route, not coping with unforeseen circumstances, a dynamic weather system was not deemed necessary and would only add confounding factors. In addition, a static weather scenario allows for easier comparison between participants and individual measurements. As a result, participants were made aware of this fact and advised that any safe route they observed in between weather would therefore remain a safe route.

G. Dependent Measures

In order to evaluate the effectiveness of a touch-based navigation display for lateral weather avoidance, several dependent measures were defined to provide objective and subjective answers to this question. The first two measures were used to determine the effectiveness of the interface in achieving the goals of a lateral weather avoidance task.

1) Total re-route delay time: The total time delay as a result of the weather avoidance trajectory was measured by subtracting the total time flown with the flight time without any route adjustments. Furthermore, aircraft position was continuously logged such that flown tracks could be drawn.

2) Passenger comfort: The root mean square of the roll angle over time, as a consequence of the commanded turns, was measured. This measure was used to determine the comfort level for passengers, given that high roll angles are perceived as unpleasant.

3) Interface inputs: All inputs given on the MCP, CDU and TND interfaces were recorded. This information was used to define and compare task execution times with those predicted by Fitts' law, as well as compare input strategies between participants and determine the length and amount of inputs necessary to complete the primary task. In addition, with the touchscreen interface, the different gestures are logged seperately, allowing for a review of the distribution in panning and zooming gestures and direct route adjustments.

4) Secondary task response time: The secondary task performance was evaluated by logging the response time in recognizing a message including the `KLM9TU` callsign by the participant.

5) Rating scale mental effort: Following each measurement run the Rating Scale Mental Effort (RSME) was used as a subjective measure of the required cognitive effort. The Dutch version of the RSME [33] was used, given Dutch nationality of all twelve participants. The RSME features a scale from 0 to 150, with nine different benchmarks allocating a verbal effort description to an effort rating.

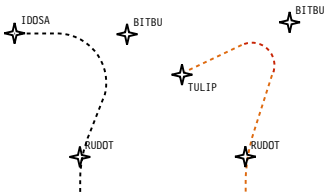
6) De-briefing: Participants commented on their control strategy, information sources, and the effect of the TND on situation awareness, complacency and operational procedures. The de-briefing was used to check the Decision Ladder mapping discussed in Section II-D, and as subjective feedback to help explain experimental observations.

H. Control Variables

In an attempt to avoid confounds, and allow for proper comparison between participants and repeated measures, various experimental factors were kept constant, summarized below.

- Participant relative location to each interface was kept constant. Due to space confinements the TND was located below the CDU, at a position usually occupied by the throttles. As a result, to allow for proper comparison of experiments, participants were moved aft when using the TND. Hence, the participants' relative location to the touchscreen was comparable to that of the MCP and CDU. Markers were installed on the cabin floor to ensure constant seat positioning.
- Participants all used the five-point seatbelt, to simulate realistic body manoeuvrability.
- Participants all used armrests, such that they could comfortably use the CDU and TND located on the pedestal.
- Participants were all right-handed, such that, given they sat on the left seat, they manipulated each interface with their dominant hand.
- Participants all received the same training and briefing, ensured by using a training and (de)briefing checklist.
- Volume, headset and amount of ambient noise to complete the secondary task were the same for all participants.
- Motion was not used during the experiment. The scope of this research was limited to an evaluation of a touchscreen interface in a motionless environment. It would only add complexity and another variable to the assessment.
- Speed and altitude were kept constant, given the focus of the experiment was limited to the lateral plane. The B777 turbulence penetration speed of 270 kts indicated airspeed (IAS) at flight level (FL) 250 was used [24].
- ND map range was limited to 5-640 nm, as in the B777.
- ND modes were limited to MAP and PLAN view.
- Outside visuals featured a field of scattered and broken clouds, with low visibility. Participants were warned of a mismatch between clouds on the outside visuals and those on the ND. The outside visuals were only used to increase the realism of the set-up.

Furthermore, the LNAV algorithm, which calculates and checks the validity of a route, only featured fly-by waypoints. These waypoints can be described as marking the intersection of two straight trajectories, where the circular transition between the two straight segments is calculated by the algorithm based on the turn radius capabilities of the aircraft. The latter was based on the ground speed and a maximum bank angle of 45 degrees. An invalid route, hence one containing a turn with an impossible turn radius, is colored orange with the invalid turn drawn in red, see Figure 8, and could not be executed.



(a) Valid route (b) Invalid route

Fig. 8. Graphical illustration of how route turn radii validity is checked and presented to the pilot

I. Apparatus

The experiment was conducted using the SIMONA Research Simulator (SRS), located at Delft University of Technology. The cabin features a $180^\circ \times 40^\circ$ three-projector DLP collimated display. Furthermore, four 4:3 format LCD displays are located on the main instrument panel, of which the left two were used to display the PFD and ND. The glareshield features a Boeing-style EFIS control panel and MCP. The pedestal features a Boeing-style CDU and large 16:9 format touchscreen display. An illustration of the set-up is given in Figure 7. Participants were placed in the left-hand seat of the SRS, which was horizontally adjustable and included an arm rest and a five-point seatbelt.

J. Procedure

Participants were placed into two groups, which differed in the order in which they used both interfaces. The full experiment lasted 3.5 hours, see Figure 9.

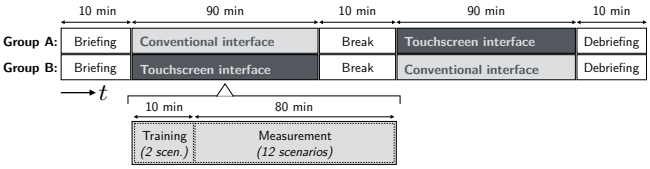
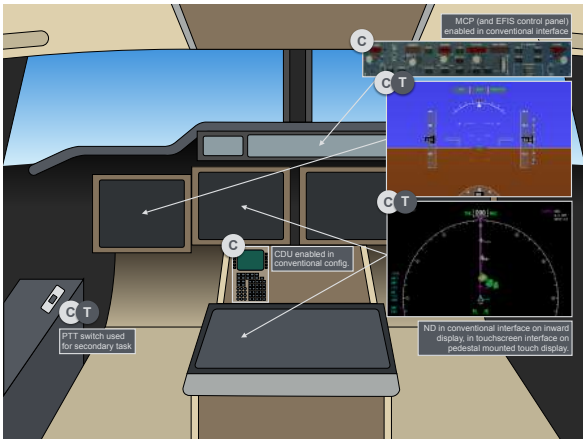


Fig. 9. Time schedule of the experiment

When starting with a new interface a training phase was used to allow the participant to get accustomed to the implementation of the interface. For the conventional interface, given operational familiarity, necessary training was found to be minimum. For the touchscreen interface sufficient training runs were conducted for each individual participant, such that they were adequately accustomed to the necessary gestures and logic behind the LNAV algorithm.

Upon commencing the measurement phase, participants were accustomed to the three different scenarios. Furthermore, following the break, and moving towards the next interface, participants already encountered the scenarios multiple times. This was found to be acceptable given the objective of evaluating interface effectiveness in a re-planning task, not its capability to cope with unforeseen circumstances. Nonetheless, weather scenarios were horizontally inverted and presented with different waypoint names in order to provide variation for the participant without influencing task difficulty. In addition to changing the order of interfaces between groups A and B, the individual sequence of measurement runs was balanced by means of a Latin square [34]. A total of four repeated measures were conducted per experimental condition, of which

(a) The SIMONA Research Simulator (SRS) at the Delft University of Technology



(b) Cabin layout of the SRS showing the available interfaces and displays for both the conventional (C) and touchscreen interface (T)

Fig. 7. Apparatus used for the experiment

there were six, resulting in twelve repetitions per interface (as illustrated in Figure 9).

Following each measurement run participants dit not receive any feedback on their dependent measure scores, as this could influence their control strategy to become more aggressive, and thus less realistic, in order to improve scores. The experiment was concluded with a verbal de-briefing, discussed in Section III-G.

K. Hypotheses

Based on the design of the experiment, reviewed literature, qualitative analysis using the Decision Ladder and the accuracy and throughput Fitts' law models hypotheses can be made. For each of the dependent measures a hypothesis is given and motivated below.

1) No significant difference in re-route time delay or passenger comfort will be observed: Each interface is expected to have an influence on how participants achieve the necessary weather avoidance route. However, it is hypothesized that the same resulting route will be achieved, albeit at a different pace and accuracy. Therefore, the resulting re-route time delay and passenger comfort should be similar.

2) The touchscreen will require less interface inputs for easy scenarios, however a lot more for difficult scenarios: Literature finds touchscreens to have the possibility to reduce workload due to their highly intuitive and direct interface [7]–[11]. Hence, for straightforward scenarios the TND is expected to require less inputs from participants. However, when faced with a challenging weather avoidance scenario, demanding a more complex and lengthy adjustment of the flight plan route, the conventional interface is expected to be much faster. This is due to two important advantages of the conventional interface, consisting of the MCP and CDU.

The MCP allows for a high amount of accuracy, as suggested by Fitts' law models (see Section II-C), for minor route adjustments through narrow corridors in between large weather systems. Executing this procedure using a TND will require a vast number of complex route adjustments, given that these corridors are rarely straight. In addition to the Fitts' law experiment [22], the rotary controller on the MCP has shown to result in high accuracy during research by Ballas, Forlines and Stanton [17], [19], [20].

The CDU is expected to be optimal in comparison to the TND when large route adjustments are necessary. The higher throughput of the CDU, as discussed in Section II-C, supports this hypothesis as a larger throughput allows for faster task execution times. Furthermore, Ballas also found an interface similar to the CDU to result in faster task execution, especially for difficult tasks [17].

Fitts' law models developed of each for these interfaces were used to predict total task execution (movement) times and the ID. The control task, using either interface, was broken down into individual necessary movements for each scenario. For each movement, the interfaces were used to determine the scale (target amplitude) and necessary accuracy (target width). For the conventional interface, basic trigonometry and the geometry of each scenario were used to determine the necessary heading commands and tolerances. For the touchscreen interface, a ND range of 20 nm was assumed to determine the amount of pixels a waypoint had to be moved across the display and with what accuracy it had to be positioned. Note that these predictions only include route adjustments, zooming and panning gestures were omitted given the difficulty of predicting how many of these gestures are necessary.

TABLE II
RESULTS OF FITTS' LAW PREDICTIONS BASED ON MODELS DISCUSSED IN SUBSECTION II-C

	$\sum$ Index of Difficulty [-]			$\sum$ Movement Time [s]		
Interface	Easy	Med.	Diff.	Easy	Med.	Diff.
Conventional	9.59	19.12	41.46	5.22	10.06	18.27
Touchscreen	5.46	17.40	52.13	4.86	11.08	31.05
Difference	*-43.1%*	*-9.0%*	*25.7%*	*-6.9%*	*10.1%*	*70.0%*

The results of these predictions, presented in Table II, further motivate this hypothesis. For the easy scenario the predicted MT and ID is indeed lower for the touchscreen, however, becomes substantially larger when looking at the difficult scenario.

3) Secondary task response times will be lower for the conventional interface: With the conventional interface the participant is providing continuous minor adjustments of the commanded heading in order to avoid weather, instead spending the beginning of each scenario analyzing possible routes and their validity. Hence, they were expected to spend more time in the lower parts of the Decision Ladder, as discussed in Section II-D, and have more spare cognitive capacity to perform the secondary task. This effect is expected to be largest during the most difficult scenario.

4) Mental effort scores will be lower for the conventional interface: A similar motivation is warranted for this hypothesis as the previous one. Spending less time in the upper parts of the decision ladder is hypothesized to result in lower mental effort scores. Both Degani [14] and Stanton [20] found improved workload ratings for non-touchscreen input devices in their research.

IV. RESULTS

A total sample size of 12 was recorded for each experiment condition. Given a within-subjects design, multiple measurements were acquired per participant and condition, as shown in Table III. Due to logistical reasons, some participants could not complete four repetitions for each condition. Overall, this resulted in two to six less measurements for five of the six conditions.

A similar statistical approach was used to check significance of the results. First, individual measurements per participant per condition were averaged and corrected for between-subjects variability. A Kolmogorov-Smirnov test was used to confirm normal distribution of the data. In case of normally distributed results, a Maulchy's test was used to check the validity of the sphericity assumption. Greenhouse-Geisser

TABLE III
SAMPLE SIZE PER EXPERIMENTAL CONDITION

Interface	Task difficulty	Sample size	Measurements
Conventional	Easy	12	43
	Medium	12	42
	Difficult	12	42
Touchscreen	Easy	12	46
	Medium	12	45
	Difficult	12	48*

*Some participants failed in avoiding weather during this condition, which led to meaningless re-route delay time results. Hence, for this variable only 35 measurements were considered (see Section IV-A).

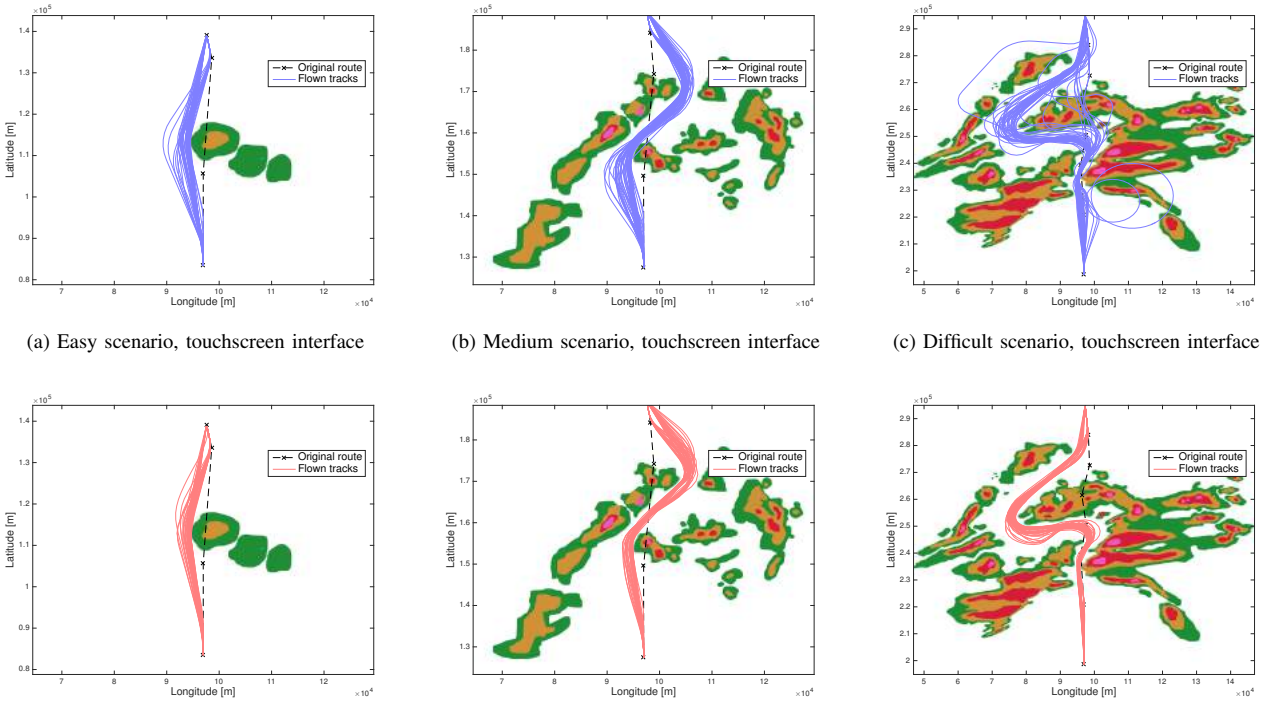
sphericity estimates were used to correct the degrees of freedom ($\epsilon < .750$) when sphericity could not be assumed. Finally, a two-way repeated measures ANOVA was used to check for significance of the results. However, for data that were not normally distributed a different approach was necessary. In these cases a Friedman's ANOVA was used to check for significance across all six conditions, and individual Wilcoxon signed rank tests were used to check pairwise significance.

A. *Re-route Delay Time*

The flown track of every measurement run for each experimental condition is presented in Figure 10. An overlay of the weather is drawn onto the same figure, allowing for a check if weather was successfully avoided. An initial observation, is that there was less variability in flown tracks for the conventional interface (see Figure 10(d), (e), (f)), compared with the touchscreen interface (see Figure 10(a), (b), (c)).

A second observation is that the difficult scenario, using the touchscreen interface, is the only condition that included measurements where participants failed at avoiding the weather. Here, during a considerable number of runs they were unable to successfully avoid the weather, and subsequent penetration of the weather is visible. When participants failed in timely modifying their flight plan route to circumnavigate the weather, they continued with a course due north, very similar to the original route and therefore resulting in a low re-route time delay. Furthermore, outliers are also observed at the other extreme, where participants logged very large re-route time delays. The cause can be observed in Figure 10(c) where a few participants accidentally executed a new route with a waypoint placed behind the aircraft's present position. As a result the LNAV algorithm generated a 360 degree turn, clearly visible, to return to this waypoint. This obviously resulted in a very large re-route time delay. Given that participants failed at executing the primary task during these cases, the individual re-route time delay measurements are meaningless, and have therefore been removed from the re-route delay time results presented in Figure 11(a).

The results for re-route delay time are shown in Figure 11(a). A clear difference can be seen between scenarios, however no differences between interfaces are evident. This finding was confirmed following a statistical analysis. Given a



(a) Easy scenario, touchscreen interface

(b) Medium scenario, touchscreen interface

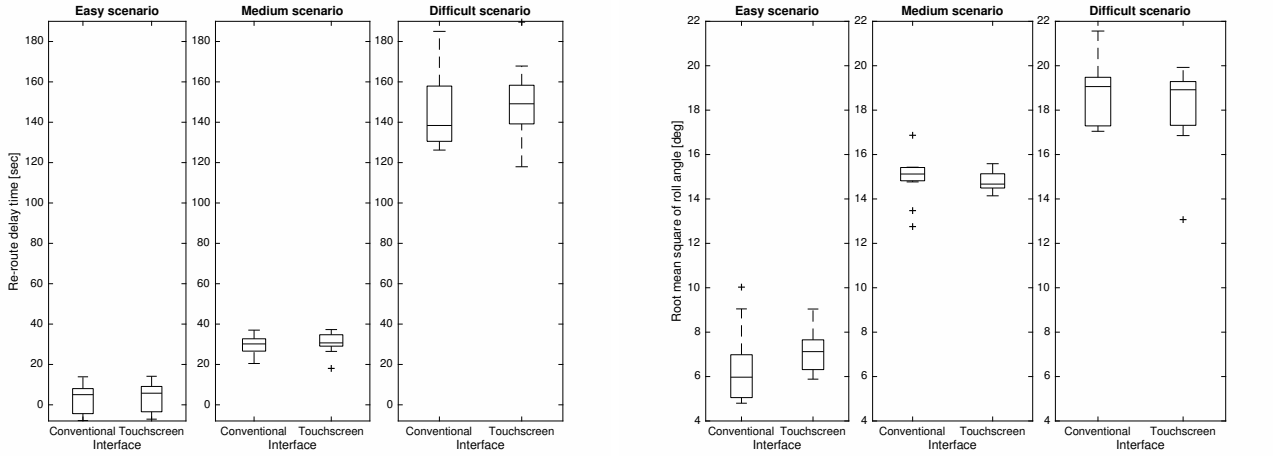
(c) Difficult scenario, touchscreen interface

(d) Easy scenario, conventional interface

(e) Medium scenario, conventional interface

(f) Difficult scenario, conventional interface

Fig. 10. Flown tracks per experimental condition



(a) Mean total re-route delay time recorded per condition per participant, corrected for between-subjects variability

(b) Root mean square of roll angle recorded per condition per participant, corrected for between-subjects variability

Fig. 11. Primary task performance

normal distribution of the data, a two-way repeated measures ANOVA was used and concludes that there was a highly significant main effect of task difficulty, $F(1.045, 11.495) = 793.005, p < .001, r = .985$. On the contrary, no significant effect is found of interface on re-route delay times, $F(1, 11) = .272, p > .05$.

The root mean square (RMS) of roll angle, used to describe passenger comfort, is presented in Figure 11(b). A similar observation can be made here. A clear difference is seen between between scenarios, where the difficult scenario resulted in, on average, the largest roll angles. This matches the steep turns observed in Figure 10(c) and Figure 10(f). Finally, no differences are observed between interfaces. This suggests that the interface had no effect on the participant's ability to achieve a particular level of passenger's comfort. These observations are confirmed following a statistical analysis. The data are not found to be normally distributed, $D(12) = .277, p < .05$, hence, Friedman's ANOVA is used. Comparing all six conditions, a significant effect of experimental condition on the RMS of roll angle is found, $\chi^2(5) = 53.810, p < .001$. Subsequently, Wilcoxon signed rank tests show no pairwise significance between interfaces for each scenario, $p > .05$.

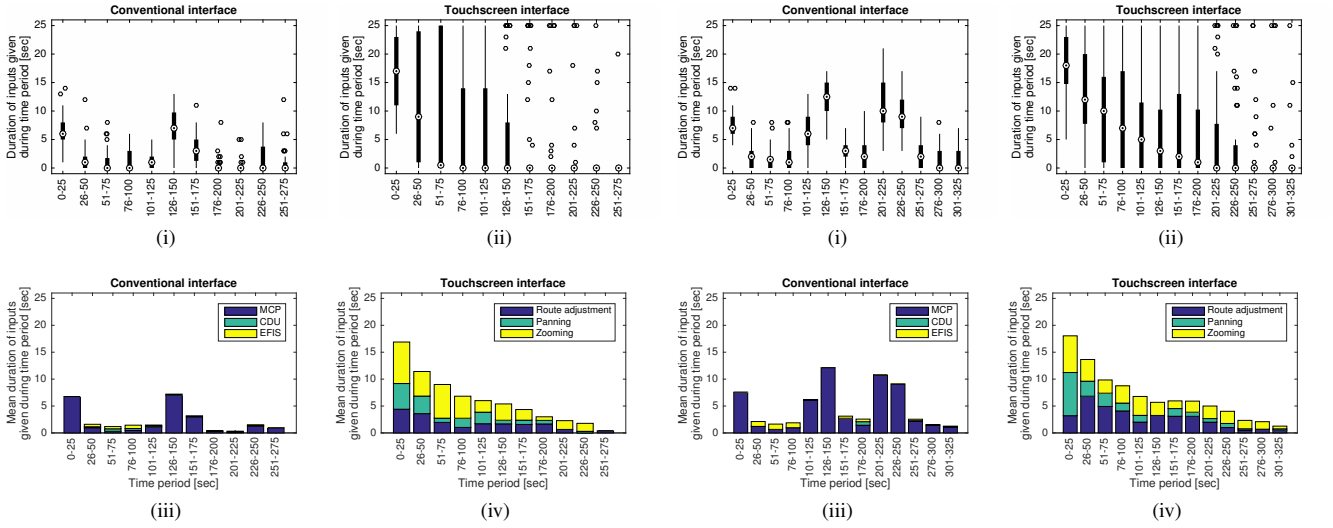
B. Interface Inputs

The timestamp and type of each input given, whether on the MCP, EFIS control panel, CDU or touchscreen was logged. Each scenario was divided into periods of 25 seconds, and for each period the total duration of recorded inputs was derived and presented in Figure 12. The difficult scenario took much longer to complete, hence, Figure 12(c) is drawn larger. Results are presented with two different representations: in a boxplot, showing the variation in recorded inputs, and on a barplot, showing the mean input duration per type of input. For the conventional interface these include heading commands using the MCP, route adjustments using the CDU and display adjustment using the EFIS control panel. For the touchscreen interface these included display adjustments (zooming and panning) in addition to direct route adjustments.

An intriguing observation can be made when closely scrutinizing the variation in input duration for the touchscreen interface with each scenario in Figure 12(a)-ii, (b)-ii and (c)-ii. A profound drop in the spread of input durations is observed during the last 125 seconds of each scenario, see Figure 12(a)-ii, (b)-ii and (c)-ii. The effect is most prominently visible during the difficult scenario, in Figure 12(c)-ii. Given a constant speed, this corresponds to the moment that the entire remainder of the route was visible within the weather radar range of 20 nautical miles, and hence all possible obstacles were visible to the participant. This thus shows a significant decrease in the amount of inputs necessary with the touchscreen interface once the complete situation was visible to the pilot. In addition, a general decline in touchscreen inputs is observed for each scenario, as illustrated in Figure 12(a)-iv, (b)-iv and (c)-iv. Furthermore, the initial time periods consist of a large number of display adjustments (panning and zooming), which shows that participants started by scanning the scenario first before adjusting the route.

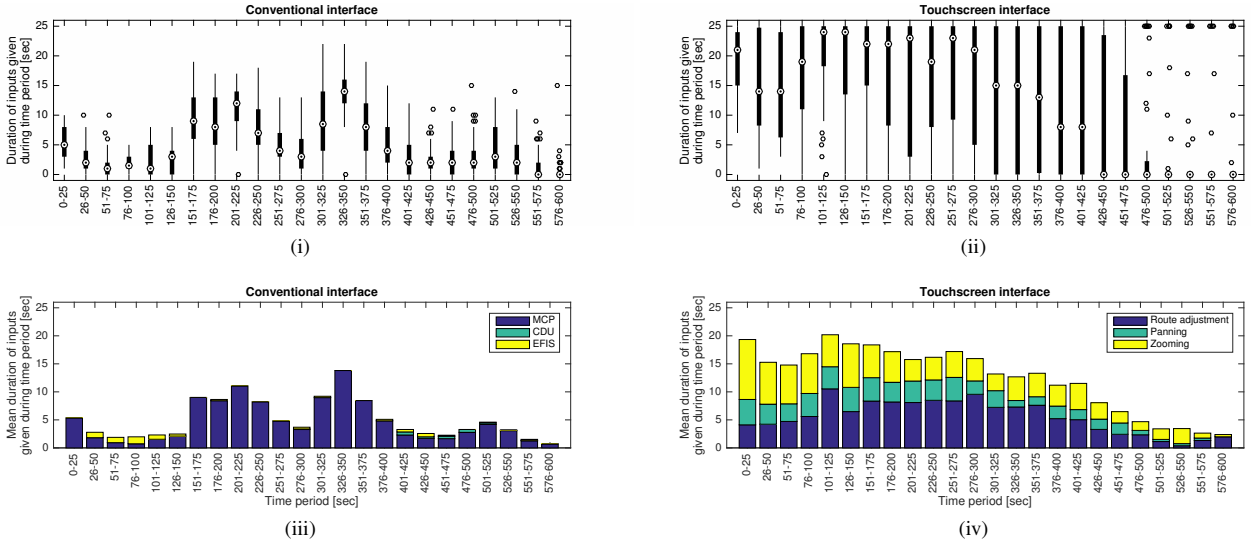
The opposite is seen for the conventional interface. During each scenario, participants immediately provided heading select inputs using the MCP, see Figure 12(a)-iii, (b)-iii and (c)-iii. As expected, and suggested by the control task analysis in Section II-D, clear peaks and troughs are seen in these figures. This effect is most visible during the easy scenario, see Figure 12(a)-iii, where only two time periods show a significant number of MCP inputs, with very little activity in the remaining periods.

Based on the type of input, the mean duration of given



(a) Easy scenario

(b) Medium scenario



(c) Difficult scenario

Fig. 12. Inputs given per interface over the course of each scenario

TABLE IV
MEAN DURATION OF INPUTS DURING ENTIRE SCENARIO

	Route adjustment [sec]			Display adjustment [sec]		
Interface	Easy	Medi.	Diff.	Easy	Medi.	Diff.
Conventional	23.63	58.24	113.43	2.49	4.26	6.81
Touchscreen	18.71	35.21	132.05	48.66	54.13	166.52
Difference	*-20.1%*	*-39.5%*	*16.4%*	-	-	-

inputs during the entire scenario can be calculated, and is presented in Table IV. The duration of route adjustments is a sum of all MCP and CDU inputs for the conventional interface. The duration of display adjustments is a sum of all EFIS inputs for the conventional interface, and a sum of all panning and zooming gestures for the touchscreen interface. As expected, the mean number of route adjustments is found to be lower using the touchscreen for the easy and medium scenario, and larger for the difficult scenario, with the touchscreen interface. The large number of display adjustments using the touchscreen, observed in Figure 12(a)-iv, (b)-iv, (c)-iv, is quantified here. A comparison between these results and the Fitts' law predictions, see in Table II, will be discussed in Section V.

A comparison between participants is provided in Figure 13, showing the mean duration of inputs given per participant for all three scenarios combined. Once again, inputs are separated

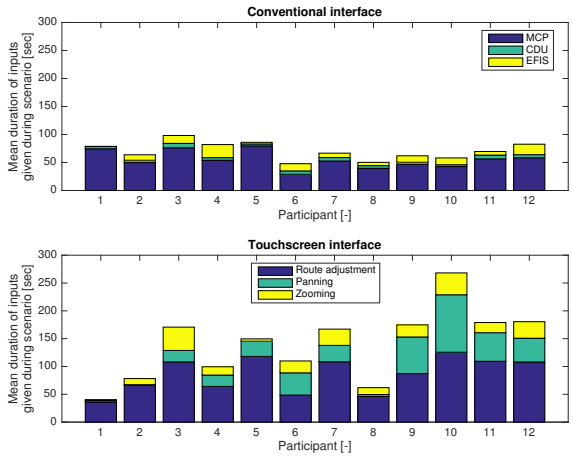


Fig. 13. Mean duration of inputs given per participant for all three scenarios

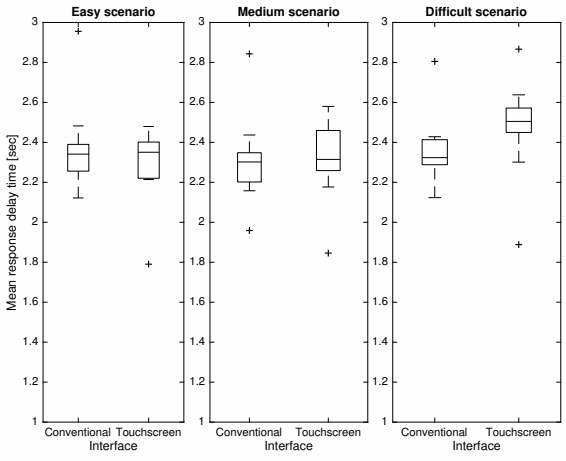


Fig. 14. Mean secondary task response delay times, corrected for between-subjects variability

into the three different types per interface. For the conventional interface, a very comparable input strategy is observed per participant, given very similar input durations. On the contrary, for the touchscreen interface large differences are observed between participants. For example, Participant 1 provided very minimal number of route adjustment inputs and nearly no zoom or pan gestures whilst Participant 10 used, on average, four times as much route adjustment inputs for the same scenarios, and a significant number of zoom and panning gestures. In general, these differences were also observed during the experiment. Certain participants were constantly attempting to optimize the route, whilst vigorously adjusting the display, whereas other participants were satisfied once the first acceptable weather avoidance route was executed.

C. Secondary Task Scores

Mean secondary task response delay times were recorded and plotted in Figure 14. For both the easy and medium scenario no difference is seen in response times, however, an average increase of 5.1% for the touchscreen is observed for the difficult scenario. This suggests that participants were busier adjusting the route to succesfully avoid the weather with the touchscreen in this scenario. As a result, they had less cognitive capacity left to monitor the ATC radio chatter and quickly respond to messages.

A violation of normality was found, $D(12) = .272, p < .05$, hence, a Friedman's ANOVA was used to provide statistical backing to these observations. When comparing all six experiment conditions, mean response delay time was found to be significantly affected by the task difficulty and interface combination, $\chi^2(5) = 14.056, p < .05$. Wilcoxon signed rank tests were used to follow-up on this finding, by comparing the effect of interface for each scenario. No significant effect of interface was found for the easy and medium scenarios, $Z = 32.00, p > .05, r = -.018$ and $Z = 34.50, p > .05, r = -.072$ respectively. However, for the difficult scenario a larger effect was found of interface on

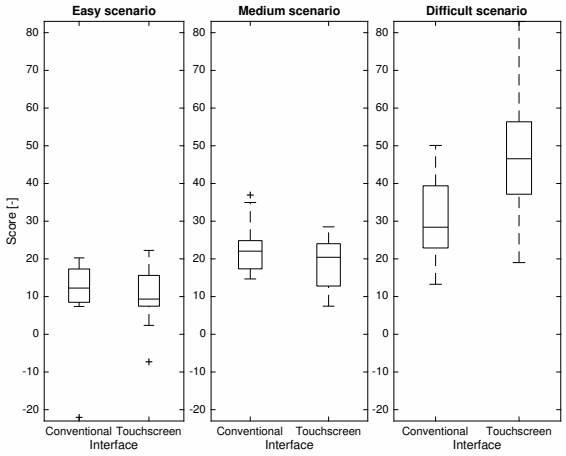


Fig. 15. Mean Rating Scale Mental Effort RSME scores, corrected for between-subjects variability

mean response delay times. Whilst the Wilcoxon test resulted in $Z = 15.00, p > .05, r = -.384$, the 2-tailed significance was found to be $p = .060$. Furthermore, the effect is found to be medium in size, $r = -.384$. This suggests that the interface had a medium-sized effect on the secondary task performance during the difficult scenario, however, not substantial in size. It is expected that, given a larger sample size, a significant effect of interface on mean response delay time for the difficult scenario would have been found.

D. Rating Scale Mental Effort

Scores given by participants using the Rating Scale Mental Effort (RSME) are shown in Figure 15. Similar scores can be observed for the easy and medium scenario, however, for the difficult scenario the average given score was 53.8% larger using the touchscreen. Nonetheless, a large spread can be seen in RSME scores for this condition. Some participants managed

to adjust the route successfully within a short period of time, resulting in a smooth passage of weather. Other participants did not succeed in doing so, and as a result experienced a significant amount of stress in safely avoiding the weather. This may have led to the large variation in scores observed between participants. It is expected that a larger sample size would decrease the spread in recorded RSME scores.

The RSME results were found to be normally distributed, hence a two-way repeated measures ANOVA was used to provide statistical backing to these observations. Overall, a highly significant main effect of task difficulty was found, $F(1.183, 13.012) = 26.036, p < .01, r = .822$. Furthermore, the difficulty $\times$ interface interaction was significant, $F(2, 22) = 10.904, p < .01, r = .672$, indicating that ratings across different levels of task difficulty were different for each interface. This confirms that the three scenarios were indeed varying in the level of difficulty. A post-hoc dependent t-test was performed on the difficult scenario alone, which shows a significant and substantial effect of the interface on the mental effort scores for the difficult scenario, $t(11) = -2.898, p < .05, r = .658$. This suggests that participants experienced the difficult scenario as more demanding when using the touchscreen interface.

V. Discussion

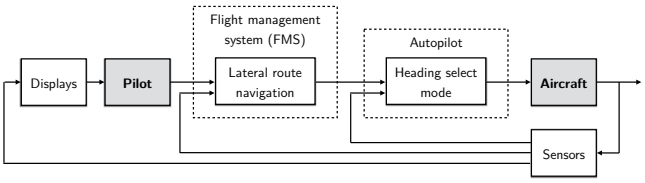
The goal of this paper was to evaluate a touch-based navigation display for a lateral weather avoidance task. The effectiveness of a touch-based navigation display was tested using an experiment, by comparing it to a conventional interface.

A. Effectiveness of a touch-based navigation display in circumnavigating weather

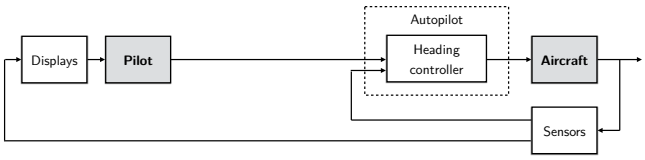
Participants were asked to re-plan their routes with the goal of minimizing re-route time delay whilst ensuring safety and passenger comfort. In general, the touchscreen interface was found to be effective in circumnavigating en-route weather, and no significant difference was found when comparing re-route time delay and the RMS of roll angle with the conventional interface. Hence, Hypothesis 1 is accepted.

However, when analyzing the flown tracks per experimental condition some differences were observed. During the difficult scenario several participants failed in safely circumnavigating weather using the touchscreen interface, resulting in flown tracks that penetrated the weather systems. Here, the first portion of the route adjustment needed to be successfully inserted within ninety seconds of the scenario, leading to substantial time pressure. In all failed attempts to avoid the weather, participants were seen to struggle with adjusting the route such that all proposed turns were valid. Often, an invalid turn further along the route would block the pilot from executing the necessary route and safely avoiding the most imminent threat. This effect was found to frustrate each participant.

Hence, all participants unanimously commented that their acceptance of the touchscreen interface was conditional on the ability to return to a more direct control of the aircraft, such as the heading select mode, during time-critical situations. This desire is illustrated graphically in Figure 16, where more direct control of the aircraft removes a control loop (Figure 16b), thereby simplifying the control task (compared to Figure 16a).



(a) Lateral navigation through `LNAV` mode by adjusting the flight plan route



(b) Lateral navigation through `HDG SEL` mode by commanding a heading

Fig. 16. Control loops during lateral navigation

In addition, ten out of twelve participants commented that in such a situation a multi-crew environment is advantageous. For example, the pilot flying reverts to the heading select mode, thereby directly commanding the necessary heading to avoid the closest obstacle. In the meantime, the co-pilot adjusts the LNAV route, without the time pressure of avoiding the closest weather cells. This suggests that a touchscreen interface would not be able to support single pilot operations, a discussion that is keeping the industry very busy today [35]–[38].

B. Ability of a touch-based navigation display to reduce task execution time and error rate

Based on Fitts' law accuracy and throughput models, it was expected that the touchscreen would require a shorter task execution time for easy scenarios. However, for more difficult scenarios a much longer task execution was expected. Based on the results, the control task took 20.1% less time and 16.4% longer for the easy and difficult scenario, respectively. Hence, Hypothesis 2 is accepted.

Overall, control task durations were measured to be longer than those predicted using the Fitts' law models. During the experiments conducted in [22], participants were asked to provide rapid and aimed movements, with the goal of acquiring a 96% success rate. In addition, they were tasked to provide simple, individual inputs. On the contrary, during the experiment discussed in this paper, the necessary inputs were part of a larger and more complex route adjustment task. Hence, a significant amount of cognitive workload is included in the measured inputs. For example, participants were observed to try multiple, different route adjustments, simply to see which resulted in an optimal trajectory. Hence, the measured input durations were longer than those predicted by Fitts' law.

Fitts' law found a large throughput for the CDU, resulting in very short expected movement times. Indeed, when scrutinizing Figure 12 the measured input duration using the CDU is very limited. The majority of inputs using the conventional interface can be attributed to the MCP, which coincides with the very low throughput score measured using Fitts' law. Furthermore, the large variation in measured touchscreen accuracy during the Fitts' law experiment helps to clarify the number of route adjustments necessary, especially during the difficult scenario. Waypoints had to be placed very precisely in order to achieve a route that both successfully avoided the weather and was valid. The size of these *hotspots*, illustrated with the diameter of the circles, where the resulting route would have viable turn radii, are shown in Figure 17. In order to achieve the necessary accuracy, participants were observed, and also commented on doing so, to continuously adjust the zoom. Zooming in provided more room to place waypoints, whilst zooming out was necessary to maintain an overview and situational awareness.

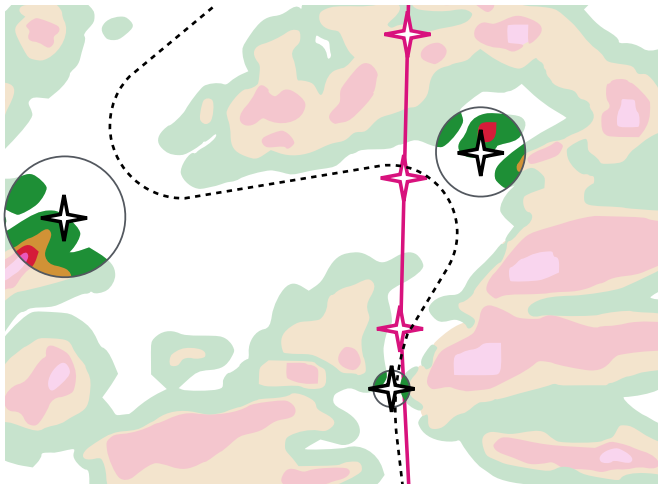


Fig. 17. Necessary accuracy in placing waypoints using the TND and difficult scenario. These hotspots are highlighted for each waypoint, where the diameter of the circles indicate their size.

Finally, participants remarked that once all obstacles were visible, i.e., were within the weather radar range, the final adjustments could be executed and no more inputs were necessary. This effect is clearly visible in Figure 12, showing a sharp decline in input duration during the last 125 seconds of each scenario, which corresponds to the portion of each scenario where the entirity of the weather is observed within the radar range of 20 nm.

Hence, the touchscreen interface clearly shows an ability to reduce task execution times for easier re-planning tasks where the complete scenario is known beforehand. However, for very fine manoeuvering tasks the touch-based navigation display results in larger task execution times, largely attributed to its lower accuracy. This supports participant feedback requesting more direct control of the aircraft when the re-planning task becomes too demanding and suggests that a TND would be more suitable for less dynamic re-planning tasks, such as planning an approach procedure. Finally, this supports the often found remark in literature that touchscreen effectiveness is highly task-dependent [14], [17]–[20].

C. Ability of a touch-based navigation display to support decision making

In Section II-D, a comparison was made between both interfaces using the decision ladder. Indeed, the continuous loop between a monitoring and revision phase for the conventional interface was observed during the experiment. Analyzing the inputs given, see Figure 12, clear spikes in input duration are observed at the necessary turns during each scenario when participants resided in the revision phase. In between these moments few inputs were given, confirming their presence in a cognitively less demanding, monitoring phase.

On the other hand, from the decision ladder analysis it was hypothesized that the usage of the touchscreen interface would remain in the cognitively demanding knowledge-based behaviour phase whilst participants were analyzing all possible routes. The experimental data indeed reflected this expectation, with a large number of touchscreen inputs at the beginning of each scenario, which decreased over time. During each scenario, a complete withdrawal from the analysis phase is only observed once the entire weather scenario was visible, hence within the radar range during the final 125 seconds of the scenario. This explains the differences in secondary task response times and mental effort scores, as illustrated in Figure 14 and Figure 15, for the difficult scenario. However, for both the easy and medium scenarios no significant differences were observed for both measures. This suggests that, when using the touchscreen, the period participants were analyzing the situation, and thus stuck in the upper part of the DL, was manageable. Hence, Hypotheses 3 and 4 are partially accepted.

D. Influence of a touch-based navigation display on pilot complacency and situation awareness

1) Complacency: The experiment discussed in this paper focused on a re-planning task within a static weather environment. Using the touchscreen interface, especially for easier scenarios, participants may become complacent once the modified route is executed. Placed in a dynamic weather scenario, with the risk of a continuously changing environment, such complacency is certainly dangerous. Whilst the avoidance route may be clear of weather upon execution, this may change as time progresses. The effect of continuous inputs using the heading select mode with the conventional interface has a positive effect on complacency, given that the pilot has to remain in the loop. Hence, a majority of participants proposed a preference for the touchscreen interface for re-planning tasks that are not subject to a dynamically changing environment, such as terrain avoidance or planning an approach to an airport.

2) Situation Awareness: Participants had mixed feelings concerning the influence of the TND on their situation awareness. A compelling advantage of the touchscreen interface is the visual representation of the desired route. This allows pilots to visualize their mental model of the desired route, and as

a result easily communicate their intentions with a co-pilot and ATC. Furthermore, due to direct adjustment of the route using the touchscreen interface, the subsequent visual feedback on aircraft performance in terms of capable turn radii was appreciated.

On the contrary, the position of the touchscreen on the pedestal led to a large amount of head-down time. Participants remarked that during real life weather avoidance, especially during daylight, the focus should be on looking outside. Hence, the MCP is conveniently located in the line of sight out the windows. However, given a multi-crew operation, having only one pilot incur substantial head-down time can be acceptable.

E. General remarks on experiment design and operational realism

Based on both experiment observations and discussions with participants the secondary task was found to be, whilst realistic, too easy. Simply recognizing a callsign, followed by pressing a switch, required little cognitive effort for participants even when vigorously interacting with either interface. This explains the lack of a statistically significant and substantial effect. A more demanding secondary task, requiring more mental effort, is recommended. For example, participants could be asked to actively respond to the content of the ATC message. This would require them to listen differently and understand the full message, instead of merely recognizing the callsign.

Furthermore, several participants had difficulty selecting waypoints due to sticky or cold fingers, which was found to be frustrating. In addition, participants remarked that when directly manipulating the ND, their hand was often obscuring the underlying waypoints, weather and route. Finally, the ability to easily operate a touchscreen within a dynamic environment such as turbulence remains questionable. These remarks confirm existing knowledge on touchscreen effectiveness [14], [16]–[20]. In regard to the LNAV algorithm, the ability to use fly-over and custom waypoints was desired by a majority of participants. Furthermore, it was suggested that the algorithm automatically remove route waypoints behind the aircraft when on heading select mode. Hence, it is expected that improvements in the software processing touch inputs will increase acceptance of a TND.

Summarizing, aircraft and equipment manufacturers are urged to carefully consider the use of touchscreen interfaces on the flight deck. For example, a TND could be used to help pilots plan departure and approach procedures prior to take-off, when ample time is available. Effective airborne use of a TND is conditional on the presence of time or designated multi-pilot procedures. For instance, the pilot flying temporarily directs the aircraft safely around weather using the heading select mode, whilst the pilot not flying uses a TND to successfully adjust the LNAV route without time pressure. Whilst this approach may work, further research is recommended on the effectiveness of a touch-based navigation display for a dynamic weather environment.

VI. Conclusion

An experiment was conducted in the SIMONA Research Simulator, where 12 commercial pilots participated in evaluating a touch-based navigation display during a lateral weather avoidance re-planning task. An explicit comparison was made with a conventional interface, including the heading select mode on the MCP and route adjustment using the CDU. Three different scenarios were designed, varying in weather avoidance task difficulty.

Results show that weather avoidance for easy scenarios, where ample time is present to adjust the route, the TND proved to be very effective. Lower task execution times were observed, whilst re-route time delay, secondary task performance and mental effort scores were similar to those of the conventional interface. However, for more complex and time-critical scenarios the TND effectiveness is questionable. In these scenarios participants sometimes failed at circumnavigating weather safely, provided a substantially larger number of inputs, scored worse on the secondary task and gave higher mental effort ratings. A direct interaction and adjustment of the LNAV route turned out to be complex and demanded too much throughput and accuracy from the touchscreen to successfully complete the task. Here, the ability to revert to a more direct control of the aircraft, such as the heading select mode, is necessary.

References

[1] B. Bulfer, *Big Boeing FMC User's Guide*. Leading Edge Publishing, 1991.

[2] SESAR Joint Undertaking, "i4D and SESAR," 2014.

[3] M. Carmona, D. Rudinskas, and C. Barrado, "Design of a flight management system to support four-dimensional trajectories," *Aviation*, vol. 19, no. 1, pp. 58–65, 2015.

[4] M. Mulder, R. Winterberg, M. M. van Paassen, and M. Mulder, "Direct Manipulation Interfaces for In-Flight Four-Dimensional Navigation Planning," *The International Journal of Aviation Psychology*, vol. 20, no. 3, pp. 249–268, 2010. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10508414.2010.487010

[5] H. Huisman, R. Verhoeven, Y. van Houten, and E. Flohr, "Crew interfaces for future ATM," pp. 33–40, 1997.

[6] B. J. A. van Marwijk, C. Borst, M. Mulder, M. Mulder, and M. M. van Paassen, "Supporting 4D Trajectory Revisions on the Flight Deck: Design of a Human-Machine Interface," *The International Journal of Aviation Psychology*, vol. 21, no. 1, pp. 35–61, 2011. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10508414.2011.537559{%}5Cnhttp://www.tandfonline.com/doi/pdf/10.1080/10508414.2011.537559

[7] G. Polek, "Boeing 777X Cockpit to Feature Touchscreen Displays," 2016. [Online]. Available: http://www.ainonline.com/aviation-news/business-aviation/2016-07-07/boeing-777x-cockpit-feature-touchscreen-displays

[8] S. Kaminani, "Human computer interaction issues with touch screen interfaces in the flight deck," *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*, 2011.

[9] B. Shneiderman, "The future of interactive systems and the emergence of direct manipulation," *Behaviour & Information Technology*, vol. 1, no. 3, pp. 237–256, 1982.

[10] W. A. Rogers, A. D. Fisk, A. C. McLaughlin, and R. Pak, "Touch a screen or turn a knob: choosing the best device for the job." *Human factors*, vol. 47, no. 2, pp. 271–88, 2005. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16170938

[11] E. L. Hutchins, J. D. Hollan, and D. A. Norman, "Direct Manipulation Interfaces," in *Human-Computer Interaction*. San Diego: Lawrence Erlbaum Associats, Inc., 1985, vol. 1, pp. 311–338. [Online]. Available: https://www.irjet.net/archives/V2/i6/IRJET-V2I6118.pdf

[12] M. Thurber, "Touchscreens clean up gulfstream symmetry flight deck," 2015. [Online]. Available: http://www.ainonline.com/aviation-news/business-aviation/2015-01-02/touchscreens-clean-gulfstream-symmetry-flight-deck

[13] GARMIN, "Garming G5000." [Online]. Available: https://buy.garmin.com/en-US/US/in-the-air/general-aviation/flight-decks/g5000-/prod90821.html

[14] A. Degani, E. A. Palmer, and K. G. Bauersfeld, ""Soft" Controls for hard displays: still a challenge," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 36, pp. 52–56, 1992. [Online]. Available: http://pro.sagepub.com/content/36/1/52.short

[15] M. Mertens, H. Damveld, and C. Borst, "An Avionics Touch Screen based Control Display Concept."

[16] S. Dodd, J. Lancaster, A. Miranda, S. Grothe, B. DeMers, and B. Rogers, "Touch Screens on the Flight Deck: The Impact of Touch Target Size, Spacing, Touch Technology and Turbulence on Pilot Performance," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, no. 1, pp. 6–10, 2014. [Online]. Available: http://pro.sagepub.com/content/58/1/6.short{%}5Cnhttp://pro.sagepub.com/lookup/doi/10.1177/1541931214581002

[17] J. A. Ballas, C. L. Heitmeyer, and M. A. Pérez, "Evaluating two aspects of direct manipulation in advanced cockpits," *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '92*, pp. 127–134, 1992. [Online]. Available: http://dl.acm.org/citation.cfm?id=142750.142770

[18] F. B. Bjorneseth, M. D. Dunlop, and E. Hornecker, "Strathprints Institutional Repository Assessing the Effectiveness of Direct Gesture Interaction for a Safety Critical Maritime Application," vol. 70, pp. 729–745, 2012.

[19] C. Forlines, D. Wigdor, C. Shen, and R. Balakrishnan, "Direct-touch vs. mouse input for tabletop displays," *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*, p. 647, 2007. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1240624.1240726

[20] N. A. Stanton, C. Harvey, K. L. Plant, and L. Bolton, "To twist, roll, stroke or poke? A study of input devices for menu navigation in the cockpit." *Ergonomics*, vol. 0139, no. February 2013, pp. 37–41, 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23384222

[21] G. Stuyven, H. Damveld, and C. Borst, "Concept for an Avionics Multi Touch Flight Deck," *SAE International Journal of Aerospace*, vol. 5, no. 1, pp. 164–171, 2012.

[22] N. van Zon, "Using Fitts ' Law to Develop an Accuracy and Throughput Model of Three Flight Deck Interfaces," 2017.

[23] J. Rasmussen, *Information Processing and Human-Machine Interaction*, series vol ed. New York: Elsevier Science Publishing Co., Inc., 1986.

[24] Boeing Commercial Aircraft, *777-200LR Operations Manual*, Everett, 2012.

[25] R. W. Soukoreff and I. S. MacKenzie, "Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI," *International Journal of Human Computer Studies*, vol. 61, no. 6, pp. 751–789, 2004.

[26] R. Jagacinski and J. Fisch, "Information Theory and Fitts' Law," in *Control Theory for Humans*, 1997, ch. 3, pp. 17–26.

[27] I. S. MacKenzie, "Fitts' law as a research and design tool in human-computer interaction," *Human-Computer Interaction*, vol. 7, no. 1, p. 48, 1992. [Online]. Available: http://portal.acm.org/citation.cfm?id=1461857

[28] C. L. MacKenzie, R. G. Marteniuk, C. Dugas, D. Liske, and B. Eickmeier, "Three-dimensional movement trajectories in Fitts' task: Implications for control," *The Quarterly Journal of Experimental Psychology Section A*, vol. 39, no. 4, pp. 629–647, 1987. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/14640748708401806

[29] I. S. MacKenzie and W. Buxton, "EXTENDING FITTS' LAW TO TWO-DIMENSIONAL TASKS INTRODUCTION Since the advent of direct manipulation," *ACM CHI'92 Conference*, pp. 219–226, 1992.

[30] W. R. Soukoreff and I. S. MacKenzie, "Theoretical upper and lower bounds on typing speed using a stylus and a soft keyboard," *Behaviour & Information Technology*, vol. 14, no. 6, pp. 370–379, 1995. [Online]. Available: http://dx.doi.org/10.1080/01449299508914656

[31] M. F. Stoelen and D. L. Akin, "Assessment of Fitts' law for quantifying combined rotational and translational movements." *Human factors*, vol. 52, no. 1, pp. 63–77, 2010.

[32] X. Bi, Y. Li, and S. Zhai, "FFitts law: Modeling Finger Touch with Fitts' Law," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, p. 1363, 2013. [Online]. Available: http://dl.acm.org/citation.cfm?id=2470654.2466180

[33] F. Zijlstra, *Efficiency in work behaviour: A design approach for modern tools*. Delft: Delft University Press, 1993.

[34] A. Field and G. Hole, *How to Design and Report Experiments*, 1st ed. London: SAGE Publications Ltd, 2003.

[35] P. C. Schutte, K. H. Goodrich, D. E. Cox, E. B. Jackson, M. T. Palmer, A. T. Pope, R. W. Schlecht, K. K. Tedjojuwono, A. C. Trujillo, R. A. Williams, J. B. Kinney, and J. S. Barry, "The Naturalistic Flight Deck System : An Integrated System Concept for Improved Single-Pilot Operations," no. December, p. 56, 2007.

[36] W. W. Johnson, D. Comerford, S. L. Brandt, R. Mogford, and V. Battiste, "NASA / CP - 2013 - 216513 NASA 's Single -Pilot Operations Technical Interchange Meeting : Proceedings and Findings," no. April, 2013. [Online]. Available: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20140008907.pdf

[37] J. Lachter and R. Collins, "Toward Single Pilot Operations : Developing a Ground Station," *Proceedings of the International Conference on Human-Computer Interaction in Aerospace*, pp. 1–9, 2014.

[38] A. Stimpson, J. Ryan, and M. L. Cummings, "Assessing Pilot Workload in Single-Pilot Operations with," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 60, no. 1, pp. 675–679, 2016. [Online]. Available: http://pro.sagepub.com/lookup/doi/10.1177/1541931213601155

Part II

IEEE Companion Article

Note: A portion of this article has already been examined under course AE4020 Literature Study

Using Fitts' Law to Develop an Accuracy and Throughput Model of Three Flight Deck Interfaces

N.C.M. van Zon
Control & Simulation, Department Control & Operations
Delft University of Technology, Delft, The Netherlands
Supervisors: C. Borst, D.M. Pool, M.M. van Paassen and M. Mulder

***Abstract*—Industry is proposing touch-based solutions to modernize the flight management system, a necessary step to cope with increasing demands on lateral navigation as airspace congestion grows. However, research evaluating touchscreen effectiveness for navigation tasks on the flight deck is lacking. An experiment was conducted with 14 participants in the SIMONA Research Simulator, aimed at creating three individual Fitts' law accuracy and throughput models of flight deck interfaces: the mode control panel, control display unit and touch-based navigation display. The former two constitute the conventional interface between the pilot and the flight management system, whilst the latter represents the industry-proposed solution to growing lateral navigation demands. Results of the experiment show the touchscreen navigation display to be the least accurate of the three interfaces, whilst the mode control panel results in the lowest throughput measure. The control display unit scores best in both accuracy and throughput, which is found to largely be attributed to the tactile and physical nature of the interface. Overall, this paper found Fitts' law to apply very successfully to flight deck interfaces, thereby contributing to this field of research.**

I. Introduction

The modern-day flight management system (FMS) was introduced on the Boeing 767 in 1982 [1] to assist pilots in both lateral and vertical navigation (LNAV, VNAV). As an interface to the FMS, the control display unit (CDU) was introduced and remains the industry standard to date. For example, whilst Boeing replaced the CDU on the Boeing 787 with a digital copy, the look and feel remains the same. According to two KLM captains, interviewed during this research, the CDU remains effective in handling present day LNAV tasks. Once the necessary commands are learned, executing various procedures can be done quickly and error-free.

However, looking ahead at future developments in LNAV procedures, the necessity to modernize the FMS interface becomes evident. The SESAR Joint Undertaking expects the number of flights in European airspace to have increased by 52% in 2035 compared to 2012 [2]. As a response, 4D trajectory based operations (TBO) is being developed amongst other mechanisms to cope with increased congestion. The basic principle of 4D TBO is to impose time constraints at which aircraft are required to arrive at specific waypoints in their flight plan [3], [4]. As a result, Huisman et al. [5] expects an increased frequency of en-route route adjustments. Hence, Van Marwijk et al. [6] call for *"a redesign of the navigation planning interface [due to] increasing punctuality in, [amongst others,] European SESAR concepts, [which will] make airborne flight plan amendment increasingly complex"*.

Touchscreens have the possibility to reduce workload whilst increasing situation awareness due to their highly intuitive and direct interface [7]–[11]. Furthermore, by directly adjusting the flight plan route using a touch-based navigation display, the LNAV mode of the FMS can remain engaged and work to ensure compliance with 4D TBO. Therefore, aircraft and equipment manufacturers have been proposing touchscreens on their newest flight decks to cope with increased complexity in navigation tasks. For example, Boeing will introduce touchscreen displays on the 777X [7], Gulfstream ships their G650 with touch-based interfaces [12] and the Garmin G5000 is now also touch-enabled [13]. However, concerns have been voiced about the loss of tactile feedback, usability in dynamic environments and physical fatigue of operation [8], [14]–[16].

Research has been done evaluating touchscreen interfaces in general and comparing them to less direct interfaces such as trackballs, trackpads and rotary controllers [14], [17]–[20]. Furthermore, Dodd et al. [16] found increased task execution time, error rates and subjective workload for touchscreen usage in turbulence and at specific cockpit positions. Furthermore, Fitts' law, first published in 1954, has been used by human-machine interface (HMI) researchers as a predictive model of movement time (MT), and a means to compare interfaces using the *throughput* measure [21]–[23]. The latter describes how many *bits* of task difficulty, defined by an index of difficulty (ID), an interface can handle per second. However, Fitts' law has not yet been applied to interfaces on the flight deck.

The goal of this paper is to use Fitts' law to develop an accuracy and throughput model of three flight deck interfaces used during LNAV. These are the mode control panel (MCP), CDU and a novel touch-based navigation display (TND), illustrated in Figure 1. The former two constitute the conventional interface between the pilot and the FMS, whilst the latter represents the industry-proposed solution to growing lateral demands. This paper is structured as follows. First, necessary background theory is discussed in Section II. Next, Section III introduces the overarching design of the three sub-experiments proposed in this paper. These are subsequently presented and discussed in Section IV, Section V and Section VI. Thereafter, Section VII attempts to provide a comparison based on throughput and accuracy of the three flight interfaces. The paper is closed with a conclusion in Section VIII.

II. Background

Fitts' law, first published in 1954, has been used by human-machine interface (HMI) researchers as a predictive model of movement time (MT), and a means to compare interfaces using the *throughput* measure [21]. The latter describes how many *bits* of task difficulty, defined by an index of difficulty (ID), an interface can handle per second. The law can be used to describe the speed-accuracy tradeoff inherent for humans engaged in rapid aimed movement tasks [22], [23].

The complete Fitts' law model to describe MT as a function of ID in a high-accuracy pointing task is presented in Eq. (1). Here, a and b are empirical linear regression constants, A is the amplitude (distance to be traversed) and W_e is the effective width of the target. The latter is empirically calculated using the standard deviation of measured endpoint coordinates.

$$MT = a + b(ID) = a + b\log_2\left(\frac{A}{W_e} + 1\right) \text{ [seconds]} \quad (1)$$

Literature finds that using W_e reflects what participants actually did, rather than what they were asked to do [24]. The spread in movement endpoints will not always align with the specified target width, resulting in inconsistent error rates across various conditions and making comparisons difficult. Participants are found to under- or overperform due to their tendency to *"to cheat on easier ID conditions"* [21] by not moving fast enough or by not covering the entire distance. Given that Fitts' law is designed for rapid aimed movements, *"participants who take their time compromise Fitts' law"* [21], hence, the W_e correction can be useful. The full background and derivation of Fitts' law has been reviewed, see [22], [23].

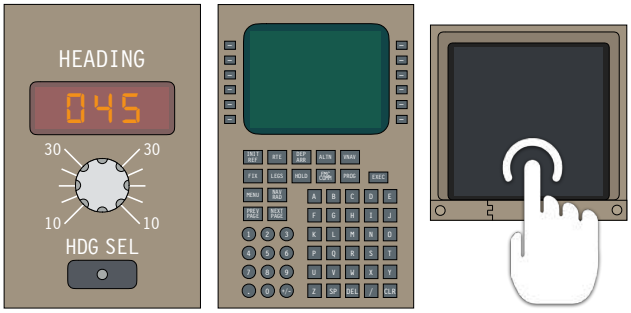
The usefulness of Fitts' law is twofold. First, it can help build models of task execution time using a particular interface. Second, it can provide a quantitative description of the FMS interface by comparing the throughput of individual interfaces. Eq. (2) defines the throughput in bits per second, which is calculated by dividing the ID by the measured MT for each participant and experimental condition. The total number of conditions and participants is defined by x and y respectively. $ID_{e_{ij}}$ defines the index of difficulty, adjusted using the effective width W_e, and MT_{ij} the movement time, both for a specific experimental condition and participant.

$$TP \text{ (Throughput)} = \frac{1}{y}\sum_{i=1}^{y}\left(\frac{1}{x}\sum_{j=1}^{x}\frac{ID_{e_{ij}}}{MT_{ij}}\right) \text{ [bits/s]} \quad (2)$$

The basic law applies to one-dimensional situations, however multiple variations to Fitts' law have been consequently researched, published and verified (see [21], [25]–[28]). Given that three different interfaces are to be researched, the most appropriate variations of the law were found in literature and are discussed below, per interface.

A. Mode Control Panel (MCP)

Research by Stoelen et al. [27] has shown that Fitts' law can be extended to rotational tasks by replacing the linear width and amplitude with an angular width ω and amplitude α,



(a) Heading control knob on the MCP (b) CDU (c) TND

Fig. 1. Three flight deck interfaces that are to be investigated.

respectively. The effective angular width ω_e can be calculated based on the standard deviation in endpoints σ_ϕ. The research in [27] found a good model fit for a smooth, continuous rotational task. The heading control knob, a conventional rotary controller, results in distrete motion inputs of which no literature has been found concerning the effectivess of Fitts' law. This model, shown in Eq. (3), was used to model the heading control knob on the MCP as an interface for the pilot to command a specified heading to the autoflight systems. Due the novelty of this approach, a reflection on its effectivess will be provided in Section IV-B.

$$MT = a + b\log_2\left(\frac{\alpha}{\omega_e} + 1\right) \text{ where } \omega_e = \sqrt{2\pi e}\sigma_\phi \quad (3)$$

B. Control Display Unit (CDU)

Research by Soukoreff and MacKenzie [26] has shown that Fitts' law can be extended to keyboard data-entry tasks, using previous research by MacKenzie [25] on 2D applications of the law. The model, shown in Eq. (4), is based on an assumption that using either the minimum height or width of the target in the computation of the ID is sufficient. MacKenzie et al. has found this to provide adequate results [25]. In the case of a key-repeat task, the amplitude is zero and thus the ID, namely $\log_2(0+1)$, will equal zero. Therefore, [26] propose an averaged repeat movement time parameter MT_{repeat} for such tasks. Furthermore, due to the physical inability to measure movement endpoints on the keys, the computation of effective width is troublesome. An alternative approach is proposed, based on an error rate, of which its effectiveness will be reflected upon in Section V-B.

$$MT_{ij} = \begin{cases} a + b\log_2\left(\frac{A_{ij}+\min(H_j,W_j)}{\min(H_j,W_j)}\right) & \text{if } i \neq j \\ MT_{\text{repeat}} & \text{if } i = j \end{cases} \quad (4)$$

C. Touch-based Navigation Display (TND)

Research by Bi et al. [28] has extended the original law to produce the Finger Fitts Law, shown in Eq. (5). Their research found it to be very effective in modelling finger input using touchscreens. Two new parameters are introduced: σ, the variation in movement endpoints and σ_a, the variation in input

device precision (e.g., the finger). The former is calculated using the distribution in endpoint coordinates during the task, where a bivariate standard deviation σ_{xy} is used for 2D tasks. The latter can be measured using a finger calibration task, where users are asked to repeatly touch an identical (in size, not location) target; exact touch locations are used to calculate the bivariate standard deviation σ_a.

$$MT = a + b\log_2\left(\frac{A}{W_e} + 1\right) \text{ where } W_e = \sqrt{2\pi e\left(\sigma^2 - \sigma_a^2\right)} \tag{5}$$

III. Experiment Design

The objective of the experiment was to develop an individual Fitts' law model for each of the interfaces using the respective models described above. Therefore, the experiment consisted of three, separate but similar experiments. The overarching design of the experiment is discussed here; subsequent sections will discuss the individual experiments. Each experiment looked at the effect of ID on observed MT, for participants engaged in an aimed movement task using the respective interface. A reflection will be given on the effectivity of each Fitts' law in fitting the data and describing the flight deck interface. Furthermore, a discussion will be held concerning the observed accuracy, error-rates and the y-intercept parameter a for each individual interface. Finally, a comparison will be attempted based on throughput measurements.

A. Participants

Given the goal of the experiment was to describe human performance in performing a precision pointing task for a specific interface using Fitts' law, prior experience with piloting aircraft was of limited relevance. The effect of previous encounters with either the MCP or CDU (for example by commercial pilots) was dealt with during the training phase, where each participant got sufficiently accustomed to the input device. Nonetheless, right-handed participants were preferred given the positioning in the left seat and thus interface operation with the right hand. A total of 14 people participated in the experiment, of which a brief profile is given in Table I.

TABLE I
Profile of participants

Profile	13 students, 1 professor
Gender	11 male, 3 female
Age	Ranging 21 to 49, averaging 24 years
Handedness	13 right-handed, 1 left-handed

In general the design was a within-subjects and balanced (by means of a latin square) experiment. The amount of repeated measures per participant and condition ranged between ten and twelve, similar to that found in literature [27]–[29].

For each individual experiment, a balanced latin square design was used, which has the ability to greatly reduce the order effect and make the participant experience a random appearance of conditions [30].

B. Apparatus

The experiment was conducted in the SIMONA Research Simulator (SRS) at the Delft University of Technical, shown in Figure 2. Motion and outside visual capabilities were not utilised, however, the interior cabin provided a realistic look and feel to the interaction between participants and the three flight deck interfaces. The DUECA realtime simulation environment was used for the technical implementation of the experiment given that an accurate measurement of time was necessary.

(a) The SIMONA Research Simulator (SRS) at the Delft University of Technology in Delft, The Netherlands.

(b) Cabin of the SRS showing each of the three flight deck interfaces.

Fig. 2. Apparatus used during the experiment.

C. Control Variables

In an attempt to avoid confounds, various experimental factors were kept constant across each sub-experiment, of which a summary is given below.

- Handedness of participants[1].
- Location and size of the touchscreen.
- Training and briefing of participants.
- Lack of motion and outside visuals during experiment.
- Participant location relative to interface (seat position).

[1] Accidentally, one left-handed participant was included

Due to space confinements in the SIMONA Research Simulator the touchscreen was located below the CDU (see Figure 2). As a result, to allow proper comparison of experiments, the participants were seated differently for the MCP/CDU and TND sub-experiments. Hence, the participant's relative location to the touchscreen was comparable to that of the CDU and MCP. Markers were installed on the cabin floor to ensure constant seat positioning (see Figure 3).

Fig. 3. Markers (yellow) on the ground indicating the necessary seat position for each experiment.

D. Procedure

Participants received a briefing document a few days prior to the experiment. An introduction was given concerning the relevance of the experiment, the task to be conducted and the expected time schedule. Prior to each experiment, following a standardized procedure, a verbal briefing was given. Most importantly, and *"essential for any Fitts' law experiment"* [21], the participant was requested to attempt a specific emphasis on *speed and accuracy* in order to achieve an approximate target hit-rate of 96% whilst providing a smooth consistent input motion. Training runs preceded data measurement to provide the participant with valuable time to master the speed-accuracy tradeoff. Furthermore, in an attempt to reduce the order effect of the interfaces to which the participants were exposed, the sequence of experiments was shuffled.

E. Hypotheses

Based on the aforementioned experiment design and the objective of this paper the following could be hypothesized. Each hypothesis is presented in *italics*, followed by a motivation.

1) Movement time will be linearly dependent on the index of difficulty.: This is the basic principle of Fitts' law, and based on years of research [21], [22] using the model it is expected to also apply, with a large coefficient of determination R^2, to the three respective interfaces.

2) The error rate will be largest for the TND, followed by the MCP and CDU.: Due to the inherent inaccuracies of the finger the error rate will be large for the TND, which will result in a larger W_e and hence smaller ID_e. Due to the discrete nature of input using the MCP and CDU their respective error rates will be lower.

3) The y-intercept (a) will be largest for the CDU, followed by the MCP and TND.: Literature has attributed a positive y-intercept, representing the required movement time for a task of zero difficulty, to additive factors such as the physical effort required to push a button [22], [23]. Hence, such additive factors are expected to be most prominent for the CDU.

4) The throughput of the MCP and CDU is expected to be larger than the TND.: Given the discrete nature of input using the MCP and CDU, an increase in index of difficulty for a specific task is expected to have a smaller effect on the resulting movement time when compared to the TND, where especially a smaller target width will increase movement time significantly. As a result, the throughput, measured in bits per second, should be larger for the MCP and CDU.

IV. Sub-Experiment I: MCP

A. Design and Procedure

1) Independent Variables: As is true for any Fitts' law experiment the main goal is to determine a model that relates interface performance to task difficulty. For Fitts' law the ID takes care of this matter. Hence, the angular amplitude α and angular width ω are the two independent variables; together they determine the ID. Finally, although literature has found direction to be a compounding factor [21], a directional variable could not be included due to limitations of the hardware used. The rotary controller shows dynamic behaviour during clockwise rotations, whilst providing linear behaviour during counter-clockwise rotations[2]. Due to the strong preference for linear rotational behaviour all movements were counter-clockwise.

$$\alpha = [10, 20, 30, 40, 50, 60] \text{ [deg]}$$
$$\omega = [2, 4, 6, 8] \text{ [deg]}$$

The choice of α and ω were such that they form a representative range of ID (see Equation 6) for the interface in question, therefore in compliance with the suggestion by [21]. Heading changes between ten and sixty degrees left or right of the current course are considered realistic, as well as a required accuracy ranging between two and eight degrees when circumnavigating complex weather systems.

$$ID_{\text{MCP}} = \log_2\left(\frac{\alpha}{\omega} + 1\right) \rightarrow 1.17 \leq ID_{\text{MCP}} \leq 4.95 \qquad (6)$$

In total the set of independent variables results in twenty four (6×4) combinations. Given that the variable α and ω are multiples of two, and the mathematical nature of the index of difficulty, a total of sixteen unique ID values result. For example, the combinations $\alpha = 10, \omega = 2$ and $\alpha = 20, \omega = 4$ produce the same ID.

[2]The exact cause for this effect is not known, however most likely attributed to a bug in the hardware

2) Dependent Variables: Fitts' law is designed to *"describe the relationship between movetime time, distance and accuracy..."* [21] of which the first and last factors are dependent measures. The movement time MT, characterized as ratio data, was measured in milliseconds. In accordance with recommendations in literature [21] only the actual time the participant moves the heading knob was measured, thereby omitting homing, dwell and reaction times. Hereby confounding factors such as cognitive effort required to understand the task and the location of the interface relative to the participant was controlled. Accuracy was the other dependent variable, characterized as interval data, and is measured by recording physical endpoints of each individual movement.

3) Task and Procedures: The experimental set-up is presented in Figure 4. On the inboard screen the navigation display shown in (see ❶) was presented. A magenta heading bug (see ❷) reflects the heading commanded on the mode control panel. At the start of each experimental condition, the bug will be reset to the north-up position. The target was shown using two cyan lines (see ❸), of which the angular distance between them represents the width ω. The angular distance between the starting position of the heading bug and the center of the target is the amplitude α. The course select knob (illustrated by ❹) on the mode control panel is a standard rotary encoder with 24 *'clicks'* per full rotation. A small LCD display above the knob reflects the commanded heading. Due to the heavy usage of the heading control knob the sensor was found degraded resulting in an inferior signal, leading to the decision to use the similar course control knob for this experiment.

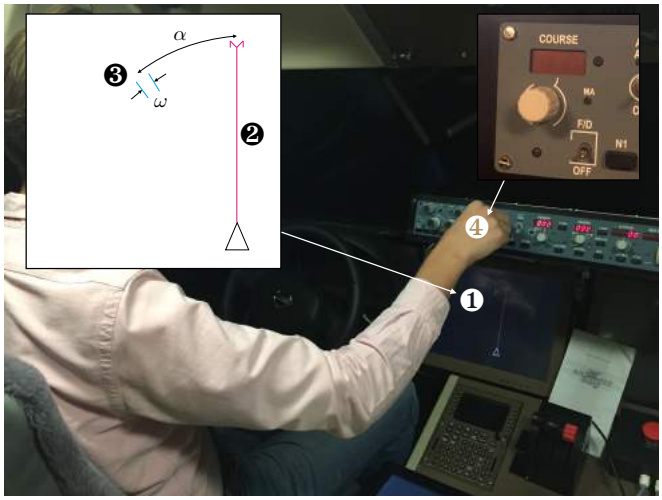


Fig. 4. Experiment procedure for MCP experiment, showing an illustration of the navigation display and heading control knob.

The task set-up in both the training and measurement phases were equal. Once the participant was ready a set of thirty-two conditions were loaded, the heading bug was reset to north-up and a new target was displayed. Measurement started when the participant began to use the input device. During the movement of the heading bug a motion derivative $d\alpha/dt$ was computed and constantly updated. The measurement stopped when no motion (ie. $d\alpha/dt = 0$) was observed for one consecutive second. Subsequently the trial conditions (α, ω), movement time and movement endpoint variables were saved and logged to a data file. Finally, the next trial was automatically loaded. Once the participant once again began to use the input device the same procedure repeated until the full set of conditions was exhausted. During the experiment the success rate in acquiring the target was displayed in the control room and communicated to the participant to provide valuable feedback on their adherance to the speed-accuracy tradeoff governing Fitts' law.

B. Results and Discussion

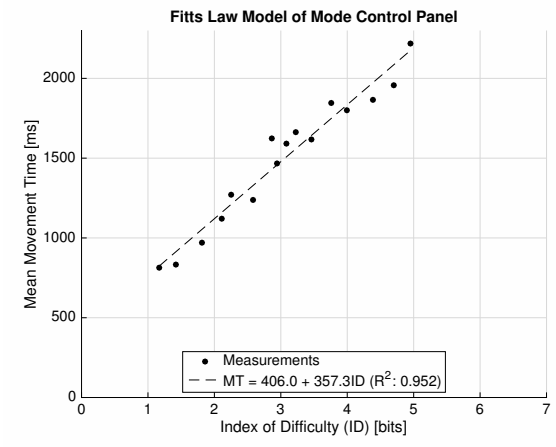
1) Data Analysis: As discussed in Section II the standard deviation in heading endpoints (σ_ϕ) was used to compute the effective width ω_e. Based on the adjusted ω_e calculated for each condition an effective index of difficulty ID_e could be derived for each condition and participant, as illustrated by Eq. (7) [22].

$$ID_e = \log_2\left(\frac{\alpha}{\omega_e} + 1\right) \text{ where } \omega_e = \sqrt{2\pi e}\sigma_\phi \qquad (7)$$

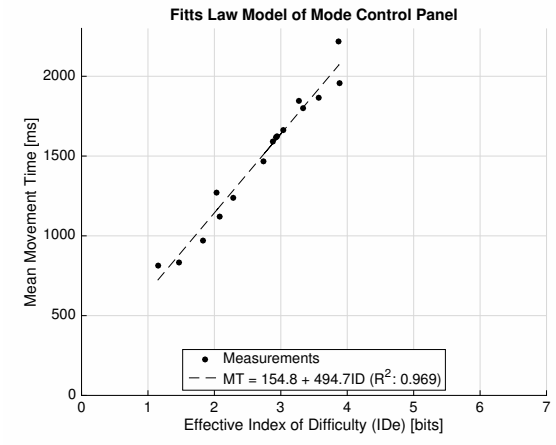
For each unique index of difficulty, of which this experiment had sixteen, a mean movement time of all respective measurements is found. These measurements are shown in Figure 5a, after which a least-squares linear regression is found and superimposed on the scatter plot. Subsequently, the proposed adjustment for accuracy is done by computing the effective width W_e based on actual distribution of movement endpoints per ID. Based on the effective width W_e an effective index of difficulty ID_e is calculated and plotted in Figure 5b, after which a new linear regression is found and superimposed on the scatter plot. Out of 2576 measurements two outliers were removed due to a very clear incorrect execution of the task by the participant. Furthermore, the mean movement times per participant is plotted in Figure 5c which shows the variability in participant performance.

An ANOVA was performed to test the significance of the results. A violation of the sphericity assumption was found using Mauchly's test, $\chi^2(90) = 132.6, p < .05$, therefore Greenhouse-Geisser sphericity estimates were used to correct the degrees of freedom ($\epsilon = 0.325$). Subsequently it could be concluded that there was a significant effect of index of difficulty (ID) on mean movement times (MT), given that $F(4.33, 56.28) = 138.47, p < .01$. Nonetheless, post-hoc Bonferonni pairwise comparisons reveal certain pairs of ID conditions with no significant effect, mainly for $ID < 1.75, ID > 3.25$. This finding can be confirmed by visual inspection of Figure 5c where the median MT per ID seems to level off at respectively small and large indices of difficulty.

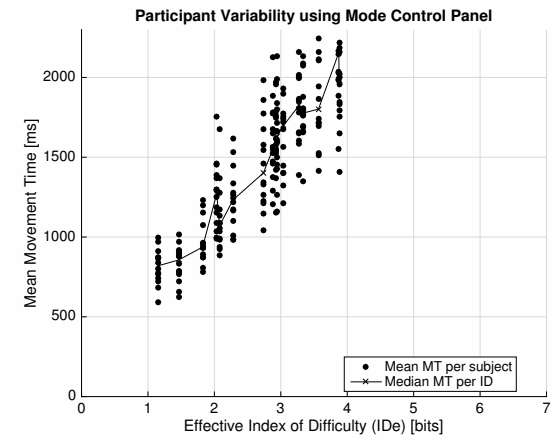
2) Results: The Fitts' law model for the mode control panel was found to be described by MT $= 154.8 + 494.7$ID with a coefficient of determination $R^2 = 0.969$. The mean achieved accuracy score across all participants was found to be 96%. As can be seen by visual inspection of Figure 5b



(a) Mean movement times against index of difficulty



(b) Mean movement times against accuracy-adjusted effective index of difficulty



(c) Mean movement times against accuracy-adjusted effective index of difficulty per participant

Fig. 5. Results for the Mode Control Panel (MCP) experiment.

with Figure 5a the accuracy-adjustment shifts the highest ID (> 3.5 bits) towards the left. The dependent measure throughput was derived, based on the accuracy-adjusted data, to be 1.80 bits/s. The y-intercept a, or expected movement time for a task of zero difficulty, was found to be 154.8 ms.

3) Discussion: The very good fit of the Fitts' law model was a pleasant surprise, given that literature lacks a study looking at the applicability of Fitts' law to a rotary controller providing discrete input signals. These results indicate that the law can also be used for such an input device. The shift in ID following the accuracy adjustment towards the left suggest an underperformance for difficult tasks. Indeed, the majority of errors were observed for challenging targets with an angular width of just two degrees. Hence, the endpoint distribution was spread out beyond the target width, resulting in a larger W_e and thus smaller ID_e. In fact, the conditions with a target width of two degrees resulted in an effective width varying between 3.6 and 4.9 degrees. Finally, the positive y-intercept can be attributed to additive factors unrelated to the ID [22], [23]. These include the necessity to provide multiple, individual rotations of the control knob to move the heading bug more than 15 to 20 degrees. This was the case for two-thirds of the conditions.

V. Sub-Experiment II: CDU

A. Design and Procedure

1) Independent Variables: This experiment in essence featured a 2D task in which a pointing device (in this case a finger) moves across the CDU interface selecting key after key (the targets). Therefore the amplitude A was characterized as the shortest distance between each key, and the width W was characterized as the minimum of either the height or width of the key. MacKenzie et. al. [25] found that the minimum height or width approach is sufficiently accurate for use in 2D Fitts' law tasks. In this experiment the independent variables A and W were defined by a set of words that needed to be entered and subsequently moved to target line select keys (LSK). One five letter word that needs to be positioned at a specified line select key constitutes five movements with respective A and W values. Furthermore, in order to complete the Fitts' law model introduced in Eq. (4), one word consisted of five consecutive repeated keys in order to determine MT_{repeat}.

$$\begin{aligned}\text{Words} =&[\text{KLM19}, \text{AET50}, \text{47MAY},\\ &\text{SSSSS}, \text{DJS73}, \text{ANW80}]\ \text{[-]}\\ \text{Target LSK} =&[\text{L1}, \text{L2}, \text{L6}, \text{R2}, \text{R4}, \text{R5}]\ \text{[-]}\end{aligned}$$

The choice of independent variables resulted in a total set of 36 different conditions, each of which consisted of five Fitts' law movements. Therefore a minimum of 180 measurements could be made during one set of initial conditions. The set of words were carefully chosen to encompass a wide range of index of difficulties, as illustrated by Eq. (8). An accurate technical drawing, including all necessary dimensions to calculate A and W, was drawn for the actual CDU used during the experiment, and resulted in Eq. (8).

$$ID_{\text{CDU}} = \log_2\left(\frac{A_{ij}}{\min(H_j, W_j)} + 1\right) \rightarrow 1.26 \leq ID_{\text{CDU}} \leq 5.17 \tag{8}$$

2) Dependent Variables: Similar to the MCP experiment, the movement time MT, excluding homing, dwell and reaction times, was measured in milliseconds. Once the participant started to type the necessary word, every subsequent keystroke (including incorrect ones) was logged and the movement time noted. The accuracy, measured as the amount of correct divided by the total amount of keystrokes, was measured and used to provide as feedback to participants. Endpoint distributions of the input device (ie. finger) on the keys could not physically be measured.

3) Task and Procedures: The experiment set-up is shown in Figure 6. A full-scale hardware CDU was used for this experiment and installed in the forward pedestal. An illustration of the CDU including the display is shown in ❶. Using the alphanumeric keyboard populates the scratchpad (see ❷). The `CLR` key could be used to backspace the scratchpad. The full content of the scratchpad could be moved to any of the 12 line select keys (see ❸) by pushing the respective key. The text subsequently moves and the scratchpad is cleared.

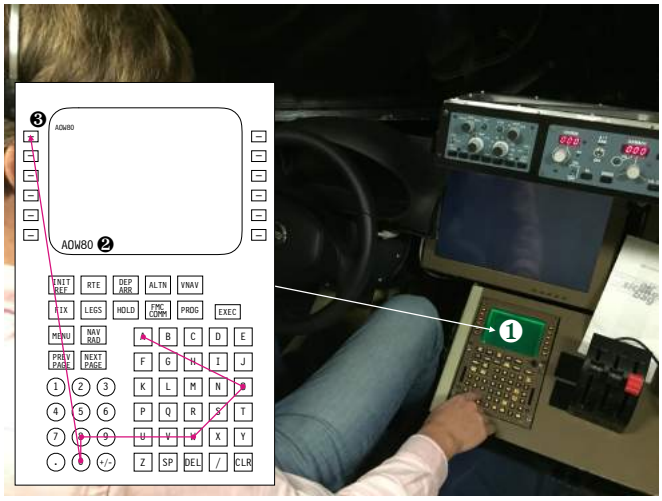


Fig. 6. Experimental procedure for CDU experiment, showing an illustration of the control display unit and location within the flight deck.

The task set-up in both the training and measurement phases was equal. Once the participant was ready a set of 36 conditions were loaded. The target word and final LSK location was presented to the participant by showing the desired word in small at the desired LSK. Figure 6 shows an example where a participant is required to enter the word `AOW80` before moving it to the top-left line select key. Furthermore, the figure illustrates how the participant would move his or her finger across the keyboard in order to accomplish the task.

The participant was advised to search and find the necessary keys prior to initiating data entry in order to keep cognitive effort during key entry at a minimum. Once the participants started to type the word each individual movement was logged. In the example shown, `A` to `O` and `O` to `W` are two consecutive, seperate movements each with their own movement time. If the user mistypped the word the `CLR` key could be used to backspace the scratchpad, all the while each movement (also the incorrect ones) were logged. The trial stopped when the participant had correctly entered the required word and moved it towards the required LSK. Subsequently, one second later, the next condition was automatically loaded. Following each trial the experimenter had the ability to provide feedback on the achieved accuracy rate to the participant, in an attempt to ensure that he or she adhered to the *speed-accuracy* principle underlying Fitts' law.

B. Results and Discussion

1) Data Analysis: Due to the physical inability to measure individual endpoints of the finger on each of the keys, an effective width could not be calculated. Soukoreff and MacKenzie propose an alternative approach which is based on the error rate to approximate the necessary adjustment for accuracy [21], presented in Eq. (9). Here, Err is the error rate of a specific condition, and $z(x)$ represents *"the inverse of the standard normal cumulative distribution, or, the z-score that corresponds to the point where the area under the normal curve is $x\%$[3]"* [21].

$$ID_e = \log_2\left(\frac{A_{ij}}{W_{ij_e}} + 1\right) \text{ where}$$

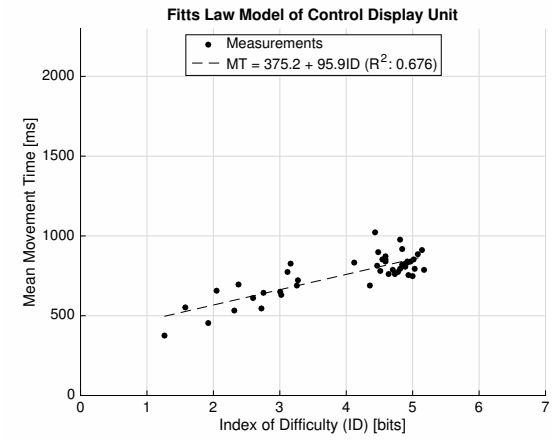
$$W_{ij_e} = \begin{cases} \min(H_j, W_j) \times \frac{2.006}{z(1-Err/2)} & \text{if } Err > 0.0049\% \\ \min(H_j, W_j) \times 0.5089 & \text{otherwise.} \end{cases} \tag{9}$$

For each unique index of difficulty a mean movement time of all respective measurements was found. These measurements are shown in Figure 7a, after which a least-squares linear regression was found and superimposed on the scatter plot. Subsequently, the proposed adjustment for accuracy was done by computing the effective width W_e based on error percentages per index of difficulty. Based on the effective width W_e an effective index of difficulty ID_e was calculated and plotted in Figure 7b, after which a new linear regression was found and superimposed on the scatter plot. Given the exact coordinates of each key, the amplitude and width of every possibile movement across the CDU could be derived.

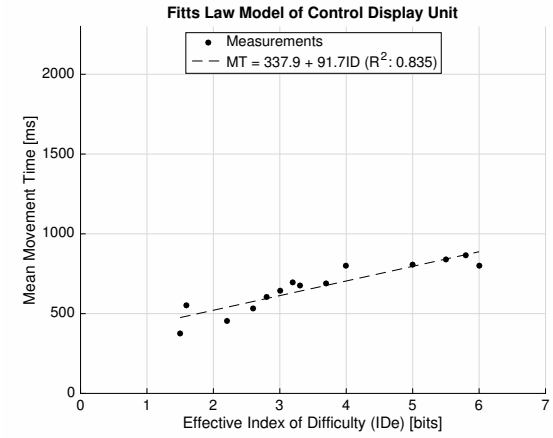
The amount of data points in Figure 7b is clearly less than in the unadjusted plot. During the experiment the participant was at times required to provide an unplanned movement, for example from a specific key to the `CLR` key to correct a mistake. These movements were also measured, however, given that their occurance was not substantial enough to provide reasonable data and were originally not planned to be included they were removed from the dataset. Nonetheless, the data measurements shown in Figure 7b include every movement originally planned to be measured. Out of these 5383 measurements only twenty-two evident outliers were removed due to improperly working and sticky keys.

An ANOVA was performed to test the significance of the results. A violation of the sphericity assumption was found

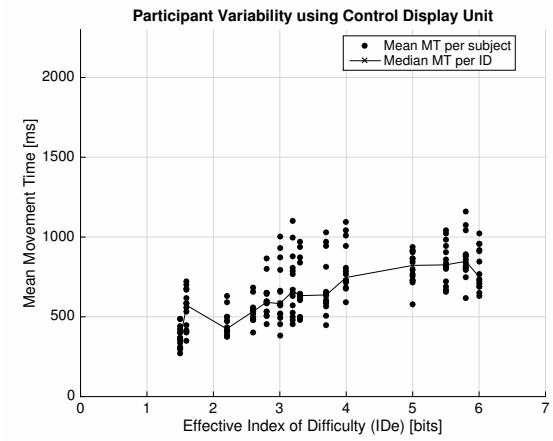
[3]This method of determining W_e can be implemented in Microsoft Excel with the formula $W \times 2.066/\text{NORMSINV}(1-Err/2)$ where Err is stored as a percentage data type (i.e., an error rate of $4\% = 0.04$, not 4.0). [21]



(a) Mean movement times against index of difficulty



(b) Mean movement times against accuracy-adjusted effective index of difficulty



(c) Mean movement times against accuracy-adjusted effective index of difficulty per participant

Fig. 7. Results for the Control Display Unit (CDU) experiment.

using Mauchly's test, $\chi^2(90) = 200.8, p < .01$, therefore Greenhouse-Geisser sphericity estimates were used to correct the degrees of freedom ($\epsilon = 0.289$). Subsequently it could be concluded that there was a significant effect of index of difficulty (ID) on mean movement times (MT), given that $F(3.76, 48.83) = 37.05, p < .01$. Nonetheless, post-hoc Bonferonni pairwise comparisons reveal certain pairs of ID conditions with no significant effect, mainly for $ID \geq 4$. This finding can be confirmed by visual inspection of Figure 7c where the median MT per ID seems to level off at large indices of difficulty.

2) Results: The Fitts' law model for the control display unit is found to be described by MT $= 337.9 + 91.7$ID, or more specifically Eq. (10), with a coefficient of determination $R^2 = 0.835$. The mean achieved accuracy score across all participants was found to be 99%. As can be seen by visual inspection of Figure 7b with Figure 7a the accuracy-adjustment shifts the highest ID (> 3.5 bits) towards the right. The dependent measure throughput is derived, based on the accuracy-adjusted data, to be 5.20 bits/s. The y-intercept a, or expected movement time for a task of zero difficulty, was found to be 337.9 ms.

$$\text{MT}_{ij} = \begin{cases} 337.9 + 91.7\text{ID} & \text{if } i \neq j \\ 267.9 & \text{if } i = j \end{cases} \text{ [ms]} \qquad (10)$$

3) Discussion: Given the significant improvement in R^2 the accuracy adjustment proposed by [23] can be concluded to work effectively. The adjustment shifts larger ID towards the right, suggesting an overperformance of participants. This is reflected by the high mean accuracy score of 99% and additionally observed during the experiment. Due to the tactile and hardware nature of the keys it was easy to achieve high accuracy. However, the nature of the keys also resulted in a large y-intercept. A significant amount of force was required to successfully depress the keys, participants were even observed to continue towards a next key whilst unsuccessfully hitting the previous. Furthermore, although participants were requested to search the necessary keys before initiating data entry an effect of cognitive effort required to find the required keys is expected to remain an influence. Finally, it is interesting to note the very gradual slope, especially at larger ID, suggesting a less significant effect of ID on MT. This may be attributed to the fact that more difficult movements on the CDU are nearly always a result of a larger amplitude. Only three different key widths were available on the CDU, whilst a very narrow target width was found to result in the most challenging task in the MCP sub-experiment.

VI. Sub-Experiment III: TND

A. Design and Procedure

1) Independent Variables: This experiment will feature a 2D task in which an object has to be moved towards a circular target. The distance to be traversed, the amplitude A, and the diameter or width W of the circular target are the main independent variables. A circular target was used because it results in an equidistance target width from any approach angle, simplifying the construction of a Fitts' law model. A representative and wide variety of variables A and

W were selected. The unit was millimeters and based on the size of an actual Boeing 777 navigation display, measuring 180×180 mm. Nonetheless, given A/W was used as a ratio in the computation of the index of difficulty the unit was of little relevance as long as they are equal. Finally, given literature has found direction to be a confouding factor [21], a direction *'heading'* variable ϕ and display rotation variable θ were introduced. These two additional IV's are illustrated in Figure 8, where the former alters the direction of movement and the latter rotates the entire reference frame of the display.

$$\begin{aligned} A &= [45, 80, 115, 150] \text{ [mm]} \\ W &= [5, 15, 25, 35] \text{ [mm]} \\ \phi &= [-25, 0, 25] \text{ [deg]} \\ \theta &= [0, 90, 180, 360] \text{ [deg]} \end{aligned}$$

The choice of independent variables results in a total set of 192 different conditions, however only sixteen (4×4) different combinations of A and W are present due to the use of directional variables. The spread in index of difficulty was found to be representative and presented in Eq. (11).

$$ID_{\text{TND}} = \log_2\left(\frac{A}{W} + 1\right) \rightarrow 1.19 \leq ID_{\text{TND}} \leq 4.95 \quad (11)$$

2) Dependent Variables: Similar to the MCP and CDU experiment, the movement time MT, excluding homing, dwell and reaction times, was measured in milliseconds. Accuracy was the other dependent variable and was measured by recording physical endpoints of each individual movement. Given the 2D nature of the task a bivariate endpoint standard deviation (σ_{xy}) was used, which has been found by literature (Wobbrock et al. cited in [28]) to better describe 2D Fitts' law tasks.

B. Task and Procedures

The experimental set-up is shown in Figure 8. A large touchscreen was installed horizontally on the center pedestal. An illustration of the display presented on the screen is shown in ❶. A white object was shown with a magenta crosshair at its center (see ❷), and could be moved around using touch-based input. The object was sized approximately 10×10 mm to be representative of a waypoint icon on a Boeing navigation display. The target was depicted using a cyan circle (see ❸) with a black crosshair. The various independent variables were used to derive the necessary x and y coordinates of both the initial object and target locations. A reference frame was used starting with $(0, 0)$ in the bottom-left and ending at $(100, 100)$ at the top-right of the display.

For the finger calibration experiment a fixed diameter magenta target, slightly larger than a typical index finger, with a white crosshair was drawn at a random (x, y) location on the display.

The task set-up in both the training and measurement phases were equal. Once the participant was ready, a set of 192 conditions were loaded, and both the object and target were reset to their respective positions. Measurement started when the participant had succesfully acquired the object and started

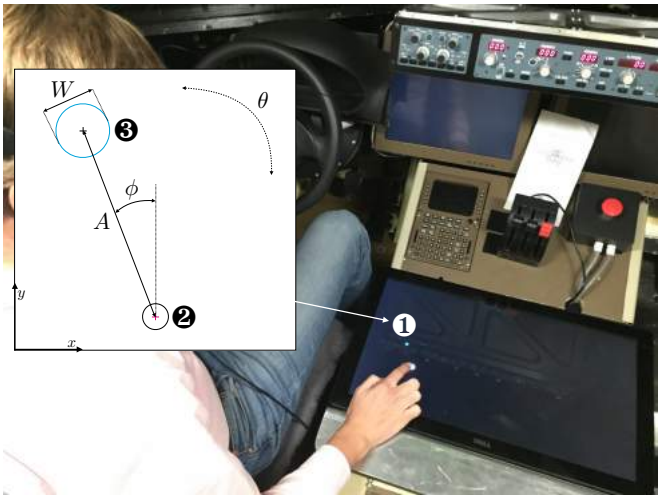


Fig. 8. Experimental apparatus for TND experiment, showing an illustration of the touchscreen display and location within the flight deck.

to move it. Object acquisition was done by providing a touch input within a touch area equal in size and location of the object. During the movement a bi-directional motion derivative dA/dt was computed and constantly updated. dA was based on both directional movements dx and dy using Pythagorean theorem, as illustrated by Eq. (12).

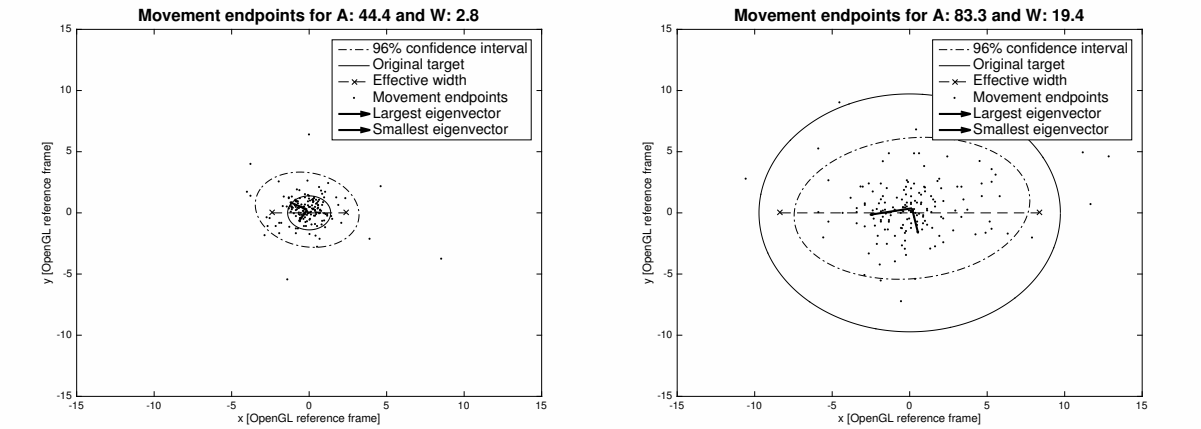
$$\frac{dA}{dt} = \frac{\sqrt{(dx)^2 + (dy)^2}}{dt} \quad (12)$$

The measurement stopped when no motion (ie. $dA/dt = 0$) was observed for one consecutive second. Subsequently the trial conditions (A and W), movement time and movement endpoints (final object and target locations) were saved and logged to a data file. Finally, the next trial was automatically loaded. When the participant once again began to use the input device, the same procedure repeated until the full set of conditions was exhausted. During the experiment the success rate in acquiring the target was displayed in the control room and communicated to the participant to provide valuable feedback on their adherance to the speed-accuracy tradeoff governing Fitts' law.

The finger calibration experiment presented participants with a target. Once a touch input was received, its location was saved, and following a one second pause a new target was shown at a random (x, y) location on the display. Subsequently, the procedure repeats until sufficient data points were conceived. The experimenter had the ability to provide feedback on the achieved hit-rate.

C. Results and Discussion

1) Data Analysis: The effective width was calculated based on the Finger Fitts' law dual distribution procedure proposed by [28]. Two bi-variate standard deviations are derived, the distribution of movement end-points σ_{xy} and the inherent accuracy of the human finger, measured during the finger calibration experiment, σ_a. The bi-variate standard deviation



(a) Endpoint distribution, original and effective width for small target (b) Endpoint distribution, original and effective width for large target

Fig. 9. Graphical illustration of effective width calculation for small and large target.

was found by first deriving the covariance matrix, subsequently finding the square root of its eigenvalues and multiplying it by two [31]. This results in two standard deviations, in the direction of the largest variance and at 90 degree angle to that direction. The minimum of these two was found to best describe the effective width. Based on the adjusted W_e, as introduced by [28] in their Finger Fitts law, calculated for each condition an effective index of difficulty ID_e can be derived for each condition and participant, as illustratied by Equation 13 [28], [31].

$$
\begin{aligned}
C &= \begin{pmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{pmatrix} \rightarrow \lambda_1, \lambda_2 \\
\sigma_{xy}, \sigma_a &= min\left(2\sqrt{\lambda_1}, 2\sqrt{\lambda_2}\right) \\
ID_e &= \log_2\left(\frac{A}{W_e} + 1\right) \text{ where } W_e = \sqrt{2\pi e\left(\sigma_{xy}^2 + \sigma_a^2\right)}
\end{aligned}
\tag{13}
$$

A graphical illustration of this procedure is given in Figure 9. For example, for the narrow target shown in Figure 9a the majority of endpoints are found beyond the nominal task target width. This is illustrated by the 96% confidence interval. As a result, the effective width was found to be substantially larger. On the contrary, for the larger target in Figure 9b every endpoint is found within the original target, thus suggesting an overperformance of the participant and necessary accuracy adjustment by reducing the effective width.

Hence, for each unique index of difficulty a mean movement time of all respective measurements was found. These measurements are shown in Figure 10, after which a least-squares linear regression was found and superimposed on the scatter plot. Subsequently, the proposed adjustment for accuracy was done by computing the effective width W_e as discussed. Based on the effective width W_e an effective index of difficulty ID_e was calculated and plotted in Figure 11a, afterwhich a new linear regression was found and superimposed on the scatter plot.

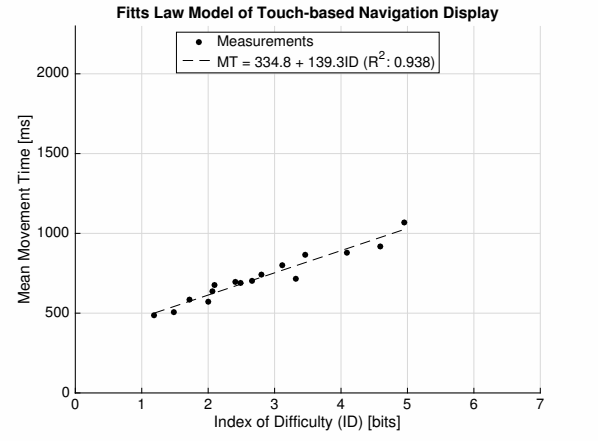
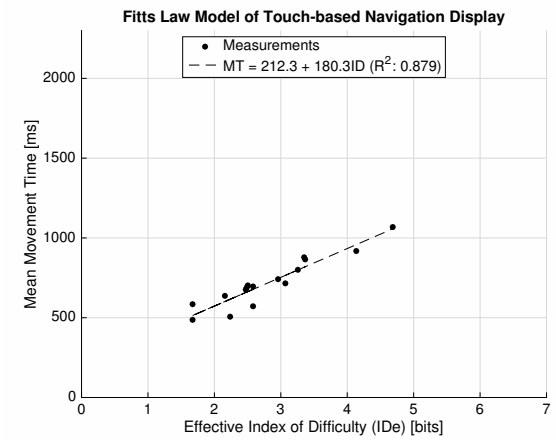


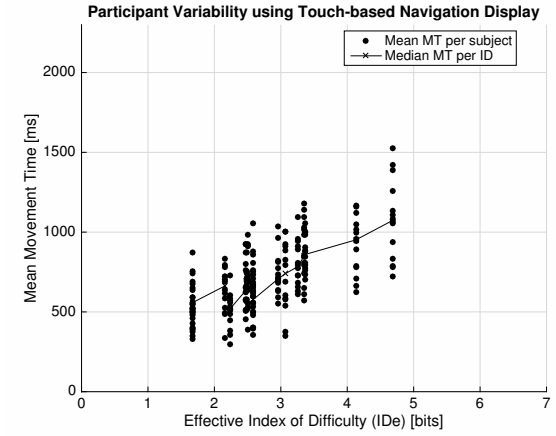
Fig. 10. Results for the Touch-based Navigation Display (TND) experiment.

An ANOVA was performed to test the significance of the results. A violation of the sphericity assumption was found using Mauchly's test, $\chi^2(90) = 156.5, p < .01$, therefore Greenhouse-Geisser sphericity estimates were used to correct the degrees of freedom ($\epsilon = 0.275$). Subsequently it could be concluded that there was a significant effect of index of difficulty (ID) on mean movement times (MT), given that $F(2.67, 34.67) = 37.20, p < .01$. Nonetheless, post-hoc Bonferonni pairwise comparisons reveal certain pairs of ID conditions with no significant effect, mainly for very similar indices of difficulty. This finding can be confirmed by visual inspection of Figure 11b where it can be seen that several sets of ID have a very close proximity to one another.

2) Results: The Fitts' law model for the touch-based navigation display is found to be described by MT $= 212.3 +$ 180.3ID with a coefficient of determination $R^2 = 0.879$. The mean achieved accuracy score across all participants was found to be 95%. As can be seen by visual inspection of 11a



(a) Mean movement times against accuracy-adjusted effective index of difficulty



(b) Mean movement times against accuracy-adjusted effective index of difficulty per participant

Fig. 11. Results for the Touch-based Navigation Display (TND) experiment.

with Figure 10, the accuracy-adjustment shifts the extreme ID (< 2 and > 4 bits) towards the center. The finger calibration parameter σ_a was found to be 3.6 mm. The dependent measure throughput is derived, based on the accuracy-adjusted data, to be 3.88 bits/s. The y-intercept a, or expected movement time for a task of zero difficulty, was found to be 212.3 ms.

3) Discussion: Although [28] found an improvement in model fit using the accuracy-adjusted Finger Fitts' law, a slight worsening of R^2 was observed. However, this can largely be explained by the large amount of ID conditions that are very closely located to one another. Combining these datasets would significantly improve the model fit. A general observation can be made that large targets were overperformed, whilst smaller targets were underperformed. This explains the shift of extreme indices of difficulty towards the center, as can be seen on the respective plots. The relatively large y-intercept may be attributed to additive factors such as the time necessary to place one's finger on the display and remove it, and was found to be similar to the y-intercept found by [28].

VII. General Discussion

The results of all three experiments show that Fitts' law is an adequate tool to develop an accuracy and throughput model for the Mode Control Panel, Control Display Unit and a Touch-based Navigation Display. Hence, Hypothesis 1 is accepted, as illustrated by Figure 12, and R^2 scores of 0.97, 0.88 and 0.85 for the MCP, TND and CDU respectively.

Furthermore, when scrutinizing the y-intercept parameter (a) of each Fitts' law model, Hypothesis 3 that a will be largest for the CDU, followed by the MCP and TND is partially accepted, as illustrated by Figure 12. The CDU indeed results in a largest expected movement time for tasks of zero-difficulty, namely 295 ms. However, the TND (not the MCP) follows with 212 ms.

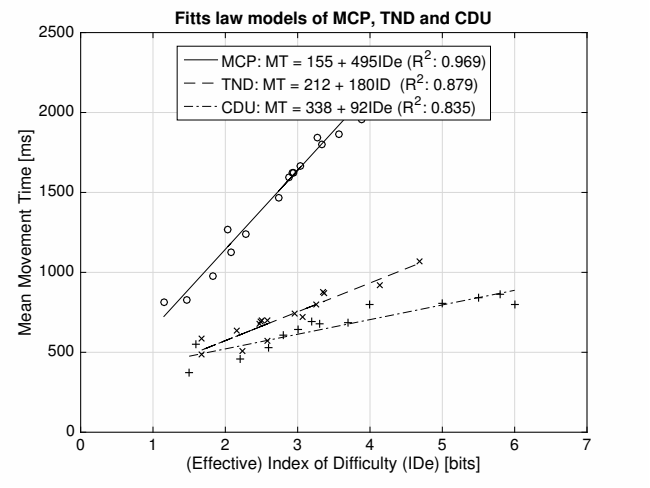


Fig. 12. Final Fitts' law models of each individual interface plotted on the same graph for comparative purposes

A. Interface Accuracy

Following experiment runs, average accuracy scores were measured per participant, per interface and are plotted in Figure 13. The CDU scores highest in terms of accuracy, with a mean at 99%, and a very narrow spread in scores. On the contrary, the touch-based ND scores worse, with a mean of 95%, and much larger spread. The MCP scores similar to the TND, however the spread in scores is much smaller for the heading control knob on the MCP. This finding is very similar to that of Stanton [20] who also compared a rotary controller with a touchscreen and found a similar result. Therefore, this confirms Hypothesis 2 presented at the beginning of this paper.

Interesting is to note that, whilst only one measurement is available, the only left-handed participant scored 70% on the TND whilst 96% and 99% on the MCP and CDU respectively. This suggests that using the traditional interfaces with a non-dominant hand is easier than with a touchscreen. This finding is intriguing, given that the position within the flight deck relative to interfaces is a given and cannot easily be adjusted per pilot, and warrants further research on the effect of handedness on flight deck performance.

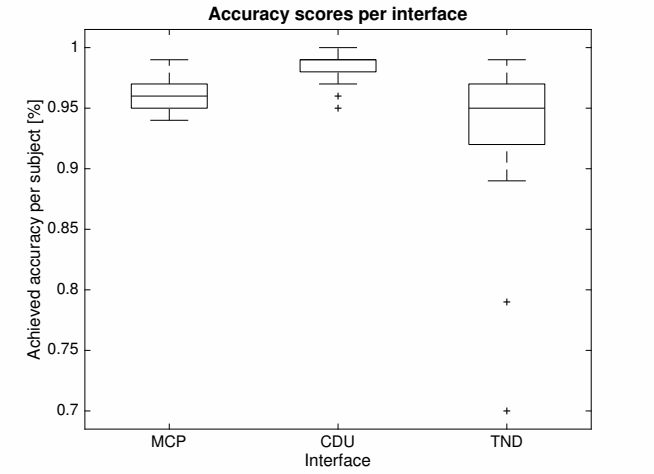


Fig. 13. Observed accuracy scores per participant per experiment.

Following discussions with participants and observations made the accuracy results are most likely attributed to tactile and discrete nature of the traditional interfaces. Due to the lack of tactile feedback and high freedom of movement with the touchscreen a high accuracy was more difficult to achieve.

B. Interface Throughput

The CDU scores highest on throughput, followed by the touch-based ND and MCP respectively. Therefore, Hypothesis 4 is partially accepted. Based on the results it can be said that for a given second of time the CDU can handle more difficult tasks compared to the other two interfaces. This may be explained by the definition of index of difficulty, which is defined by the movement amplitude and target width. On the CDU the target width remained constant, given that the keys had a pre-defined size. Hence, the difficulty in movements was reflected in the distance to be moved. Moving a larger distance was observed to be easier than acquiring a very narrow target. In addition, the physical keys on the CDU make it fairly easy to acquire the target successfully. On the contrary, with the MCP and TND, target difficulty varied both by amplitude and width. For the latter, it was observed on both interfaces, that a very narrow target slowed down participants and required them to provide much higher accuracy. Finally, the MCP scores substantially worse compared to the other two interfaces. This may be attributed to the *"latency and non-linear movement of the heading control knob"* as hypothesized by several participants. Research by Stanton [20] also found a rotary controller to lead to longer task times compared to a touchscreen interface.

C. General Remarks

Whilst the control display unit scores highest on both accuracy and throughput, this does not imply that it is therefore the most optimal interface with the flight management interface. During the experiment participants were asked to locate the necessary keys prior to key entry to keep cognitive effort at a minimum. Hence, the CDU performs well given that the user is fully aware of the necessary task. However, during a complex lateral navigation task a substantial amount of cognitive effort is expected to be needed into to determine the necessary actions using the CDU. A question remains whether, when used during a realistic navigation task, the CDU still scores better than a touch-based interface.

TABLE II
MEASURED THROUGHPUT OF EACH INTERFACE

Interface	Throughput [bits/s]
MCP	1.80
CDU	5.20
TND	3.88

VIII. CONCLUSION

Three separate, but similar, experiments were conducted with the aim of developing an accuracy and throughput model of three flight deck interfaces. These were the mode control panel, control display unit, and a novel touch-based navigation display. The experiments were conducted in the SIMONA Research Simulator with a total of fourteen participants engaged in rapid aimed movements using each of respective interfaces.

Fitts' law was found to provide a very adequate description of each interface, resulting in individual models for accuracy and throughput. The CDU is found to score best in both accuracy and throughput, which based on observations and participant feedback is largely attributed to the tactile and physical nature of the itnerface. This seems to confirm thoughts pilots presented during interviews conducted in preparation of this research. Furthermore, the inherent inaccuracy of the touchscreen as found in literature (see [14], [17]–[20]) is confirmed during this experiment. Finally, the mode control panel whilst accurate contains some discrepancies resulting in a very low throughput.

The results of this paper provides very useful insight in the inherent performance of the flight deck interfaces when used during a simple, aimed movement. However, more research is required to fully understand their effectiveness when used during a complex lateral navigation re-planning task.

REFERENCES

[1] B. Bulfer, *Big Boeing FMC User's Guide*. Leading Edge Publishing, 1991.
[2] SESAR Joint Undertaking, "i4D and SESAR," 2014.
[3] M. Carmona, D. Rudinskas, and C. Barrado, "Design of a flight management system to support four-dimensional trajectories," *Aviation*, vol. 19, no. 1, pp. 58–65, 2015.
[4] M. Mulder, R. Winterberg, M. M. van Paassen, and M. Mulder, "Direct Manipulation Interfaces for In-Flight Four-Dimensional Navigation Planning," *The International Journal of Aviation Psychology*, vol. 20, no. 3, pp. 249–268, 2010. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10508414.2010.487010
[5] H. Huisman, R. Verhoeven, Y. van Houten, and E. Flohr, "Crew interfaces for future ATM," pp. 33–40, 1997.

[6] B. J. A. van Marwijk, C. Borst, M. Mulder, M. Mulder, and M. M. van Paassen, "Supporting 4D Trajectory Revisions on the Flight Deck: Design of a Human-Machine Interface," *The International Journal of Aviation Psychology*, vol. 21, no. 1, pp. 35–61, 2011. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10508414.2011.537559{%}5Cnhttp://www.tandfonline.com/doi/pdf/10.1080/10508414.2011.537559

[7] G. Polek, "Boeing 777X Cockpit to Feature Touchscreen Displays," 2016. [Online]. Available: http://www.ainonline.com/aviation-news/business-aviation/2016-07-07/boeing-777x-cockpit-feature-touchscreen-displays

[8] S. Kaminani, "Human computer interaction issues with touch screen interfaces in the flight deck," *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*, 2011.

[9] B. Shneiderman, "The future of interactive systems and the emergence of direct manipulation," *Behaviour & Information Technology*, vol. 1, no. 3, pp. 237–256, 1982.

[10] W. A. Rogers, A. D. Fisk, A. C. McLaughlin, and R. Pak, "Touch a screen or turn a knob: choosing the best device for the job." *Human factors*, vol. 47, no. 2, pp. 271–88, 2005. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16170938

[11] E. L. Hutchins, J. D. Hollan, and D. A. Norman, "Direct Manipulation Interfaces," in *Human-Computer Interaction*. San Diego: Lawrence Erlbaum Associats, Inc., 1985, vol. 1, pp. 311–338. [Online]. Available: https://www.irjet.net/archives/V2/i6/IRJET-V2I6118.pdf

[12] M. Thurber, "Touchscreens clean up gulfstream symmetry flight deck," 2015. [Online]. Available: http://www.ainonline.com/aviation-news/business-aviation/2015-01-02/touchscreens-clean-gulfstream-symmetry-flight-deck

[13] GARMIN, "Garming G5000." [Online]. Available: https://buy.garmin.com/en-US/US/in-the-air/general-aviation/flight-decks/g5000-/prod90821.html

[14] A. Degani, E. A. Palmer, and K. G. Bauersfeld, ""Soft" Controls for hard displays: still a challenge," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 36, pp. 52–56, 1992. [Online]. Available: http://pro.sagepub.com/content/36/1/52.short

[15] M. Mertens, H. Damveld, and C. Borst, "An Avionics Touch Screen based Control Display Concept."

[16] S. Dodd, J. Lancaster, A. Miranda, S. Grothe, B. DeMers, and B. Rogers, "Touch Screens on the Flight Deck: The Impact of Touch Target Size, Spacing, Touch Technology and Turbulence on Pilot Performance," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, no. 1, pp. 6–10, 2014. [Online]. Available: http://pro.sagepub.com/content/58/1/6.short{%}5Cnhttp://pro.sagepub.com/lookup/doi/10.1177/1541931214581002

[17] J. A. Ballas, C. L. Heitmeyer, and M. A. Pérez, "Evaluating two aspects of direct manipulation in advanced cockpits," *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '92*, pp. 127–134, 1992. [Online]. Available: http://dl.acm.org/citation.cfm?id=142750.142770

[18] F. B. Bjorneseth, M. D. Dunlop, and E. Hornecker, "Strathprints Institutional Repository Assessing the Effectiveness of Direct Gesture Interaction for a Safety Critical Maritime Application," vol. 70, pp. 729–745, 2012.

[19] C. Forlines, D. Wigdor, C. Shen, and R. Balakrishnan, "Direct-touch vs. mouse input for tabletop displays," *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*, p. 647, 2007. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1240624.1240726

[20] N. A. Stanton, C. Harvey, K. L. Plant, and L. Bolton, "To twist, roll, stroke or poke? A study of input devices for menu navigation in the cockpit." *Ergonomics*, vol. 0139, no. February 2013, pp. 37–41, 2013. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23384222

[21] R. W. Soukoreff and I. S. MacKenzie, "Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI," *International Journal of Human Computer Studies*, vol. 61, no. 6, pp. 751–789, 2004.

[22] R. Jagacinski and J. Fisch, "Information Theory and Fitts' Law," in *Control Theory for Humans*, 1997, ch. 3, pp. 17–26.

[23] I. S. MacKenzie, "Fitts' law as a research and design tool in human-computer interaction," *Human-Computer Interaction*, vol. 7, no. 1, p. 48, 1992. [Online]. Available: http://portal.acm.org/citation.cfm?id=1461857

[24] C. L. MacKenzie, R. G. Marteniuk, C. Dugas, D. Liske, and B. Eickmeier, "Three-dimensional movement trajectories in Fitts' task: Implications for control," *The Quarterly Journal of Experimental Psychology Section A*, vol. 39, no. 4, pp. 629–647, 1987. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/14640748708401806

[25] I. S. MacKenzie and W. Buxton, "EXTENDING FITTS' LAW TO TWO-DIMENSIONAL TASKS INTRODUCTION Since the advent of direct manipulation," *ACM CHI'92 Conference*, pp. 219–226, 1992.

[26] W. R. Soukoreff and I. S. MacKenzie, "Theoretical upper and lower bounds on typing speed using a stylus and a soft keyboard," *Behaviour & Information Technology*, vol. 14, no. 6, pp. 370–379, 1995. [Online]. Available: http://dx.doi.org/10.1080/01449299508914656

[27] M. F. Stoelen and D. L. Akin, "Assessment of Fitts' law for quantifying combined rotational and translational movements." *Human factors*, vol. 52, no. 1, pp. 63–77, 2010.

[28] X. Bi, Y. Li, and S. Zhai, "FFitts law: Modeling Finger Touch with Fitts' Law," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, p. 1363, 2013. [Online]. Available: http://dl.acm.org/citation.cfm?id=2470654.2466180

[29] J. Accot and S. Zhai, "Beyond Fitts' law: models for trajectory-based HCI tasks," *Proceedings of the ACM SIGCHI Conference on ...*, no. 1, pp. 295–301, 1997. [Online]. Available: http://dl.acm.org/citation.cfm?id=258760

[30] A. Field and G. Hole, *How to Design and Report Experiments*, 1st ed. London: SAGE Publications Ltd, 2003.

[31] V. Spruyt, "How to draw a covariance error ellipse?" *Computer vision for dummies*, 2014. [Online]. Available: http://www.visiondummy.com/2014/04/draw-error-ellipse-representing-covariance-matrix/

Part III

Literature Review

Note: This part has already been examined under course AE4020 Literature Study

Chapter 2

Contemporary Lateral Navigation

State-of-the-industry

The first topic of this literature review will focus on the state-of-the-art in the field of enroute lateral navigation. The goal is to gain the necessary understanding of the CDU and FMS in use today, combined with operational experiences of its use, complemented with an insight in future developments. As mentioned in the thesis scope (see section 1) this thesis will focus on the Boeing-style interface. More specifically, the Boeing 777 has been selected as a basis for this review, given its modern and representative flight management system and avionics suite.

First, an introduction will be given into the basic operating principles of the flight management system (FMS) and autoflight systems in section 2-1. Second, a detailed review will be presented in section 2-2 of the lateral navigation functionalities available to pilots using the control display unit (CDU). This review will be followed by a summary of operational experiences with the CDU in section 2-3, based on conducted pilot interviews. Finally, this chapter will conclude with a glance into the future of flight management systems, presented in section 2-4, and how operational procedures for lateral navigation will evolve.

2-1 The Flight Management System

During normal operation a large commercial airliner, such as the Boeing 777, will fly its entire mission under command of the automatic flight system. In the first autoflight systems, as illustrated by Figure 2-1, the pilot directly interacts with the system by commanding the necessary airspeed, heading and altitude to achieve a specific task. Airspeed and pitch, bank and yaw control are traditionally decoupled into separate systems; the auto-throttle and autopilot respectively. Both systems are designed to translate the pilot's commands into the necessary control states, such as thrust setting and elevator and aileron deflection, used to directly control the aircraft. Air data sensors are used to close the loop, and provide feedback

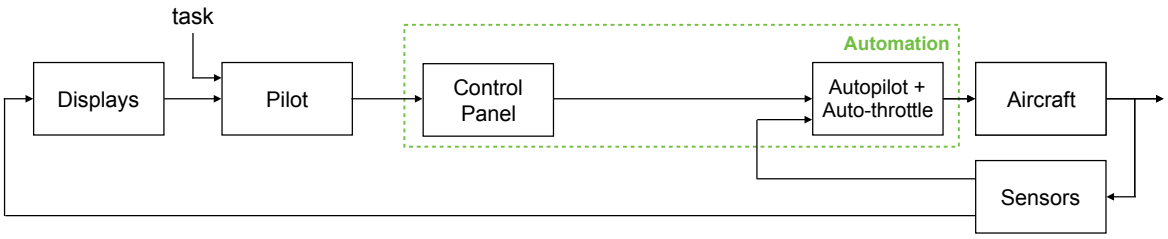


Figure 2-1: Automatic flight prior to the introduction of the flight management system. The pilot commands a heading, altitude and vertical speed and airspeed to the autopilot and auto-throttle respectively, which in turn flies the aircraft [25].

to the autoflight systems as well as the pilot through visual representations on the flight deck displays.

The first version of the modern day flight management system was introduced in 1982 on the Boeing 767 [1], of which the basic principles are still in use today. During the majority of normal operation, the crew will now interface with the autoflight systems though the FMS. Essentially, the FMS is constituted of three parts. First, a database containing aircraft performance and navigational data is stored in the flight data storage unit (FDSU), necessary to compute optimal flight profiles and other parameters. The control display unit (CDU) embodies the primary interface between the pilot and the FMS. The CDU is used by the pilot to enter the desired lateral and vertical flight trajectory, including several performance measures.

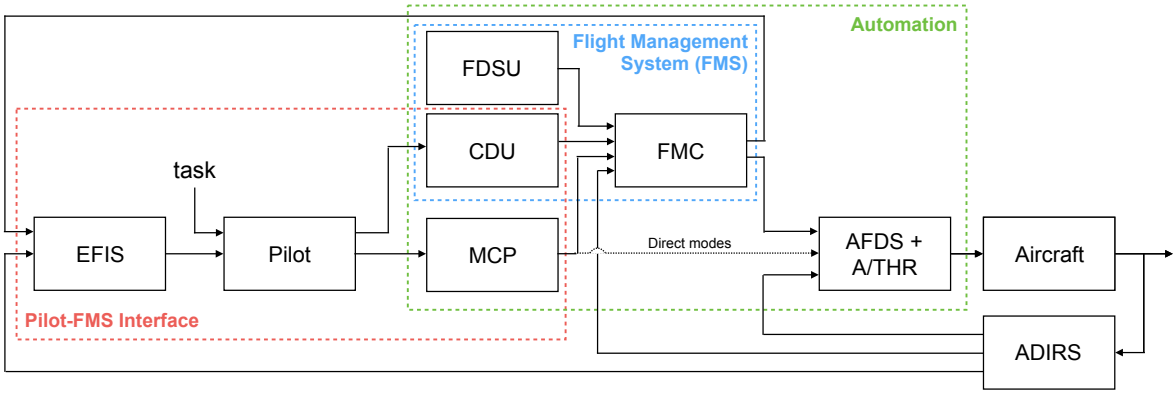


Figure 2-2: Automatic flight on the Boeing 777 [25]. The pilot commands an airspeed and lateral and vertical flight profile to the flight management system (FMS) via the control display unit (CDU) and mode control panel (MCP). The FMS will subsequently compute the necessary heading, altitude, vertical speed and airspeed to command the autopilot flight director system (AFDS) and auto-throttle (A/THR) which flies the aircraft [25]. However, pilots still have the possibility to directly command a specified heading, altitude, vertical speed and airspeed. The air data inertial reference system (ADIRS) keeps track of aircraft position and air data, which is fed back to the automation and the pilot via a visual representation on the electronic flight instrumentation system (EFIS).

Subsequently, the flight management computer (FMC) can compute the necessary airspeed, heading and altitude profiles. Once the crew activates the FMS lateral and vertical naviga-

tional functionalities on the mode control panel (MCP) the FMS will command the autoflight system, which consists of an autopilot flight direct system (AFDS) commanding roll, pitch and yaw and the auto-throttle (A/THR) system commanding thrust. These two systems will in turn control the aircraft. However, crews still have the ability to revert to a direct control of the autoflight system. The same mode control panel which is used to activate the FMS can also be used to directly command a specified airspeed, heading, altitude and vertical speed. Furthermore, a comprehensive air data inertial reference system (ADIRS) is installed on the Boeing 777 to keep precise track of the aircraft's present horizontal and vertical position as well as air data which includes airspeed, heading, vertical speed and altitude [25]. The data is used by the FMS, AFDS and A/THR as feedback to succesfully control the aircraft, as well as presented to the pilot through the electronic flight instrument system (EFIS), consisting of the various flight deck displays.

The FMS has become an effective tool to command full automatic flight of a complete airspeed, lateral and vertical flight profile from takeoff to landing. The combination of the CDU, MCP and EFIS form the interface between the pilot and the FMS, and are further elaborated upon in Figure 2-3, Figure 2-4 and Figure 2-5.

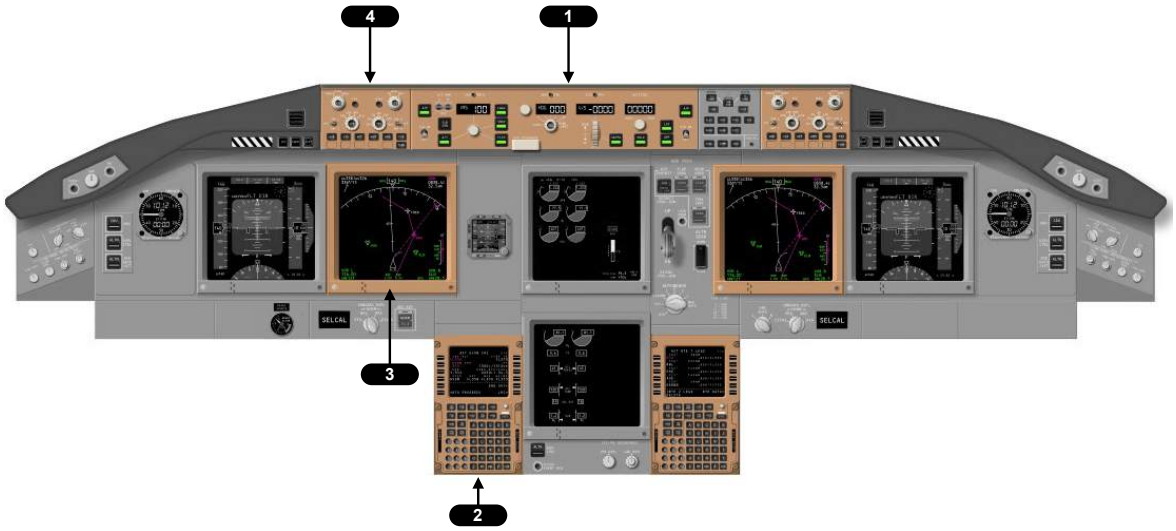


Figure 2-3: As shown in Figure 2-2 the CDU, MCP and EFIS together form the interface between the pilot and the FMS. The ❶ **mode control panel** (MCP) is located on the glareshield, where it can easily be seen and operated. The MCP is further discussed in Figure 2-4. Two ❷ **control display units** (CDU) are installed on the center pedestal, one for each pilot, operate independently and are easily reached. The ❸ **navigation display** (ND) is located infront of each pilot, adjacent to the primary flight display (PFD) and displays the aircraft present position together with the lateral flightroute programmed in the FMS. The display is controlled by the ❹ **EFIS control panel**. Both the CDU and ND are further elaborated upon in Figure 2-5 [25][1].

[1]Background image: flightvectors.com

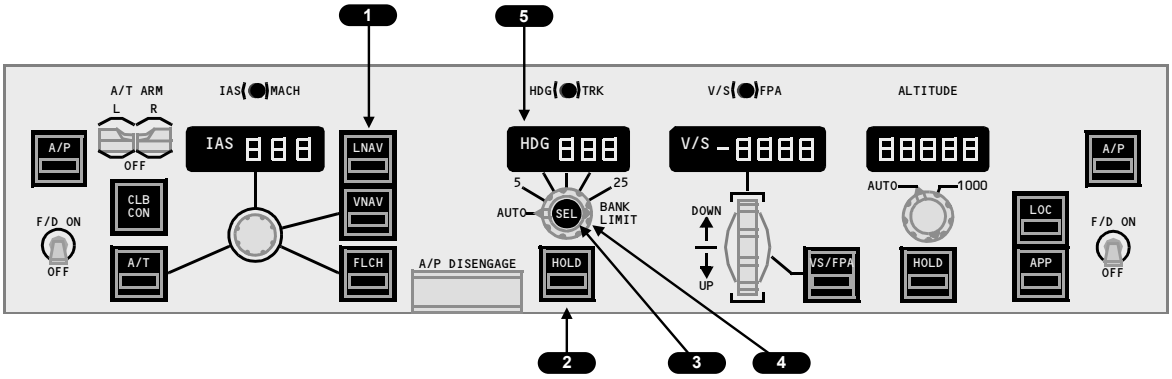


Figure 2-4: *The mode control panel (MCP)* allows the pilot to select between the various autoflight modes. To activate the FMS lateral navigation functionality the ❶ `LNAV` **mode** must be enabled. Subsequently the autopilot will fly the flight plan programmed in the FMS. Direct modes can also be activated, and as a result the pilot can bypass the FMS and command a specified airspeed, heading, altitude or vertical speed. For example, activating the ❷ **heading** `HOLD` **mode** will command the autopilot to level the wings and hold and maintain the current heading. Activating the ❸ **heading** `SEL` **mode** will command the autopilot to fly the heading selected using the ❹ **heading select knob** which is reflected in the ❺ **heading display**. Until the pilot re-activates the `LNAV` mode the autopilot will ignore the flight plan route programmed in the FMS [25].

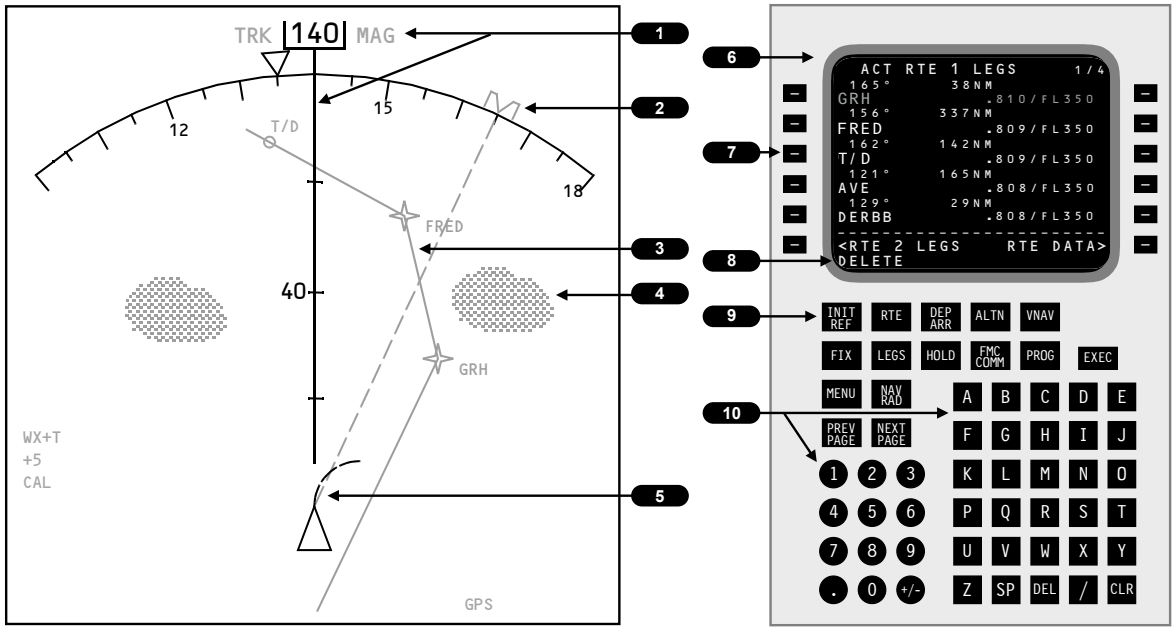


Figure 2-5: *The navigation display (ND)* gives an intuitive, top-down track-up view of the world, present aircraft position, route and surrounding items of interest. Consequently, objects which are to the left of the aircraft will be shown on the left on the display. At the top, the ❶ **present heading** of the aircraft is given. When the heading SEL mode is activated the ❷ **heading bug** is shown which will rotate around the aircraft based on the selected heading on the MCP. Activating the LNAV mode will command the autopilot to follow the ❸ **flight-plan route** which is displayed. The map can be enriched by adding for example an overlay of ❹ **weather**, however one can also depict terrain, surrounding traffic, waypoints, stations, and airports. At the bottom of the navigation display one finds the ❺ **own ship symbol** including a heading trend line. *The control display unit (CDU)* allows the pilot to directly interface with the FMS, and in the Boeing 777 features a color-enabled ❻ **LCD display**. A set of six ❼ **line-select keys** on either side of the display allow the user to select a specific line or enter text from the ❽ **scratchpad** into one of the twelve lines. At the bottom a set of quick-access ❾ **menu keys** can be used to switch between CDU pages and the ❿ **alpha-numeric keyboard** is used for data entry [25].

Flight plan definition

The lateral flight plan is primarily based on the worldwide airway network, which consists of radio beacon stations, GPS waypoints and airways connecting the dots. Prior to flight commencement the crew can enter the lateral flight plan based on the constituent airways and waypoints. An example for a flight from Amsterdam to Hamburg is shown in Figure 2-6. After takeoff a direct route is flown to the ANDIK waypoint and hence programmed as such. Subsequently the route follows the UN873 airway to the GRONY waypoint and so forth.

Complementing existing waypoints, several different pilot-defined waypoints can be programmed into the flight path, lifting the necessity to solely rely on existing airways and waypoints. For example, the custom waypoint ABC01 in Figure 2-7 can be constructed in three ways. First, a position bearing and distance can be specified. ABC01 is located at a bearing of 135° and 13 nautical miles from the radio station ABC, therefore the CDU command ABC135/13 will insert this point in the route. The same point can also be constructed

by specifying the intersection of two airways, for example the intersection between the 135° radial from ABC with the 040° radial from DEF using the CDU command ABC135/DEF040. However, in cases where there are no nearby existing fixes to use as a construction point, the unique GPS coordinate of ABC01 can also be inserted using the CDU. Finally, an along-track waypoint can be programmed which will position itself at a specified location along the flight plan route. The point XYZ01 in Figure 2-7 is located 15 miles prior to the XYZ fix and programmed using the command XYZ/ − 15.

Figure 2-6: Example lateral flight plan definition on the ❶ RTE **page** for a route from Amsterdam to Hamburg, entirely based on airways and waypoints [25].[2]

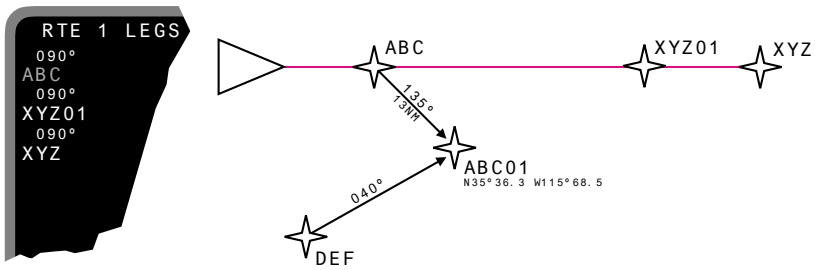


Figure 2-7: An overview of a few of the pilot-defined waypoints that can be programmed in the FMS [1], [25].

2-2 Lateral Navigation Functionalities

The previous section described the basic working principles of the flight management system, and the methodology for programming the lateral flight route using the CDU. This section will focus on standard operating procedures for enroute lateral navigation. Four methods using the CDU and one using the MCP have been identified and will be described. Route adjustments using the CDU are executed on one, or a combination of two pages: the ACT RTE and RTE LEGS page. The former presents the lateral route as a combination of airways and waypoints, and is mainly used prior to takeoff for route definition (see Figure 2-6). The latter

[2]Flight plan map: skyvector.com

presents the route broken down into the individual waypoints that together constitute the chosen airways. Furthermore, the RTE LEGS page also presents information about the vertical flightpath, however that remains beyond the scope of this review.

Airborne Route Adjustment

A direct edit of the lateral flight plan can be done on the CDU RTE LEGS page, which shows an overview of the active and upcoming waypoints enroute, as shown in Figure 2-8a. Here, new waypoints can be inserted at any given location and existing waypoints can be deleted. Note that airways used to construct the flight plan prior to takeoff are no longer visible. The RTE LEGS page will show every single waypoint along the route, which is furthermore presented on the navigation display. Following an adjustment of the flight plan, whether that be inserting or deleting a waypoint the FMS will not automatically re-connect waypoints or assume what will be the next point along the flight path. Instead, it will display a ROUTE DISCONTINUITY on both the CDU and the ND which will have to be solved by the pilot. Although this may seem as a cumbersome process, it ensures that the pilot is fully aware of his or her actions without the automation taking unnoticed steps. To assist the pilot, the navigation display will show a preview of the route modifications as a dashed line, alongside the current active route in magenta. As such, the impact of the route adjustment can be observed very easily.

The FMS onboard the Boeing 777 has the capability to store two fully independent routes. The active route is presented at the top of the CDU page, as shown in Figure 2-8a by RTE 1. This functionality allows the crew to fully adjust the second, inactive route without impacting the active lateral navigation of the FMS and autoflight systems. It can also be used program an expected diversion or re-route, such that when it is necessary it can be activated immediately [25].

Route Offset

Alongside directly adjusting the route waypoints the FMS also features an offset functionality, often used by ATC to ensure traffic separation on congested airways. The ACT RTE page (not to be confused with the RTE LEGS page which shows individual waypoints and is used for enroute modification of the route) is not only used by a pilot to enter the flight plan prior to takeoff, but can also be used whilst airborne to command a route offset. This is done very easily by entering either L or R in the scratchpad, corresponding to an offset to the left or right of the route respectively, followed by the required distance in nautical miles. In Figure 2-9a a route offset is commanded of 6 nautical miles to the right of the course, which is reflected by the dashed magenta line on the ND in Figure 2-9b.

Direct-to

At any point during a flight a pilot can command the FMS to deviate from an airway and navigate directly to a specified waypoint. Offering *"directs"* is common practice for ATC to help shorten flight times, improving efficiency and punctuality, and can vary in range. A direct-to command is inserted on the RTE LEGS page, and executed by selecting the desired waypoint to fly to and placing it at the top of the list of waypoints. This will make it the active

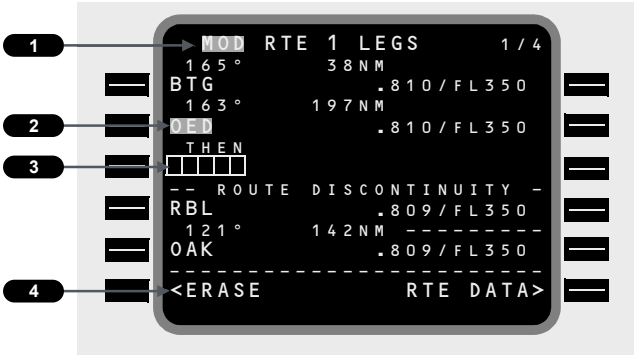
waypoint, and remove all points before it. By default, the FMS will propose the most direct route to this waypoint as shown by the dashed line ❹ in Figure 2-10b. However, a specified course to intercept the route towards this waypoint can also be commanded. The FMS will then fly the aircraft such that the specified inbound course is intercepted. The dashed line ❸ in Figure 2-10b shows an example where an intercept course of 150° is commanded to the `BTG` waypoint.

Airway intercept

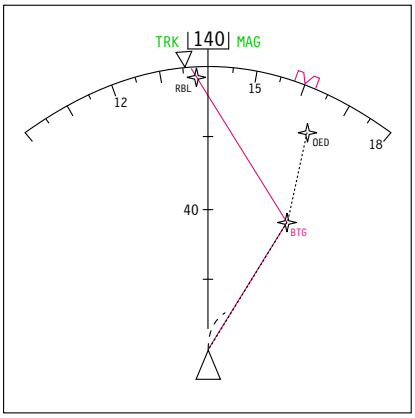
Besides the possibility to intercept a commanded direct-to routing, the FMS also has the ability to intercept another airway. This can be used following a request by ATC to discontinue following a present airway and transition to a new one, for example in an attempt to avoid weather, traffic, terrain or military operations. A pilot will enter the new required airway and route in the CDU, and subsequently the FMS will compute the closest waypoint abeam to the present position, as well as a logical position along the new airway to intercept. As in the direct-to functionality the intercept course can be adjusted to the pilot's liking.

Heading select

Besides the aforementioned lateral flight trajectory adjustments that can be commanded using the CDU and FMS, a pilot can at any moment during a flight revert to a more direct control using the heading select and hold mode. In this mode the pilot will select a specified heading, instead of request the FMS to follow the lateral flight plan. The autoflight systems will continue to maintain the heading until the pilot changes it or the FMS `LNAV` functionality is re-activated. As shown in Figure 2-12b a magenta heading bug will activate, giving a visual representation of the selected heading. Heading select is often used by ATC to provide *"radar vectors"* to a runway's instrument landing system, a temporary deviation of flight path due to traffic or by the flight crew in an attempt to avoid weather. Heading select may also be used in combination with the route offset functionality. For example, if a crew is cleared to deviate a maximum of thirty nautical miles due east of the flight trajectory, in an attempt to avoid weather, the pilot can insert the offset in the CDU which will in turn display the dashed magenta line on the navigation display. Subsequently the pilot can manually command a heading using heading select, and use the offset line on the ND as an indication of the cleared *'solution space'* within which the pilot is allowed to manoeuvre.

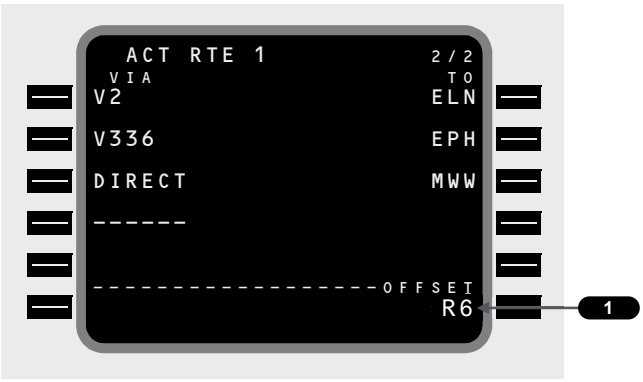


(a) Control display unit showing the RTE LEGS page [25]

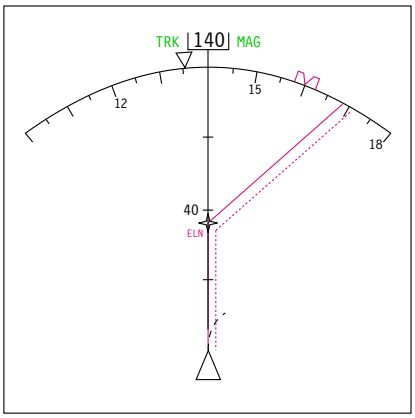


(b) Navigation display [25]

Figure 2-8: Whilst airborne, a pilot can **adjust his or her route** on the ❶ RTE LEGS **page** by simply typing in the identifier of the required waypoint into the scratchpad and inserting it where needed. In this example, the waypoint ❷ OED **is inserted** in between BTG and RBL, which is reflected on the navigation display (Figure 2-8b). A ❸ **route discontinuity** is shown, since the FMS does not automatically assume that RBL will be the waypoint succeeding OED. If this is the pilot's intention he or she can select RBL and insert in on the third line to re-connect the route and resolve the discontinuity. Notice that the navigation display (Figure 2-8b) will show a preview of any route modifications with a dashed line. The entire operation can be ❹ **cancelled by pressing** ERASE. *The navigation display* shows the active route in and the modified route in a dashed white line (shown here in grey for clarity). Pressing the EXEC button on the CDU will activate the modified route, commanding the autoflight system to follow it, an action that will immediately be reflected on the ND.

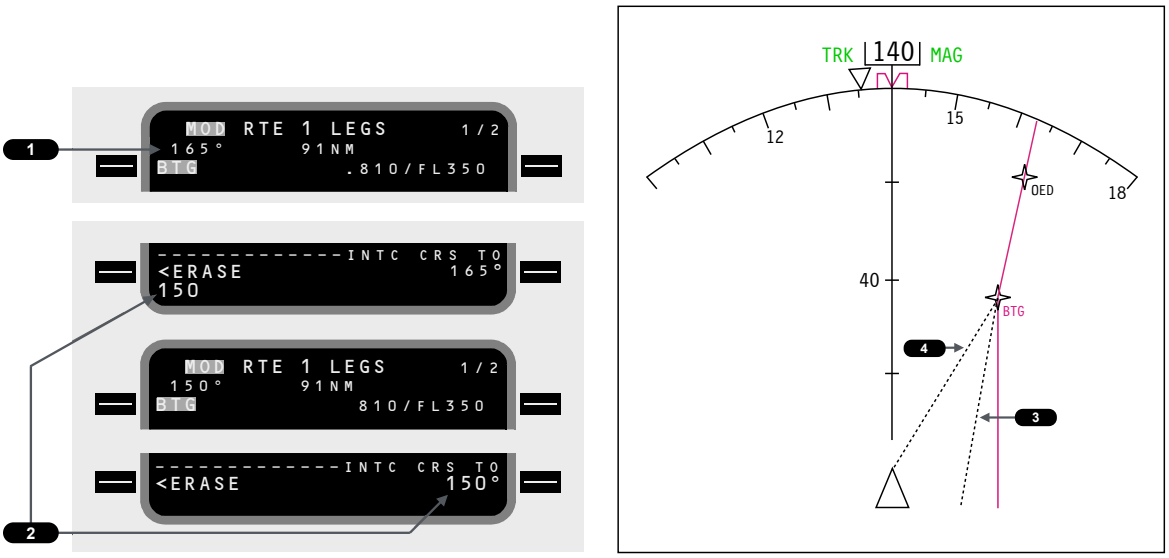


(a) Control display unit showing the ACT RTE page [25]



(b) Navigation display [25]

Figure 2-9: A ❶ **route offset** can be entered on the RTE page by specifying the direction together with the distance in nautical miles. Here R06 commands a route offset of six nautical miles to the right of the flightpath, which is reflected by a dashed magenta line on the navigation display.



(a) Control display unit showing the RTE LEGS page [25]

(b) Navigation display [25]

Figure 2-10: A **direct-to** is commanded on the RTE LEGS page by placing the ❶ **desired waypoint** at the top of the list. The FMS will propose ❹ **the shortest routing**. In case a ❷ **specified intercept course** is desired this can be inserted and the FMS will reflect the ❸ **intercept routing** on the ND. Upon execution of the direct-to command the dashed lines will become magenta, confirming that the routing has been activated.

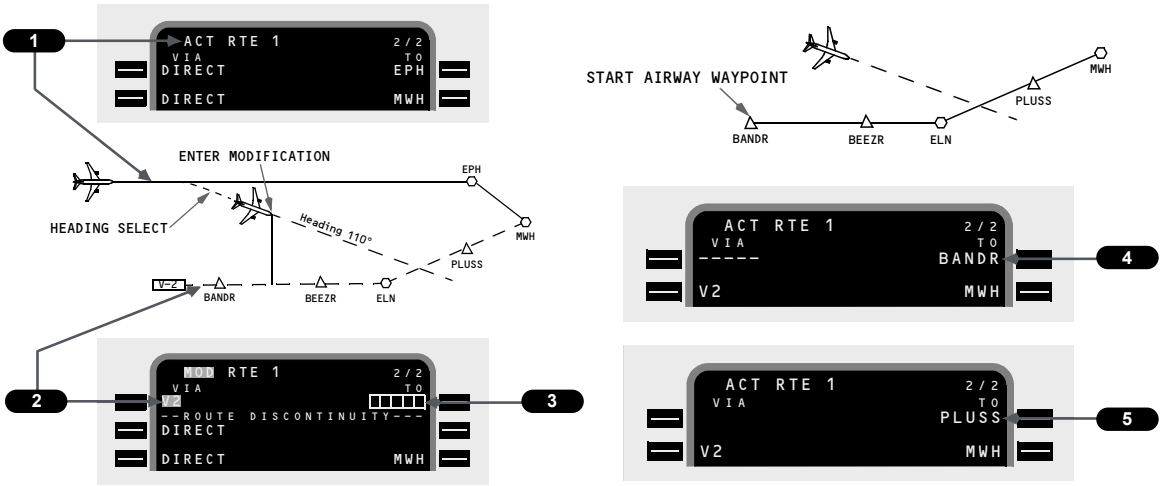
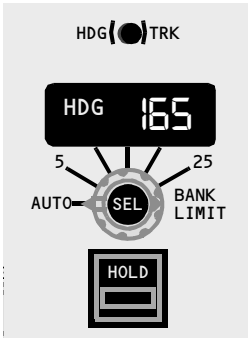
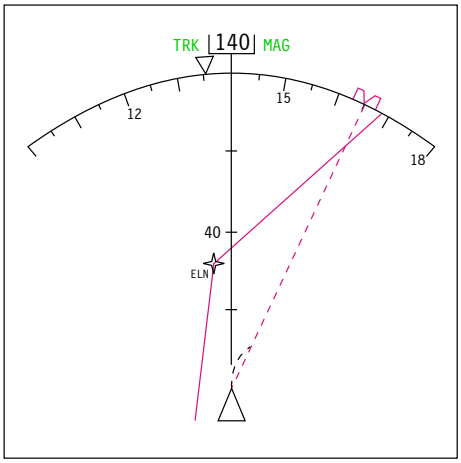


Figure 2-11: The ❶ **active route page** displays the current flight trajectory flying direct to EPH and subsequently to MWH. In this example, ATC requests the pilot to turn right and maintain a heading of $110°$ to intercept the V – 2 airway to MWH. To comply, the pilot will first command the autoflight system to maintain the desired heading and deviate from the flight route (see Figure 2-12a). The ❷ **desired airway** is inserted and subsequently the ❸ **desired airway exit point** is inserted in the boxes that appear. As a result, the FMS will identify BANDR as being the ❹ **closest abeam waypoint** on the new airway. Subsequently the pilot will select the new airway as the active route and uses the intercept procedure described in Figure 2-10a to intercept V – 2 at a course of $110°$. The FMS will subsequently identify PLUSS as the ❺ **new active waypoint** given its closest proximity to the new airway [25].



(a) Mode control panel (MCP) showing the heading select and hold interface [25]



(b) Navigation display [25]

Figure 2-12: The heading select mode will deactivate the FMS LNAV (lateral navigation) mode and command a pilot-specified heading which it will hold until it is altered. In this example, a heading of 165° is commanded, reflected on both the mode control panel and the navigation display by the magenta heading bug.

2-3 Operational Experiences

Following the review of CDU functionality and operational procedures, various pilot's were interviewed with the goal of gaining a better understanding of the actual usability of the CDU interface and features. Furthermore, pilots were asked which lateral navigation tasks were encountered often, and which very rarely. Finally, the use of touchscreen technology on the flight deck was proposed, in order to gather an operational point of view. A mixed selection of commercial pilots, the majority with extensive experience using a Boeing-style CDU, and research pilots were interviewed. Pilots will be cited anonymously, as *'Pilot A, B, C...'*, however an indication of their experience is given in Table 2-1.

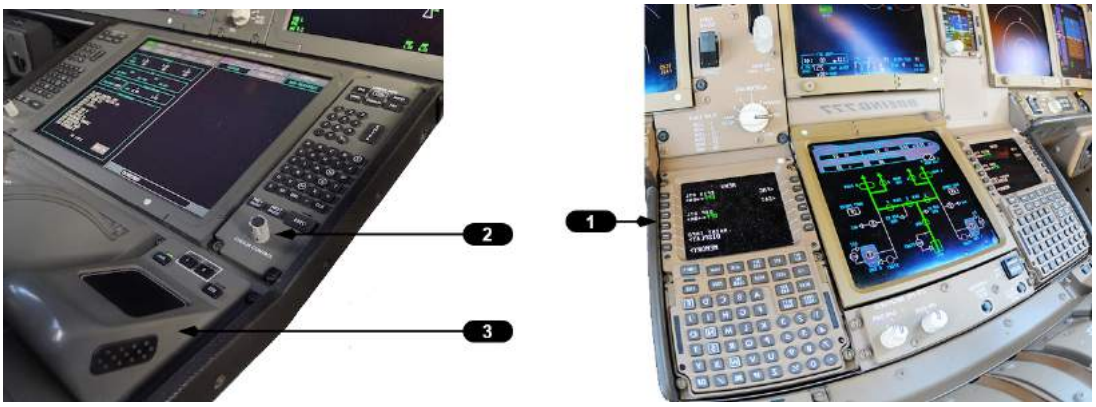
Table 2-1: Overview of interviewed pilots

Pilot	Type	Experience
A	Retired airline pilot	20,000 flying hours logged, mostly on DC-9, Boeing 737 and 747
B	Retired airline/research pilot and professor	9,000 flying hours logged, mostly on Cessna Citation, Boeing 737, 757 and 767
C	Active airline pilot	19,000 flying hours logged, mostly on Boeing 767, 777 and 787 and MD-11
D	Active research pilot and flight instructor	1,800 flying hours logged, mostly on Cessna Citation and Fairchild Metroliner
E	Active research pilot	3,500 flying hours logged, mostly on Cessna Citation and general aviation

Using the CDU

In general, operational experiences with the control display unit were found to be very positive. The system is very robust and, given the necessary training, easy to use. Commands are kept very short and easy to memorize. For example, inserting a direct-to command is quickly done, with only a few clicks on the CDU's keyboard. Furthermore, the availability of a dual route on the modern Boeing FMS and the ability to create and store custom waypoints is a feature greatly appreciated by Pilots A and C. Whilst working on editing a route, the dashed preview presented on the navigation display on the Boeing 747, 777 and 787 is always used by Pilots A and C to preview the modification, and helps them increase their situational awareness. The lack of this feature on the Cessna Citation FMS, was mentioned by Pilot D as being very unfortunate. However, Pilot D remarks that whilst the CDU may be easy to use for an experienced user, it remains a very non-intuitive interface, especially during highly stressful situations when pilots tend to speed up their decision making. The crash of American Airlines Flight 965 into the mountains surrounding Cali, Colombia is often used as an example of the CDU's inherent design flaw. Due to the textual (instead of graphical) representation, the crew of Flight 965 accidentally re-routed the aircraft to an incorrect waypoint, became confused by the information presented on the CDU, and failed to notice their mistake. The result led to a loss of situational awareness, and ultimately a controlled flight into the terrain surrounding the valley of Cali, Columbia [26].

Pilot C holds, at time of writing, a dual type certificate for both the Boeing 777 and 787 and flies both aircraft interchangeably, allowing him to reflect on the differing CDU interfaces (see Figure 2-13a and Figure 2-13b. The interface with the Boeing 777 CDU consists of twelve line-select keys, as described in Figure 2-5 and shown by ❶ in Figure 2-13b, together with an alphanumeric keyboard. The Boeing 787 also features an alphanumeric keyboard, however now the display can been moved to the large widescreen center display, and as a result the line-select keys have been removed. They have been replaced by two interfaces: a rotary controller and cursor control device (CCD), both of which can be used to select the lines and page buttons of the CDU. Based on his first experiences, Pilot C concluded his preference for the more conventional line-select buttons of the 777. The CCD was over-sensitive, increasing the error-rate and overal workload. When flying the Boeing 787, Pilot C tries to avoid the CCD and prefers using the rotary controller instead. Furthermore, Pilot C remarks that altough it may look very basic, the alphanumeric keyboard works very effectively. Since the typical CDU entry is no longer than ten to fifteen characters, the necessity for a more familiar QWERTY-style keyboard is not present. In fact, the airline Pilot C works for provided its crews with a loose QWERTY keyboard, but soon removed it from the flight deck since it was rarely used.



(a) CDU interface of the Boeing 787 consists of a ❷ **rotary controller**, ❸ **cursor control device** and alphanumeric keyboard.

(b) CDU interface of the Boeing 777 conists of ❶ **line-select keys** and alphanumeric keyboard.

Figure 2-13: Comparison between 787 and 777 CDU interface which Pilot C uses interchangeably.[3]

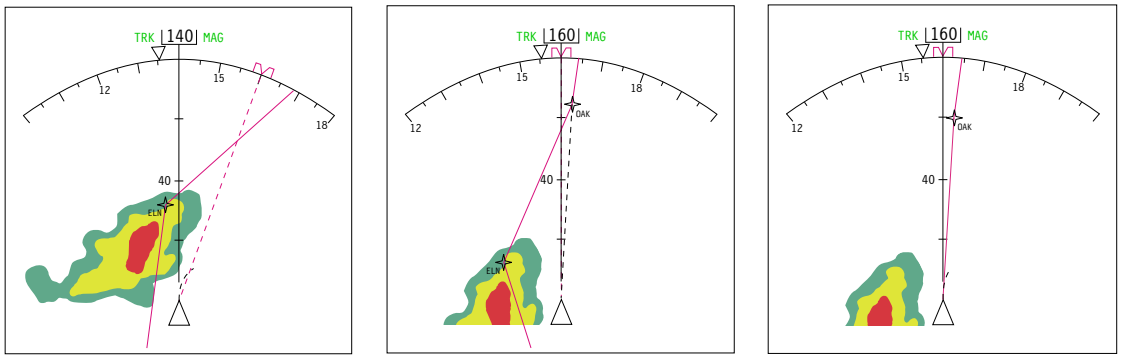
En-route re-planning tasks

Prior to departure commercial pilots will enter the flight plan based on a set of airways and waypoints as designed by the airlines' dispatch office and filed with air traffic control (ATC). However, once airborne Pilots A and C rarely re-programmed or changed the airways in their programmed route. In fact, especially in modern airspaces such as North America and Europe, commercial pilots are much less adhering to traditional airways as was the case a few decades ago. Air traffic control, especially in light of SESAR and FAA NextGen programs are continuously trying to give more direct routings to improve efficiency and punctuality

[3]Photographer: R. Ngo (left), Tek (right) (note the intentional horizontal inversion of the image)

[6]. Only overhead more traditional airspaces, such as continental Africa, Russia and China do commercial aircraft fly a specific airway for the entire flight. Pilots A and C remarked that it is highly unusual to receive an ATC clearance for an entirely different airway when en-route. Operationally, Pilots A and C have only experienced this overhead the United States, during attempts by ATC to re-route traffic around very large weather systems, such as hurricanes, however this only occurs approximately once every twenty flights. All interviewed pilots concluded that the vast majority of airborne route adjustments consistent of direct-to clearances given by air traffic control.

Airborne route adjustments, from the flight deck perspective, are mainly conducted in an attempt to avoid weather. All pilots mentioned that weather avoidance is always, with the aforementioned exception involving very large systems overhead the United States, using the heading select mode of the autopilot (see section 2-2). This is done because the weather radar onboard of commercial aircraft scans weather by looking ahead, and as a result will only show the initial set of weather returns, hiding potential weather behind it. This makes it very difficult for airline crews to perform long-term weather re-routing. Furthermore, weather can be very dynamic. Requesting ATC to deviate from the flight path using heading select is a very effective approach as it allows pilots to fly a dynamic, easy to change route around weather systems and respond quickly to new and updated weather radar returns. Finally, the chance that a waypoint exists exactly at the position where a crew wants to direct their aircraft towards is very small, if not non-existent. Following the safe passage of weather ATC will in the majority of cases issue a direct-to clearance to a logical subsequent waypoint en-route, in a minority of cases an airway intercept clearance is issued. Summarising, the common practice for weather avoidance is illustrated graphically using the navigation display in Figure 2-14.



(a) The heading bug on the ND activates, and reflects the commanded heading of 165°

(b) Following safe passage ATC cleares a direct route to OAK which is inserted into the CDU

(c) The direct route is activated, which is in turn reflected on the ND

Figure 2-14: Common practice for enroute weather avoidance [25] and Pilots A, C and D

To conclude, Pilot A made an interesting observation regarding weather avoidance techniques. He remarked, that given the likelyhood of satellite weather datalink in the future, which will provide the pilot with up-to-date weather information for the entire flight an FMS-based route-adjustment approach could likely replace the heading-select method. The satellite weather datalink will enable flight crews to proceed with longer-term flight planning, making an FMS-based route adjustment approach more suitable.

Touchscreens on the flight deck

Upon proposing the idea of introducing touch on the flight deck, responses were mixed. Overall, the interviewed pilots were hesitant, especially given the lack of tactile feedback. Being able to locate a switch by means of feeling or a quick glance and subsequently being able to activate it without looking and still be sure that one exact click was registered is an immense advantage of a button, says Pilot D. Touchscreens currently lack effective tactile feedback and require the user to find a switch based only on visual cues. A comparison can be made to operating a navigation device in your car whilst driving. One can easily find the physical rotary knob used to zoom in and out without looking, however properly locating and activating the zoom buttons on a touchscreen requires the driver to momentarily glance away from the road. Furthermore, aircraft manufacturers have designed cockpit switches to each have a unique shape and size. For example, a pilot can, without looking, grab for the flap handle and based on the shape and size of the handle he or she encounters confirm whether it is indeed the flap handle.

On the contrary, Pilot A predicted that younger next-generation pilots which have been growing up with touch-enabled mobile devices will find it easier to embrace and get used to touch on the flight deck. Furthermore, the potential of further enlarging cockpit displays featuring a touch interface could help improve situational awareness. For example, it can at times be difficult to understand an air traffic controller when he or she clears a specific waypoint. Being able to interact directly with the navigation display, panning around to find the desired waypoint will not only have the potential of speeding up the flight planning task but also increase the situational awareness of the crew, according to Pilots B and D.

2-4 Future Developments

Although the CDU seems to meet the demands of the lateral navigation tasks pilots face today, one must look at the sustainability of the CDU interface in the context of future developments in the field of lateral navigation. Huisman et. al. expect an increased frequency of en-route route adjustments in the future, mainly due to the introduction of 4D trajectory based operations [5]. Carmona et. al. describes 4D trajectory based operations (TBO) as follows: *4D [TBO] is based on establishing in advance a sequence for all aircraft converging to a specific point in a congested area. This is achieved using trajectory predictions computed by ATM ground and airborne systems. The main idea is to provide each aircraft with a time constraint to get a specific merging point while allowing this aircraft to perform an autonomous flight in order to achieve this point in a given time* [3]. An example of such a required time of arrival constraint is shown in Figure 2-15a and Figure 2-15b, both extracts of a SESAR demonstration flight. Furthermore, the possibility of further developments of *"free flight"*, where responsibility for safe separation of traffic is moved back to the flight deck, can further increase demands [5]. Van Marwijk et. al. call for *"a redesign of the navigation planning interface"* due to *"increasing punctuality in future ATM environments, in particular the American NextGen and European SESAR concepts, [which will] make airborne flight plan amendment icnreasingly complex"* [6]. In light of these developments Mulder et. al. voices concerns for the current weather avoidance procedure, since the manual heading select mode will make it increasingly difficult to adhere to the required time of arrival (RTA) requirements.

Weather re-routes will have to be executed using the FMS, such that it can recalculate the necessary speed, lateral and vertical profiles to ensure RTA compliance [4].

(a) Airbus A320 controller-pilot data link communications (CPDLC) display showing an incoming request to cross the CH993 waypoint at 16:23:00 with a tolerance of only 10 seconds. [2]

(b) Required time of arrival (RTA) constraint reflected on the navigation display [2]

Figure 2-15: SESAR demonstration flight of technology required for 4D trajectory-based operations

2-5 Requirements for the Next Generation

This first literature review topic has focussed on the state-of-the-art in the field of airborne lateral navigation. An introduction of the flight management system (FMS) in general, and the specific implementation on the Boeing 777 was presented. Combining knowledge of the CDU operating procedures and operational experience of a group of commercial and research pilots, the CDU seems to be an effective tool for the enroute lateral navigational tasks pilots face today. These tasks, according to the interviewed pilots, boil down to a heading-select approach for weather avoidance and direct-to approach for ATC-cleared route adjustments. However, a review of future developments in the field of lateral navigation, which include 4D trajectory-based operations and free flight, question the sustainability of the current CDU interface.

The following two chapters will present a literature review of how a next-generation FMS interface, using touchscreen technology, could possibly meet the identified future demands on lateral navigation.

Chapter 3

Next Generation Lateral Navigation

Evaluating the Speed-Accuracy Tradeoff

As discussed in chapter 1 one of the requirements for a next generation lateral navigation interface is that it should decrease task execution time and error rate. This chapter focusses on the manual use of the Pilot-FMS interface, and how different input methods can be compared and evaluated. Optimizing the combination of task execution time and error rate can be seen as a speed-accuracy tradeoff, and forms the basis of the Fitts' law model described in section 3-1. The Fitts' law will be introduced, and several relevant extensions are discussed, complemented with a reflection on the usability of the model during this thesis. Second, the theory behind direct manipulation will be introduced in section 3-2, of which its relevancy can be attributed to the predominant advantage of directness in touch-based interfaces. Third, an introduction to touchscreen technology, complemented with recommendations for its use is presented in section 3-3. Finally, this chapter concludes with relevant evaluations found in literature of the use of touch, in section 3-4.

3-1 The Speed-Accuracy Tradeoff

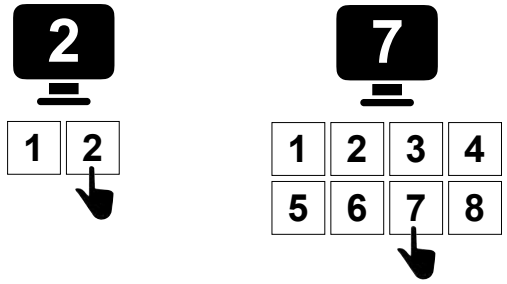
The Fitts' law, as first published in 1954, has been used by HMI researchers ever since as a predictive model, and a means to compare interfaces and input devices be means of the *throughput* measure [21]. In fact, the law has become so popular that it was introduced as the basis the ISO 9241-9 standard for performance testing of ergonomic requirements for nonkeyboard computer input devices [27]. The law published in 1954 was able to predict movement times, for humans engaging in an aimed movement task, and successfully managed to incorporate the speed-accuracy tradeoff [22]. The basic law applies to 1D situations, however multiple variations to the Fitts' law have been consequently researched, published and verified.

This section will start by introducing the theoretical basis of the Fitts' law, as introduced in 1954 and updated in 1992, in subsection 3-1-1. Next, the extensions and variations which are

applicable to this thesis are presented in subsection 3-1-2. Finally, this section will conclude with recommendations made for the use of the Fitts' law in HMI research in subsection 3-1-3.

3-1-1 The Fitts' Law

In order to gain a proper understanding of the Fitts' law, a brief introduction of information theory is a prerequisite. Consider the very simple choice reaction task (CRT) in Figure 3-1. A computer monitor displays a number, and subsequently the user is required to respond with the appropriate key. For example, in Figure 3-1a the user is presented with two alternatives, the numbers one and two, and given two is displayed he or she will press the button labeled '2'. The task shown in Figure 3-1b is slightly more complicated, presenting the user with eight different alternatives. Merkel performed a very similar task to the one illustrated in Figure 3-1 and found that the response time would increase at a constant rate, for every doubling of the number of alternatives (cited in [22]).



(a) A simple CRT with two alternatives

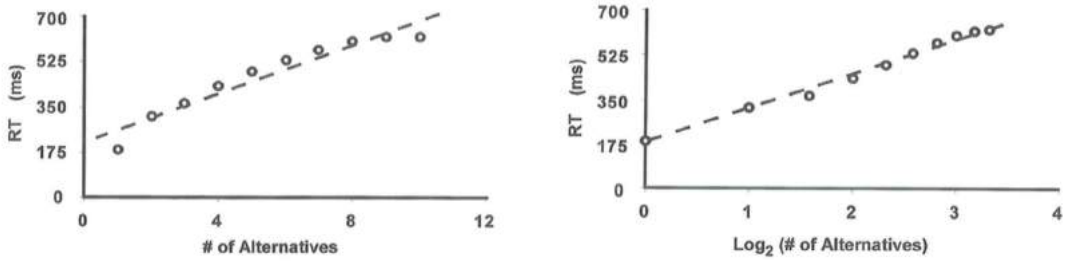
(b) A more complex CRT with eight alternatives

Figure 3-1: An illustration of a choice reaction task, where a stimulus presented on a display requires an appropriate response by means of a button press from the user [22].

Shannon and Weaver proposed a logarithmic representation, with base two, of the number of alternatives (cited in [22]) and referred to it as *"bits"*. The number of bits could be described as the average number of yes and no questions required to identify the correct response. In the example illustrated in Figure 3-1b, eight different alternatives are presented, which corresponds to a total of three bits ($\log_2(8) = 3$). The *"yes-no"* analogy can be shown here. Given that the number seven was the stimulus, the three questions *"is the number odd, is the number larger than four and is the number five?"* suffice to identify the correct response. The results of Merkel's experiment, including the logarithmic representation, is presented in Figure 3-2.

Furthermore, Hyman proposed in 1953 to describe the choice reaction task using probability by proclaiming that the *"response time depends on what could occur, not what actually occured"* (cited in [22]). As such, a task with a larger number of alternatives will take longer because there is a larger amount of different possibilities for the human to consider. The information presented in the task $H(x)$ can described using Equation 3-1 where $p(x_i)$ is the probability of alternative i.

$$H(x) = \sum_i p(x_i) \log_2 \left(\frac{1}{p(x_i)} \right) \tag{3-1}$$



(a) Absolute number of alternatives

(b) Logarithmic representation of number of alternatives

Figure 3-2: Results of Merkel's CRT experiment [22].

Hick further constructed on the principle introduced by Hyman by describing the information transmission during a specific event as a measure of uncertainty reduction (cited in [22]). This uncertainty reduction is described by Hick as the transmittion of one bit of information, and can be described mathematically by Equation 3-2.

$$\text{Information transmitted} = -\log_2\left(\frac{\text{Number of alternatives } \textbf{after} \text{ event}}{\text{Number of alternatives } \textbf{before} \text{ event}}\right) \text{ [bits]} \quad (3\text{-}2)$$

In the previous example of eight alternatives, the uncertaintity would be three bits. Imagine the user being told that the number is larger than four, the uncertaintity would reduce to two bits. As a result, a total of one bit of information is transmitted, as shown in Equation 3-3 which uses Equation 3-2.

$$\text{Information transmitted} = -\log_2\left(\frac{4}{8}\right) = 1 \text{ bit} \quad (3\text{-}3)$$

Original Fitts' Law

In 1954 Fitts performed three experiments, of which the most notable a reciprocal tapping task. In an attempt to describe the *"speed-accuracy"* tradeoff of human motor performance subjects were asked to *"emphasize accuracy versus speed to achieve an approximate error rate of 4%"* for a task involving continuous tapping using a stylus between two targets. The experimental setup is shown in Figure 3-3, where W represents the width of the target and A the distance, or amplitude between them.

Fitts used the information analogy in order to explain the observations made during his experiment. The number of alternatives is translated to possible positions of the subject's stylus during the experiment. As such, the uncertainty in the position after the event (stylus movement) is described by the target width (W), given the assumption that the subject will place the stylus on the target. The uncertainty before the movement is described by twice the distance between targets, or amplitude ($2A$). The addition of the 2 is arbitrary, and required to ensure that the result is larger than zero [22]. The number does not impact the relationship of Fitts' model, but merely adds one bit. Finally, Fitts proposes the information transmission

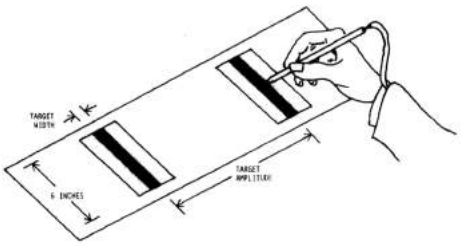


Figure 3-3: Reciprocal tapping task experiment, as executed by Fitts [22].

relation in Equation 3-2 as a description of the index of difficulty of his experiment, as presented in Equation 3-4.

$$ID \text{ (Index of Difficulty)} = -\log_2\left(\frac{W}{2A}\right) = \log_2\left(\frac{2A}{W}\right) \text{ [bits]} \tag{3-4}$$

The results of Fitts' experiment are presented in Figure 3-4 as movement times per index of difficulty. Interestingly, a similar observation is made as in Merkel's experiments, namely that movement time increases linearly as a function of information transferred, measured by Fitts as index of difficulty. As a result, the Fitts' law is as follows (Equation 3-5):

$$MT \text{ (Movement Time)} = a + b(ID) = a + b\log_2\left(\frac{2A}{W}\right) \tag{3-5}$$

The constants a and b can be found using a linear regression on data obtained through an experiment measuring movement time whilst adjusting width (W) and amplitude (A). Two observations can be made based on the proposed law. First, the y-intercept a is ideally found to be zero, such that a task of zero difficulty takes no time to complete. However, literature has shown non-zero intercepts to occur, which are often attributed to additive factors unrelated to the index of difficulty, such as the application pressure required in pushing a button [22], [23]. Second, the inverse of the slope b is described in literature as an index of performance, given that its unit is bits/s [23]).

Revised Fitts' Law

MacKenzie proposed a revised adjustment of the original Fitts' law in 1992, which he proclaims to be *"more faithful to Shannon's Theorem 17* [see Equation 3-6] *, which gives the information capacity for a communication channel (C) as a function of signal power (S) and noise power (N)"* and formed the theoretical basis of the original information analogy [22], on which Fitts' law is constructed.

$$C = B\log_2\frac{S+N}{N} \tag{3-6}$$

An oversimplified, but effective way, to understand Shannon's 17th Theorem is presented by Jagacinski [22]. The number of alternatives before a signal is sent can be seen as the sum of all possible signals (S) and all possible noise (N). Subsequently, once the signal has been

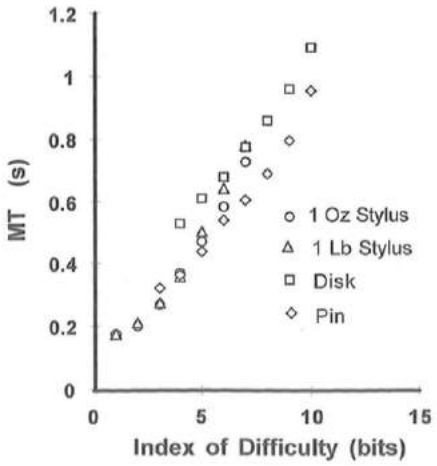


Figure 3-4: Movement times for all three tasks used by Fitts in his experiment [22].

sent, the only remaining variation results from the noise (N). Using this analogy, MacKenzie's revision of the Fitts' law (see Equation 3-7 [22], [23]) is as follows, referred to as the *"Shannon Formulation"*:

$$MT = a + b\log_2\left(\frac{A+W}{W}\right) = a + b\log_2\left(\frac{A}{W}+1\right) \tag{3-7}$$

The advantages of the Shannon Formulation are threefold. First, it is found to have a slightly better fit of the original data collected by Fitts. Second, it exactly mimics the information theory on which the Fitts' law is based. Third, the formulation always results in a positive index of difficulty. This is beneficial, given that a negative difficulty has no real meaning [23]. However, there remains a limitation given that *"users actual pointing precision may be different from the nominal task specifications"* [28]. The human error in reaching the target width is not yet accounted for.

As a response, Crossman (cited in [22]) proposed a further modification, endorsed by MacKenzie, allowing for Fitts' law to consider the impact of error in the model of movement time. The Fitts' law is designed to describe the *"speed-accuracy"* tradeoff, and as such subjects are asked to balance both measures to achieve a positive *hit-rate* around 96%. Therefore, if subjects under- or overperform this target an adjustment of the width is necessary in order to ensure the reliability of the model. This necessity is further confirmed given the noise perturbation assumption in the Shannon Formulation, which requires *"the endpoints of the movements [to] be normally distributed about the target"* [22]. In order to achieve the 96% hitrate, as illustrated in Figure 3-5, and adhere to the Shannon assumption the effective width is found to be equal to the uncertainty in a normal distribution, $\log_2\left(\sqrt{2\pi e}\sigma\right)$, where σ is the standard deviation in the endpoints [22].

Given this statement, a remark can be made regarding the human motor performance. Namely, if the experimental error is not equal to 4%, then the defined width W is not *"an appropriate index of channel noise"* or description of the actual width used by the human in the execution of the task [22]. For example, if the experimental error is much larger than

4%, the effective width will is found to be larger than the actual width in order to fulfill the requirement that *"the endpoints of the movements [should be] be normally distributed about the target"* [22]. Given the necessary success-rate of 96% the effective width W_e can be calculated using Equation 3-8, where σ_x is the standard deviation of the endpoints logged as x-coordinates along the axis of approach to the target [22].

$$W_e \text{ (Effective Width)} = \sqrt{2\pi e}\sigma_x = 4.133\sigma_x \tag{3-8}$$

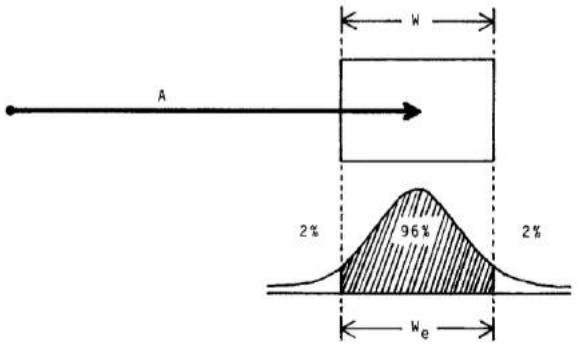


Figure 3-5: In order to properly describe the *speed-accuracy* tradeoff and accept the Shannon Formulation of Fitts' law, the distribution of the endpoint coordinates is expected to be a normal distribution, with an error rate of 4%. For experiments where this is not the case, an effective width W_e is introduced [23].

Literature finds that using the effective width in Fitts' law, as presented in Equation 3-9, more accurately reflects what experimental subjects actually did, rather than what they were asked to do. According to MacKenzie, using W_e *"more faithfully embodies the speed-accuracy tradeoff"* [23].

Summarising, the complete Fitts' law model to describe movement time (MT) as a function of index of difficulty (ID) in a high-accuracy pointing task is presented in Equation 3-9, where a and b are linear regression constants, A is the amplitude, or distance to be traversed during the movement and W_e is the effective width of the target, calculated using the standard deviation of the endpoint coordinates.

$$MT = a + b(ID) = a + b\log_2\left(\frac{A + W_e}{W_e}\right) \text{ [seconds]} \tag{3-9}$$

3-1-2 Fitts' Law Extensions and Variations

The original Fitts' law was found to be extremely robust, during many subsequent evaluations [29], however remains limited to 1D pointing tasks. In order to succesfully use Fitts' law during this thesis as an evaluative tool of the flight management system interface, several variations and extensions of the model needed to be addressed. A literature review of variations identifies numerous extensions, such as to 2D, 3D, trajectory-based and rotational tasks as well as adaptions for the use of touchscreen and keyboard entry tasks. This section will focus on the most important variations, that are applicable to the scope of this thesis.

Fitts' Law Extended to 2D Tasks

MacKenzie and Buxton 1992 [30] proposed an extension of the Fitts' law to succesfully apply to 2D tasks. Since the Fitts' law was derived for a simple 1D task, the identification of the width W is unambiguous. For two-dimensional tasks, unless the target is a circle, the width W will vary based on the approach angle, as illustrated in Figure 3-6. In their paper, the authors propose and evaluate two models to correctly model a two-dimensional tasks using Fitts' law. The first is referred to as the $\texttt{SMALLER} - \texttt{OF}$ model, which simply takes the minimum of the height H and width W as the parameter describing the target size. The second is referred to as the $\texttt{W}'(\texttt{W} - \texttt{PRIME})$ model and calculates the actual width of the target, based on the approach angle, as illustrated by Figure 3-6.

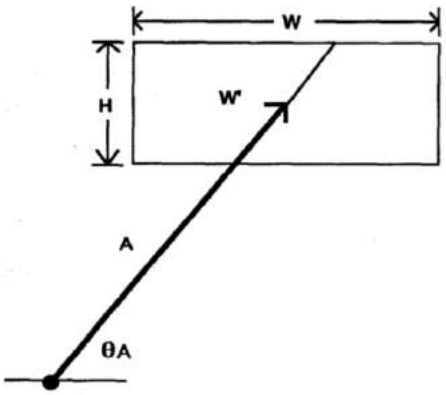


Figure 3-6: Definition of the width W is no longer straightforward in a 2D task, given that it changes based on the approach angle θ_A [23].

Both the $\texttt{SMALLER} - \texttt{OF}$ and $\texttt{W}'$ models are concluded to be empirically superior to the 1D Fitts' law, with a coefficient of determination (R^2) of 0.9501 and 0.9333 respectively, compared to 0.8097 [30]. The $\texttt{SMALLER} - \texttt{OF}$ model is concluded to be advantageous given its ease of application, whilst inherently limited to rectangular targets. The $\texttt{W}'$ model is found to be theoretically attractive, given that it preserves the one-dimensionality of Fitts' law however requires more complex calculations. Summarising, the 2D variation of Fitts' law as proposed by MacKenzie and Buxton is presented in Equation 3-10 and requires choosing between the $\texttt{SMALLER} - \texttt{OF}$ and $\texttt{W}'$ model.

$$MT = a + b\log_2\left(\frac{A + \min(H, W)}{\min(H, W)}\right) \qquad MT = a + b\log_2\left(\frac{A + W'}{W'}\right) \tag{3-10}$$

Fitts' Law Extended to Keyboard Entry Tasks

Soukoreff and MacKenzie 1995 [31] proposed using the 2D $\texttt{SMALLER} - \texttt{OF}$ extension for describing keyboard data entry tasks using Fitts' law. The proposed model is presented in Equation 3-11 and predicts the movement time between keys i and j. In the equation, A_{ij} represents the distance between the two keys and $\min(H_j, W_j)$ the minimum dimension of the target key. Since there exists a possibility that the same key is pressed consecutively,

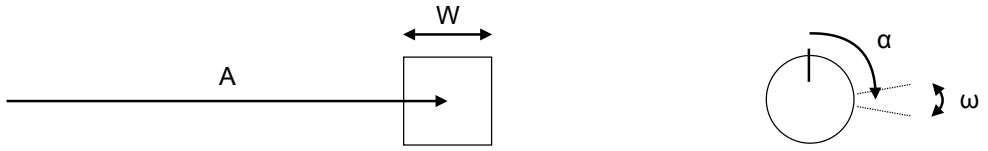
resulting in a zero amplitude A_{ij} the model refers to a pre-determined MT_{repeat} value when $i = j$. The key-repeat movement time is found using a simple experiment, asking subjects to repeatedly press the same key and subsequently measuring the time inbetween presses.

$$MT_{ij} = \begin{cases} a + b \log_2 \left(\frac{A_{ij} + \min(H_j, W_j)}{\min(H_j, W_j)} \right) & \text{if } i \neq j \\ MT_{\text{repeat}} & \text{if } i = j \end{cases} \tag{3-11}$$

This model can, and was used in both IEEE articles, to predict movement times for data entry using the alphanumeric keyboard of the control display unit.

Fitts' Law Extended to Rotational Tasks

Stoelen and Akin 2010 [27] Knight and Crossman (cited in [27]) both found similar indices of performance for rotational and translational tasks, and concluded that Fitts' law can also effectively be applied to rotational tasks. Both researchers proposed and succesfully were able to use a very simple translation of width W and amplitude A to the circular domain. Consider the the rotary controller in Figure 3-7b. The amplitude A is defined as the total angular distance α required to traverse, whilst the width W is defined as the angular tolerance ω, illustrated by the dashed lines. In this example, the values are 90° and 20° respectively.



(a) Width W and Amplitude A in a translational task

(b) Angular width ω and angular amplitude α in a rotational task

Figure 3-7: Knight and Crossman (cited in [27]) found that a simple analogy of width and amplitude worked very effectively in using Fitts' law to describe rotational tasks.

Stoelen subsequently built on their research by proposing the hypothesis that a combined rotational and translational task could be modelled using one linear regression, as shown in Equation 3-12.

$$MT_{\text{combined}} = a + b\,(ID_{\text{rotational}} + ID_{\text{translational}}) \tag{3-12}$$

$$ID_{\text{rotational}} = \log_2 \left(\frac{\alpha + \omega}{\omega} \right) \tag{3-13}$$

$$ID_{\text{translational}} = \log_2 \left(\frac{A + W}{W} \right) \tag{3-14}$$

His hypothesis was tested by performing three experiments. In the first and second, experimental subjects were asked to perform pure translational and pure rotational tasks respectively. In the third, they were asked to perform combined translational and rotational tasks.

Table 3-1: Results from Stoelen's experiment [27], showing coefficient of determination R^2 and Fitts' law coefficients a and b.

Experiment	r^2	a	b
1. Pure translational	0.984	0.220	0.310
2. Pure rotational	0.930	0.270	0.320
3. Combined translational and rotational	0.817	0.460	0.320
Addition of experiment 1 + 2	0.953	0.500	0.320

Subsequently, Stoelen compared the correlation coefficients of the Fitts' law model describing the combined experiment, and the model describing the addition of the translational and rotational experiment. The results of his experiment is shown in Table 3-1. Interestingly, the addition of the results from experiment 1 and 2 yields a much better correlation (0.953) compared to the combined task experiment (0.817). This confirms his hypothesis that for combined tasks, one can perform separate translational and rotational experiments, and subsequently sum movement times and index of difficulties to develop the Fitts' law model [27].

The research performed by Knight, Crossman and Stoelen can and was used in both IEEE articles to evaluate the performance of the heading select mode using the mode control panel.

Fitts' Law Extended to Touchscreens

Bi et. al. 2013 [28] Given this thesis has the objective of evaluating the performance of touch-based interfaces on the flight deck, a touchscreen extension to Fitts' law can not be omitted. Bi, Li and Zhai developed the Finger Fitts Law (or FFitts Law) in 2013, a variation designed for evaluating touchscreen interaction.

The biggest difference between pointing tasks on a touchscreen, compared to more traditional input devices, is that the interface (your finger) is often larger than the desired target. Furthermore, users tend to rely on visual features of their fingers for placing touch inputs, whilst the device will tend to use the centroid of the touch area as the input point. As such, the registered point does not always coincide with the perceived touch point. The variability of end points which is considered using the, earlier discussed, effective width methodology is hypothesized to be insufficient for touch-based interfaces.

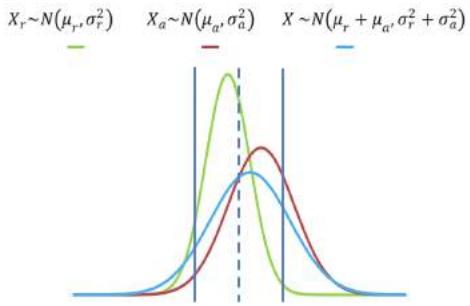


Figure 3-8: Dual variance shown in acquisition of the target. Target edges illustrated by solid blue line, center by dashed blue line. The green, red and blue lines show distribution of touch points for χ_r, χ_a and χ respectively [28].

Bi et. al. propose in their paper to adopt a dual variation in the calculation of the effective width, presented in Equation 3-15 and Figure 3-8. χ_R is defined as the relative component of the variation, and depends on the desire to hit the target (the speed-accuracy tradeoff, e.g. the variation previously used in the compution of effective width). It therefore reflects the precision relative to the movement amplitude. χ_A is a new addition, and defined as independen of the speed-accuracy tradeoff and reflects the absolute precision of the input device (finger) and internal human-motor system (human).

$$\begin{aligned} \chi &= \chi_R + \chi_A \approx N\left(\mu_r + \mu_a, \sigma_r^2 + \sigma_a^2\right) \\ \sigma^2 &= \sigma_r^2 + \sigma_a^2 \\ \sigma_r &= \sqrt{\sigma^2 - \sigma_a^2} \end{aligned} \tag{3-15}$$

Based on his dual variation hypothesis, the paper computes the variation of the relative component as a function of the standard deviations of the total and absolute parts. Building upon the original calculation of effective width W_e in Equation 3-8, Bi proposes an updated effective width, derived as follows and presented in Equation 3-16.

$$\begin{aligned} W_e &= \sqrt{2\pi e}\sigma_r && (= 4.133\sigma_r) \\ W_e &= \sqrt{2\pi e}\sqrt{\sigma^2 - \sigma_a^2} && \\ W_e &= \sqrt{2\pi e\left(\sigma^2 - \sigma_a^2\right)} && \left(= 4.133\sqrt{\sigma^2 - \sigma_a^2}\right) \end{aligned} \tag{3-16}$$

Summarising, the updated calculation of effective width as proposed by Bi et. al. for the evaluation of touch-based devices is now dependent on σ, the variation in endpoint coordinations (e.g. previously already a component of Fitts' law) and the σ_a, the variation of the input device precision (e.g. the finger). The former calculated using the variance in endpoint coordinates during the task, where a bivariate variance σ_{xy} is used for 2D tasks. The latter can be measured using a finger calibration task, where users are asked to repeatly touch the same target; exact touch locations is used to calculate the variance σ_a. The exact Finger Fitts Law is shown in Equation 3-17 where W_e is substituted with Equation 3-16.

$$MT = a + b\log_2\left(\frac{A + \sqrt{2\pi e\left(\sigma^2 - \sigma_a^2\right)}}{\sqrt{2\pi e\left(\sigma^2 - \sigma_a^2\right)}}\right) \tag{3-17}$$

Bi et. al. tested their revised model during three experiments, namely a 1D, 2D and typing task using a touchscreen, and found the Finger Fitts model to show better correlation for all experiments [28]. The model can and was used in both IEEE articles to develop a predictive model of movement time for tasks using the touch-based navigation display.

3-1-3 Using Fitts' law in HMI Research

Given the intention to use Fitts' law as an evaluative and predictive tool during this thesis, a review of Soukoreff and MacKenzie's paper in 2004 [21], stating several recommendations for

the use of Fitts' Law in HMI research is useful. Both gentlemen have extensive experience in the field, and therefore the recommendations hold significant value. The majority of these recommendations coincide with the description of Fitts' law in the ISO-standard [21]. A summary of these recommendations will now be presented, for a more extensive discussion and reasoning behind each of them the reader is referred to [21].

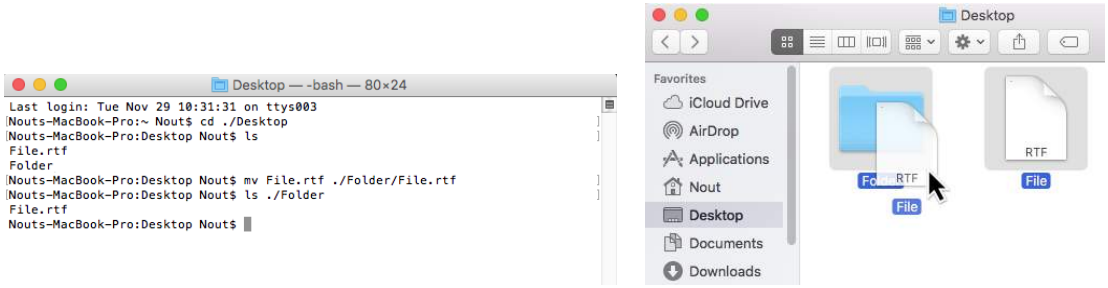
They advise the use of the Shannon Formulation (Equation 3-9) to model the experimental data. Furthermore, a wide and representative range of indices of difficulty (ID) is to be used. As the Fitts' law consists a linear regression of experimentally found data, it loses its predictive power outside of the bounds of ID used during the experiment. Therefore one should limit predictions using the constructed Fitts' law to the range of chosen ID values. Furthermore, a recommendation is made to measure the variation of endpoint position, and respective error-rates in order to compute the effective width (W_e). This adjustment is crucial to ensure a proper fit and predictive reliability of the model. The model itself should be developed using a linear regression with constants a and b, where the y-intercept a should be verified to be zero. In case of a non-zero intercept, a discussion must be held on the presence of additive factors affecting the task performance. Finally, if the Fitts' law is to be used in comparing input devices and interfaces, the dependent variable *throughput* can be effective in doing so. Presented in Equation 3-18, it calculates the throughput in bits per second by dividing the index of difficulty by the measured movement time for each subject and experimental condition. The variables x and y define the total number of experimental conditions and subjects respectively, $ID_{e_{ij}}$ the index of difficulty, adjusted using the effective width W_e, for a specific experimental condition and subject, and MT_{ij} the movement time for a specific experimental condition and subject.

$$TP \text{ (Throughput)} = \frac{1}{y}\sum_{i=1}^{y}\left(\frac{1}{x}\sum_{j=1}^{x}\frac{ID_{e_{ij}}}{MT_{ij}}\right) \text{ [bits/s]} \tag{3-18}$$

Concluding, the usefullness of the Fitts' law for this thesis is found to be twofold. First, it can provide a quantitative description of the flight management interface by comparing the throughput of the input devices under evaluation. Second, it can be used to construct a predictive model of task execution time using a particular interface. The 2D extension of Fitts' law, proposed by [30], will be used in combination with the extension for keyboard entry, proposed by [31], to evaluate data entry using the control display unit. The rotational task extension, as proposed by [27], will be used to evaluate the heading select mode using the rotary controller on the mode control panel. Finally the 2D extension, combined with the Finger Fitts law, as proposed by [28], will be used to evaluate the touch-based navigation display interface.

3-2 Direct Manipulation Theory

Direct manipulation has been praised as an intuitive and effective approach to interface design, and is centered around the principle that representations of objects are visualized, behave and can be interacted with as if they are the objects themselves. For example, manipulating a file on your computer, being able to select it and drag it into a desired folder is an example of direct manipulation (see Figure 3-9a and Figure 3-9b).



(a) Manipulating files on your computer using the terminal command-line interface, is a typical example of indirect manipulation.

(b) Intuitively being able to select, drag and drop files into folders using the graphical user interface, is a typical example of direct manipulation.

Figure 3-9: Direct manipulation explained by example of computer file management

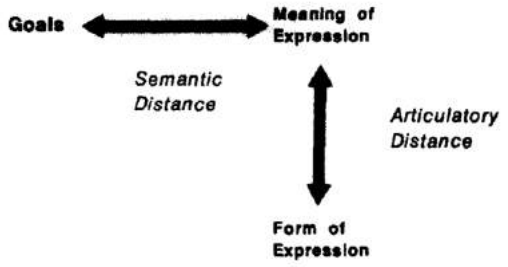
Shneiderman et. al. first introduced the term *direct manipulation* in 1982, allocating the term to interfaces where *"physical actions or labeled buttons [replace] complex syntax and textual commands"* [9]. Objects of interest are often displayed graphically, in a manner that reflects *"the way one thinks about the problem"* [11]. Hutchins et. al. laid the foundation of direct manipulation theory in 1985, identifying *"the root of the approach [being] the assumption that the feeling of directness results from the commitment of fewer cognitive resources"* [11]. In his paper, he describes a direct manipulation interface (DMI) as a combination of two concepts, namely direct *distance* and direct *engagement*. Furthermore, Cook proposes describing a direct interface as a *verite system* [32]. All three concepts are described in subsequent sections.

3-2-1 Concept of Distance

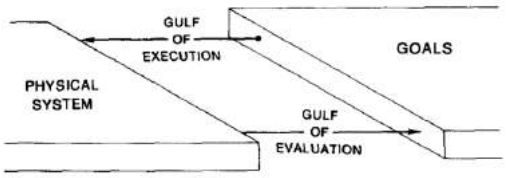
Distance is described as the *"[seperation] between one's thoughts and the physical requirements of the system under use"* [11] and separated into two forms: *semantic* and *articulatory*. The former reflects the relationship between the meaning of the expression in the interface language and the intention of the user. The latter reflects the relationship between the physical form of an expression in the interface language and its meaning. An example of an interface with large semantic distance is one that does not display information in the same way a user thinks about them. An example of large articulatory distance is if the command "cat" has to be used to display a file; there is no relationship between the command name and intended use [17]. Semantic and articulatory distance together form the "gulfs of execution and evaluation" [11] which Hutchins uses to describe the gap between the goals of the user and the physical system itself. Figure 3-10a and Figure 3-10b presents a graphical representation of Hutchins' theory on distance.

3-2-2 Concept of Engagement

Engagement is described as *"the qualitative feeling that one is directly manipulating the objects of interest"* [11]. Hutchins presents two metaphors of human-machine interaction, being a conversation and model-world metaphor. The latter *"creates the sensation in the user of*



(a) Semantic distance represents the relation between goals with the meaning of the expression and articulatory distance the relation between the form of expression with the same meaning [11]



(b) Gulfs of execution and evaluation span the gap between the user goals and the physical system. The gulfs are unidirectional and include a combination of both semantic and articulatory distance [11].

Figure 3-10: Hutchins' basic principle behind *distance* as an aspect of *directness*

acting upon the objects of the task domain themselves" [11]. For example, if a theater production presents a model-world interaction *"the members of the audience ... wilfully suspend their beliefs that the players are actions and become directly engaged in the content of the drama"* [11]. Ballas argues that to achieve engagement a user and computer system need to share a common communications medium, often a visual display with presents a graphical and explicit view of the task domain [17].

3-2-3 Concept of Verite Systems

Cook et. al. in 1996 provided an additional explanation of directness by segmenting human-machine systems into three forms: verite, abstraction and ordinateur, which in turn imply true, remote and virtual respectively [32]. In verite systems, the control and display are the same object. For example, in a small piston-powered aircraft the yoke is connected directly to the control surfaces and as a result the position of the yoke is a direct representation of the position of the ailerons and elevators. In abstraction systems, the response of an object is not apparent in the control itself, but in a dedicated display. For example, a flap lever can be used to command an extension, which is subsequently confirmed on a respective flap position display. However, in case of a failure in the extension mechanism this will only be apparent in the display. To return to the principle of a verite system a pilot should not be able to move the lever if the actual flap cannot move either [32]. Finally, ordinateur systems are freed from one-to-one control/display relationships. For example, pressing the `ENTER` key on a computer keyboard can lead to multiple outputs and is thereby not restricted. Fly-by-wire control systems found on modern flight decks are examples of ordinateur systems [32].

However, creating a verite system needs to be achieved without exposing the operator to physically unworkable or dangerous tasks [33]. For example, extending the landing gear by directly moving the gear itself (see Figure 3-11) would be physically too dangerous. As a result Lintern [33] and Cook [32] plea designers to create a thorough understanding of how pilots mentally visualize certain functionalities and processes, in order to incorporate more of verite systems' valuable features. Development of a comprehensive abstraction hierarchy is, according to Lintern [33], the best methodology to ensure that all relevant properties are

represented in the interface.

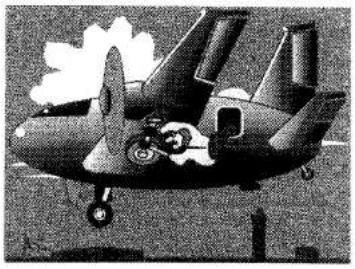


Figure 3-11: The most primitive verite (true) system. The landing gear is literally directly manipulation in order to extend or retract. A malfunction will immediately be apparent by the inability to operate the control [33].

Concluding, the use of direct manipulation in a next generation lateral navigation input devices promises improve the ease and intuitiveness of the interface. Touchscreen technology is the most predominant embodiment of direct manipulation in hardware, which will be further discussed in the subsequent section.

3-3 Touchscreen Technology

In the past decade touchscreen technology has become an integrated part in everyday life. A vast majority of the world now carriers a multi-touch capable smart phone around everyday. Babies are growing up with iPad's and manage to use the interface without a lot of explanation. The potential of multi-touch technology is tremendous. Jeff Han [34] introduced the world to the multi-touch gestures during his TED talk in 2005. Soon thereafter, in 2007 Apple launched the iPhone introducing touch technology to a mobile device. Essentially, three types of touchscreen technology exist, each with respective (dis)advantages.

Resistive touchscreen technology is often used for simple, single-contact systems such as the interface with your printer. Capacitive technology is used whenever high quality, multi-touch interfaces are needed such as a mobile phone, tablet or in this case avionics. Figure 3-12 shows the basic principle of capacitive touchscreens. Touching the glass using a conductive element, such as your finger, will cause a change in capacitance and oscillation frequency in the conductive pattern. Using software, the contact point is translated into an absolute position on the screen. Since this is possible at multiple positions in the pattern, the technology is capable of registering multi-touch gestures [15].

3-3-1 Benefits and Challenges

The specific use of touchscreen technology on board of the flight deck has been discussed within literature, most notably by Degani et. al. [14] and Kaminani et. al. [8], [36], [37]. These specific benefits and drawbacks of the technology itself are summarized below. section 3-4 presents a review of touchscreen evaluations in the context of task performance.

Table 3-2: Overview of the three major touchscreen technologies [35]

Type	Functionality	Benefits	Drawbacks
Resistive	Several thin transparent, electrically-resistive layers separated by a thin space. Touching the display will connect the layers, registering a signal	Resistant to liquids, contaminants, low cost, can be used with gloves	Poor contrast, significant pressure needed
Capacitive	Coated with a thin transparent conductive material. Touch registered when there is a measurable change in capacitance (makes use of the fact that your skin/body is conductive.	Little pressure is needed, smooth interaction, multi-touch possible	Gloves will not work
Infrared	Uses infrared beams across the active surface to detect input. Disrupting a beam registers touch.	Can detect essentially any input device	Sensitive to dirt and dust, easy to accidentally press since little contact pressure is needed.

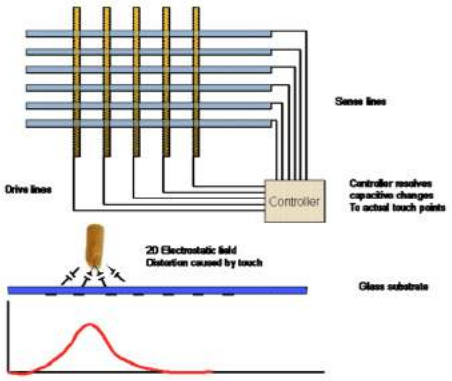


Figure 3-12: Principle of capacitive touchscreens: when the glass is touched by a conducting element (i.e. a finger), a change in capacitance is registered and translated to an absolute screen position, and capable of registering multi-touch gestures [15]

Touchscreen benefits and challenges

Potential Benefits

- Potential of reducing flight deck workload by introducing easy manipulation capabilities [8]
- Potential to declutter the flight deck, since one touchscreen display can be used for multiple interfaces and thus uses less space [10]
- Potential to reduce cognitive effort, search time and motor movement [38]
- Supports direct hand-eye coordination, no need to memorise commands reducing training efforts [10]

Challenges

- Difficulty of selecting small targets and doing precise pointing; *"the fat finger problem"* [19]
- Occlusion of the target object since the hand will be infront of the screen [19]
- Lack of tactile feedback, decreasing the ease of using soft controls compared to hard controls [8], [14]
- Loss of dedicated, geographical location of controls and resulting pilot muscle memory [8], [14]
- Necessity to navigate between displays during emergency situations [8], [14]
- Response time and display update rate of complex systems [8], [14]
- Reliability of soft controls and interfaces in high-risk domains [8], [14]
- Environmental effects such as vibration and glare on the visual and tactile usability [8], [14]
- Accidental touches [8], [14]
- Physical fatigue of controlling a touchscreen and inability to rest arm on display [8], [14]
- Fingerprint residues, dirt from gloves [8], [14], [15]

3-3-2 FAA Regulatory and Guidance Material

As a response to a growing interest in the use of touchscreen technology, the Federal Aviation Administration has developed regulatory and guidance material [35]. The main regulatory item was to *"ensure that touch screens do not result in unacceptable levels of workload, error rates, speed and accuracy ... [as well as ensuring that they] resist scratching, hazing, or other damage that can occur through normal use ... [and continue] to provide acceptable performance after long-term use and exposure to skin oils, perspiration, environmental elements (e.g., sun), impacts (e.g., clipboard), chemical cleaners that might be used in the flight deck, and any liquids that be brought onboard by flightcrew members (e.g., coffee)"* [35].

Recommendations on the use of Touchscreens

Beringer 1985 [39] Beringer performed a study on human performance in the use of an infrared touch-input device. Recommendations include to design touch-based displays such that an area without touch-recognition is placed around a soft key to avoid accidental activation of other functionalities. Furthermore, touchscreens should be installed perpendicular to the user's line of sight to keep parallax errors at a minimum. Right- and left-handedness was also found to effect performance.

Cardosi and Murphy 1995 [40] As part of an extensive study on the design and evaluation of human factors in air traffic control systems Cardosi and Murphy recommend to avoid touch screens for tasks that require pilots to have their arms raised and unsupported for long periods of time.

Lewis 1993 [41] As part of an extensive literature review on touch-screen research Lewis identified the optimal touchscreen position to be at a 30° angle from the horizontal and no greather than 45° from horizontal, greater angles may induce significant fatigue in the user.

NASA 1995 [42] NASA's Man-System Integration Standards have set several recommendations for the use of touch. First, if a touchscreen is used in combination with other input devices, the user should not be required to switch frequently and if required it should not impose additional workload. Second, input feedback should be given within 100 milliseconds, either tactile, auditory or visual. Finally, touch-sensitive areas should be clearly indicated on the display.

Yeh 2003 [35] As part of an FAA report Yeh made several recommendations for the use of touch-screens on the flight deck. First, a soft key should only register an input is the user presses it and subsequently releases. If a button is pressed without releasing (e.g., by dragging your finger away before releasing the soft key) the resulting action should not activate. This recommendation is based on the lift-off touch logic introduced by Sears and Shneiderman in 1992 [43]. Second, if a touchscreen requires a minimum pressure to register input this amount should not lead to fatigue during prolonged periods of use. Third, soft key's on a display should be positioned such that a user will not obscure critical information when reaching out to press the button.

3-4 Relevant Concepts of Direct Manipulation and Touchscreen in Literature

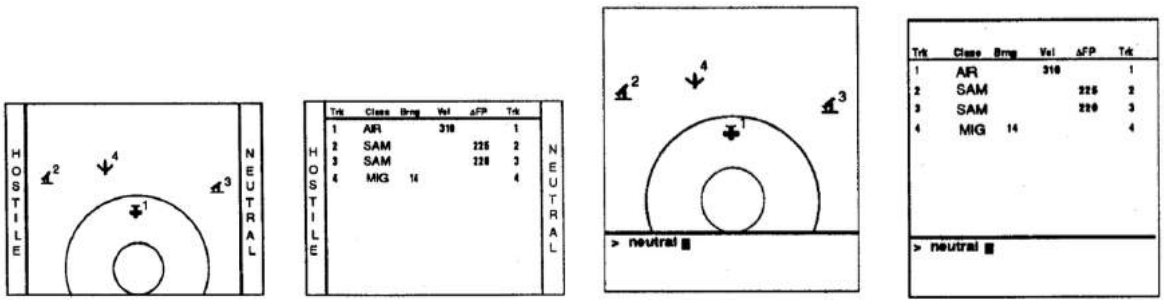
Following a review of Fitts' law, direct manipulation theory and touchscreen technology a literature review of relevant concept evaluations is presented in this section. Although the body of research regarding Fitts' law is extensive, literature discussing the direct application of the law on the flight deck is scarce. The majority of relevant literature discuss the use of direct manipulation and touchscreen on a more general level, of which the most interesting papers are now discussed.

Ballas 1992 et al. [17]

Ballas et al. performed an interesting study in 1992 *"evaluating two aspects of direct manipulation in advanced cockpits"*. Experimental subjects were given two tasks: track and identify surrounding aircraft and label them as hostile or neutral. To support this task four displays were developed, varying in both direct engagement and semantic distance, shown in Figure 3-13. A very interesting parallel can be made between the tabular displays Ballas developed and the CDU, as well as between the graphical display and a touch-based navigation display. Furthermore, the direct manipulation theory as presented in section 3-2 is evaluated.

Figure 3-13a is a graphical display which is touch-enabled, allowing for direct manipulation of the objects presented. This display consisted the lowest semantic distance and most direct engagement for the user in performing the task. A comparison can be made with a touch-based navigation display on the flight deck, allowing a crew to directly manipulate the flight route. Figure 3-13b is also touch-enabled however the information is presented in a tabular instead of graphical form, similar to the control display unit. Although it supports direct engagement, the semantic distance is much larger due to the textual representation of the model world. Figure 3-13c achieves lower semantic distance by presenting a graphical display, however input is given using keypad commands. This input method is comparable to the alphanumeric keyboard and scratchpad on the control display unit. Finally, Figure 3-13d has the most indirect engagement using both keypad commands and tabulary display, resulting in a large semantic distance. This is comparable to a CDU, but without the availability of graphical output on the ND.

Several parallels have been made between the displays in Ballas' experiment and the lateral navigation tools on the flight deck. However, it is important to note that the current CDU/ND combination is effectively a combination of both the displays in Figure 3-13c and Figure 3-13d since a pilot can use both the CDU and ND as output. In the experiment by Ballas subjects were not given both a graphical and tabular interface; only one of these interfaces were available.



(a) Graphical display with touchscreen input (direct manipulation)

(b) Tabular display with touchscreen input

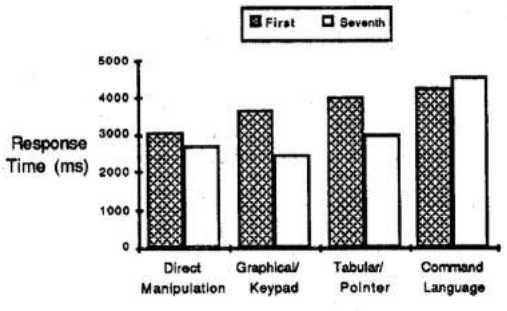
(c) Graphical display with keypad command input

(d) Tabular display with keypad command input

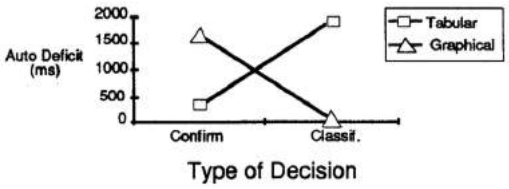
Figure 3-13: Four displays used in Ballas' evaluation of direct manipulation [17]

Results show that the direct manipulation interface (Figure 3-13a) performed best in terms of response time for the first runs of the experiment, see Figure 3-14a. This suggests that the interface is indeed the most intuitive and easy to use. However, after performing more trials,

which allowed subjects to get accustomed to the interfaces, the graphical display combined with the keypad input resulted in the best performance (Figure 3-13c). Furthermore, it was concluded that the difference in display performance is task-dependent, as illustrated Figure 3-14b. A tabular display provided better support in the tracking task, whilst the classification task was better performed using the graphical display. Ballas concluded that this effect can be attributed to the hypothesis that *"simple decisions [are] better served by the tabular display, [whilst] complex decisions [were better served] by the graphical display"* [17]. In ecological interface design, discussed in section 4-2, the goal is transform a *cognitive [and complex] task to a perceptual task* [44]. Ballas' conclusion seems to confirm the necessity of this goal, by showing that the graphical interface performs better for complex, cognition-intensive tasks. As for the simpler, tracking task subjects only needed to know the value of a single parameter, instead of the complete set of state variables. Ballas used this fact to describe the higher performance of the tabular display, which presented these parameters as values directly, for simple tasks.



(a) Results show that the direct manipulation performed best at first, and the graphical/keypad display better after several trials.



(b) Results show that the performance of a graphical or tabular display highly depends on the task to be performed.

Figure 3-14: Results of Ballas' evaluation of direct engagement and semantic distance on task response time [17]

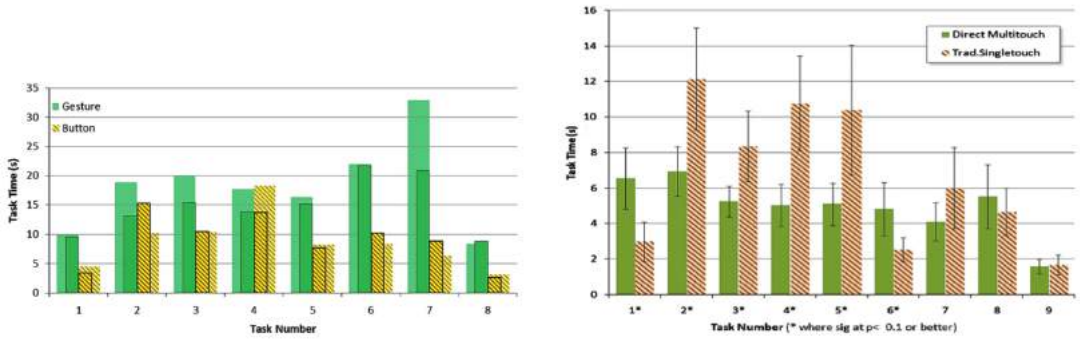
Summarising, Ballas found that a direct manipulation interface, with a graphical output and touch-based input, resulted in the quickest response times during the first trials, whilst the graphical output and keypad input display performed best after several trials. This shows that direct manipulation allows for an intuitive interaction which is easy to learn, however keypad command-line input seems to outperform touch-based input once the human has accustomed to it. Finally, Ballas also showed the task-dependency of interface performance, concluding that graphical displays performed best for complex tasks whilst tabular displays did so for simple tasks.

Bjørneseth 2012 et al. [18]

In his research, focused on a maritime application, Bjørneseth looked at the performance of direct manipulation in general and the difference between single- and multi-touch input gestures in particular. Marine human-machine interfaces have some interesting similarities to those found on a flight deck: both have very high safety and reliability requirements, suffer from dynamic environments (rough seas vs. air turbulence) and challenging and complex

tasks (e.g. docking a ship vs. landing an aircraft). Touchscreens are already widely used on the bridge of a ship, but often limited to a standard button-and-menu interface. Bjørneseth hypothesized that task execution time would decrease if multi-touch gestures were also possible. Subjects were assigned to performed eight different tasks, varying from zooming the display to surging and swaying the ship, using both direct multi-touch gestures and more traditional single-touch buttons.

The most interesting results of his research are presented in Figure 3-15. The measurements of execution times per task show that the *"traditional button interface, despite several shortcomings, unexpectedly performed better than the gesture alternative [for certain tasks], despite its careful design"* [18]. His final conclusion included that the performance of single- and multi-touch gestures highly depended on the specific task to be executed. This supports the task-dependency conclusion, already discovered by Ballas et al [17].



(a) Results on execution time per task (front: static condition, rear: motion)

(b) Results on execution time per task

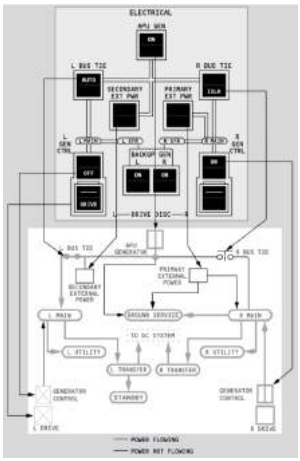
Figure 3-15: Results of Bjørneseth's evaluation of single and multi-touch gestures [18]

Degani 1992 et al. [14]

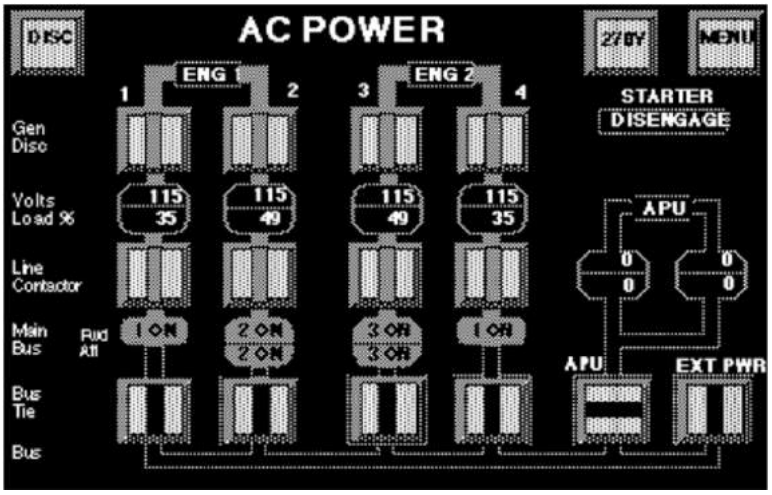
The research performed by Degani et al. in 1992 involved an evaluation of touch-based aircraft system display, in an attempt to evaluate the effectivess of direct manipulation. Modern flight decks have replaced complex flight engineer and system panels with graphical, intuitive system displays however still use traditional button and switch inputs on an overhead panel for system management [25]. The traditional, input/output seperated interface, is illustrated in Figure 3-16a and the touch-based display designed by Degani in Figure 3-16b.

Degani attempted to design a display incorporating a model-world representation of the task domain. His interface allowed users to directly engage with the system, in this case the electrical power system. For example, a user could simply push on one of the electrical lines to open or close the specific power supply. Furthermore, the value would display its actual position in a very intuitive graphical manner. Clickable areas were clearly visualized and aural feedback was given following succesful input recognition.

Although the visual representation of the systems on the display were highly appreciated, the touch-based inputs were less so. Degani summarised his results on the basis of a questionaire, of which the results are shown in Figure 3-17a and Figure 3-17b. The largest challenge was

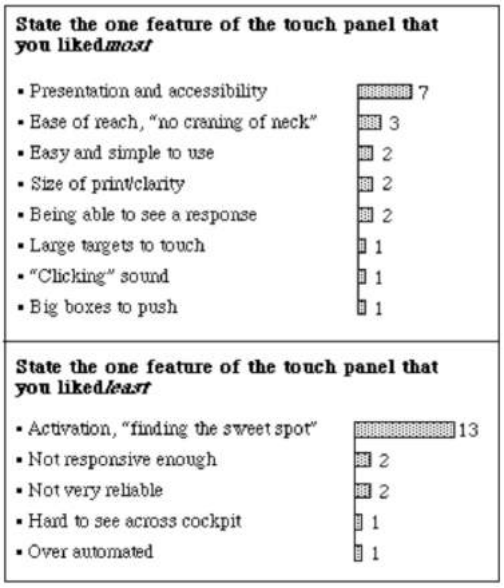


(a) System input and output separated on the Boeing 777: switches on the overhead panel vs. system status on a display [25]

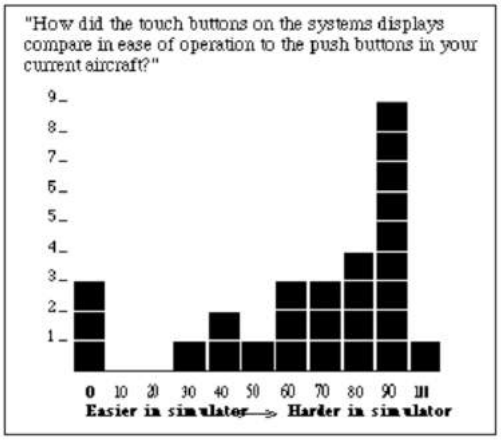


(b) Touch-based system display designed by Degani. Electrical power lines and generators can be managed directly on the display [14].

Figure 3-16: Comparison between traditional system management and the touch-based display designed by Degani.



(a) Survey results following a questionaire of most and least appreciated touchscreen features.



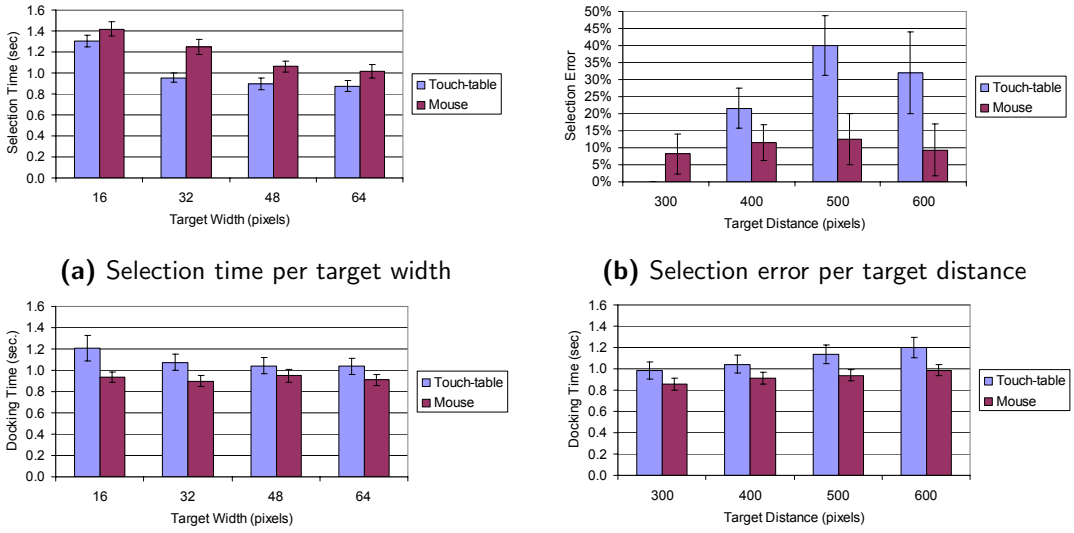
(b) Survey results for the question "How did the touch buttons on the system displays compare in ease of operation to the push buttons in your current aircraft?"

Figure 3-17: Survey conducted post evaluation of the touch-based system display to assess usability [14]

found to be *"finding the sweet spot"* in order to activate a specific soft key, and overal test subjects found the touchpanel more difficult to operate compared to the traditional switches they were used to.

Forlines 2007 et al. [19]

Forlines performed a study in 2007 comparing touch with mouse inputs for both unimanual and bimanual tasks during a simple task of selecting, dragging and docking an object to a target area. Independent variables consisted of the selection and docking time, as well as the selection error rate. Dependent variables included the target width and distance, both measured in pixels. The results are shown in Figure 3-18. For the target selection task, the direct touch input resulted in lower selection times however the indirect mouse input yielded a lower selection error rate. During the subsequent dragging and docking task, the mouse outperformed the touch input in terms of docking times for both variations in target width and distance. Forlines concludes that indirect mouse input is faster and more accurate for unimanual tasks. Furthermore, as illustrated by Figure 3-19, an interesting observation can be made about the difference between touch-input on a horizontal and vertical display. Due to a larger reach distance between object locations on horizontal displays the touched area increases in size. This effect is subsequently recommended to be taken into account in the design of touchscreen interfaces.



(a) Selection time per target width

(b) Selection error per target distance

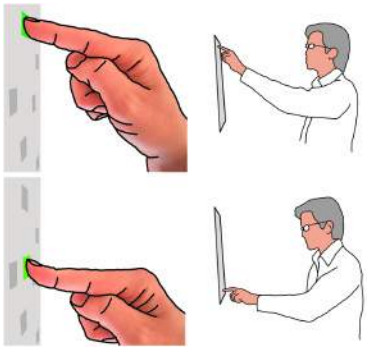
(c) Docking time per target width

(d) Docking time per target distance

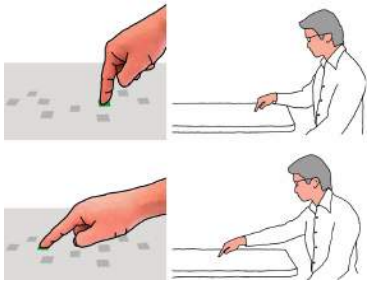
Figure 3-18: Results of Forlines' experiment comparing a touch-table and mouse for a selection, dragging and docking task of a simple square object [19].

Stanton 2013 et al. [20]

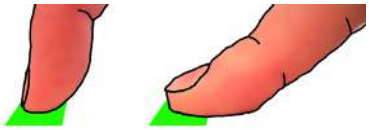
In Stanton's study four input devices were assessed on task execution time, error rate, workload, subjective usability and hand- and body discomfort. Two tasks were assessed, of which the second is of most interest to the scope of this thesis. Stanton let users perform an *"action*



(a) User selecting an object on a vertical touchscreen



(b) User selecting an object on a horizontal touchscreen



(c) Difference in touch area between vertical and horizontal touchscreen

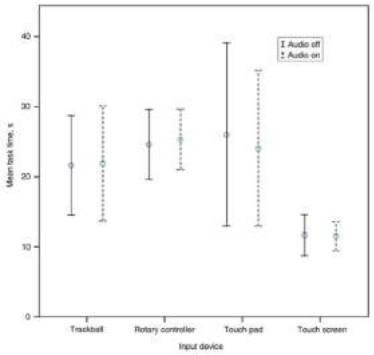
Figure 3-19: Interesting observation showing the difference between vertical and horizontal touchscreens. Touching vertical displays will be more likely to result in a similar touch area on various parts of the display than horizontal displays [19].

menu-based" task, which involved selection of a target on a screen, following by a selection of surrounding targets. The four input devices that were tested were a trackball, rotary controller, touchpad and touchscreen. Furthermore, the effect of auditory feedback was analysed.

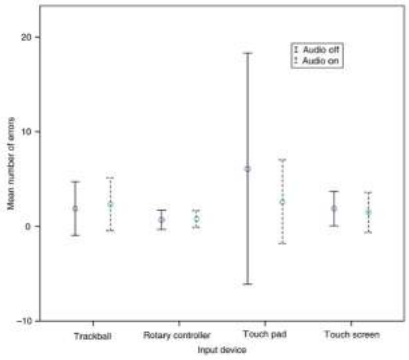
Measurements of mean task time and error rate are shown in Figure 3-20. The touchscreen was found to result in the shortest task execution time, and in addition with the lowest variance in the results. In terms of error rate, the touchscreen performed comparable to the trackball and rotary controller. Notably, the touchpad resulted in a very large variance in error rate, whilst the rotary controller performed best. However, the touchscreen performed the worst in terms of median body discomfort, as shown in Figure 3-20c. Finally, an hollistic comparison of independent variables is presented using a star diagram in Figure 3-20d. Based on all six parameters, the touchscreen performed best.

Stanton proposes the choice of an input device as a tradeoff between cognitive and physical performance. Direct devices enable an interaction that will be easier to understand however the absence of a device which translates physical movements from the user into system actions is likely to result in a higher physical burden on the operator. Furthermore, the assertion of Ballas [17] and Bjørneseth [18] that touchscreen performance is highly task dependent is confirmed by Stanton. It was found that touchscreens were superior for quick, alphanumeric inputs whereas rotary controllers outperformed the other input devices for scrolling and continuous adjustment tasks, such as setting the volume.

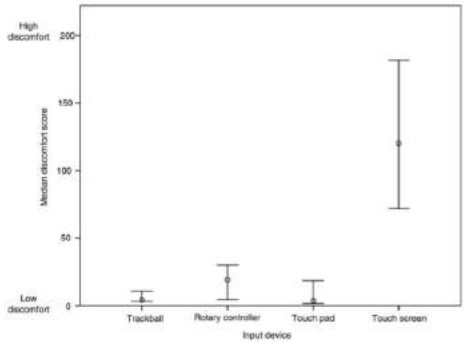
Finally, Stanton reflected on the fact that his subjects were not strapped into their seats during the experiment, and as a result a recommendation was made to take into account the additional difficulty of reaching out and operating a touchscreen when the pilot is strapped in his or her seat.



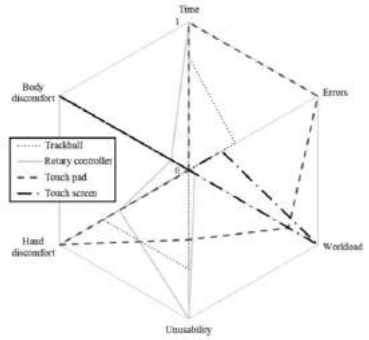
(a) Mean task time for each of the four input devices



(b) Mean number of errors for each of the four input devices



(c) Median body discomfort score for each of the four input devices



(d) Star diagram summarising the experimental results per input device. A small star, staying closer to the center represents a positive, good score

Figure 3-20: Results of Stanton's experiment comparing a trackball, rotary controller, touch pad and touch screen for an action menu-based task [20]

3-5 Summary

This chapter started with an introduction of Fitts' law, which was found to be very robust and widely appraised by HMI researchers. As a result, it will be used during this preliminary thesis as a quantitative evaluation of a touch-based navigation display, compared to the state-of-the-art. Furthermore, direct manipulation theory and its embodiment in touchscreen technology was introduced and concepts evaluations reviewed. The prevailing observation describes the effectiveness of touch-based input as being highly task dependent, whilst acknowledging the promise of increase intuitiveness of touchscreen technology. This further comfirms the necessity to evaluate touch-based input for the specific tasks encountered during enroute lateral navigation. The *directness* attributed to touchscreen technology, as described in section 3-2 and section 3-3, not only holds a promise to decrease task execution time and error rate. Being able to directly interface with the lateral flight plan on a navigation display could potentially improve the support for decision-making by the pilots during the task to be executed. This topic will subsequently be discussed in the following chapter.

Chapter 4

Next Generation Lateral Navigation

Evaluating Decision Making Support

A second goal of the next generation lateral navigation interface was defined during the introduction of this thesis as improving the support for decision making. In executing a specific task, pilots will use a multitude of resources available to them, and as a result the design of an interface can impact the ability to make quick and effective decisions. This chapter aims to gain an understanding on how humans think about processes and make decisions, as well as how the design of an interface can support to crew in doing so.

First, the often cited human performance and behaviour models by Rasmussen are described in section 4-1. His models help describe the different levels of human performance and behaviour, cognitive representations of the world and inherent decision processes. Subsequently, interfaces can be evaluated based on their adherence and support of human cognition based on these models. Ecological interface design (EID) is an example of a methodology that attempts to incorporate these methods with the goal to ease decision making by transforming a *"cognitive to a perceptual task"*. Although the scope of this thesis is not the design of a novel interface, the taxonomies introduced by Rasmussen and the principles behind EID can be used to perform a qualitative comparison between a touch-based navigation display and the state-of-the-art. It is therefore found to be relevant, and discussed in section 4-2. Finally, a few examples of the application of EID and Rasmussen's models have been found in literature, which are summarised in subsection 4-2-3.

4-1 Human Performance and Behaviour

In 1983 Jens Rasmussen published a human performance model which has since then formed the basis for further studies into human-machine interface design [45]. A proper discussion of HMI design can therefore not be held without an introduction to the model developed by Rasmussen, presented in this section.

Rasmussen starts with the realisation that humans are not simple deterministic input-output beings but are driven by goals and teological by nature [45]. Rosenbluth, Wiener and Bigelow define teological behaviour as behaviour that *"is modified during its course by signals from the goal"* [46] (cited in [45]). Teological behaviour is as much based on past experiences as feedback during its course. Polanyi expanded this concept by introducting the distinction between causes of physical events and reasons for physical functions [47]. Reasons can control functions by selection, e.g. experience from the past whilst cause can control a function through the physical function of a system [45]. Rasmussen states that human efficiency in handling complex situations is largely due to the availability of vast amounts of past experiences and different mental representations of the environment. Rasmussen developed two taxonomies in an attempt to describe human performance, behaviour and the mental representations of the environment he identified. The skill-, rule- and knowledge-based behaviour taxonomy is described in subsection 4-1-1, the abstraction hierarchy in subsection 4-1-2 and the decision ladder in subsection 4-1-3.

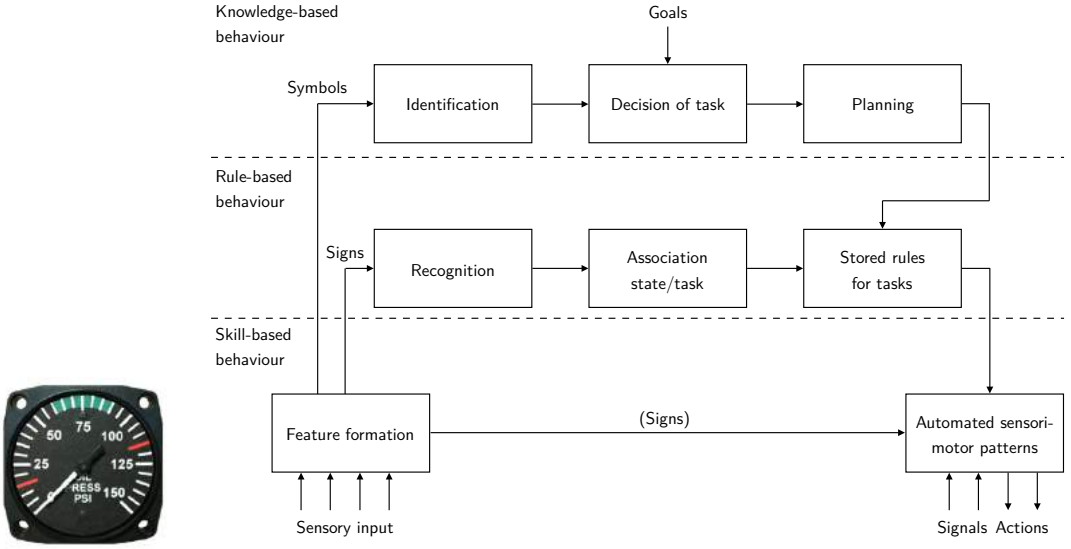
4-1-1 Skills, Rules and Knowledge

Rasmussen segments human performance into three levels: skill-, rule- and knowledge-based (SRK) performance as shown in Figure 4-1b. At the most basic level is skill-based behaviour (SBB). This level of behaviour represents sensory-motor performance and occurs without conscious control. Sensory inputs are immediately transformed into actions, e.g. riding a bicycle. At the next level, rule-based behaviour (RBB) is the execution of a stored rule of procedure. Based on sensory input, the human will recognise the situation and based on a set of stored rules devise a set of resulting actions. On the flight deck the execution of a checklist is a good example of RBB. At the highest level, and with the highest cognitive workload, is knowledge-based behaviour (KBB), used when there are no stored rules or procedures to solve a specific situation. KBB is associated with unfamiliar situations, and those where higher level goals must be considered to come to a solution. Examples are solving a puzzle or making a go/no-go decision for a flight [45]. KBB often relies on past experiences in the identification, decision of task and planning of a response to a certain input.

Signals, Signs and Symbols

All three levels of behaviour are dependent on how the human performer observes information in the surrounding environment. Rasmussen proposes three levels of information, signals, signs and symbols, each of which is linked to one of the three levels of human performance [45] and thus also presented in Figure 4-1b. Figure 4-1a shows a conventional oil pressure gauge and can be used to clarify the difference between signals, signs and symbols. Signals are used in SBB and will induce an automated response, as is reflected in Figure 4-1b. For example, a signal from the guage would be that the oil pressure is currently 0 psi. Signals have no significance or meaning, except as direct physical time-space data [45]. Signs offer more context, and are used in RBB as a trigger to initiate a procedure. In this example, the user will receive a sign that the oil pressure is outside of the 'green' normal operating area, trigging action to follow an engine check procedure. Finally, a symbol can express the true meaning of the information presented, and is consequently used in KBB. For example, should the engine be running properly a symbol can be that the instrument must be broken,

given that a properly running engine will have a positive oil pressure. The true understanding of a symbol will require an increased level of congitive effort to identify the true contextual meaning of the information.



(a) Signals, signs and symbols on a gauge[1]

(b) The three levels of human performance and behaviour as identified by Rasmussen [45]

Figure 4-1: Rasmussen's skill-, rule- and knowledge-based behaviour taxonomy on human performance and behaviour [45]

4-1-2 Abstraction Hierarchy

Where the SRK taxonomy is used to develop a representation of human performance and behaviour, the abstraction hierarchy (AH) as proposed by Rasmussen [45] can be used to develop representations of complex systems, work environments and domains [48]. The AH is constructed using several layers, each of which describes the system and work domain in terms of a unique set of attributes. Higher abstraction levels are used to describe the work domain in terms of its purpose and function. Lower abstraction levels describe the work domain in terms of the physical implementation [45]. In essence, the AH is a combination of different models at varying levels that together form a representation of the entire system [48].

Consider the abstraction hierarchy proposed by Borst to describe the work domain of aircraft terrain-avoidance in Figure 4-3. At the highest level, is the functional purpose of the system described as being the highest level system objective and desired state of the environment. In this example, productivy, efficiency and safety together from the final goal in the terrain-avoidance task. One level below is the abstract function, which describes the governing laws in the work domain [45]. In terrain-avoidance, aircraft energy management and physical separation govern the work environment [49]. Next, the generalised function, is a set of

[1]Photo: aircraftspruce.com

functions that describe the process, in essence the solution to the problem. Borst describes the generalised function as the set of aerodynamic forces, kinematics and dynamics [49], together they form the solution to avoiding terrain. At the bottom of the AH the physical function level describes the functions of the physical parts and hardware used to achieve the generalised function, and the physical form describes the actual location and appearance of these components [45], [49].

Rasmussen, Pejtersen and Goodstein in 1994 propose a *"why-what-how"* and *"means-end"* relationship to describe the relation between the different levels of the AH [50], see Figure 4-3. Returning to Figure 4-3, enter the work domain at the generalised function, e.g. maneuvering. This now defines the *"what"*. The level above, the abstract function, describes *"why"* we want to maneuver: *"because it serves the ends of energy management and separation"* [49]. Moving one level below, the physical function, describes *"how"* we will maneuver: by means of deflecting the control surfaces and adjusting engine thrust. Furthermore, in this example the physical function provides a means (control surfaces) to achieve the end (energy management and separation) [49], [50].

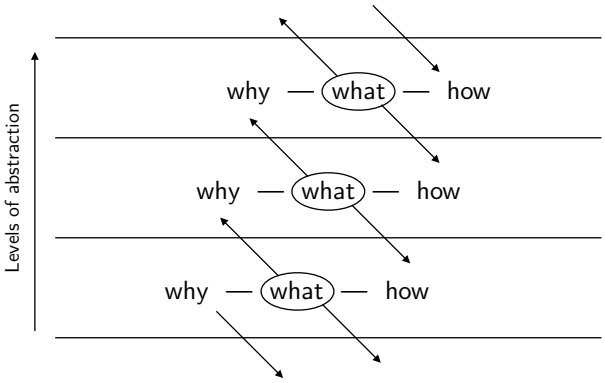


Figure 4-2: The *"why-what-how"* relationship can be used to describe the relationship between the different AH levels [50], [51].

4-1-3 Decision Ladder

Whilst the abstraction hierarchy was used by Rasmussen to describe and analyse the work domain, his decision ladder (DL) can be used to look specifically at the execution and analysis of a control task. Rasmussen describes the decision ladder as a *"normative model of the necessary decision phases ... for control of a physical system"* [52], and is presented in Figure 4-4.

A user starts along the decision ladder at the bottom left, by detecting the need for intervention. Transitioning to a state of alert, the user will look around to *"observe some important data in order to have direction for subsequent activities"* [52]. Subsequently, he or she will have to use the set of observations to form an understanding of the current state of the system, and evaluate the *"possible consequences with reference to the established operational goals"*. What follows is the most congition-intensive state, namely the evaluation and selection of possible actions that could result in satisfying the goals of the system, and achieving

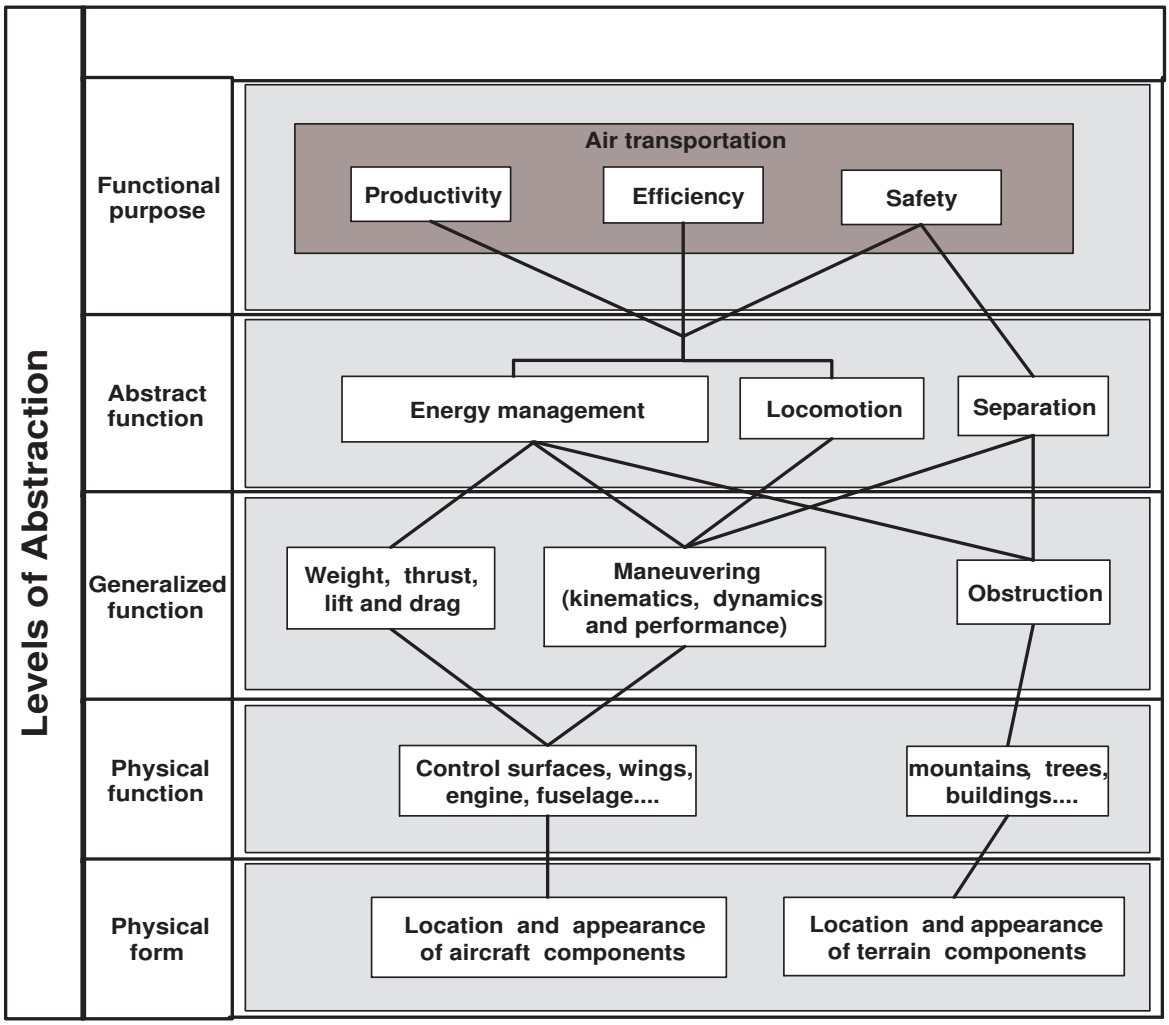


Figure 4-3: Example of the AH applied to terrain avoidance [49].

a specific target state. The operator now transitions from an analytical to a planning and execution phase, down the right side of the ladder. A set of tasks and strategy to acquire the new target state is subsequently defined, formulated and finally executed.

As is illustrated by Figure 4-4 the various steps along the decision ladder can be segmented into the aforementioned skill-, rule- and knowledge-based taxonomy. The relevancy and power of the decision ladder in evaluating an interface is by using it to identify possible shortcuts that can be made. For example, the use of a standardized checklist or procedure for an engine failure onboard an aircraft can be used to jump from the identification of the current state of the system (failed engine) to a procedure and sequence of actions to be executed (checklist). The congitive-intense knowledge-based domain is therefore skipped.

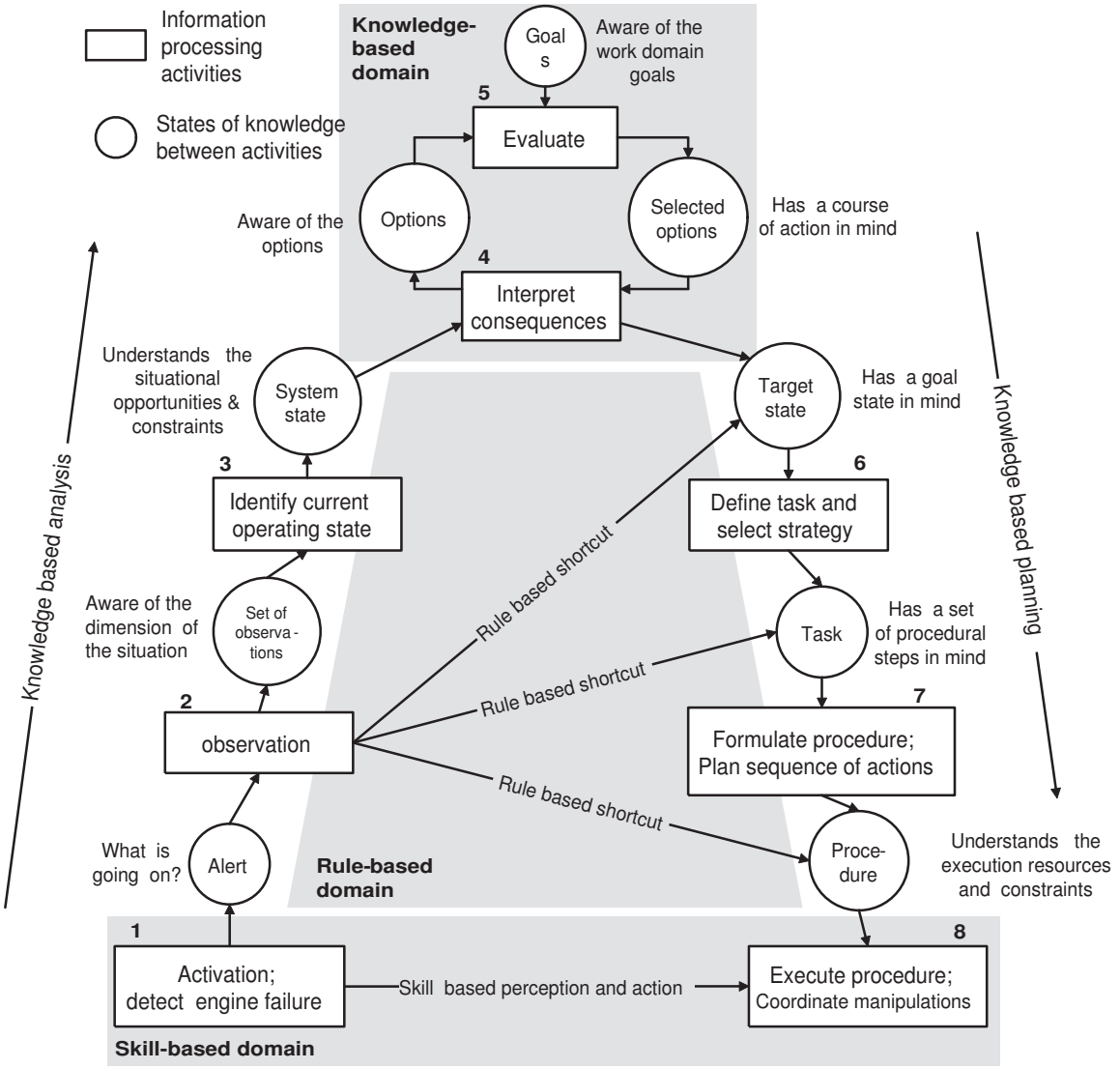
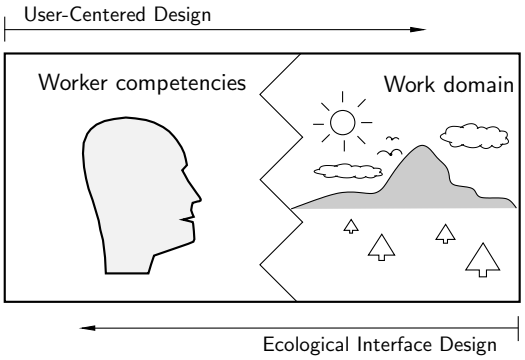


Figure 4-4: Decision Ladder [49].

4-2 Ecological Interface Design

Humans tend to view the world in a very functional manner, in terms of actual functional capabilities of the system [53]. For example, a pilot's attempt to determine the available glide range following an engine failure, will unlikely consist of a very detailed calculation of distances, vectors and angles. In fact, during basic flying training on a Cessna 172 a pilot is instructed to judge the available glide distance as being any object within approximately one-third of the wing strut, viewed from the pilot's seat (see Figure 4-5b). Ecological Interface Design (EID) uses this as a starting point, and approaches the human-machine interface from the work domain, as illustrated by Borst in Figure 4-5a. This makes the approach differ from the more traditional user-centered approach, which designs the interface based on worker abilities and competencies. Ecological interfaces map the functional goals and states,

and displays information that is actually meaningful to pilot. It attempts to *"[transform] a cognitive task to a perceptual task"* [44]. A review of EID serves to better understand the interface between the pilot and the flight management system, not only from an input but also from an output perspective.



(a) EID uses the work domain as starting point, as opposed to a more traditional user-centered design starting with the worker competencies [49].

(b) Example of the work domain as starting point: a Cessna 172 pilot using a one-third point along the wing strut as a reference for the available glide range following an engine failure[2]

Figure 4-5: Starting point of Ecological Interface Design

The ecological approach was first introduced by Rasmussen and Vicente [54], and their aim was to facilitate an effective cooperation between man and machine by developing an interface that reflects the underlying principes of the work domain in ways that support skill-, rule-, and knowledge-based behaviour. According to Borst et. al. a method to achieve this goal is through metaphorical design and direct manipulation [55]. Figure 4-6 shows the relative position of the interface that is strived for in the ecological approach. Furthermore, EID partly stems from the realisation the humans, especially in the aerospace domain, maintain an essential role in high-tech systems, given their ability to adapt and use their creativity during ordinary and unanticipated events, thereby increasing the resilience of the system [55]. Captain Sullenberger's creative, heroic and successful landing of his Airbus A320 on the Hudson River following a twin engine failure at low altitude highlights the unique abilities of the human controller [56]. Ecological interface design places an emphasis on finding ways technology can facilitate human creativity and adaptability in coping with unanticipated events instead of replacing weaknesses with more automation [55]. Figure 4-6 illustrates the interaction of an ecological interface with the human, machine and governing ecology.

4-2-1 Two Goals of Ecological Interface Design

Summarising, there are two goals that are strived for in ecological design. First, the interface must aim to transform cognitive tasks into perceptual ones in order to reduce mental workload for the user. Figure 4-7 presents a separation assurance overlay added to the navigation display which provides a very graphical representation of the airspeed and heading affordances in order to ensure safe separation [57]. Here, a pilot is relieved of the cognitive task of calculating a safe airspeed and heading and in turn is presented with a more intuitive solution space representation that allows the pilot to immediately percieve safe airspeed and headings. The

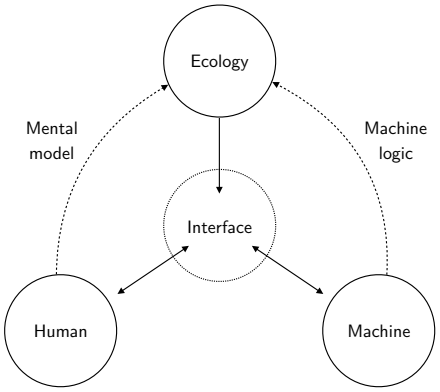


Figure 4-6: Ecological interface design embodies a triadic approach, including the human, machine and governing ecology in the cognitive system [55].

second goal of EID is to foster problem-solving, creatitivity and adaptability by supporting skill-, rule-, and knowledge-based behaviour. Dinadis offers the following methodology to achieve this goal [58]. SBB can be fostered by allowing a user to directly manipulate the display and its objects. RBB can be supported by providing a consistent one-to-one mapping between the work domain and the signs provided by the interface. Finally, Dinadis proposes a work domain representation in the form of an abstraction hierarchy to support KBB. This representation will serve as an externalised mental model designed to help problem-solving [58].

4-2-2 EID Relevance to Flight Deck Design

The current prevailing operating philosophy within aviation seems to be very rule- and procedure-based. Airline crews are expected to diligently execute steps to satisfy mission goals and rarely dispose of a thorough understanding of underlying systems [58]. Airlines emphasize *"button-pushing rather than knowledge of systems"* [59]. However, partially also triggered by the Hudson landing of US Airways flight 1549 and crash of Air France flight 447 [56], [60] a strong argument is made for a more human-centric flight deck. Dinadis argues that EID is a very effective methodology to bridge the gap towards a flight deck that fully supports crews in problem-solving, creativity and adaptability during flight [58].

4-2-3 Using Ecological Interface Design to Support Decision Making

Research conducted by Borst and Van Marwijk et al. present relevant examples of the use of Rasmussen's models and EID to proactively support decision making. Hence, they will briefly be reviewed and discussion in this subsection.

Borst 2009 [49]

Borst looked at using Ecological Interface Design in designing an interface for terrain awareness and avoidance. A decision ladder, illustrated in Figure 4-8, was used to compare two

[2]Photographer: N. van Zon

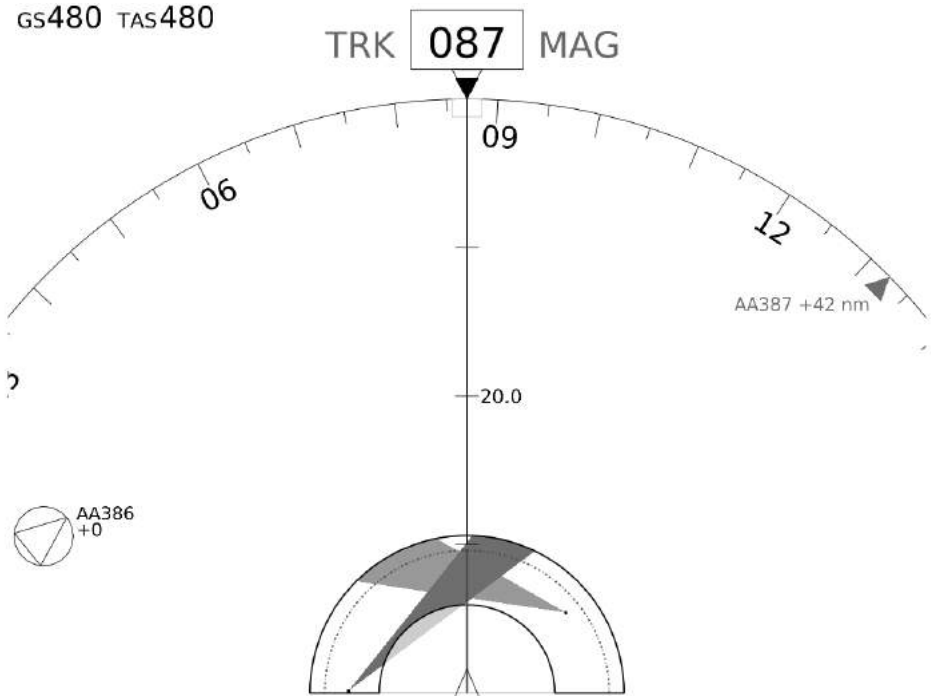


Figure 4-7: A separation assurance overlay is added to the navigation display, which provides a functional presentation of the airspeed and heading affordances of the aircraft in order to ensure separation [57].

terrain awareness and avoidance systems. The first, an Ecological Synthetic Vision Display (ESVD), *"allows pilots to detect a terrain collision ahead and the current state of risk, understand the urgency of the risk relative to the current operating state of the aircraft and the required state as dictated by the terrain, and predict and evaluate the escape options to circumvent the hazard by means of the perceptual cues"*. The support for knowledge-based analysis using the ESVD is shown in Figure 4-8 by the grew areas. As shown, the planning of tasks and strategies are not supported by the ESVD. Therefore, Borst concludes that the display is *"primarily suited for skilled pilots who know how to execute optimal escape maneuvers instead of novice pilots"* [49]. The second display, a Synthetic Vision Ground Collision Avoidance System (SVGCAS), automates the knowledge-based level and instantly offers a rule-based shortcut *"from the observation of the alert to the target state in whicht he aircraft should be steered"* [49]. Instead of showing the possible solution space, the SVGCAS determines the most optimal solution and commands the necessary response to avoid the terrain. Borst notes that this will most likely result in a decrease of situational awareness of the pilot, but an increase in pilot decision-making as compared to the ESVD.

Van Marwijk 2011 et al. [6]

The research by Van Marwijk is directly linked to the challenges identified and described in chapter 2. Van Marwijk proposes to use direct manipulation of the flight plan during airborne trajectory revision, in combination with ecological interface design as a solution to growing

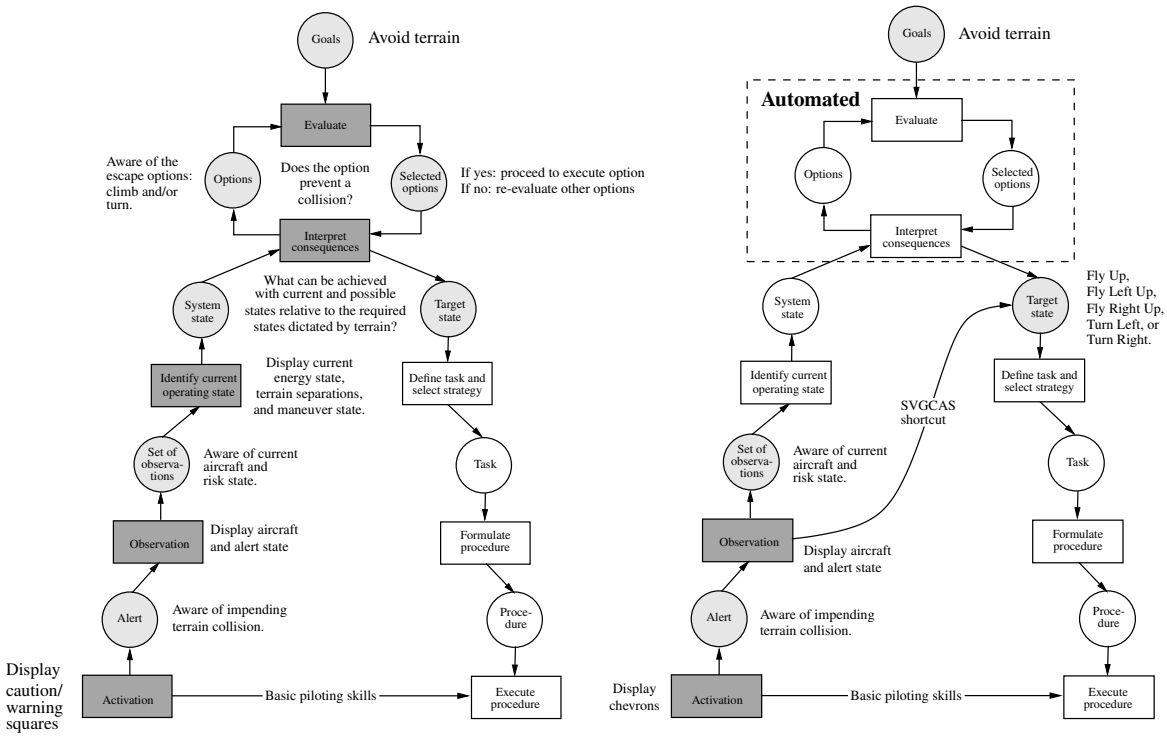


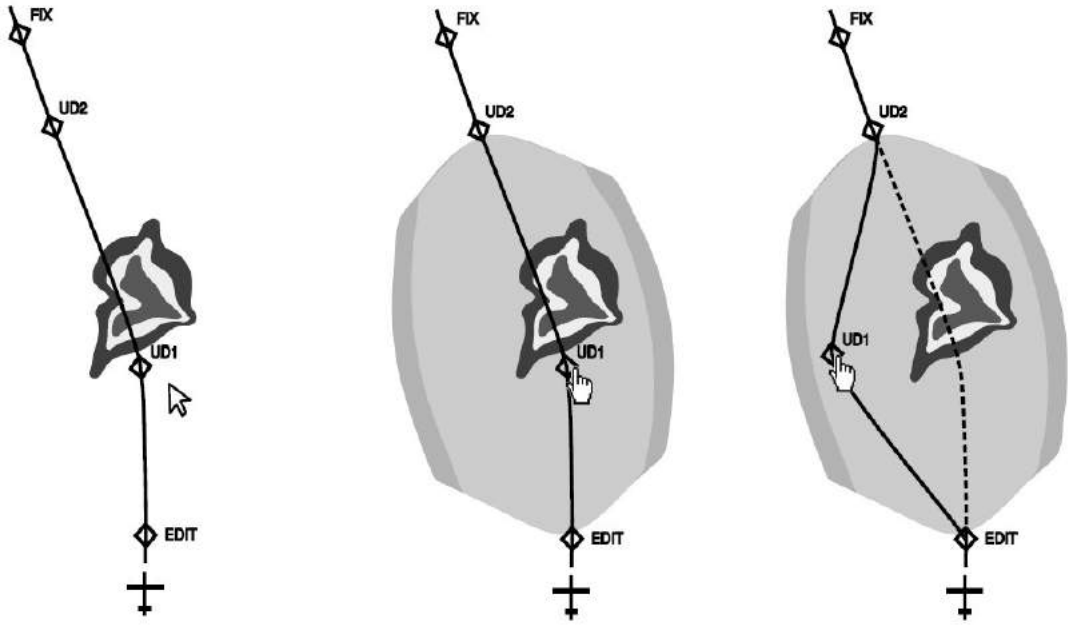
Figure 4-8: Application of a Decision Ladder to compare decision-making using an Ecological Synthetic Vision Display (EVSD, left) and a Synthetic Vision Ground Collision Avoidance System (SVGCAS, right). The former supports knowledge-based problem solving by showing affordance zones for safe terrain avoidance. The latter automates this step and provides a shortcut by directly commanding a response strategy to the pilot. [49].

demands on lateral navigation performance, such as 4D trajectory based operations.

Van Marwijk describes the navigation interface using the abstraction hierarchy, by identifying that *"existing navigation and planning interfaces enable pilots to perceive the flight plan as a sequence of altitudes, speeds, waypoints, and other navigation-related information (generalised function level), but do not explicitly represent what effects a modification of the flight plan would have on higher lever goals, such as efficiency and punctuality (functional purpose level)"* [6]. In his design, these higher level goals are displayed on the lateral and vertical navigation displays (see Figure 4-9a and Figure 4-9b) as coloured affordance zones.

Secondly, van Marwijk effectively uses the decision ladder methodology to determine which tasks should be allocated to either the automation or the pilot. In depth calculation of the impact of a re-route on fuel efficiency and time punctuality are left to the automation, whereas the selection of an alternate route is left to the pilot. Third, van Marwijk also uses the skill-, rule- and knowledge-based behaviour taxonomy to perform a worker compentency analysis to further map the cognitive workload of the interface for the user. The use of direct manipulation was identified as supporting skill-based behaviour due to its intuitive nature. The projection of affordance zones directly on the navigation display helps transfer the task of choosing an alternate route from knowledge- to rule-based behaviour. Van Marwijk hypothesizes that the use of direct route manipulation instead of the more traditional CDU keypad entry will help reduce cognitive load by translating the task from RBB to SBB.

Concluding, the cognitive task has effectively been translated to a perceptual one, adhering to the basic principles of ecological interface design.

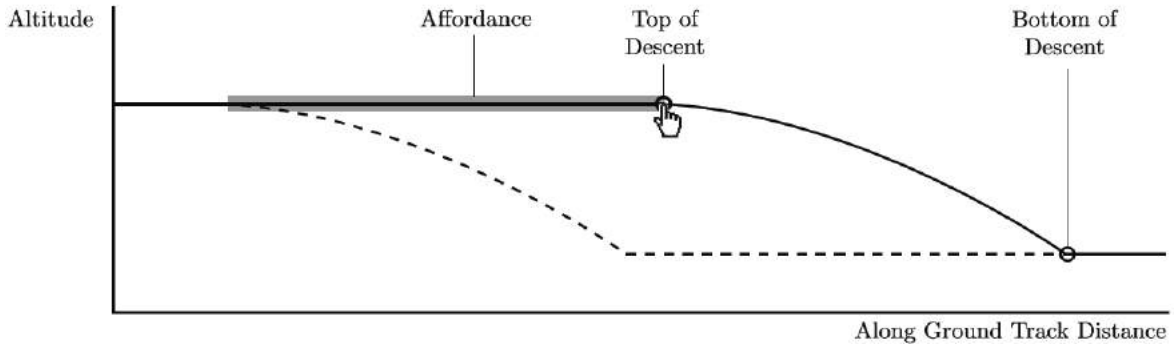


Original trajectory, obstructed by weather cell.

Upon selection of UD1, its affordances are highlighted.

UD1 is replaced within its affordance zone.

(a) van Marwijk's proposed lateral navigation interface showing affordance zones of higher level goals such as efficiency and punctuality



(b) van Marwijk's proposed vertical navigation interface showing affordance zones of higher level goals such as efficiency and punctuality

Figure 4-9: Prototype developed of an EID-inspired navigation display [6].

Furthermore, van Marwijk identifies several internal and external constraints that are of interest when considering a re-planning task.

- **Flight envelope:** Re-planning is limited by aircraft speed limits, which in turn impact the possibilities for adhering to a time constraint at a certain waypoint. Required time of arrival constraints of next generation ATM systems will mandate a specific ground

speed, which in combination with high altitude and strong winds may be difficult to comply with.

- **Aircraft dynamics:** If a re-plan includes rapid changes in speed, heading or altitude the dynamics of the aircraft (maximum possible acceleration, bank angle) can limit the possibilities.
- **Fuel:** Flying very length deviations can be limited by the amount of fuel available. International regulations stipulate the requirement for 45 minutes of reserve fuel, hence deviations in excess of three quarters of an hour will push limits on fuel reserves.
- **Obstructions:** The presence of other traffic, weather systems, terrain, prohibited airspace will limit possibilities for a re-plan.
- **Operational regulations:** Several regulations can limit the possibilities for an effective re-route, such as maximum speeds, altitude or very limited airspace corridors such as overhead Afghanistan.
- **Required time of arrival:** Especially in next generation ATM systems pilots' will see increased requirements on required time of arrival at specific metering fixes, such as the initial approach fix or the north atlantic track starting fix.

4-3 Summary

This chapter has discussed to use of the skill-, rule- and knowledge-based taxonomy together with an abstraction hierarchy and decision ladder to qualitatively evaluate the decision-making support provided by a specific interface. Ecological interface design was introduced and discussed, given its goal to embody the aforementioned models and active support of pilot cognition. Although the design of a display or interface falls beyond the scope of this thesis, the taxonomies and models discussed can be used to compare the touch-based navigation display with the state-of-the-art. Whilst both interfaces will include an identical navigation display, the possibility to directly manipulate with the route could potentially improve the decision-making of the crew during a lateral weather avoidance task. Furthermore, touch-screen interaction with the flight management system could help achieve the triadic approach embodied by ecological interface design (see Figure 4-6).

Part IV

Appendices

Appendix A

Movement Time Predictions / Simulation using Fitts' Law Models

During the research described in part one, Fitts' law models developed and discussed in part two were used to simulate each weather avoidance scenario. The results of these simulations were predictions of movement times (MT) and an index of difficulty (ID) for each individual task. The calculation is further elaborated upon in this appendix, and presented in Figure A-1.

For the conventional interface, respective necessary actions are listed. Subsequently, the necessary movement amplitude and width were determined and used to calculate ID and MT. For a MCP movement each scenario was simulated to determine required heading changes and their respective tolerances. For CDU movements a technical drawing (see Figure A-2) of the device was used to determine movement amplitudes and key widths. An assumption was made that 1000 milliseconds were needed to depress a single button on the MCP, for instance when activating the heading select mode. For CDU actions the entire sequence of keys is displayed, of which the total ID and MT is presented in a single row, in Figure A-1.

For the touchscreen interface, 1000 milliseconds are counted for initiating a route adjustment given that this was the time delay programmed in selecting waypoints to start a route modification. For each waypoint adjustment the initial (x_i, y_i) and final (x_f, y_f) waypoint coordinates was determined in pixels and converted to mm using the published pixel density of the display used. Based on these coordinates, a linear movement was assumed, resulting in the target amplitude. Subsequently the tolerance, or target width, in placing each waypoint without violating turn radius capabilities of the aircraft was determined by trial-and-error and close scrutiny of the LNAV algorithm. Combined, the MT and ID was derived for each individual movement. Finally, 1000 milliseconds was assumed for depressing the soft `EXEC` key on the display. Note the same value of 1000 ms used both depressing `EXEC` on the TND, or one of the MCP keys.

Fitts Law prediction of MT and ID

Easy scenario
Conventional interface

Action	A [deg]	W [deg]	ID [-]	MT [ms]
Depress HDG SEL	-	-	-	1,000.00
HDG left 7 degrees	7.00	4.00	1.46	877.42
HDG right 22 degrees	22.00	8.00	1.91	1,098.91
CDU <LS3> <LS1> <EXEC>	-	-	6.22	1,246.50
Depress LNAV	-	-	-	1,000.00
		Sum	**9.59**	**5,222.83**
		Mean	*1.92*	*1,044.57*

Easy scenario
Touchscreen interface

Action	xi [px]	yi [px]	xf [px]	yf [px]	A [mm]	W [mm]	ID [-]	MT [ms]
Activate route adjustment	-	-	-	-	-	-	-	1,000.00
Adjust waypoint	516	357	471	260	28.34	12.00	5.46	2,857.42
Depress EXEC	-	-	-	-	-	-	-	1,000.00
						Sum	**5.46**	**4,857.42**
						Mean	*1.82*	*1,619.14*

Medium scenario
Conventional interface

Action	A [deg]	W [deg]	ID [-]	MT [ms]
Depress HDG SEL	-	-	-	1,000.00
HDG left 7 degrees	7.00	2.00	2.17	1,229.11
HDG right 47 degrees	47.00	2.00	4.61	2,439.28
HDG right 7 degrees	7.00	2.00	2.17	1,229.11
HDG left 80 degrees	80.00	8.00	3.46	1,867.42
CDU <LS4> <LS1> <EXEC>	-	-	6.70	1,290.30
Depress LNAV	-	-	-	1,000.00
		Sum	**19.12**	**10,055.23**
		Mean	*2.73*	*1,436.46*

Medium scenario
Touchscreen interface

Action	xi [px]	yi [px]	xf [px]	yf [px]	A [mm]	W [mm]	ID [-]	MT [ms]
Activate route adjustment	-	-	-	-	-	-	-	1,000.00
Adjust waypoint	538	480	510	411	19.73	10.00	5.78	3,014.17
Adjust waypoint	586	212	679	263	28.11	10.00	5.90	3,074.13
Adjust waypoint	521	130	553	130	8.48	10.00	5.73	2,991.67
Depress EXEC	-	-	-	-	-	-	-	1,000.00
						Sum	**17.40**	**11,079.97**
						Mean	*3.48*	*2,215.99*

Difficult scenario
Conventional interface

Action	A [deg]	W [deg]	ID [-]	MT [ms]
Depress HDG SEL	-	-	-	1,000.00
HDG left 3 degrees	3.00	1.00	2.00	1,145.00
HDG right 26 degrees	26.00	1.00	4.75	2,508.67
HDG left 131 degrees	131.00	2.00	6.06	3,152.36
HDG right 100 degrees	100.00	4.00	4.70	2,481.72
HDG right 50 degrees	50.00	4.00	3.75	2,013.67
HDG right 5 degrees	5.00	4.00	1.17	734.11
HDG left 35 degrees	35.00	8.00	2.43	1,356.00
CDU <NP> <LS2> <PP> <LS1> <EXEC>	-	-	16.60	2,873.60
Depress LNAV	-	-	-	1,000.00
		Sum	**41.46**	**18,265.14**
		Mean	*4.15*	*1,826.51*

Difficult scenario
Touchscreen interface

Action	xi [px]	yi [px]	xf [px]	yf [px]	A [mm]	W [mm]	ID [-]	MT [ms]
Activate route adjustment	-	-	-	-	-	-	-	1,000.00
Adjust waypoint	517	390	500	254	36.32	4.00	7.03	3,632.44
Depress EXEC	-	-	-	-	-	-	-	1,000.00
Adjust waypoint	545	163	586	188	12.73	8.00	6.11	3,180.01
Adjust waypoint	534	155	110	313	119.91	4.00	7.07	3,655.37
Adjust waypoint	585	315	609	247	19.11	4.00	7.20	3,720.05
Adjust waypoint	107	318	77	386	19.70	4.00	4.79	2,528.24
Adjust waypoint	521	434	509	474	11.07	4.00	7.04	3,637.91
Adjust waypoint	618	259	615	273	3.79	4.00	7.28	3,758.98
Depress EXEC	-	-	-	-	-	-	-	1,000.00
Adjust waypoint	599	318	541	335	16.02	12.50	5.61	2,933.11
Depress EXEC	-	-	-	-	-	-	-	1,000.00
						Sum	**52.13**	**31,046.12**
						Mean	*4.34*	*2,587.18*

Figure A-1: Movement time predictions and simulation of each scenario using the Fitts' law models

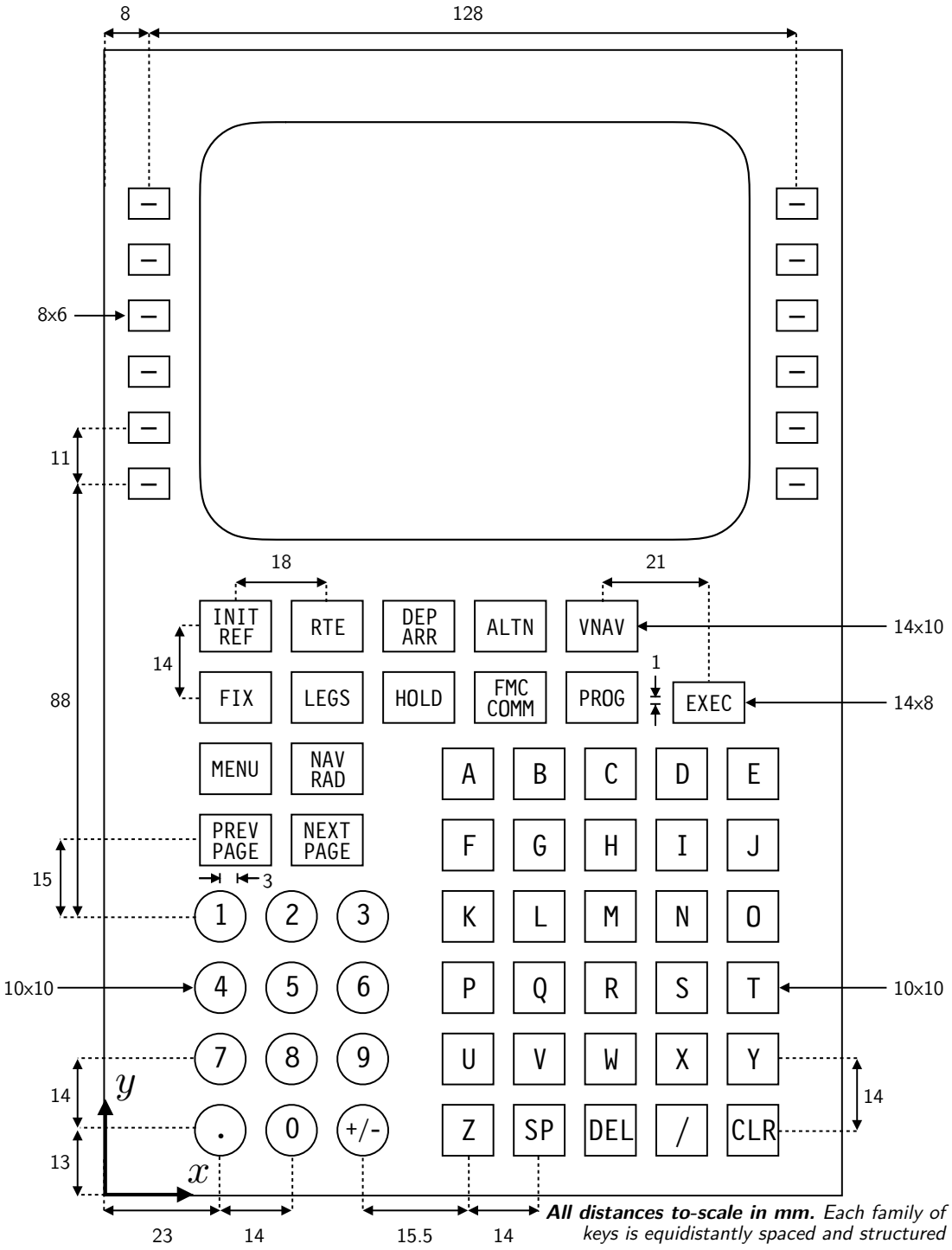


Figure A-2: Technical drawing of the Boeing 777 CDU showing scaled distances in mm, used to measure amplitude A and W for each movement.

Appendix B

Illustration of a Lateral Weather Avoidance Route

As an illustration of both the weather avoidance task and the design of the interfaces the procedure for a straightforward weather scenario is demonstrated using both the TND and MCP/CDU interface set-up. Note that the weather scenario displayed here is different to those discussed in the IEEE article. The illustration in this appendix serves as a general demonstration.

Scenario using the TND Figure B-1 provides an illustration of the weather avoidance procedure using the TND. First, the `RUDOT` waypoint is selected and moved towards the left, snapping automatically to the `IDARO` waypoint when in the vicinity (see ❶). Subsequently, `EXEC` is pushed the execute this command and adjust the flightplan. This is reflected on the navigation display with a magenta line (see ❷). The resulting flightplan, via `BITBU`, is clearly not very efficient. Hence, the `BITBU` waypoint can be selected and moved along the track from `IDARO` to `IDOSA` (see ❸). This is a custom waypoint, hence the naming `BITBU01`. The `EXEC` button is pressed to confirm the route adjustment, and subsequently this is reflected on the navigation display (see ❹).

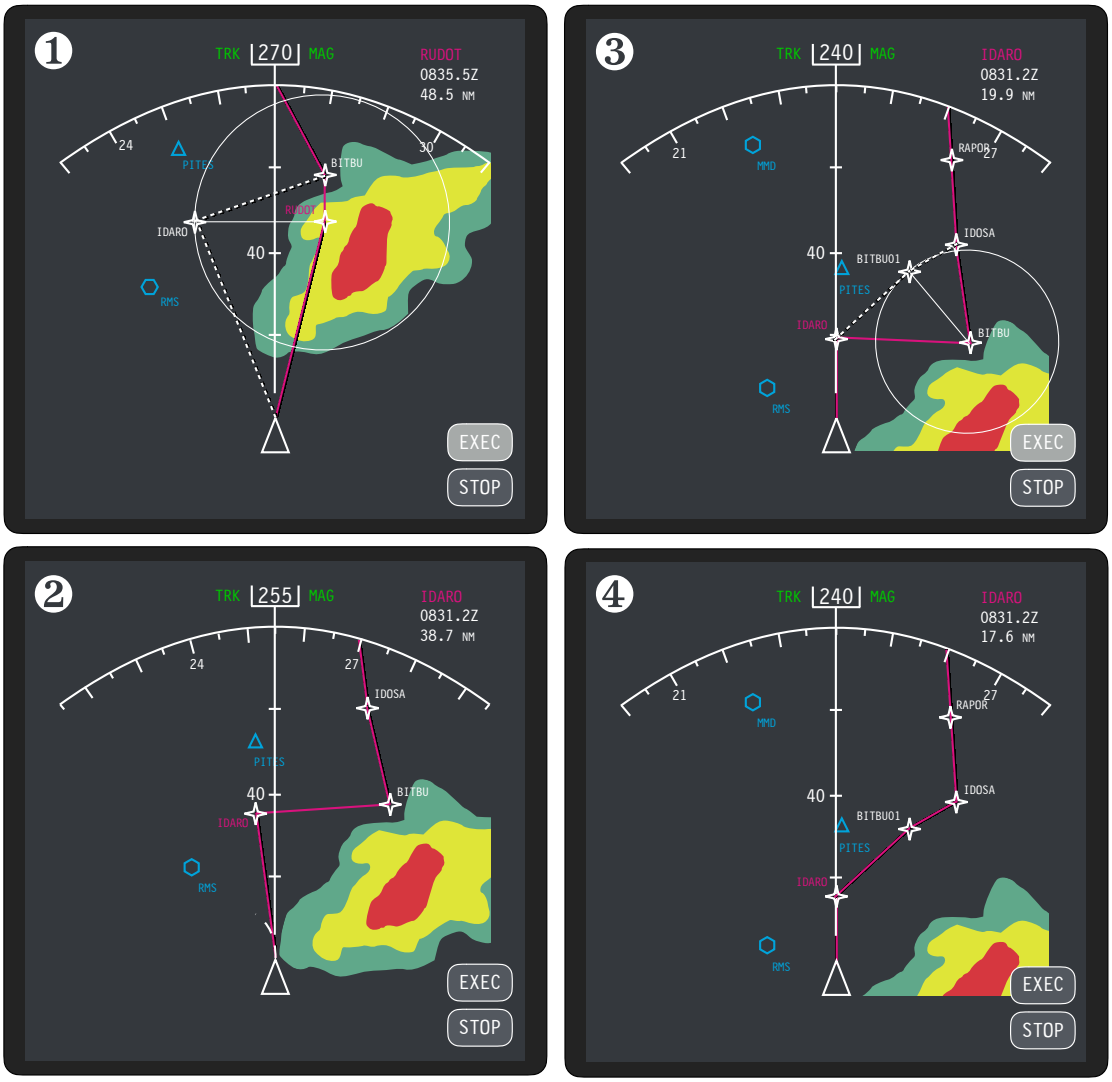


Figure B-1: Example using a TND to circumnavigate weather.

Easy Scenario using the MCP and CDU Figure B-2 and Figure B-3 provide an illustration of the weather avoidance procedure using the MCP and CDU. First, the heading select mode is activated on the MCP by pressing the `HDG SEL` button (see ❶). The heading bug on the ND activates and is rotated left fifteen degrees to 255° due North. The autopilot turns the aircraft left to establish the commanded heading (see ❷).

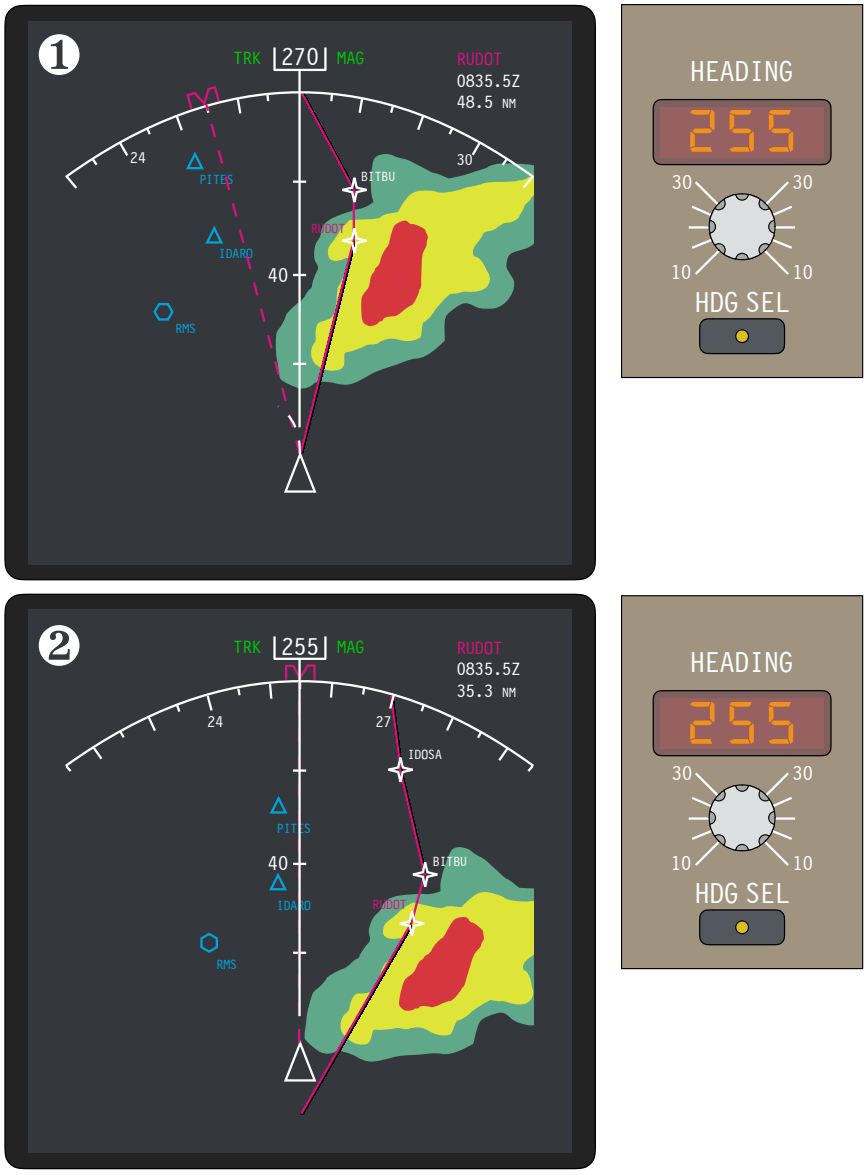


Figure B-2: Example using the MCP and CDU to circumnavigate weather(part 1).

Following safe passage of the weather system a sensible waypoint is chosen to which a direct route is desired, in this case the IDOSA waypoint (see ❸). Using the CDU, the respective adjacent line select key on the RTE page (LS3) is selected followed by the LS1 key to move the desired waypoint to the top of the list. The route adjustment is reflected on the navigation display and activated following a press of the EXEC key (see ❹). The direct-to command is executed, removing waypoints prior to IDOSA and commanding the autoflight system to navigate towards it.

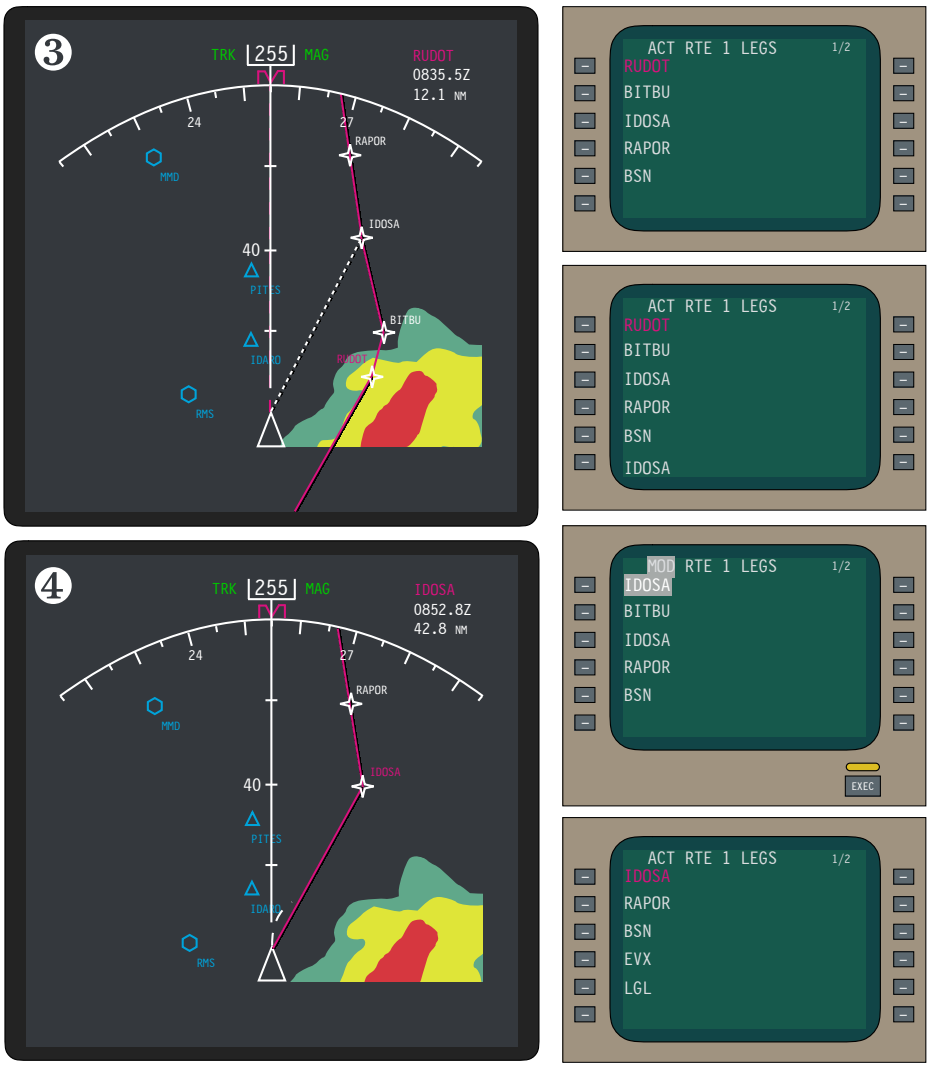


Figure B-3: Example using the MCP and CDU to circumnavigate weather (part 2).

Easy Scenario using only the CDU Note that an alternative procedure using only the CDU is also possible. The IDARO waypoint can be inserted as the first waypoint in the flightplan, commanding the autopilot to fly directly towards is. Next, the RUDOT and BITBU waypoints can be deleted such that following IDARO the aircraft continues to IDOSA and the remainder of the original flight plan. The resulting route will be similar to that achieved using the TND.

Appendix C

Additional Results for the Lateral Weather Avoidance Experiment

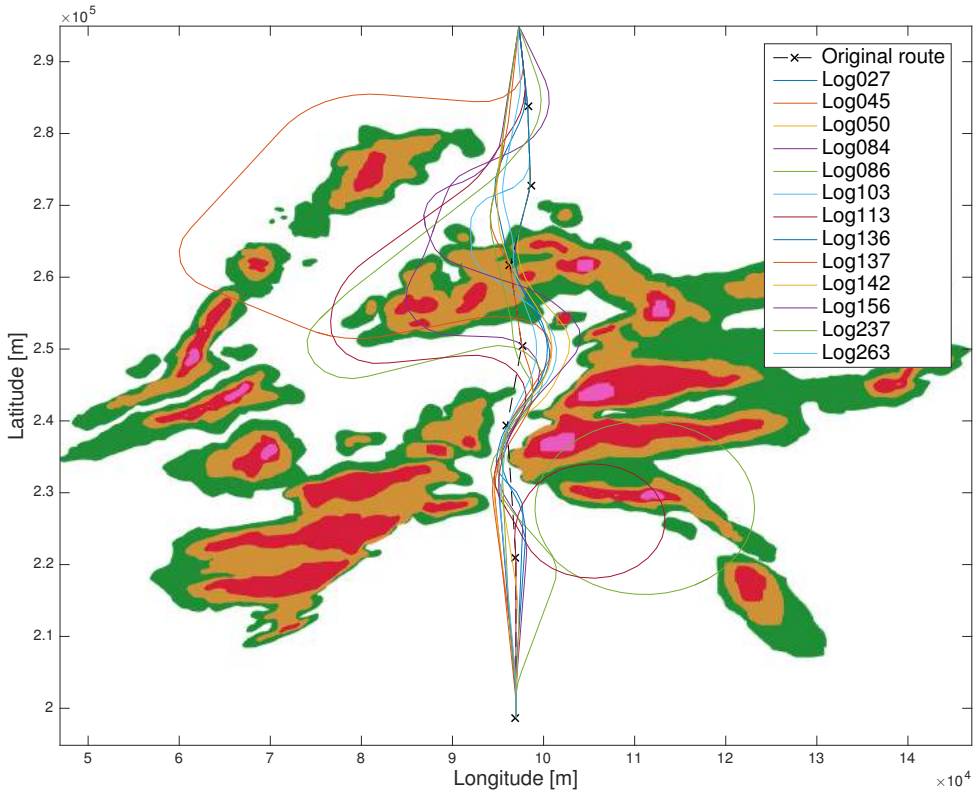
Additional results, omitted from the IEEE article, are presented in this appendix.

C-1 Outliers in Re-Route Time Delay

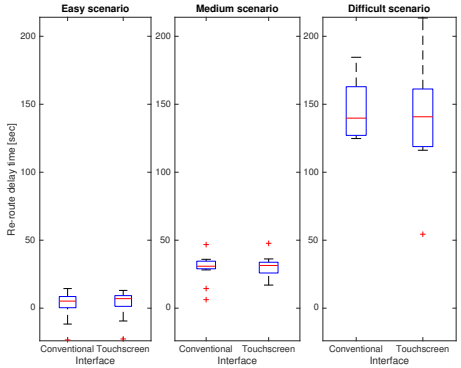
As discussed in the IEEE article, a few participants were not able to successfully avoid weather using the touchscreen during the difficult scenario. This led to meaningless re-route time delays. The flown tracks of these outliers are shown, in addition to the re-route time delay, corrected for between-subject variability, with and without the outliers. Note that the statistical results with and without the outliers remained the same. All of these outliers occurred using the touchscreen interface during the difficult scenario. A total of thirteen measurements were removed for the touchscreen interface, difficult scenario condition in the calculation of re-route delay time results. This is illustrated in Table C-1.

Table C-1: Sample size per experimental condition

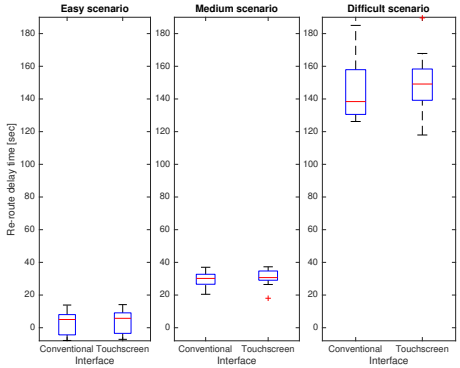
Interface	Task difficulty	Sample size	Measurements for re-route delay time	Measures for remaining variables
Conventional	Easy	12	43	43
	Medium	12	42	42
	Difficult	12	42	42
Touchscreen	Easy	12	46	46
	Medium	12	45	45
	Difficult	12	35	48



(a) Flown tracks which led to the outliers (all occurred using the touchscreen during the difficult scenario)



(b) Corrected for between-subject variability, showing outliers (see touchscreen, difficult scenario)

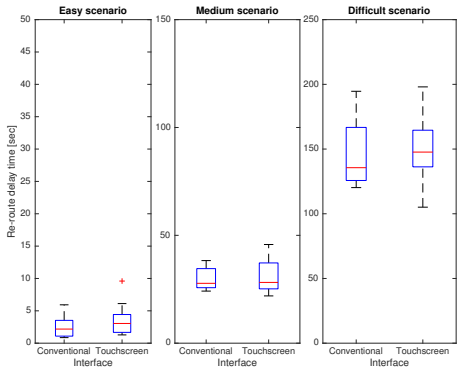


(c) Corrected for between-subject variability, without outliers (see touchscreen, difficult scenario)

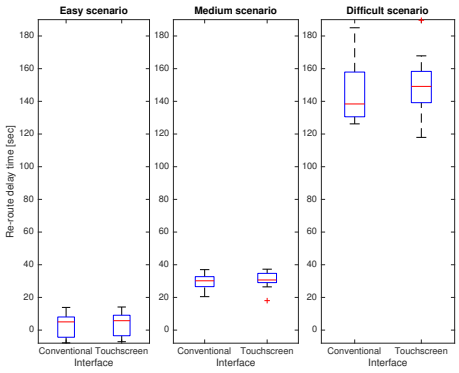
Figure C-1: Individual measurements where participants failed to successfully avoid weather, leading to meaningless re-route time delays scores.

C-2 Between-Subject Variability

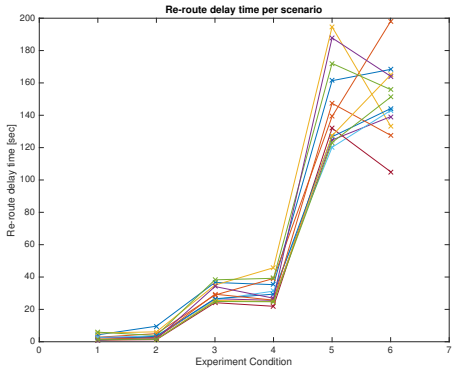
As discussed in the IEEE article, each of the three experiment scores were corrected for between-subject variability. The individual scores per experimental conditions were plotted, showing significant between-subject variability for re-route time delay, roll angles, mental effort scores and secondary task time delays. Hence, a between-subjects correction was necessary, of which the results are presented, per dependent measure, in this section.



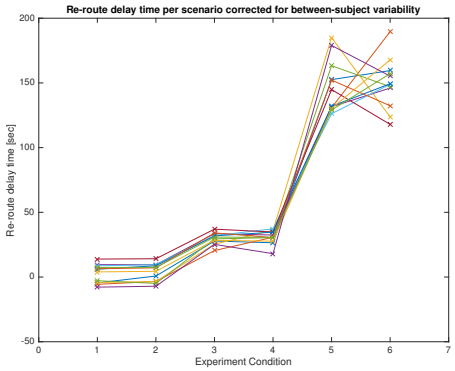
(a) Raw measurements



(b) Corrected for between-subject variability

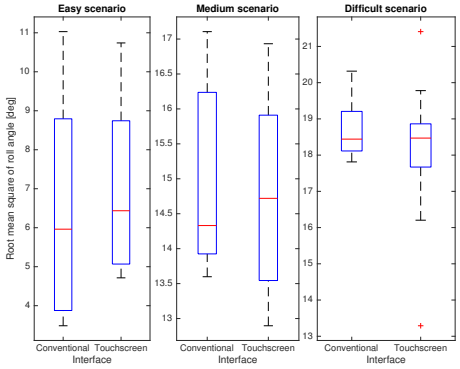


(c) Raw measurements per participant

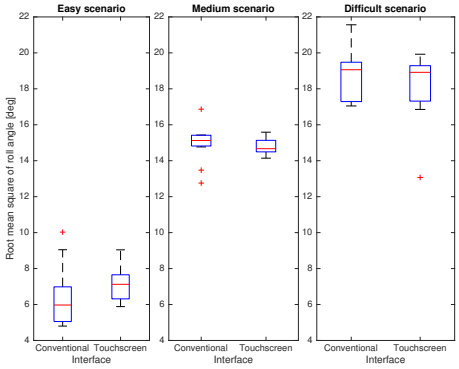


(d) Corrected for between-subject variability per participant

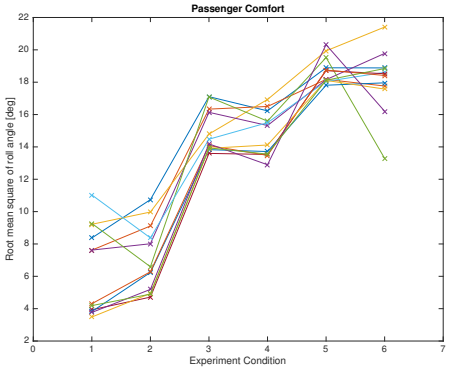
Figure C-2: Re-route time delay (outliers removed), including corrections for between-subject variability



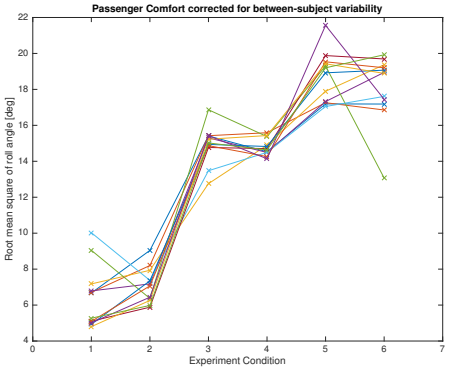
(a) Raw measurements



(b) Corrected for between-subject variability

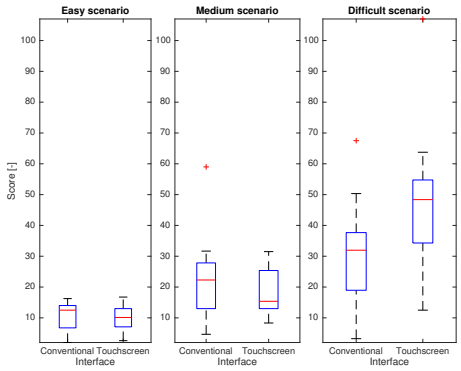


(c) Raw measurements per participant

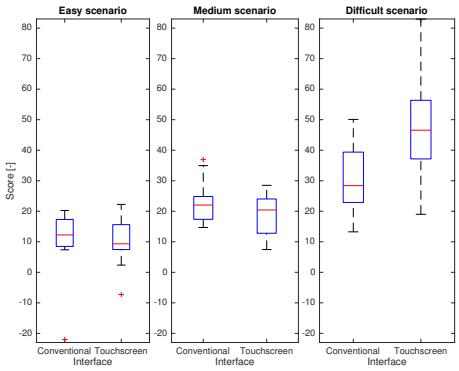


(d) Corrected for between-subject variability per participant

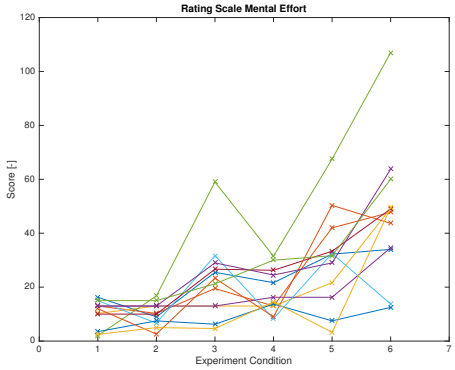
Figure C-3: Root mean square of roll angle, including corrections for between-subject variability



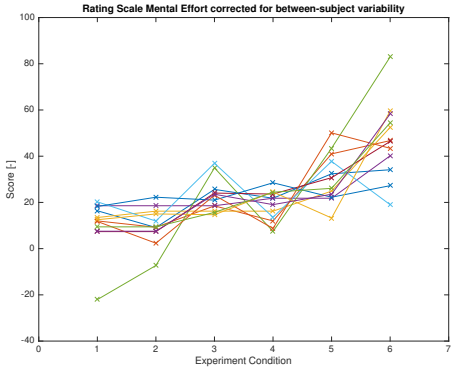
(a) Raw measurements



(b) Corrected for between-subject variability

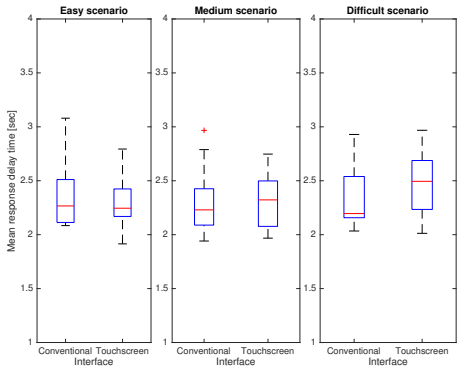


(c) Raw measurements per participant

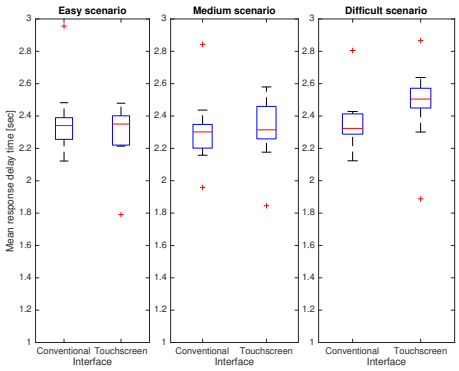


(d) Corrected for between-subject variability per participant

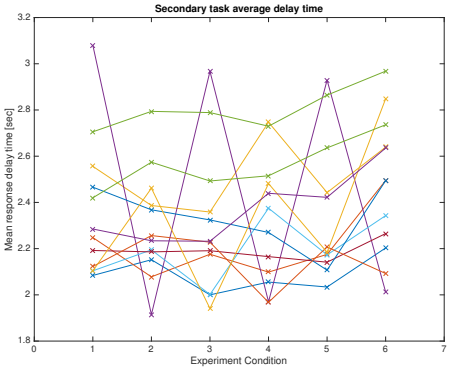
Figure C-4: Rating Scale Mental Effort scores, including corrections for between-subject variability



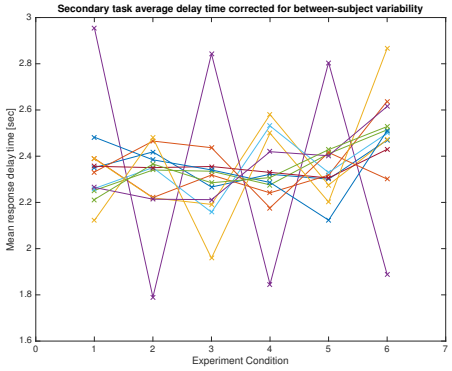
(a) Raw measurements



(b) Corrected for between-subject variability



(c) Raw measurements per participant

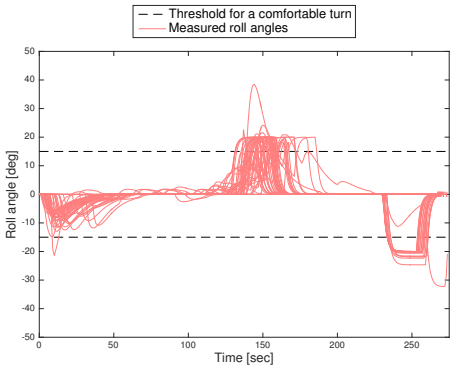


(d) Corrected for between-subject variability per participant

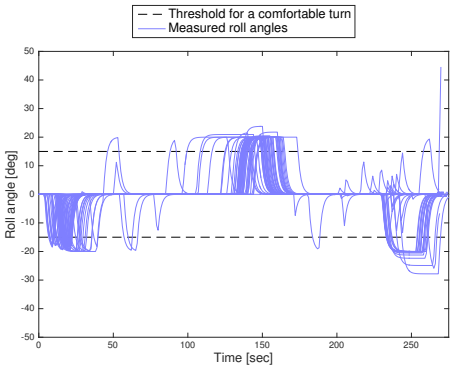
Figure C-5: Secondary task response delay times, including corrections for between-subject variability

C-3 Roll Angles per Scenario

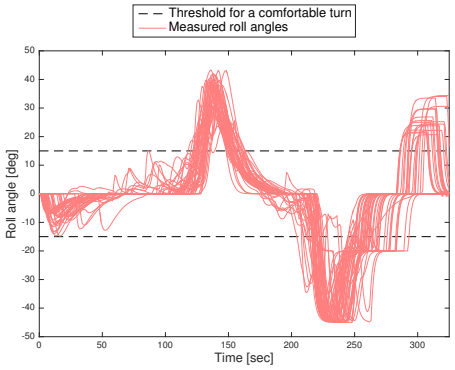
The root mean square of roll angle, discussed in the IEEE article, was derived from measured roll angles. The comfortable turn threshold of 15° is based on interviews with KLM pilots.



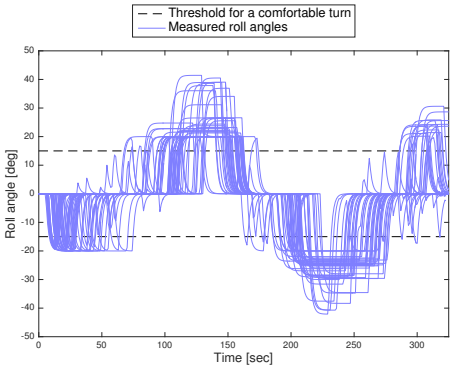
(a) Easy scenario using the conventional interface



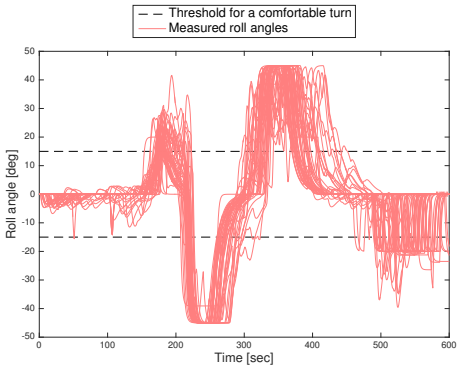
(b) Easy scenario using the touchscreen interface



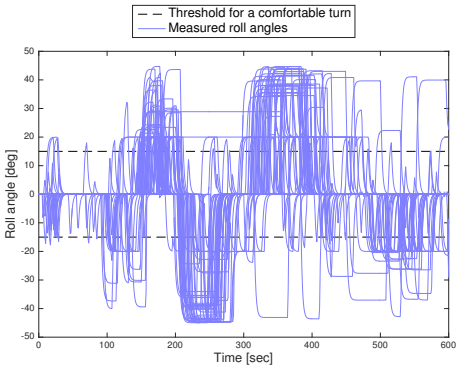
(c) Medium scenario using the conventional interface



(d) Medium scenario using the touchscreen interface



(e) Difficult scenario using the conventional interface

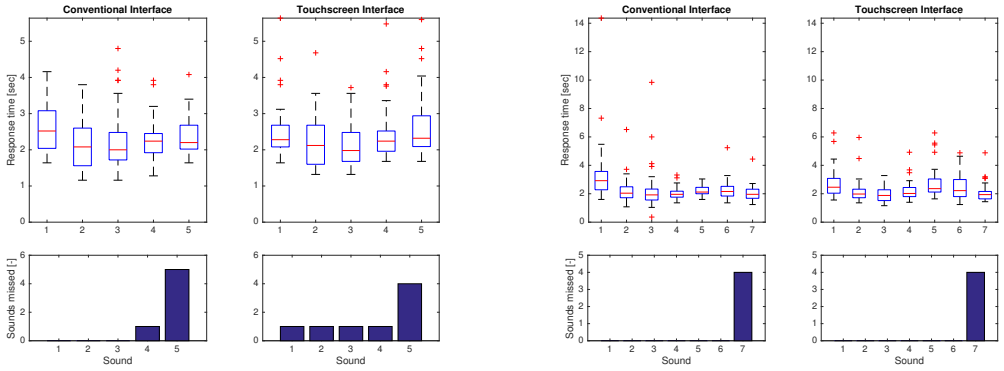


(f) Difficult scenario using the touchscreen interface

Figure C-6: Roll angles measured over time for each measurement

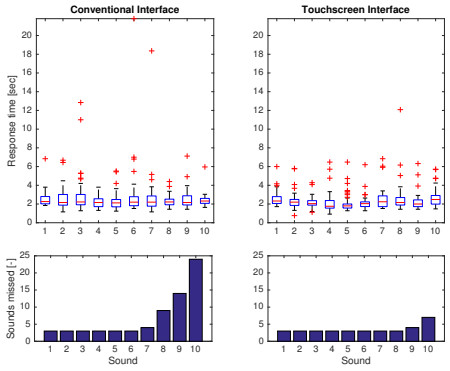
C-4 Secondary Task Scores per Individual Sound

Overall secondary task scores were discussed in the IEEE article, however these can also be analysed for each individual sound. This has been done, and is presented in this section.



(a) Easy scenario

(b) Medium scenario

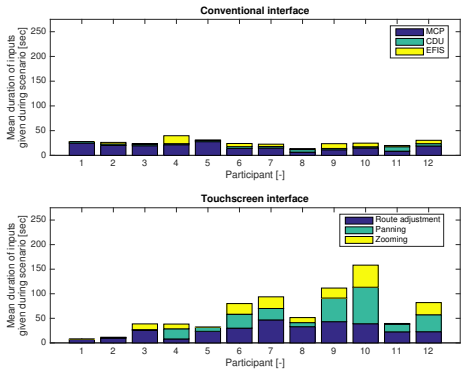


(c) Difficulty scenario

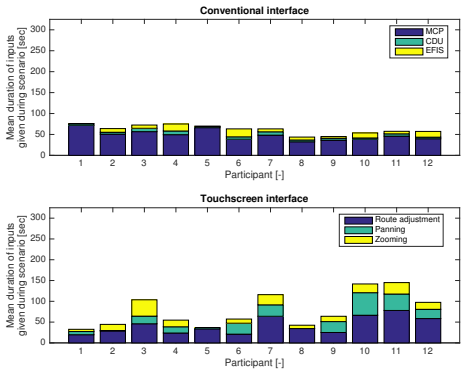
Figure C-7: Secondary task scores per individual `KLM9TU` message per scenario

C-5 Participant Variability in Interface Usage

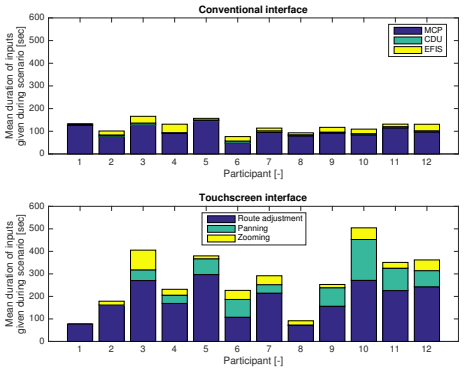
A figure was discussed in the IEEE article, presenting the variability in interface usage per participant for all three scenarios combined. However, individual analysis per scenario is also possible and presented in this section.



(a) Easy scenario



(b) Medium scenario



(c) Difficulty scenario

Figure C-8: Amount of type of inputs given per participant per scenario

Appendix D

Balancing the Lateral Weather Avoidance Experiment

In order to balance the order of experimental conditions experienced per participant during the lateral weather avoidance experiment, a latin square [61] was used. The resulting order of conditions per participant is presented in Figure D-1. Participant 0 did not take part in the final experiment, but helped to refine the design by cooperating in a full-scale test run of the procedures.

Schedule

Number	Interface	Schedule (part one)													
0	Conventional	Training	Diff. Bi	Easy A	Med. Bi	Med. A	Easy Bi	Diff. A	Break	Diff. B	Easy Ai	Med. B	Med. Ai	Easy B	Diff. Ai
1	Conventional	Training	Easy A	Med. A	Diff. Bi	Diff. A	Med. Bi	Easy Ai	Break	Easy Bi	Med. Ai	Diff. B	Diff. Ai	Med. B	Easy B
2	Conventional	Training	Med. A	Diff. A	Easy A	Easy Ai	Diff. Bi	Med. Ai	Break	Med. Bi	Diff. Ai	Easy Bi	Easy B	Diff. B	Med. B
3	Conventional	Training	Diff. A	Easy Ai	Med. A	Med. Ai	Easy A	Diff. Ai	Break	Diff. Bi	Easy B	Med. Bi	Med. B	Easy Bi	Diff. B
4	Conventional	Training	Easy Ai	Med. Ai	Diff. A	Diff. Ai	Med. A	Easy B	Break	Easy A	Med. B	Diff. Bi	Diff. B	Med. Bi	Easy Bi
5	Conventional	Training	Med. Ai	Diff. Ai	Easy Ai	Easy B	Diff. A	Med. B	Break	Med. A	Diff. B	Easy A	Easy Bi	Diff. Bi	Med. Bi
6	Conventional	Training	Diff. Ai	Easy B	Med. Ai	Med. B	Easy Ai	Diff. B	Break	Diff. A	Easy Bi	Med. A	Med. Bi	Easy A	Diff. Bi
7	Next-Generation	Training	Easy B	Med. B	Diff. Ai	Diff. B	Med. Ai	Easy Bi	Break	Easy Ai	Med. Bi	Diff. A	Diff. Bi	Med. A	Easy A
8	Next-Generation	Training	Med. B	Diff. B	Easy B	Easy Bi	Diff. Ai	Med. Bi	Break	Med. Ai	Diff. Bi	Easy Ai	Easy A	Diff. A	Med. A
9	Next-Generation	Training	Diff. B	Easy Bi	Med. B	Med. Bi	Easy B	Diff. Bi	Break	Diff. Ai	Easy A	Med. Ai	Med. A	Easy Ai	Diff. A
10	Next-Generation	Training	Easy Bi	Med. Bi	Diff. B	Diff. Bi	Med. B	Easy A	Break	Easy B	Med. A	Diff. Ai	Diff. A	Med. Ai	Easy Ai
11	Next-Generation	Training	Med. Bi	Diff. Bi	Easy Bi	Easy A	Diff. B	Med. A	Break	Med. B	Diff. A	Easy B	Easy Ai	Diff. Ai	Med. Ai
12	Next-Generation	Training	Diff. Bi	Easy A	Med. Bi	Med. A	Easy Bi	Diff. A	Break	Diff. B	Easy Ai	Med. B	Med. Ai	Easy B	Diff. Ai

Number	Interface	Schedule (part two)													
0	Next-Generation	Training	Easy Ai	Med. Ai	Diff. A	Diff. Ai	Med. A	Easy B	Break	Easy A	Med. B	Diff. Bi	Diff. B	Med. Bi	Easy Bi
1	Next-Generation	Training	Med. Ai	Diff. Ai	Easy Ai	Easy B	Diff. A	Med. B	Break	Med. A	Diff. B	Easy A	Easy Bi	Diff. Bi	Med. Bi
2	Next-Generation	Training	Diff. Ai	Easy B	Med. Ai	Med. B	Easy Ai	Diff. B	Break	Diff. A	Easy Bi	Med. A	Med. Bi	Easy A	Diff. Bi
3	Next-Generation	Training	Easy B	Med. B	Diff. Ai	Diff. B	Med. Ai	Easy Bi	Break	Easy Ai	Med. Bi	Diff. A	Diff. Bi	Med. A	Easy A
4	Next-Generation	Training	Med. B	Diff. B	Easy B	Easy Bi	Diff. Ai	Med. Bi	Break	Med. Ai	Diff. Bi	Easy Ai	Easy A	Diff. A	Med. A
5	Next-Generation	Training	Diff. B	Easy Bi	Med. B	Med. Bi	Easy B	Diff. Bi	Break	Diff. Ai	Easy A	Med. Ai	Med. A	Easy Ai	Diff. A
6	Next-Generation	Training	Easy Bi	Med. Bi	Diff. B	Diff. Bi	Med. B	Easy A	Break	Easy B	Med. A	Diff. Ai	Diff. A	Med. Ai	Easy Ai
7	Conventional	Training	Med. Bi	Diff. Bi	Easy Bi	Easy A	Diff. B	Med. A	Break	Med. B	Diff. A	Easy B	Easy Ai	Diff. Ai	Med. Ai
8	Conventional	Training	Diff. Bi	Easy A	Med. Bi	Med. A	Easy Bi	Diff. A	Break	Diff. B	Easy Ai	Med. B	Med. Ai	Easy B	Diff. Ai
9	Conventional	Training	Easy A	Med. A	Diff. Bi	Diff. A	Med. Bi	Easy Ai	Break	Easy Bi	Med. Ai	Diff. B	Diff. Ai	Med. B	Easy B
10	Conventional	Training	Med. A	Diff. A	Easy A	Easy Ai	Diff. Bi	Med. Ai	Break	Med. Bi	Diff. Ai	Easy Bi	Easy B	Diff. B	Med. B
11	Conventional	Training	Diff. A	Easy Ai	Med. A	Med. Ai	Easy A	Diff. Ai	Break	Diff. Bi	Easy B	Med. Bi	Med. B	Easy Bi	Diff. B
12	Conventional	Training	Easy Ai	Med. Ai	Diff. A	Diff. Ai	Med. A	Easy B	Break	Easy A	Med. B	Diff. Bi	Diff. B	Med. Bi	Easy Bi

Figure D-1: A latin square was used to balance the experiment

Appendix E

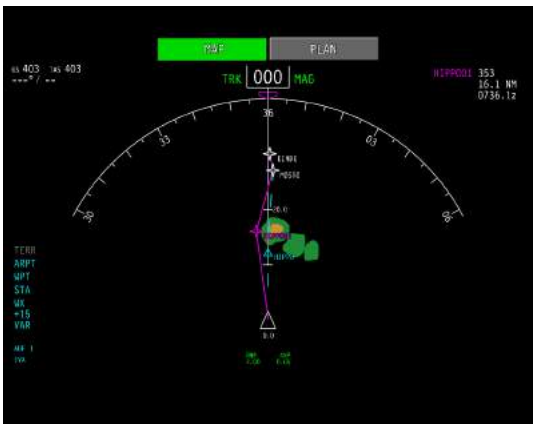
Design of Primary Flight- and Navigation Displays

This appendix provides additional insight in the design of the primary flight- and navigation display used during the lateral weather avoidance experiment. The ND is presented in Figure E-1 once per scenario, with a possible weather avoidance route in magenta and the original flight plan route drawn in cyan. Furthermore, the PFD is presented in Figure E-1.

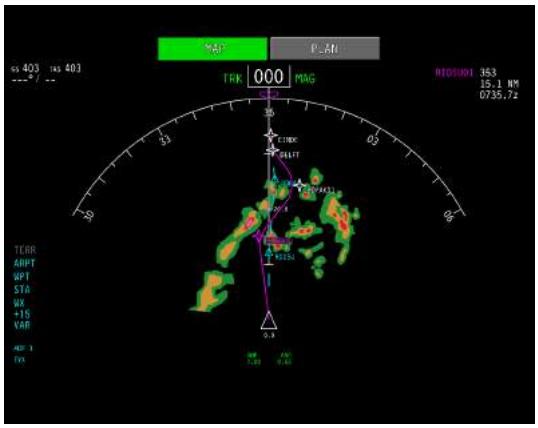
Both displays have been designed based on the Boeing implementation in the 777 and 787. Hence, the active flight plan route is drawn in magenta, a modified route in a stippled white line and a second, in-active route in cyan. The latter was used to display the original route to give participants an indication of their deviation from the original path. True airspeed and ground speed, as well a wind indication (although not used in the experiment) was presented in the top-left of the ND. The active waypoint was presented in the top-right with an expected time of arrival and distance to go indication. The two large MAP/PLAN buttons could be used with the touchscreen interface, or a knob on the EFIS panel with the conventional interface to switch between MAP en PLAN mode. In MAP mode the ND centered on the aircraft's present position and rotated with the current track of the aircraft. In PLAN mode the view orientation and ND centered position remained constant. With the touchscreen interface, the latter could be adjusted by panning. Finally, the PFD presented participants with standard information, such as airspeed, altitude and attitude. The flight mode annunciator at the top was used by participants to check which mode (LNAV or HDG SEL) was activated.

In addition, Figure E-2 and Figure E-3 present photographs[1] of the real Boeing 777 and 787 cockpit respectively. The different interfaces studied in this thesis are highlighted.

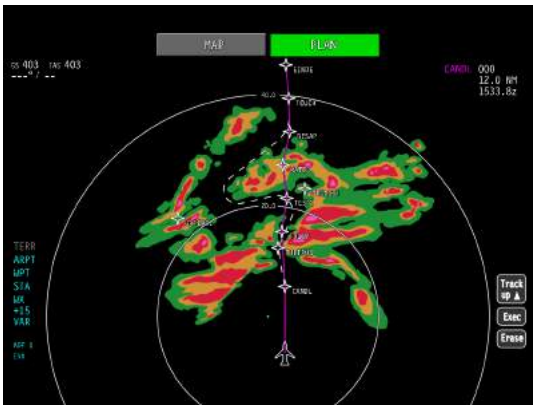
[1]Photographer: KLM 777/787 Captain



(a) Easy scenario on the navigation display in MAP mode, with weather avoidance route in magenta, original route in cyan.



(b) Medium scenario on the navigation display in MAP mode, with weather avoidance route in magenta, original route in cyan.



(c) Difficult scenario on the navigation display in PLAN mode, with weather avoidance route in white (not yet executed), active route in magenta.

(d) Primary flight display

Figure E-1: Displays used during the lateral weather avoidance experiment

(a) ❶ (Rotary controller) commanding the heading select mode on the MCP. ❷ ND on the main instrument panel.

(b) ❷ ND on the main instrument panel. ❸ CDU on the pedestal.

Figure E-2: Photographs from the Boeing 777 cockpit displaying the MCP, ND and CDU studied in this thesis.

Figure E-3: Photographs from the Boeing 787 cockpit displaying the ❶ PFD, ❷ ND and a ❸ *virtual* CDU.

Appendix F

Autopilot Control Law

The scope of this thesis was on the lateral plane, where a heading select mode was compared to a touch-based lateral navigation (LNAV) mode. Both of these modes were part of the autopilot, which commanded the aircraft to fly a specific heading or track. The logic behind these modes is discussed here.

F-1 Heading select mode

The heading select mode was implemented using a simple ramp control law. Based on the difference between the commanded and the current heading, a roll angle between 0 and 45 degrees was commanded, see Figure F-1. The size of the ramp was tuned such that it produced similar turn performance as the LNAV mode.

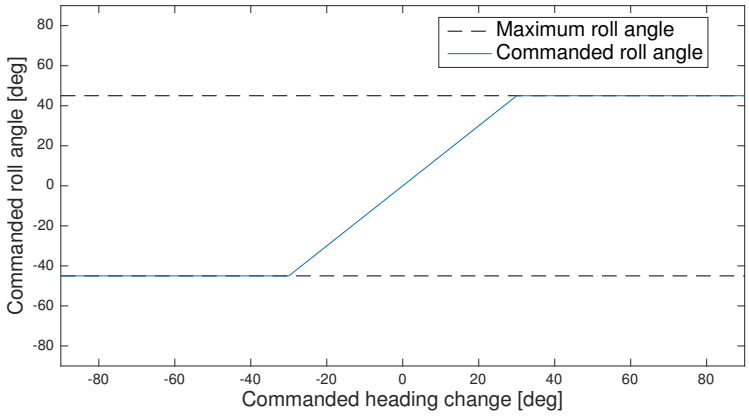


Figure F-1: A simple ramp was used for the heading select mode roll angle control law.

F-2 LNAV mode

The LNAV algorithm attempted to create a turn with a roll angle of 20 degrees. If unable, it would increase the roll angle until a geometric fit was achieved. The algorithm itself did not include a maximum roll angle, however, if the result of the LNAV algorithm included turns that exceeded 45 degrees of roll the route could not be executed. In these cases the route designed by the algorithm would be drawn in a orange/red color on the ND, indicating its invalidity. Valid routes were drawn with a white line, indicating to the pilot that he or she can execute it. Following execution, the modified route becomes magenta and the autopilot starts to follow it.

The flight plan route was saved as a set of segments and sub segments. Segments connected the different route waypoints. Each segment was built out of sub segments, which were either straight of curved, and were created by the LNAV algorithm.

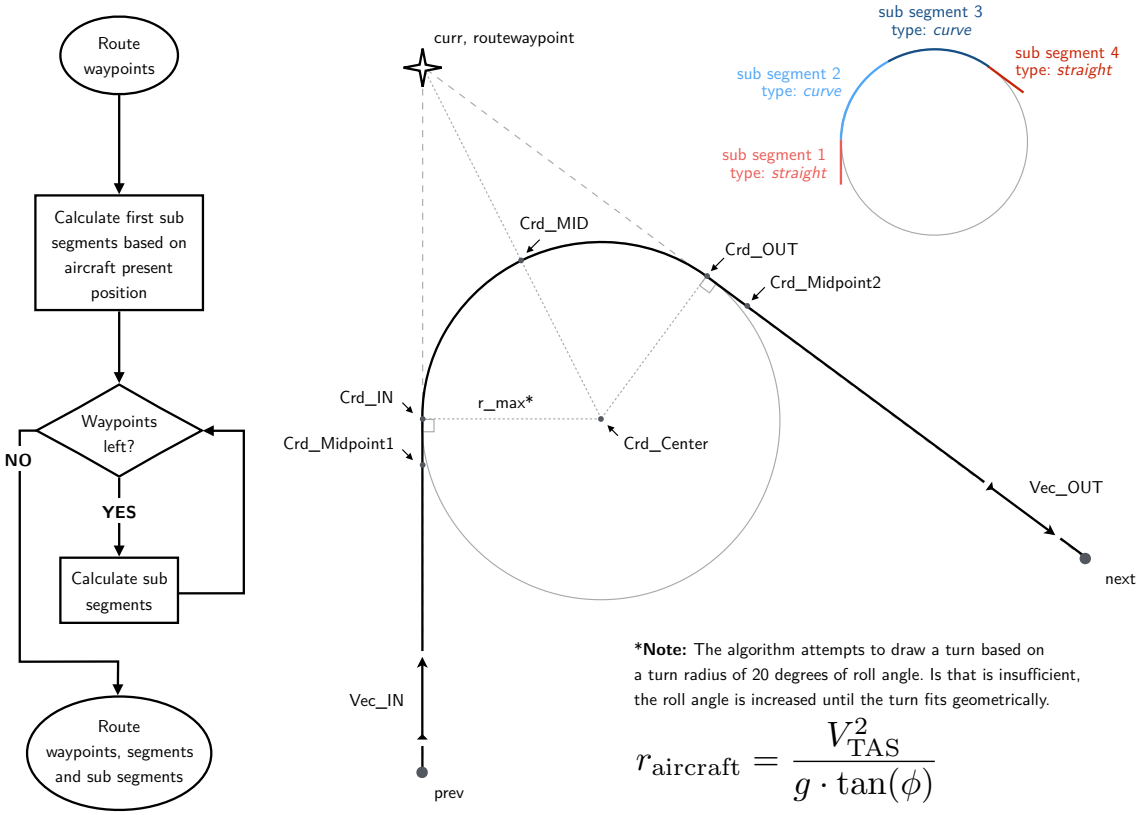


Figure F-2: The LNAV algorithm attempts to create a turn with a roll angle of 20 degrees. If unable, it will increase the roll angle until a geometric fit was achieved. The algorithm returned straight and curved subsegments, which were saved in the FMS, and drawn on the ND.

Appendix G

Software Implementation of Both Experiments in DUECA

In this appendix the technical implementation of both experiments is discussed. Both were conducted in the SIMONA Research Simulator and hence used the DUECA realtime programming platform for the software implementation and management of the experiment. Figure G-1 and Figure G-2 present a block diagram of the software developed for the lateral weather avoidance and Fitts' law experiments respectively. DUECA works based on modules and communication channels linking modules together, as illustrated. Furthermore, due to the implementation of the channels the software could run across various computers, which is also illustrated.

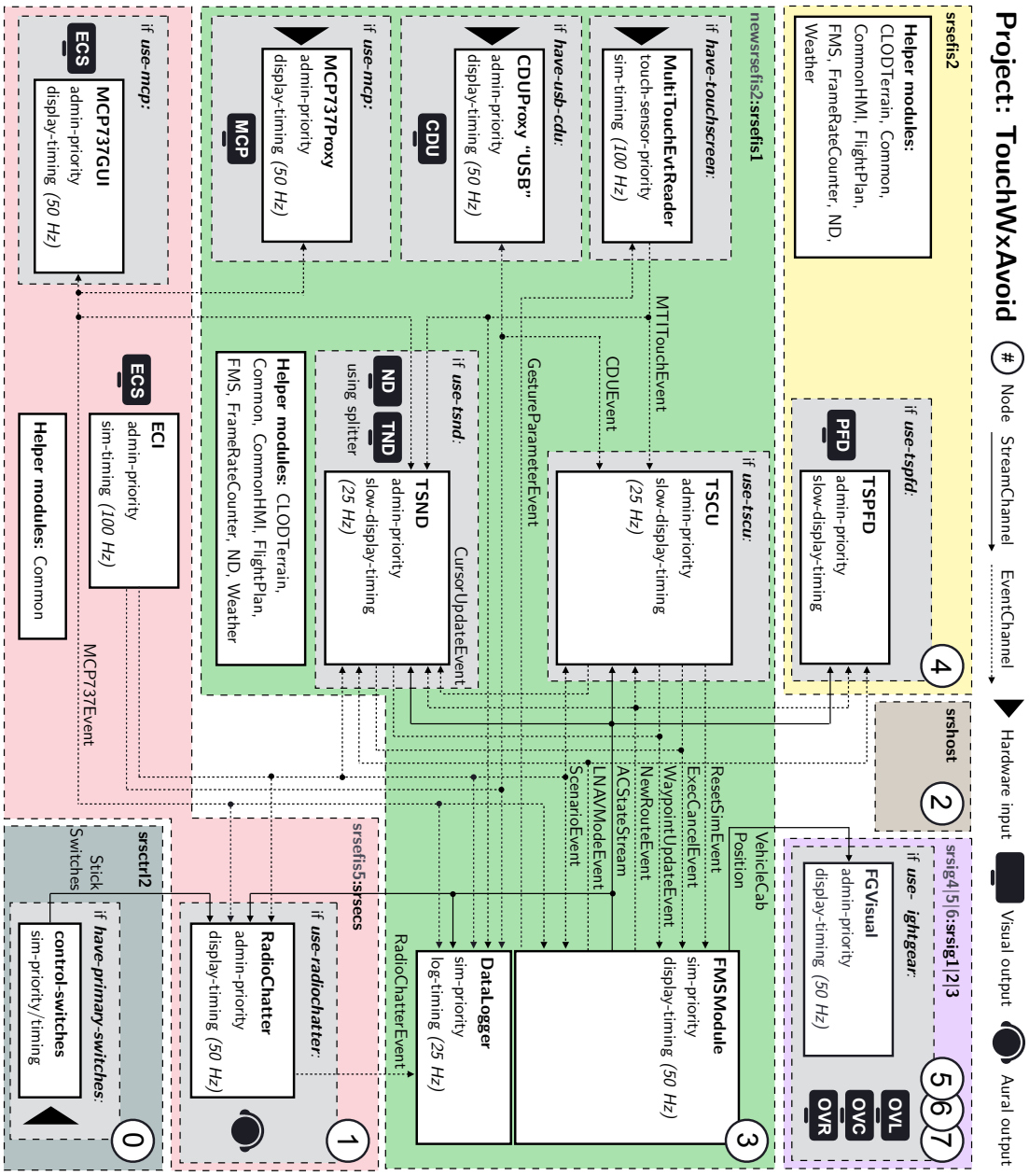


Figure G-1: Technical implementation of the lateral weather avoidance experiment in DUECA.

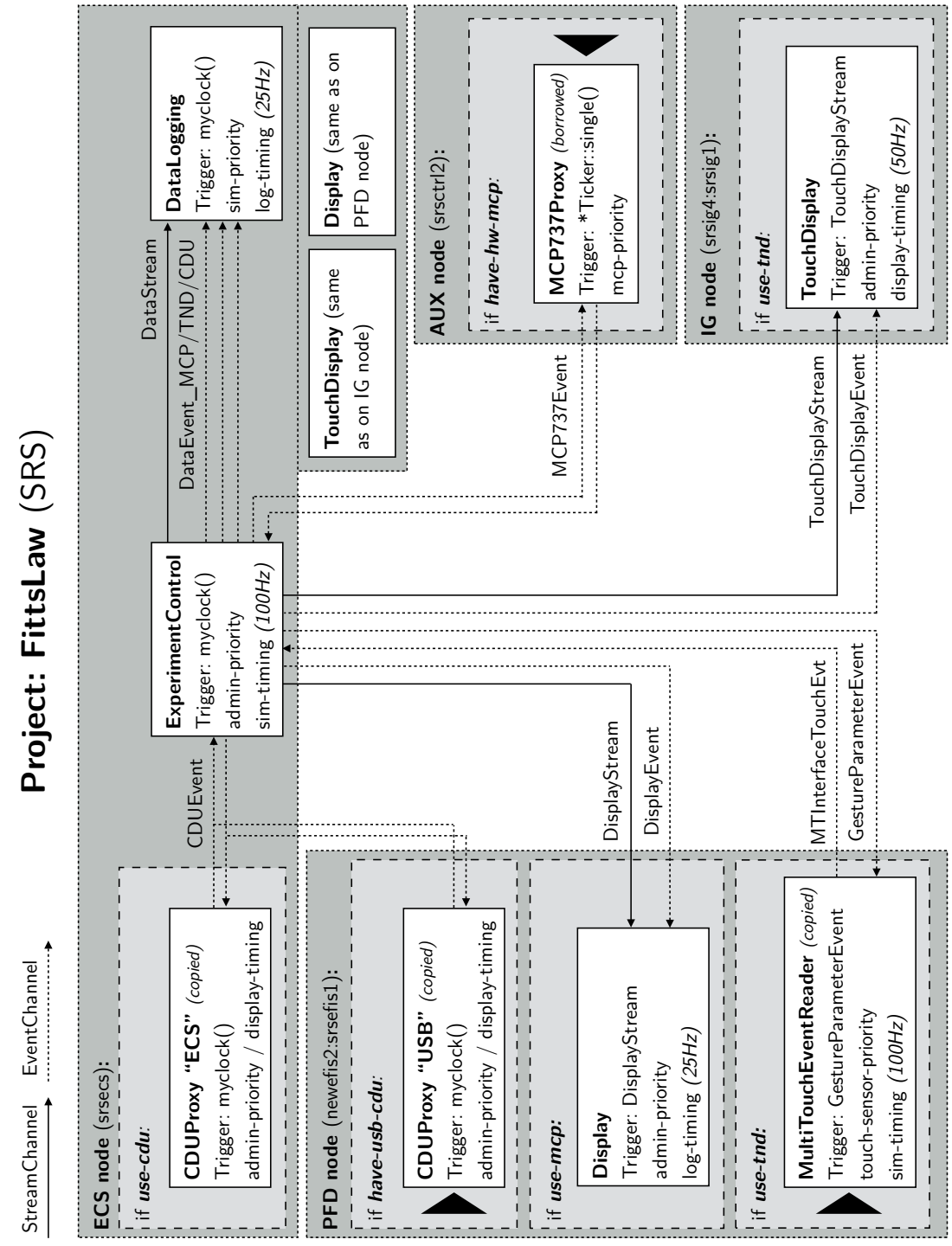


Figure G-2: Technical implementation of the Fitts' law experiment in DUECA.

Appendix H

Experiment Briefing, Checklists and Datalogs

For both experiments a briefing document was written for participants to read a week prior to the experiment. Furthermore, a checklist was made for the verbal (de)briefing during the experiment, to ensure that each participant received the same information. Finally, a datalog was created for the researcher to use during the experiment. These are all presented in this appendix, in the following order.

1. Lateral weather avoidance experiment (IEEE article) — Briefing
2. Lateral weather avoidance experiment (IEEE article) — Checklist and datalog
3. Fitts' law experiment (IEEE companion article) — Briefing
4. Fitts' law experiment (IEEE companion article) — Checklist and datalog

Experiment Briefing

N.C.M. (Nout) van Zon
MSc Control & Simulation

1 Background

Given increasing demands on lateral navigation as airspace congestion grows, a proper re-design of the interface between the pilot and the flight management system is required. Whilst industry proposes touch-based solutions, research in this area is lacking. An experiment is proposed, aimed at comparing a conventional lateral navigation (LNAV) interface with its next-generation equivalent. The former includes a mode control panel (MCP), control display unit (CDU) and primary flight- and navigation displays (PFD, ND). The latter constitutes of a touch-enabled navigation display (TND). This research is a necessary and crucial step to ensuring a sustainable, safe and effortless Pilot-Flight Management System (FMS) interface for the future.

Aim The aim of this experiment is to investigate and compare a touch-based navigation display with the conventional Pilot-FMS interface, consisting of the MCP and CDU, within the scope of a lateral weather avoidance task. Furthermore, the effects of varying task difficulty in combination with the ability to perform two tasks simulatenously will be investigated.

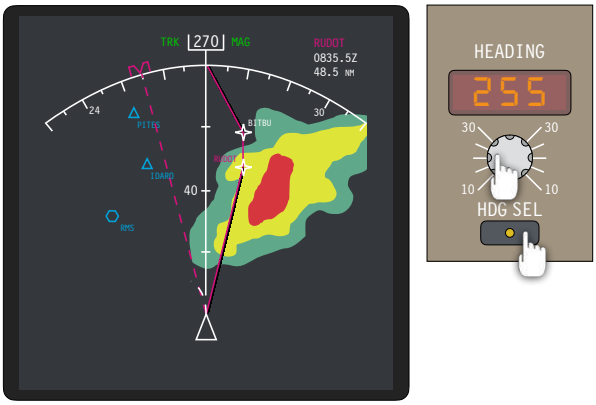
2 Your Task

In this experiment you are operating a Boeing 777-200 cruising on autopilot at flight level 250 and an indicated airspeed of 270 knots. You will operate the aircraft on autopilot during the entire experiment, with the goal of executing two tasks, each of which is described below. In executing these tasks the primary should take precedence.

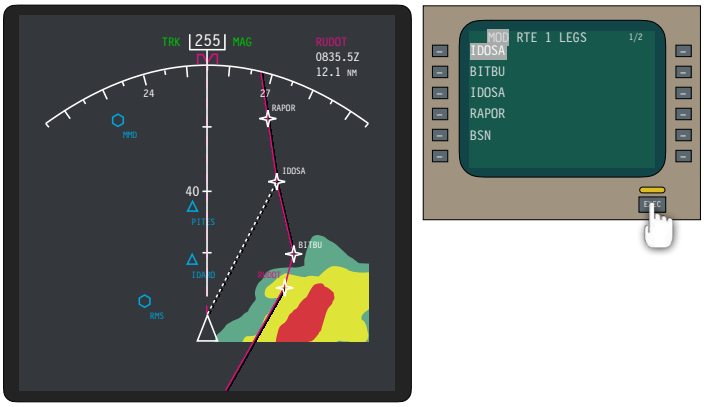
Primary Task Your primary task is to use the available interface to circumnavigate any weather you encounter en-route. Your goal should be to avoid weather systems at all costs, whilst trying to minimize your total re-route time delay. Your ND contains a weather radar, with a radial refresh rate of 360° per second and maximum range of 20 NM.

When using the conventional interface you are requested to use the heading select (`HDG SEL`) mode to circumnavigate any weather you encounter. Note that the heading hold mode, encountered on amongst others the Boeing 747 and 777, is not available. Following safe passage you are requested to use the control display unit to execute a direct-to maneuver towards a logical waypoint and proceed on the `LNAV` mode. When using the next-generation interface you will remain on `LNAV` mode, however have the ability to move existing, or create custom waypoints along your route using your finger. An illustration of these procedures is given in Figure 1, and will be further explained during the briefing and training phase of the experiment.

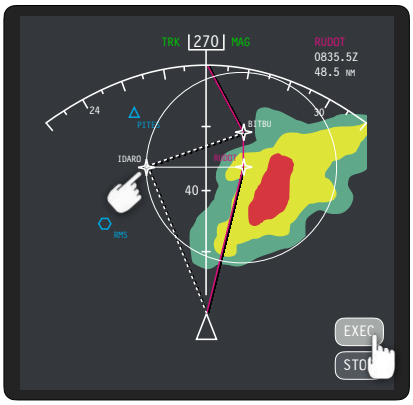
Secondary Task During each measurement run you will hear a continuous radio chatter in your headset. Your task is to identify messages which are destined for you (those containing your callsign: **KLM 9TU**) and subsequently press a designated push-to-talk (PTT) switch to confirm that you have received the message. You are not asked to verbally respond to the message.



(a) **Conventional interface:** The heading select mode can be used to command a specified heading, using the mode control panel.



(b) **Conventional interface:** A direct-to maneuver can be executed using the control display unit.



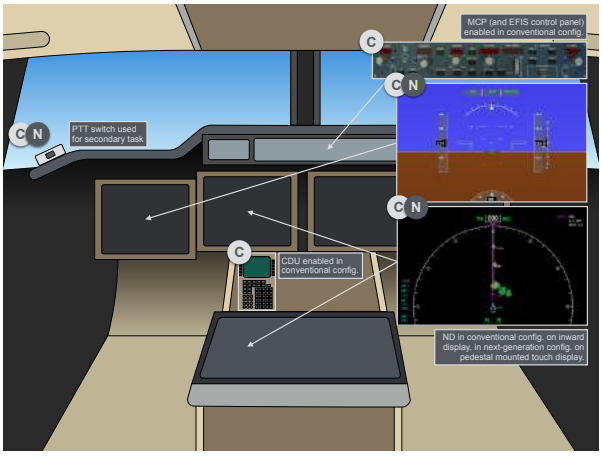
(c) **Next-generation interface:** The touch-enabled navigation display can be used to insert a new waypoint or move an existing waypoint.

Figure 1: Illustration of the primary task for each interface

3 Apparatus

The experiment will be conducted using the SIMONA Research Simulator (SRS), located at the Delft University of Technology. Although the simulator features a motion platform it will not be used during the experiment. The cabin features a collimated 180-degree horizontal by 40-degree vertical outside field of view. Furthermore, four LCD displays (4:3 format) are located on the main instrument panel. The pedestal features a Boeing-style control display unit[1] and large touchscreen display (16:9 format). An illustration of the set-up is given in Figure 2.

(a) The SIMONA Research Simulator (SRS) at the Delft University of Technology



(b) Cabin layout of the SRS showing the available interfaces and displays for both the conventional (C) and next-generation interface (N)

Figure 2: Apparatus used for the experiment

4 Experimental Procedures

The experiment is expected to take 3.5 hours, including briefing, de-briefing and breaks. The schedule is illustrated in Figure 3. During the briefing we will discuss this document and verbally discuss the safety briefing and your tasks for the forthcoming experiment. Subsequently I will ask you to sign a consent form, the content of which is discussed in section 5. Finally you will take a seat in the SIMONA Research Simulator.

The experiment is split into two, given that you will be using both the conventional and the next-generation interface. The order of these interfaces varies per participant, however the procedure remains the same. First a training and familiarisation phase will allow you to get accustomed to the apparatus and the interface. Although designed to reflect the avionics suite found on the Boeing 777 there are a few simplifications and differences to illustrate. Following the familiarisation phase we will proceed to measurement, during which various weather scenarios will be presented, differing in length between five and eight minutes. During the measurement phase you will be asked to execute both the primary and secondary task, previously discussed in section 2. Subsequent to each measurement run you will be asked to rate your mental effort required to perform the combination of both tasks.

Following a coffee break we will continue with an identical training and measurement phase, now using the other interface. The experiment will conclude with a verbal de-briefing.

[1] Produce of CPFlight

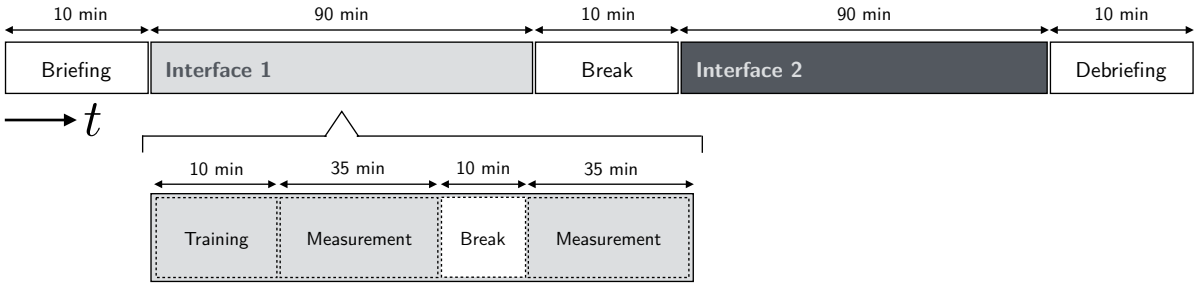


Figure 3: Time schedule of the experiment

5 Your Rights

Participation in the experiment is voluntarily, which means that you can chose to discontinue your collaboration at any time, including during the experiment. All measurement data during the experiment will remain completely anonymous and confidential, such that only the experimenter can link the data back to you. By participating in this experiment you agree that data measured during your experiment may be published. Within any such documentation or publication experiment participants are referred to by number, not by name. Concluding, you are kindly requested to not discuss the experiment with anyone until the entire experiment has been completed to avoid fellow participants becoming biased.

6 Safety Guide for SIMONA Research Simulator

The motion platform of the SRS will **not** be used during this experiment, motion-related safety instructions are therefore omitted from this section. Nonetheless, there are several important safety instructions regarding cabin occupants in the SRS.

Seat belts Given the experimental scenario consists of a turbulent weather scenario, seat belts will be worn during the experiment regardless of the fact that the SRS will not be moving. Each seat contains a 5-point seat belt, with two shoulder straps, two waist straps and a fifth between your legs. Rotating the fastener will release all five belts.

Seat adjustment The seat can be moved horizontally and vertically using the available mechanical handles. During the experiment the seat will be placed at specified positions to make sure each participant is seated in the same position. Nonetheless, it is advised to move the seat fully aft for (de-)boarding.

Intercom and video surveillance An open intercom and video connection is present between the cabin and the control room of the SRS. Video surveillance is not recorded, but used by the experimenter in the control room to monitor your safety. You will be provided a headset with which a continuous communication is possible, without having to press a *push-to-talk switch*. Furthermore the headset is used for the secondary task, as such the volume will be kept at the same position for each participant.

Fire Extinguishers Two fire extuinguishers are attached to the simulator access bridge. Additional extinquishers can be found on the left behind the motion control cabinet and on the 3rd floor balcony in front of the control room. Operating instructions are attached to the devices.

Emergency Cabin Evacuation Given that the motion-system will not be used and as a result the bridge will remain connected to the cabin the egress procedure is straight forward.

1. Communicate intentions with control room (before removing headset)
2. Remove seat belts, push seat fully aft, deboard
3. Open the cabin door, exit simulator using the access bridge
4. Move towards the building's emergency exits as necessary (see Figure 4)

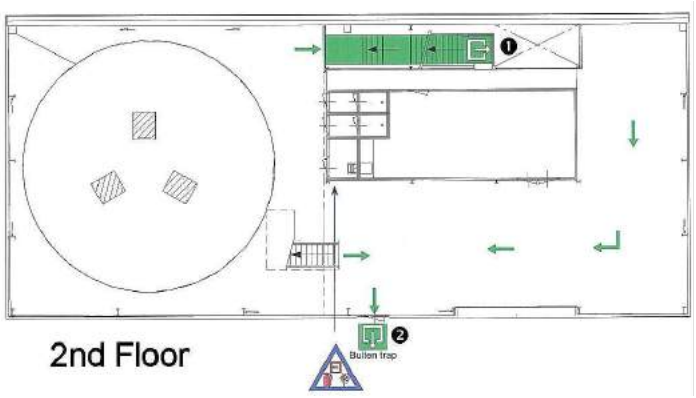


Figure 4: Emergency evacuation route on the second floor (on which the SRS is situated)

General Emergency Instructions General TU Delft instructions in case of an emergency are presented below:

Emergency number 112	HOW TO HANDLE IN CASE OF AN EMERGENCY	Emergency number 112

GENERAL RULES
- become familiar with the map shown above this instruction;
- know the nearest (emergency) exit(s).
- know where fire extinguishers and firealarm buttons are located;
- don't smoke where not allowed;
- keep escape routes and (emergency) exits free;
- keep acces to fire extinguishers free;
- the sound of the evacuation alarm is:
- note of how to handle with: ACCIDENTS, FIRE and EVACUATION;
- follow the instructions of the emergency responders.

ACCIDENT

Immediately call number 112, in case of accident or medical assistance is needed and report as follows:
- your name;
- give exact location of the accident (building, floor, room);
- report the nature of the accident, also report if an ambulance is requested (your indication);
- reassure the victim(s);
- never move an injured person unnecessarily;
- inform emergency responders and follow their instructions.

FIRE
- stay calm, panic is often worse than the fire itself;
- think first, than act;
- activate the nearest fire alarm (break the glass);
- call number 112 and report:
  - your name;
  - give the exact location (building, floor, room, etc.);
  - report what's on fire (furniture, hazardous materials etc.);
- if possible, extinguish the fire with the nearby extinguisher(s);
- if extinguishing is no longer possible or save, stop whatever your were doing, close windows, immediatly leave the room and close all doors behind you;
- warn others nearby on your way out.

EVACUATION

When the evacuation alarm sounds, act as follow:
- follow strictly the instructions of your own manager or the emergency responders;
- close cabinets, drawers, windows and doors and disconnect computers, instrumentation, machines, etc;
- make sure to leave the office or workarea immidiately and follow the instructions of the emergency responders;
- in smoke or heat keep your head as low as possible to the ground (crawl), leave the room immidiately and take the stairway to safety. Never use elevator(s);
- go to the assembly point and wait for further instructions.

Here you are | (Emergency) exit | Assembly point | Fire hose | Extinguisher | Danger of electrical tension | Alarm button | Escape exit | Evacuation route | TU Delft

Experiment Checklist and Datalog

N.C.M. (Nout) van Zon
MSc Control & Simulation

Subject number.....

Date and time.....

PREFLIGHT BRIEFING

☐ ☐ **Discuss safety briefing:** egress procedure, smoke hood, emergency stop

☐ ☐ **Discuss primary and secondary task:**

1. Use the **given interface to avoid weather** encountered en-route with **minimum re-route delay** whilst **considering safety and comfort** of your passengers.
2. **Respond to ATC** messages including **your callsign** `KLM 9TU` with the PTT switch.

☐ ☐ **Primary task is executed using the disposable interface with the right hand:**

- Caution maximum bank angle set to 45 degrees
- Caution 20 NM planning horizon (short for experiment purposes)
- Always end at the final waypoint `EINDE`
- Active weather avoidance route (`RTE1`) in `MAGENTA`, original cleared route (`RTE2`) in `CYAN`
- Both MAP and PLAN views of the ND available, of which the latter can be panned with the TND

Conventional Interface Briefing

☐ **Procedure:** HDG SEL to avoid weather, thereafter direct-to (CDU) to re-join route. Re-join route before last waypoint!

☐ **Consists of:** MCP, CDU, PFD and ND

☐ **MCP:** LNAV mode, HDG SEL mode/knob. HDG HOLD (ie. 777/747) is not available

☐ HDG SEL knob *jumpy* for fast movements

☐ LNAV mode arms (feedback on PFD), when close to route LNAV activates

☐ No necessity to activate `HDGSEL` at end of route

☐ **CDU:** Route adjustments, direct-to cmd

☐ Direct-to always from present position

☐ Custom waypoints (PBD) not possible

Next-Generation Touch Interface Briefing

☐ **Procedure:** Re-position existing waypoints to avoid weather

☐ **Consists of:** TND, PFD and ND

☐ One finger used to re-position waypoints or select soft buttons

☐ Waypoint select delay time is 1 second

☐ `EXEC/ERASE` keys show when modifying rte

☐ If modified route is invalid (due to turnradius limitations), route becomes orange/red and execution is not possible

☐ Two fingers used to zoom or pan

☐ View is re-centred using the `TRACK UP` key

☐ **Sufficient training** including `MOD = ACT`, avoiding loops, various gestures, etc...

☐ ☐ **Secondary task is executed using your headset and the PTT switch with the left hand**. Maintain your left hand on the switch, such that dwell and homing time are omitted in your response time

☐ ☐ **Discuss weather radar:**

- Maximum range is 20 nm and refresh rate is 1 sec
- Presents a top-down view (similar to *buienradar*), therefore *'what you see is what is'*
- Weather is static, and therefore does not change, if you see an opening the opening is there
- The outside visuals do not reflect the weather radar
- *Groen is niet doen*

☐ ☐ **Discuss *Rating Scale Mental Effort* for post-scenario subjective rating**

☐ ☐ **Set seat position for experiment interface**

☐ ☐ **Use the right hand armrest**

☐ ☐ **Put on your seatbelts and headset**

☐ ☐ **Set lights as required:** flood lights, MCP/CDU backlighting, chart/map light for RSME

POSTFLIGHT DE-BRIEFING

☐ (with permission, note anonymity) Initiate audio recording of entire de-briefing

How did you use the **conventional interface** to decide and execute your weather avoidance route? (hint to decision ladder: where was skill-, rule- and knowledge-based behaviour used? How were decisions made, based on high-level goals or lower-level inputs, etc...)

How did you use the **touchscreen interface** to decide and execute your weather avoidance route?

How do you compare both interfaces in terms of **situational awareness**?

How do you compare both interfaces in terms of **complacency**?

Which interface do you prefer? Which was easier to use?

To improve the touch interface, what features could be added?

Based on your experience today, do you see a touch-based interface as a promising candidate to replace the CDU?

What operational challenges do you see using a touchscreen interface?

Scenario	Log #	Re-route delay [s]	ATC success rate	ATC avg delay [s]	RSME score	Comments
Interface 1:						
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
Interface 2:						
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						

Experiment Briefing

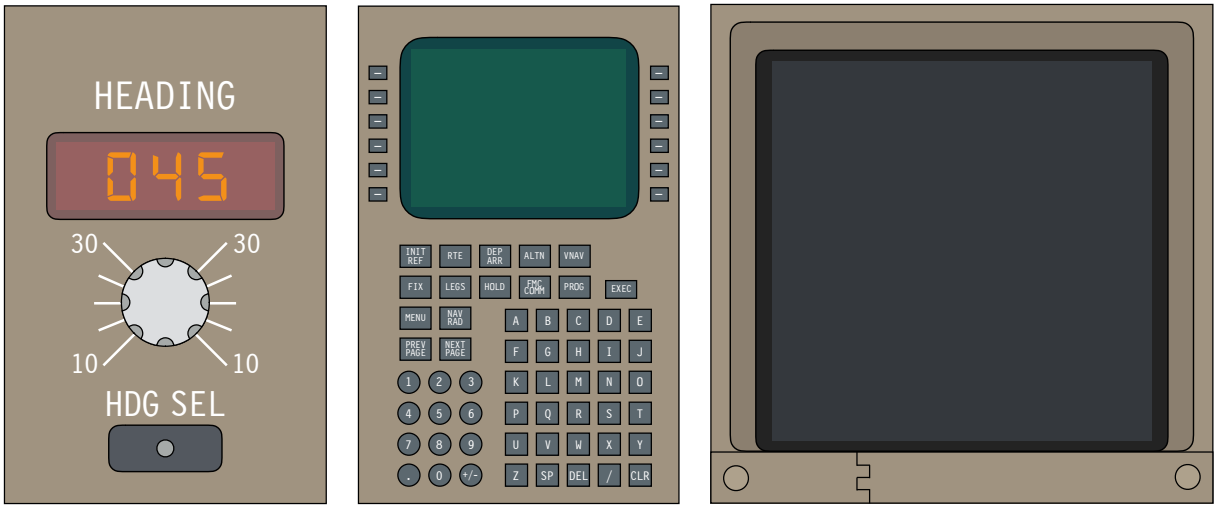
N.C.M. (Nout) van Zon
MSc Control & Simulation

1 Background

Given increasing demands on lateral navigation as airspace congestion grows, a proper re-design of the interface between the pilot and the flight management system is required. Whilst industry proposes touch-based solutions, research in this area is lacking. An experiment is proposed, aimed at creating three individual quantitative models of Pilot-FMS interfaces: the Mode Control Panel (MCP), Control Display Unit (CDU) and a Touch-based Navigation Display (TND). The former two constitute the state-of-the-art flight deck interface between the pilot and the FMS, whilst the latter represents the industry-proposed solution to growing lateral demands. This research is a necessary and crucial step to ensuring a sustainable, safe and effortless Pilot-Flight Management System (FMS) interface for the future.

2 Flight Deck Interfaces

Three flight deck interfaces will be investigated during this experiment, each of which is illustrated in Figure 1. The heading control knob on the mode control panel (MCP), in combination with the control display unit (CDU) form an integral part of modern day flight management systems. Using the MCP a pilot can command a specific heading to the autoflight system, which in turn will fly and maintain the mandated value. Furthermore, the pilot has the ability to adjust his or her flight plan route using the CDU, for example by deleting certain waypoints to proceed direct to one further en-route. The third interface evaluated during this experiment, the touch-based navigation display (TND), represents the industry proposed solution to growing demands in lateral navigation. By allowing pilots to directly manipulate a graphical representation of their route flight deck engineers hope to improve the ease and speed of executing lateral navigation tasks.



(a) Heading control knob on the mode control panel (MCP)

(b) Control display unit (CDU)

(c) Touch-based navigation display (TND)

Figure 1: Three flight deck interfaces that are to be investigated.

3 Experiment Design

This section will provide a general overview of the experiment, followed by more detailed descriptions of each sub-experiment in subsequent subsections.

Goal The goal of this experiment is to collect data on movement times and input accuracy for each of the aforementioned flight deck interfaces. This experiment forms an integral part of a broader thesis study looking at the use of touch-based interfaces on the flight for lateral navigation tasks.

Design Given that three independent interfaces are the subject of this study the experiment will consist of three respective sub-experiments, each focussing on one of three interfaces. The order of the sub-experiments differs based on the group you are allocated to.

Experimental Conditions The main independent variables that will change throughout each experiment are the necessary movement- and target sizes. Movement time will be measured in combination with the distribution of your movement endpoints as a description of your achieved accuracy.

Task Your task will be use each interface to execute rapid aimed movements, of which the exact task will be described in the following subsections. You are requested to emphasize *speed and accuracy* in order to achieve an approximate target hit-rate of 96% whilst providing a smooth input motion.

Apparatus This experiment will be conducted in the SIMONA Research Simulator at the Delft University of Technology. The outside visuals and motion capabilities of the simulator will not be used. Nonetheless, the cabin provides a realistic look and feel to the interaction with the different flight deck interfaces.

Time Schedule Figure 2 provides insight into the expected time necessary to complete the experiment. For each experiment the measurement phase will be foregone by a training phase to allow you to get accustomed to the interface. Each sub-experiment will take approximately twenty minutes to conduct, hence including breaks the total time will amount to 1.5 hours.

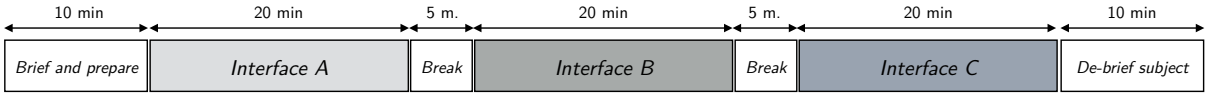


Figure 2: Time schedule of the full experiment for each participant group. The order of interfaces (MCP, CDU or TND) is mixed randomly per participant.

3.1 Mode Control Panel (MCP) Experiment

The experiment set-up is presented in Figure 3. On the inboard screen a navigation display shown, of which an illustration is given in ❶. A magenta heading bug reflects the commanded value on the mode control panel, and is reset to the 0° north-up position prior to each trial. Your task will be to move the heading bug, in a smooth and prompt manner, towards the target depicted by the two cyan lines. The angular distance α towards the task and the angular width (size) ω of the target will change throughout the experiment. A small LCD display above the control knob reflects the commanded heading. Once again, you are requested to move in rapid aimed movements. If you overshoot and miss the target, you are asked not to attempt a correction. As mentioned in the introduction you should aim for a hit-rate of 96%, and therefore a small margin for error is possible. Feedback on the actual hit-rate will be given. Hence, the most important take-aways are:

- Your task is to use the control knob to move the heading bug towards the target.

- Provide a rapid, aimed and smooth input motion, aim for a hit-rate of 96%.
- Over- or undershooting the target is acceptable, don't attempt to correct. Feedback on the actual achieved hit-rate will be given.

3.2 Control Display Unit (CDU) Experiment

The experiment set-up is presented in Figure 4. The control display unit is located in the forward pedestal ❶ and features an alphanumeric keyboard and small display. Your task will be to enter specified words and subsequently move them to a requested location on the display. Only use one finger, preferably your index-finger. Using the alphanumeric keyboard will populate a *scratchpad* which you can find at the ❷ bottom of the display. You can use the `CLR` key to backspace the scratchpad in case you make a mistake. One of the twelve line select keys ❸ adjacent to the display can then be used to move the entire content of the scratchpad to the respective location. The required word and final location is communicated by displaying the word in small at the required final location. For example, in this case the word `AOW80` is required at the top-left of the display. Hence, you will enter the word into the scratchpad using the alphanumeric keyboard and subsequently push on the top-left line select key to move the text. You are advised to take your time in finding the necessary keys before starting to type. Once you start typing the clock will begin and you are asked to move in a series of rapid aimed movements. If you press an incorrect key, promptly use the `CLR` key to backspace and continue typing. Hence, the most important take-aways are:

- Your task is to type the desired alphanumeric word and move it to the desired line-select key location on the display. This is communicated by displaying the word in small at the desired location.
- Try to locate all the necessary keys before starting to type. If you make a mistake whilst typing, promptly use the `CLR` key to backspace and continue.
- Only type with one finger, preferably your index-finger.
- Whilst typing, attempt to provide rapid aimed movements in order to achieve a hit-rate (e.g. punching the correct keys) of 96%. Feedback on the actual achieved hit-rate will be given.

3.3 Touch-based Navigation Display (TND) Experiment

The experiment set-up is presented in Figure 5. A large touch-screen is installed on the center console within the flight deck. A display ❶ is presented on it which includes a white object featuring a magenta crosshair. You can use your finger to select the object and move it around the display. The target is presented by a cyan circle with a white crosshair. The shortest distance A and target size W as well as directional variables ϕ and θ are changed throughout the experiment. Your task is to select the object and move it in a rapid, aimed and smooth motion towards the target. Once again, overshooting the target is acceptable given that you are requested to achieve a hit-rate of 96%, and therefore a small margin for error is possible. Feedback on the actual hit-rate will be given. A finger calibration experiment will be conducted, during which your task is to tap on targets randomly drawn on the display. Hence, the most important take-aways are:

- Your task is to use your finger to move the white object towards the within cyan circular target.
- Provide a rapid, aimed and smooth input motion, aim for a hit-rate of 96%.
- Over- or undershooting the target is acceptable, don't attempt to correct. Feedback on the actual achieved hit-rate will be given.

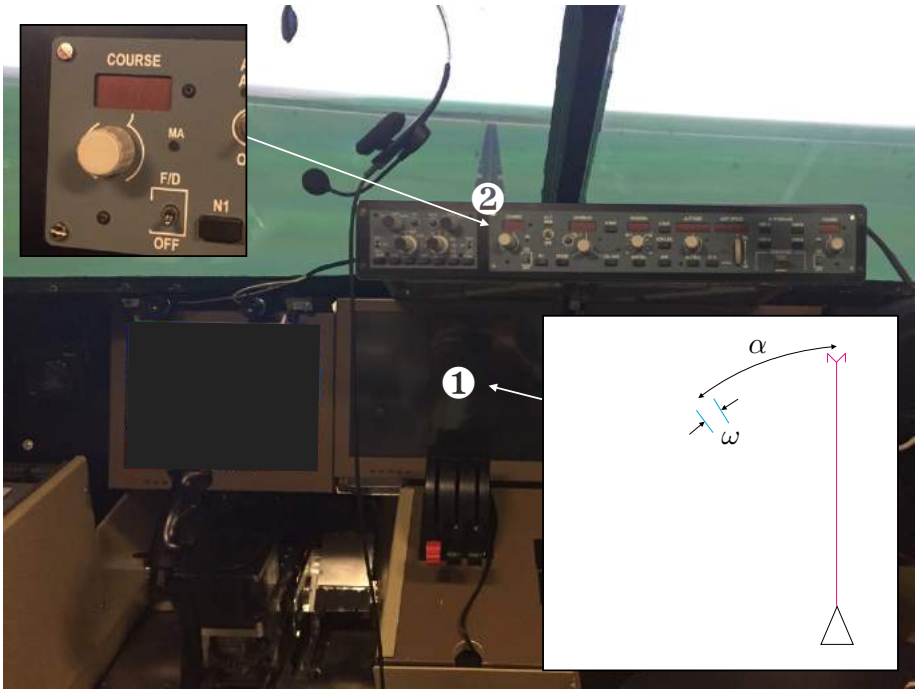


Figure 3: Experimental apparatus for MCP experiment, showing an illustration of the navigation display and heading control knob.

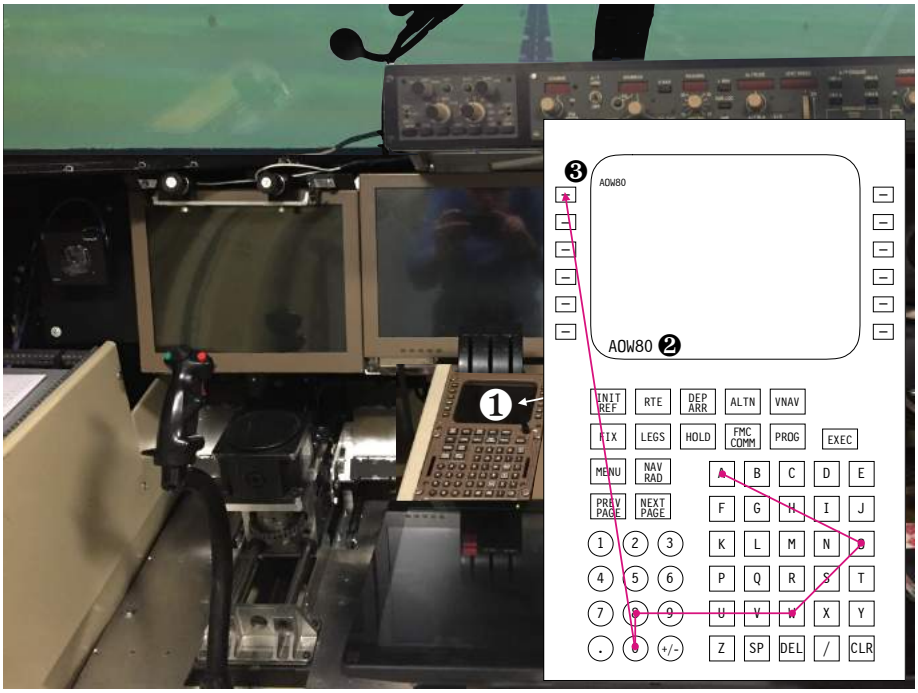


Figure 4: Experimental apparatus for CDU experiment, showing an illustration of the control display unit and location within the flight deck.

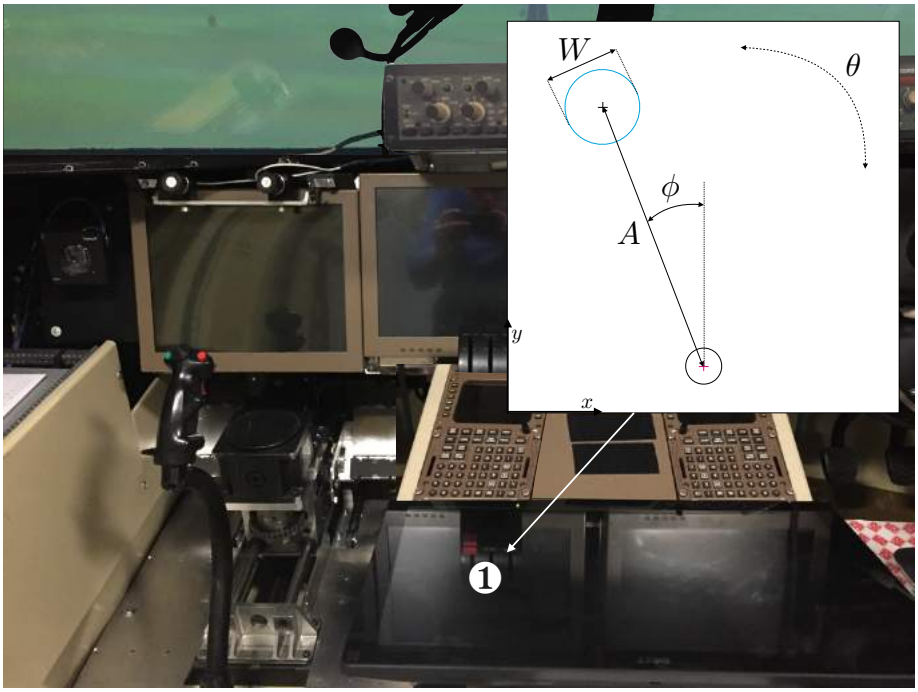


Figure 5: Experimental apparatus for TND experiment, showing an illustration of the touchscreen display and location within the flight deck.

Experiment Checklist and Datalog

N.C.M. (Nout) van Zon
MSc Control & Simulation

Subject number.....

Date and time.....

Checklist MCP	**Checklist CDU**	**Checklist TND**
Before training phase: ☐ **Set seat in forward position** ☐ **Explain course select knob** ☐ **Explain navigation display** ☐ **Discuss control strategy:** ☐ Use right index finger and thumb, position on the knob ☐ Multiple intermediate rotations as required ☐ Smooth, rapid and aimed movement towards target ☐ Accept an under- or overshoot, do not attempt to correct ☐ Aim for a hit-rate of 96% Before measurement phase: ☐ **Expect** 8×24 **movements**, approximately 20 minutes ☐ **Pauses are possible any time**, given that between each trial the clock is paused ☐ **Hit-rate updates** are given between each of the 8 sets	Before training phase: ☐ **Set seat in forward position** ☐ **Check contrast 65** (EXEC + S) ☐ **Explain CDU functionality:** ☐ Alphanumeric keypad (note numbers are opposite to keyboard, similar to iPhone) ☐ Scratchpad (including CLR functionality) ☐ Line select keys ☐ **Discuss control strategy:** ☐ Use right index finger only ☐ Search for necessary keys before starting to type ☐ Keys require some force ☐ Once typing, move promptly ☐ Aim for a hit-rate of 96% Before measurement phase: ☐ **Expect** 2×36 **movements**, approximately 20 minutes ☐ **Pauses are possible any time** ☐ **Hit-rate updates** between sets	Before training phase: ☐ **Set seat in rear position** ☐ **Explain display:** *'cursor'*, cyan target, req. for target hit ☐ **Discuss control strategy:** ☐ Use right index finger only ☐ Smooth, rapid and aimed movement towards the target ☐ Accept an under- or overshoot, do not attempt to correct ☐ Aim for a hit-rate of 96% Before calibration phase: ☐ **Explain display:** magenta target, random position ☐ **Discuss control strategy:** single index-finger tap on target, accuracy is only importance (no timing) Before measurement phase: ☐ **Expect** 1×192 **movements**, approximately 20 minutes ☐ **Pauses are possible any time**

Group (A/B/C) .....

Order of interfaces .....

Interface	**Phase** [T/C/M[1]]	**Latin square**	**Success rate**	**Log file**	**Comments**
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					

[1]Experimental phase, namely **T**raining, **C**alibration (only applicable to finger fitts law) or **M**easurement

Appendix I

Rating Scale Mental Effort

The Rating Scale Mental Effort, as developed by Zijlstra [62], was introduced in the IEEE article and used during the lateral weather avoidance experiment. Due to dutch nationality of all twelve participants, the respective version was used and is presented in Figure I-1. The RSME features a scale from 0 to 150, including nine descriptions of mental effort used to guide the participant.

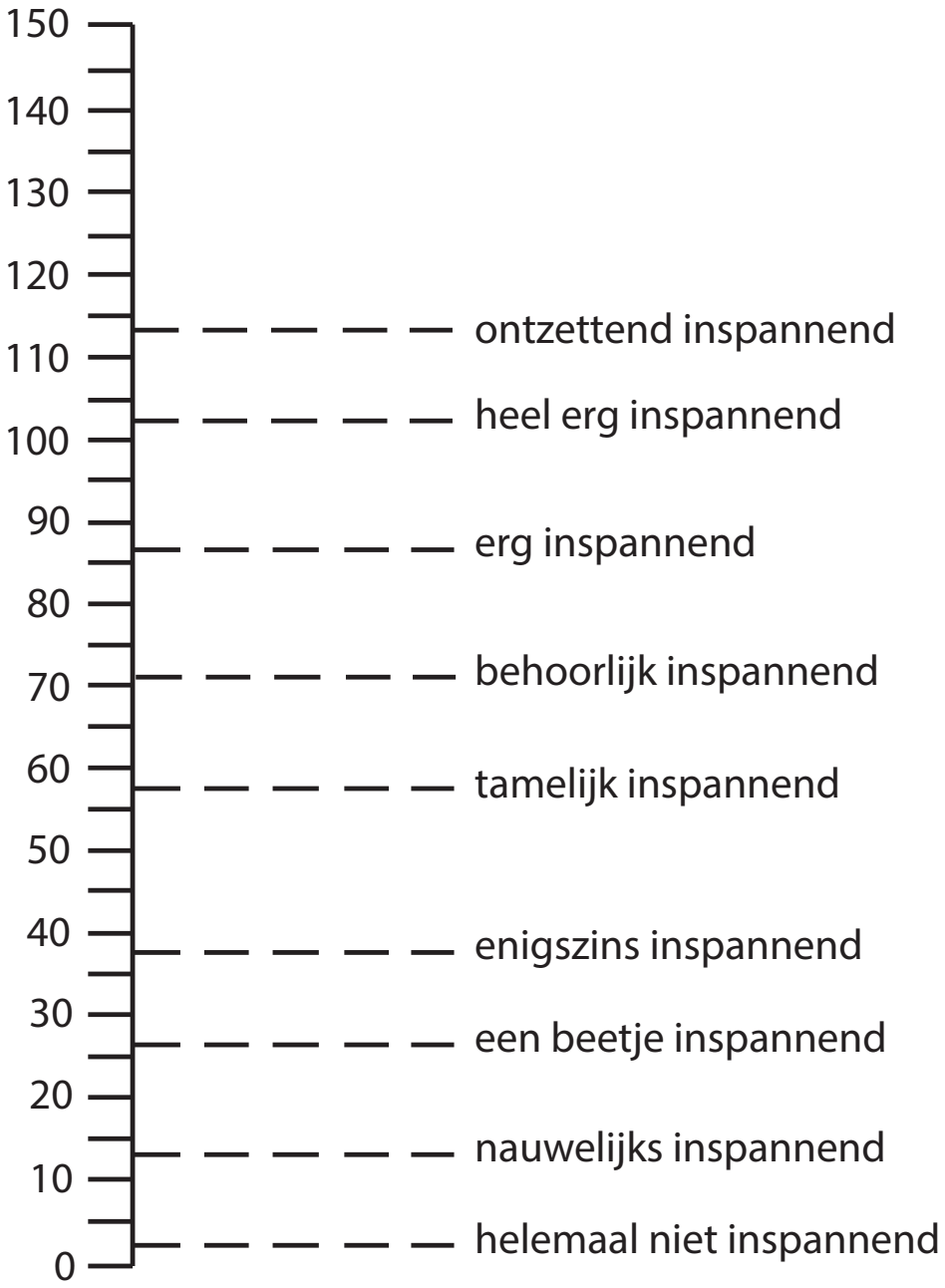


Figure I-1: Rating Scale Mental Effort, using dutch calibration descriptions

Appendix J

Statistical Analysis using SPSS

The results of both experiments were analysed using SPSS, and discussed in both IEEE articles. This appendix presents the direct SPSS output, on which statistical results were based.

J-1 IEEE Article

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
EC	.209	12	.153	.853	12	.040
ET	.204	12	.182	.845	12	.031
MC	.229	12	.082	.873	12	.071
MT	.200	12	.200	.901	12	.162
DC	.208	12	.160	.858	12	.046
DT	.189	12	.200*	.917	12	.262

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Figure J-1: Kolmogorov-Smirnov test of normality for re-route time delay

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser
Difficulty	.086	24.523	2	.000	.522
Interface	1.000	.000	0	.	1.000
Difficulty * Interface	.079	25.381	2	.000	.521

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Epsilon[b] Huynh-Feldt	Lower-bound
Difficulty	.529	.500
Interface	1.000	1.000
Difficulty * Interface	.527	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: Difficulty + Interface + Difficulty * Interface

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Figure J-2: Maulchy's test of sphericity for re-route time delay

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square
Difficulty	Sphericity Assumed	284810.190	2	142405.095
	Greenhouse-Geisser	284810.190	1.045	272549.676
	Huynh-Feldt	284810.190	1.059	269007.576
	Lower-bound	284810.190	1.000	284810.190
Error(Difficulty)	Sphericity Assumed	3950.683	22	179.576
	Greenhouse-Geisser	3950.683	11.495	343.692
	Huynh-Feldt	3950.683	11.646	339.225
	Lower-bound	3950.683	11.000	359.153
Interface	Sphericity Assumed	56.145	1	56.145
	Greenhouse-Geisser	56.145	1.000	56.145
	Huynh-Feldt	56.145	1.000	56.145
	Lower-bound	56.145	1.000	56.145
Error(Interface)	Sphericity Assumed	2271.646	11	206.513
	Greenhouse-Geisser	2271.646	11.000	206.513
	Huynh-Feldt	2271.646	11.000	206.513
	Lower-bound	2271.646	11.000	206.513
Difficulty * Interface	Sphericity Assumed	17.989	2	8.994
	Greenhouse-Geisser	17.989	1.041	17.278
	Huynh-Feldt	17.989	1.054	17.072
	Lower-bound	17.989	1.000	17.989
Error(Difficulty*Interface)	Sphericity Assumed	4056.778	22	184.399
	Greenhouse-Geisser	4056.778	11.452	354.227
	Huynh-Feldt	4056.778	11.591	350.005
	Lower-bound	4056.778	11.000	368.798

Figure J-3: Analysis of variance (ANOVA) for re-route time delay (part 1)

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		F	Sig.
Difficulty	Sphericity Assumed	793.005	.000
	Greenhouse-Geisser	793.005	.000
	Huynh-Feldt	793.005	.000
	Lower-bound	793.005	.000
Error(Difficulty)	Sphericity Assumed		
	Greenhouse-Geisser		
	Huynh-Feldt		
	Lower-bound		
Interface	Sphericity Assumed	.272	.612
	Greenhouse-Geisser	.272	.612
	Huynh-Feldt	.272	.612
	Lower-bound	.272	.612
Error(Interface)	Sphericity Assumed		
	Greenhouse-Geisser		
	Huynh-Feldt		
	Lower-bound		
Difficulty * Interface	Sphericity Assumed	.049	.952
	Greenhouse-Geisser	.049	.839
	Huynh-Feldt	.049	.841
	Lower-bound	.049	.829
Error(Difficulty*Interface)	Sphericity Assumed		
	Greenhouse-Geisser		
	Huynh-Feldt		
	Lower-bound		

Figure J-4: Analysis of variance (ANOVA) for re-route time delay (part 2)

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
EC	.277	12	.012	.851	12	.038
ET	.177	12	.200*	.910	12	.213
MC	.224	12	.098	.827	12	.019
MT	.197	12	.200*	.908	12	.202
DC	.247	12	.042	.869	12	.064
DT	.226	12	.090	.895	12	.136

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Figure J-5: Kolmogorov-Smirnov test of normality for the RMS of roll angle

Test Statistics[a]

N	12
Chi-Square	53.810
df	5
Asymp. Sig.	.000

a. Friedman Test

Figure J-6: Friedman's ANOVA for the RMS of roll angle

Ranks

		N	Mean Rank	Sum of Ranks
ET - EC	Negative Ranks	2[a]	11.50	23.00
	Positive Ranks	10[b]	5.50	55.00
	Ties	0[c]		
	Total	12		
MT - MC	Negative Ranks	8[d]	6.25	50.00
	Positive Ranks	4[e]	7.00	28.00
	Ties	0[f]		
	Total	12		
DT - DC	Negative Ranks	7[g]	6.00	42.00
	Positive Ranks	5[h]	7.20	36.00
	Ties	0[i]		
	Total	12		

a. ET < EC
b. ET > EC
c. ET = EC
d. MT < MC
e. MT > MC
f. MT = MC
g. DT < DC
h. DT > DC
i. DT = DC

Test Statistics[a]

	ET - EC	MT - MC	DT - DC
Z	-1.255[b]	-.863[c]	-.235[c]
Asymp. Sig. (2-tailed)	.209	.388	.814

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.
c. Based on positive ranks.

Figure J-7: Wilcoxon signed rank tests for the RMS of roll angle

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
EC	.180	12	.200*	.865	12	.056
ET	.150	12	.200*	.971	12	.923
MC	.184	12	.200*	.898	12	.149
MT	.125	12	.200*	.934	12	.419
DC	.272	12	.014	.851	12	.038
DT	.120	12	.200*	.973	12	.935

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Figure J-8: Kolmogorov-Smirnov test of normality for secondary task time delay

Test Statistics[a]

N	12
Chi-Square	14.056
df	5
Asymp. Sig.	.015

a. Friedman Test

Figure J-9: Friedman's ANOVA for secondary task time delay

Ranks

		N	Mean Rank	Sum of Ranks
ET - EC	Negative Ranks	5[a]	6.80	34.00
	Positive Ranks	6[b]	5.33	32.00
	Ties	1[c]		
	Total	12		
MT - MC	Negative Ranks	6[d]	5.75	34.50
	Positive Ranks	6[e]	7.25	43.50
	Ties	0[f]		
	Total	12		
DT - DC	Negative Ranks	2[g]	7.50	15.00
	Positive Ranks	10[h]	6.30	63.00
	Ties	0[i]		
	Total	12		

a. ET < EC
b. ET > EC
c. ET = EC
d. MT < MC
e. MT > MC
f. MT = MC
g. DT < DC
h. DT > DC
i. DT = DC

Test Statistics[a]

	ET - EC	MT - MC	DT - DC
Z	-.089[b]	-.353[c]	-1.883[c]
Asymp. Sig. (2-tailed)	.929	.724	.060

a. Wilcoxon Signed Ranks Test
b. Based on positive ranks.
c. Based on negative ranks.

Figure J-10: Wilcoxon signed rank tests for secondary task time delay

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
EC	.214	12	.136	.848	12	.035
ET	.167	12	.200*	.975	12	.958
MC	.183	12	.200*	.886	12	.106
MT	.193	12	.200*	.914	12	.240
DC	.189	12	.200*	.959	12	.769
DT	.211	12	.147	.894	12	.134

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Figure J-11: Kolmogorov-Smirnov test of normality for RSME scores

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser
Difficulty	.309	11.736	2	.003	.591
Interface	1.000	.000	0	.	1.000
Difficulty * Interface	.827	1.905	2	.386	.852

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Epsilon[b] Huynh-Feldt	Lower-bound
Difficulty	.621	.500
Interface	1.000	1.000
Difficulty * Interface	.993	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: Difficulty + Interface + Difficulty * Interface

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Figure J-12: Maulchy's test of sphericity for RSME scores

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square
Difficulty	Sphericity Assumed	10019.372	2	5009.686
	Greenhouse-Geisser	10019.372	1.183	8470.143
	Huynh-Feldt	10019.372	1.242	8065.789
	Lower-bound	10019.372	1.000	10019.372
Error(Difficulty)	Sphericity Assumed	4233.058	22	192.412
	Greenhouse-Geisser	4233.058	13.012	325.321
	Huynh-Feldt	4233.058	13.664	309.790
	Lower-bound	4233.058	11.000	384.823
Interface	Sphericity Assumed	282.823	1	282.823
	Greenhouse-Geisser	282.823	1.000	282.823
	Huynh-Feldt	282.823	1.000	282.823
	Lower-bound	282.823	1.000	282.823
Error(Interface)	Sphericity Assumed	1699.423	11	154.493
	Greenhouse-Geisser	1699.423	11.000	154.493
	Huynh-Feldt	1699.423	11.000	154.493
	Lower-bound	1699.423	11.000	154.493
Difficulty * Interface	Sphericity Assumed	1452.104	2	726.052
	Greenhouse-Geisser	1452.104	1.704	852.001
	Huynh-Feldt	1452.104	1.985	731.527
	Lower-bound	1452.104	1.000	1452.104
Error(Difficulty*Interface)	Sphericity Assumed	1464.844	22	66.584
	Greenhouse-Geisser	1464.844	18.748	78.134
	Huynh-Feldt	1464.844	21.835	67.086
	Lower-bound	1464.844	11.000	133.168

Figure J-13: Analysis of variance (ANOVA) for RSME scores (part 1)

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		F	Sig.
Difficulty	Sphericity Assumed	26.036	.000
	Greenhouse-Geisser	26.036	.000
	Huynh-Feldt	26.036	.000
	Lower-bound	26.036	.000
Error(Difficulty)	Sphericity Assumed		
	Greenhouse-Geisser		
	Huynh-Feldt		
	Lower-bound		
Interface	Sphericity Assumed	1.831	.203
	Greenhouse-Geisser	1.831	.203
	Huynh-Feldt	1.831	.203
	Lower-bound	1.831	.203
Error(Interface)	Sphericity Assumed		
	Greenhouse-Geisser		
	Huynh-Feldt		
	Lower-bound		
Difficulty * Interface	Sphericity Assumed	10.904	.001
	Greenhouse-Geisser	10.904	.001
	Huynh-Feldt	10.904	.001
	Lower-bound	10.904	.007
Error(Difficulty*Interface)	Sphericity Assumed		
	Greenhouse-Geisser		
	Huynh-Feldt		
	Lower-bound		

Figure J-14: Analysis of variance (ANOVA) for RSME scores (part 2)

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	DC	30.6042	12	17.74986	5.12394
	DT	47.0642	12	24.67065	7.12180

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	DC & DT	12	.613	.034

Paired Samples Test

		Paired Differences				
					95% Confidence Interval of the Difference	
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper
Pair 1	DC - DT	-16.46000	19.67522	5.67975	-28.96104	-3.95896

Paired Samples Test

		t	df	Sig. (2-tailed)
Pair 1	DC - DT	-2.898	11	.014

Figure J-15: Dependent t-test for RSME scores during the difficult scenario

J-2 IEEE Companion Article

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser
ID	.000	132.553	90	.012	.325

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Epsilon[b] Huynh-Feldt	Lower-bound
ID	.502	.077

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: ID

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Figure J-16: Maulchy's test of sphericity for the Fitts' law MCP experiment

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F
ID	Sphericity Assumed	36815901.2	15	2454393.42	138.469
	Greenhouse-Geisser	36815901.2	4.329	8503518.39	138.469
	Huynh-Feldt	36815901.2	6.760	5446118.73	138.469
	Lower-bound	36815901.2	1.000	36815901.2	138.469
Error(ID)	Sphericity Assumed	3456413.54	195	17725.198	
	Greenhouse-Geisser	3456413.54	56.283	61410.915	
	Huynh-Feldt	3456413.54	87.880	39330.912	
	Lower-bound	3456413.54	13.000	265877.965	

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Sig.
ID	Sphericity Assumed	.000
	Greenhouse-Geisser	.000
	Huynh-Feldt	.000
	Lower-bound	.000
Error(ID)	Sphericity Assumed	
	Greenhouse-Geisser	
	Huynh-Feldt	
	Lower-bound	

Figure J-17: Analysis of variance (ANOVA) for the Fitts' law MCP experiment

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser
ID	.000	200.788	90	.000	.289

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Epsilon[b] Huynh-Feldt	Lower-bound
ID	.421	.077

Figure J-18: Maulchy's test of sphericity for the Fitts' law CDU experiment

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F
ID	Sphericity Assumed	4161924.74	13	320148.057	37.045
	Greenhouse-Geisser	4161924.74	3.756	1107932.40	37.045
	Huynh-Feldt	4161924.74	5.473	760432.944	37.045
	Lower-bound	4161924.74	1.000	4161924.74	37.045
Error(ID)	Sphericity Assumed	1460540.89	169	8642.254	
	Greenhouse-Geisser	1460540.89	48.834	29908.140	
	Huynh-Feldt	1460540.89	71.150	20527.548	
	Lower-bound	1460540.89	13.000	112349.300	

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Sig.
ID	Sphericity Assumed	.000
	Greenhouse-Geisser	.000
	Huynh-Feldt	.000
	Lower-bound	.000
Error(ID)	Sphericity Assumed	
	Greenhouse-Geisser	
	Huynh-Feldt	
	Lower-bound	

Figure J-19: Analysis of variance (ANOVA) for the Fitts' law CDU experiment

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b] Greenhouse-Geisser
ID	.000	156.542	90	.000	.275

Mauchly's Test of Sphericity[a]

Measure: MEASURE_1

Within Subjects Effect	Epsilon[b] Huynh-Feldt	Lower-bound
ID	.393	.077

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. Design: Intercept
Within Subjects Design: ID

b. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

Figure J-20: Maulchy's test of sphericity for the Fitts' law TND experiment

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F
ID	Sphericity Assumed	5223061.36	15	348204.091	37.193
	Greenhouse-Geisser	5223061.36	2.666	1958876.18	37.193
	Huynh-Feldt	5223061.36	3.419	1527733.01	37.193
	Lower-bound	5223061.36	1.000	5223061.36	37.193
Error(ID)	Sphericity Assumed	1825602.30	195	9362.063	
	Greenhouse-Geisser	1825602.30	34.663	52667.740	
	Huynh-Feldt	1825602.30	44.445	41075.717	
	Lower-bound	1825602.30	13.000	140430.946	

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Sig.
ID	Sphericity Assumed	.000
	Greenhouse-Geisser	.000
	Huynh-Feldt	.000
	Lower-bound	.000
Error(ID)	Sphericity Assumed	
	Greenhouse-Geisser	
	Huynh-Feldt	
	Lower-bound	

Figure J-21: Analysis of variance (ANOVA) for the Fitts' law TND experiment

Appendix K

Fun Facts

- 2,850 grams of chocolate given to participants as a token of appreciation for their contributions.
- Participants flew a total of 22,163.31 kilometers or 29 hours, 41 minutes and 43 seconds. This is the equivalent of roughly 55% of the Earth's circumference.
- Amount of code for the lateral weather avoidance experiment (project name: TouchWxAvoid):
  - 357 files
  - 11,763 blank lines
  - 17,697 commented lines
  - 46,893 lines of (mostly C++) code
  - 188,999,205 bytes
- Amount of code for the Fitts' law experiment (project name: FittsLaw):
  - 55 files
  - 1,773 blank lines
  - 2,966 commented lines
  - 7,679 lines of (mostly C++) code
  - 2,350,337 bytes
- 3,181 lines of MATLAB code necessary to analyse results of both experiments.
- 886,425 bytes of data logged for the Fitts' law experiment.
- 296,020,702 bytes of data logged for the lateral weather avoidance experiment.
- 171 articles, documents, papers, reports and books read as part of this thesis' literature review.

Bibliography

[1] B. Bulfer, *Big Boeing FMC User's Guide*. Leading Edge Publishing, 1991, pp. 1–14,19–27,35–38,61–76,95–104,149–166,185–196,205–.

[2] SESAR Joint Undertaking, *I4d and sesar*, 2014.

[3] M. Carmona, D. Rudinskas, and C. Barrado, "Design of a flight management system to support four-dimensional trajectories", *Aviation*, vol. 19, no. 1, pp. 58–65, 2015, ISSN: 18224180 16487788. DOI: 10.3846/16487788.2015.1015284.

[4] M. Mulder, R. Winterberg, M. M. van Paassen, and M. Mulder, "Direct manipulation interfaces for in-flight four-dimensional navigation planning", *The International Journal of Aviation Psychology*, vol. 20, no. 3, pp. 249–268, 2010, ISSN: 1050-8414. DOI: 10.1080/10508414.2010.487010. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10508414.2010.487010.

[5] H. Huisman, R. Verhoeven, Y. van Houten, and E. Flohr, "Crew interfaces for future atm", pp. 33–40, 1997.

[6] B. J. A. van Marwijk, C. Borst, M. Mulder, M. Mulder, and M. M. van Paassen, "Supporting 4d trajectory revisions on the flight deck: design of a human-machine interface", *The International Journal of Aviation Psychology*, vol. 21, no. 1, pp. 35–61, 2011, ISSN: 1050-8414. DOI: 10.1080/10508414.2011.537559. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/10508414.2011.537559{\%}5Cnhttp://www.tandfonline.com/doi/pdf/10.1080/10508414.2011.537559.

[7] G. Polek, *Boeing 777x cockpit to feature touchscreen displays*, 2016. [Online]. Available: http://www.ainonline.com/aviation-news/business-aviation/2016-07-07/boeing-777x-cockpit-feature-touchscreen-displays (visited on 10/31/2016).

[8] S. Kaminani, "Human computer interaction issues with touch screen interfaces in the flight deck", *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*, 2011, ISSN: 21557195. DOI: 10.1109/DASC.2011.6096098.

[9] B. Shneiderman, "The future of interactive systems and the emergence of direct manipulation", *Behaviour & Information Technology*, vol. 1, no. 3, pp. 237–256, 1982, ISSN: 0144-929X. DOI: 10.1080/01449298208914450.

[10] W. A. Rogers, A. D. Fisk, A. C. McLaughlin, and R. Pak, "Touch a screen or turn a knob: choosing the best device for the job.", *Human factors*, vol. 47, no. 2, pp. 271–88, 2005, ISSN: 0018-7208. DOI: 10.1518/0018720054679452. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/16170938.

[11] E. L. Hutchins, J. D. Hollan, and D. A. Norman, "Direct manipulation interfaces", in *Human-Computer Interaction*, vol. 1, San Diego: Lawrence Erlbaum Associats, Inc., 1985, pp. 311–338, ISBN: 9781848003552. DOI: 10.1207/s15327051hci0104. [Online]. Available: https://www.irjet.net/archives/V2/i6/IRJET-V2I6118.pdf.

[12] M. Thurber, *Touchscreens clean up gulfstream symmetry flight deck*, 2015. [Online]. Available: http://www.ainonline.com/aviation-news/business-aviation/2015-01-02/touchscreens-clean-gulfstream-symmetry-flight-deck (visited on 10/31/2016).

[13] GARMIN, *Garming g5000*. [Online]. Available: https://buy.garmin.com/en-US/US/in-the-air/general-aviation/flight-decks/g5000-/prod90821.html.

[14] A. Degani, E. A. Palmer, and K. G. Bauersfeld, ""Soft"controls for hard displays: still a challenge", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 36, pp. 52–56, 1992, ISSN: 01635182. DOI: 10.1177/154193129203600114. [Online]. Available: http://pro.sagepub.com/content/36/1/52.short.

[15] M. Mertens, H. Damveld, and C. Borst, "An avionics touch screen based control display concept",

[16] S. Dodd, J. Lancaster, A. Miranda, S. Grothe, B. DeMers, and B. Rogers, "Touch screens on the flight deck: the impact of touch target size, spacing, touch technology and turbulence on pilot performance", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 58, no. 1, pp. 6–10, 2014, ISSN: 1541-9312. DOI: 10.1177/1541931214581002. [Online]. Available: http://pro.sagepub.com/content/58/1/6.short{\%}5Cnhttp://pro.sagepub.com/lookup/doi/10.1177/1541931214581002.

[17] J. A. Ballas, C. L. Heitmeyer, and M. A. Pérez, "Evaluating two aspects of direct manipulation in advanced cockpits", *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '92*, pp. 127–134, 1992. DOI: 10.1145/142750.142770. [Online]. Available: http://dl.acm.org/citation.cfm?id=142750.142770.

[18] F. B. Bjorneseth, M. D. Dunlop, and E. Hornecker, "Strathprints institutional repository assessing the effectiveness of direct gesture interaction for a safety critical maritime application", vol. 70, pp. 729–745, 2012.

[19] C. Forlines, D. Wigdor, C. Shen, and R. Balakrishnan, "Direct-touch vs. mouse input for tabletop displays", *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*, p. 647, 2007, ISSN: 10892680. DOI: 10.1145/1240624.1240726. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1240624.1240726.

[20] N. A. Stanton, C. Harvey, K. L. Plant, and L. Bolton, "To twist, roll, stroke or poke? a study of input devices for menu navigation in the cockpit.", *Ergonomics*, vol. 0139, no. February 2013, pp. 37–41, 2013, ISSN: 1366-5847. DOI: 10.1080/00140139.2012.751458. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/23384222.

[21] R. W. Soukoreff and I. S. MacKenzie, "Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci", *International Journal of Human Computer Studies*, vol. 61, no. 6, pp. 751–789, 2004, ISSN: 10715819. DOI: 10.1016/j.ijhcs.2004.09.001.

[22] R. Jagacinski and J. Fisch, "Information theory and fitts' law", in *Control Theory for Humans*, 1997, ch. 3, pp. 17–26.

[23] I. S. MacKenzie, "Fitts' law as a research and design tool in human-computer interaction", *Human-Computer Interaction*, vol. 7, no. 1, p. 48, 1992, ISSN: 0737-0024. DOI: 10.1207/s15327051hci0701_3. [Online]. Available: http://portal.acm.org/citation.cfm?id=1461857.

[24] G. Stuyven, H. Damveld, and C. Borst, "Concept for an avionics multi touch flight deck", *SAE International Journal of Aerospace*, vol. 5, no. 1, pp. 164–171, 2012, ISSN: 19463855. DOI: 10.4271/2012-01-2120.

[25] Boeing Commercial Aircraft, *777-200LR Operations Manual*. Everett, 2012.

[26] Flight Safety Foundation, "Boeing 757 cfit accident at cali, colombia, becomes focus of lessons learned", *Flight Safety Digest*, pp. 1–31, 1998.

[27] M. F. Stoelen and D. L. Akin, "Assessment of fitts' law for quantifying combined rotational and translational movements.", *Human factors*, vol. 52, no. 1, pp. 63–77, 2010, ISSN: 0018-7208. DOI: 10.1177/0018720810366560.

[28] X. Bi, Y. Li, and S. Zhai, "Ffitts law: modeling finger touch with fitts' law", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, p. 1363, 2013. DOI: 10.1145/2470654.2466180. [Online]. Available: http://dl.acm.org/citation.cfm?id=2470654.2466180.

[29] X. Zeng, a. Hedge, and F. Guimbretiere, "Fitts' law in 3d space with coordinated hand movements", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, no. 1, pp. 990–994, 2012, ISSN: 1541-9312. DOI: 10.1177/1071181312561207. [Online]. Available: http://pro.sagepub.com/lookup/doi/10.1177/1071181312561207.

[30] I. S. MacKenzie and W. Buxton, "Extending fitts' law to two-dimensional tasks introduction since the advent of direct manipulation", *ACM CHI'92 Conference*, pp. 219–226, 1992, ISSN: 0897915135. DOI: 10.1145/142750.142794.

[31] W. R. Soukoreff and I. S. MacKenzie, "Theoretical upper and lower bounds on typing speed using a stylus and a soft keyboard", *Behaviour & Information Technology*, vol. 14, no. 6, pp. 370–379, 1995, ISSN: 0144-929X. DOI: 10.1080/01449299508914656. [Online]. Available: http://dx.doi.org/10.1080/01449299508914656.

[32] R. I. Cook, "Verite, abstraction, and ordinateur systems in the evolution of complex process control", *Proceedings Third Annual Symposium on Human Interaction with Complex Systems. HICS'96*, vol. 1970, pp. 38–42, 1996. DOI: 10.1109/HUICS.1996.549490. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=549490.

[33] G. Lintern, T. Waite, and D. A. Talleur, "Functional interface design for the modern aircraft cockpit", *International Journal of Aviation Psychology*, vol. 9, no. 3, pp. 203–223, 1999, ISSN: 10508414. DOI: 10.1207/s15327108ijap0903.

[34] J. Han, *The radical promise of the multi-touch interface*, 2006. [Online]. Available: https://www.ted.com/talks/.

[35] M. Yeh, Y. J. Jo, C. Donovan, and S. Gabree, "Human factors considerations in the design and evaluation of electronic flight bags (efbs)", p. 206, 2003.

[36] S. Kaminani, "Touch screen technology in flight deck-how far is it helpful?", *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*, vol. 3000, no. Figure 2, pp. 1–5, 2012, ISSN: 21557195. DOI: 10.1109/DASC.2012.6382965.

[37] ——, "Foundations for deriving empirical model to predict selection times with soft keyboard in the flight deck", *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*, pp. 1–5, 2013, ISSN: 21557209. DOI: 10.1109/DASC.2013.6712596.

[38] B. Shneiderman, "Direct manipulation for comprehensible, predictable and controllable user interfaces", *Proceedings of the 2Nd International Conference on Intelligent User Interfaces*, pp. 33–39, 1997, ISSN: 00010782. DOI: 10.1145/238218.238281. [Online]. Available: http://doi.acm.org/10.1145/238218.238281.

[39] D. B. Beringer and J. G. Peterson, "Underlying behavioral parameters of the operation of touch-input devices: biases, models, and feedback", *The Journal of the Human Factors and Ergonomics Society*, vol. 27, no. 4, pp. 445–458, 1985. DOI: 10.1177/001872088502700408.

[40] K. Cardosi and E. Murphy, "Human factors in the design and evaluation of air traffic control systems", Federal Aviation Administration, Tech. Rep., 1995.

[41] J. R. Lewis, "Ibm technical report: literature review of touch-screen research from 1980 to 1992", IBM, Boca Raton, Florida, Tech. Rep., 1993.

[42] NASA, "Touch-sensitive display design requirements", in *Man-Systems Integration Standards*, 1995, p. 9.3.3.4.7.

[43] A. Sears and B. Shneiderman, "High precision touchscreens:design strategies and comparisons with a mouse", *International Journal of Man–Machine Studies*, vol. 34, pp. 593–613, 1991, ISSN: 09600779. DOI: 10.1016/S0960-0779(02)00434-4. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0960077902004344.

[44] J. M. Flach, F. Tanabe, K. Monta, J. Rasmussen, and K. J. Vicente, "An ecological approach to interface design", *Proceedings of the ...*, pp. 295–299, 1998, ISSN: 1071-1813. DOI: 10.1177/154193129804200324. [Online]. Available: http://pro.sagepub.com/content/42/3/295.short.

[45] J. Rasmussen, "Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models", *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-13, no. 3, pp. 257–266, 1983, ISSN: 21682909. DOI: 10.1109/TSMC.1983.6313160.

[46] A. Rosenbluth, N. Wiener, and J. Bigelow, "Behaviour, purpose and teleogogy", in *Philosophy of Science*, Volume 10, 1943, pp. 18–24.

[47] M. Polanyi, "Personal knowledge. towards a post-critical philosophy", in, London: Routledge & Kegan Paul, 1958.

[48] A. Bisantz and K. Vicente, "Making the abstraction hierarchy concrete", *International Journal of Human Computer Studies1*, 1992.

[49] C. Borst, *Ecological Approach to Pilot Terrain Awareness*. 2009, ISBN: 9789053351963. [Online]. Available: http://repository.tudelft.nl/view/ir/uuid:6bcb6aa4-2a0f-4ca2-93cb-e2fe7abdbd41/.

[50] J. Rasmussen, A. Pejtersen, and L. Goodstein, *Cognitive Systems Engineering*. New York: Wiley-Interscience, 1994.

[51] C. Borst, *Ae4308 supervisory control & cognitive systems lecture slides: cognitive systems engineering*, 2015.

[52] J. Rasmussen, *Information Processing and Human-Machine Interaction*, Series Vol. New York: Elsevier Science Publishing Co., Inc., 1986.

[53] J. Flach, *Guest lecture: aerospace human-machine systems*, Delft, 2016.

[54] K. J. Vicente and J. Rasmussen, "The ecology of human-machine systems ii: mediating 'direct perception' in complex work domains", *Ecological Psychology*, vol. 2, no. 3, pp. 207–249, 1990, ISSN: 1040-7413. DOI: 10.1207/s15326969eco0203_2. [Online]. Available: http://www.informaworld.com/10.1207/s15326969eco0203{\%}7B{_}{\%}7D2.

[55] C. Borst, J. M. Flach, and J. Ellerbroek, "Beyond ecological interface design: lessons from concerns and misconceptions", *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 2, pp. 164–175, 2015, ISSN: 21682291. DOI: 10.1109/THMS.2014.2364984.

[56] National Transportation Safety Board, "Loss of thrust in both engines after encountering a flock of birds and subsequent ditching on the hudson river", p. 213, 2009. DOI: NTSB/AAR-10/03.

[57] J. Ellerbroek, M. M. van Paassen, and M. Mulder, "Evaluation of a seperation assistance display in a multi-actor experiment", *Proceedings of the 16th International Symposium on Aviation Psychology*, 2011.

[58] N. Dinadis and K. J. Vicente, "Designing functional visualizations for aircraft systems status displays", *International Journal of Aviation Psychology*, vol. 9, no. 3, pp. 203–223, 1999, ISSN: 10508414. DOI: 10.1207/s15327108ijap0903.

[59] W. Scott, "United training stresses cockpit discipline", *Aviation Week & Space Technology*, pp. 50–51, 1995.

[60] Bureau d'Enquêtes et d'Analyses pour la sécurité de l'aviation Civile, "Final report on the accident on 1st june 2009 to the airbus a330-203 registered f-gzcp operated by air france flight af 447 rio de janeiro - paris", no. June 2009, pp. 1–223, 2012, ISSN: 0029-6570. DOI: 10.2172/875800. [Online]. Available: https://www.bea.aero/docspa/2009/f-cp090601.en/pdf/f-cp090601.en.pdf.

[61] A. Field and G. Hole, *How to Design and Report Experiments*, First. London: SAGE Publications Ltd, 2003, ISBN: 9780761973829.

[62] F. Zijlstra, *Efficiency in work behaviour: A design approach for modern tools*. Delft: Delft University Press, 1993, pp. 1–186.