Display Formats for Smart Glasses to Support Pilots in General Aviation

Anzeigeformate für Smartglasses zur Unterstützung von Piloten in der allgemeinen Luftfahrt

Paul Manuel Haiduk

Dissertation D17
Darmstadt 2017

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Maschinenbau
Institut für Flugsysteme und Regelungstechnik

Display Formats for Smart Glasses to Support Pilots in General Aviation

Anzeigeformate für Smartglasses zu Unterstützung von Piloten in der allgemeinen Luftfahrt

Vom Fachbereich Maschinenbau an der
Technischen Universität Darmstadt
zur
Erlangung des Grades eines Doktor der Philosophie (Dr.phil.)
genehmigte

Dissertation

vorgelegt von

Paul Manuel Haiduk, M.Sc.

aus Frankfurt am Main

Berichterstatter: Prof. Dr.-Ing. Uwe Klingauf
Mitberichterstatter: Prof. Dr. rer. medic. Josef Wiemeyer

Tag der Einreichung: 02. Mai 2017
Tag der mündlichen Prüfung: 28. Juni 2017

Darmstadt 2017

D 17

Abstract

This dissertation develops and evaluates various display formats for smart glasses which could provide information to support pilots in general aviation on flights under visual flight rules. The aim of a new display format is the reduction of pilot task load and the increase of pilot situation awareness.

Under visual flight rules, pilots apply rules known as see-and-avoid. However, the monitoring of airspace conflicts with information acquisition from head-down instrumentation. Conventional displays may drive the pilot's attention head-down at the expense of monitoring the scene outside, which has the potential to lead to breakdowns in task management. One of the main reasons for accidents is human error (84% in GA), which is associated with an increased workload resulting in a loss of situation awareness. One way to prevent accidents is to reduce workload to an adequate level and to increase situation awareness; the projection of supporting information in the head-up area could be one to do so. A proposed solution is the use of smart glasses, which project the most important information directly into the field of view. This dissertation is the only research work in the field that scientifically investigates the feasibility and utility of display formats for smart glasses for use in the cockpit of general aviation.

The EPSON Moverio BT-200 smart glasses are selected based on set requirements for integration within the research flight simulator Diamond DA 40-180 at the Institute of Flight Systems and Automatic Control. Four different display formats are implemented and tested with regard to subjective- workload and usability in a preliminary simulator study with $N = 7$ participants. The results of the preliminary investigation show that the developed Primary Flight Display format has the highest usability and is therefore selected for further development.

The Primary Flight Display format is further developed with consideration of the user feedback from the preliminary study. A new flight guidance symbology for lateral guidance, called Lateral Guidance Line (LGL), is designed and added to the format. A magenta colored line in the center of the format supports the pilot in maintaining track. The lateral guidance symbology is designed to show when to initiate a turn and when the turn should be completed in order to minimize deviations from a desired track (e.g. traffic pattern). In the final evaluation, the LGL format is evaluated with $N = 20$ pilots. In addition to assessing the subjective- usability and workload, the lateral deviations from a given flight path are recorded.

Spatial awareness is operationalized through eye-tracking and a secondary reaction task using visual signals. Pilots fly twice, once with the LGL format on the smart glasses and once with conventional instruments without smart glasses at two different airfields in a balanced order.

The effectiveness of the Lateral Guidance Line display format can be confirmed. The lateral deviations from the target trajectory are significantly lower in the group using the format compared to the group using conventional instruments (without smart glasses), while task load remained the same. An increase in *eyes-out* time as well as fewer missed signals on the secondary task proves the potential of the display format to increase spatial awareness compared to conventional instruments. The subjective suitability of the Lateral Guidance Line format was rated 73 (on a scale of 0 to 100) which corresponds to a good subjective usability and is not significantly different from the evaluations of the previously implemented prototypes.

Overall, the results of the investigation show that smart glasses have the potential to support pilots in general aviation and to potentially reduce accident rates. Only few hardware challenges remain in the development of this format. The work draws on recommendations from the feedback of various general aviation interest groups and points out future research questions.

Kurzfassung

Die vorliegende Dissertation entwickelt und evaluiert verschiedene Anzeigeformate auf Datenbrillen, dessen Informationen den Piloten in der allgemeinen Luftfahrt bei Flügen nach Sichtflugregeln unterstützen. Ziel eines neuen Anzeigeformates ist die Reduktion der Arbeitsbelastung und die Erhöhung des Situationsbewusstseins des Piloten.

Bei Flügen nach Sichtflugregeln gilt das Prinzip Sehen und gesehen werden. Die Luftraumbeobachtung während des Flugs steht dabei im Konflikt mit der Beobachtung der Instrumente, welche im Cockpitinnenbereich angeordnet sind. Dies führt vor allem in der allgemeinen Luftfahrt nach Sichtflugregeln wiederholt zu Unfällen. Eine der häufigsten Unfallursachen sind menschliche Fehler (56%), die mit erhöhter Arbeitsbelastung und folglich einem Verlust des Situationsbewusstseins verbunden sind. Eine Möglichkeit, die Arbeitsbelastung zu reduzieren und das Situationsbewusstsein zu erhöhen, ist die Darstellung von unterstützenden Informationen im head-up Bereich. Eine Realisierungsmöglichkeit ist die Nutzung von Datenbrillen, welche die wichtigsten Informationen direkt in das Blickfeld projizieren. Die vorliegende Dissertation ist bisher die einzige Forschungsarbeit auf dem Gebiet zur Untersuchung der Umsetzbarkeit und des Nutzens von Anzeigeformaten auf Datenbrillen zur Anwendung im Cockpit der allgemeinen Luftfahrt.

Anhand von Anforderungen für die Integration einer Datenbrille innerhalb eines Forschungssimulators Diamond DA 40-180 am FSR wird die Datenbrille EPSON Moverio BT-200 ausgewählt. Vier verschiedene Anzeigeformate werden implementiert und in einer Voruntersuchung durch Probanden mit $N = 7$ hinsichtlich subjektiver Gebrauchstauglichkeit und Arbeitsbelastung getestet. Die Ergebnisse der Voruntersuchung zeigen, dass das Format Primary Flight Display in der Gebrauchstauglichkeit die besten Ergebnisse erzielt. Daher wird dieses für die Weiterentwicklung ausgewählt.

Das Format Primary Flight Display wird mit Berücksichtigung des Nutzerfeedbacks aus der Vorstudie weiterentwickelt. Eine neuartige Flugführungsanzeige, Lateral Guidance Line, zur lateralen Führung wird gestaltet und in das Format eingebunden. Die magentafarbene, pfeilartige Darstellung zeigt dem Piloten an, wann eine Kurve ein- und auszuleiten ist, um die laterale Abweichung zu einer Solltrajektorie (z.B Platzrunde) zu minimieren.

In der Hauptuntersuchung wird das Lateral Guidance Line Format mit $N = 20$ Piloten im Vergleich zu konventionellen Instrumenten evaluiert. Neben der

subjektiven Gebrauchstauglichkeit und Arbeitsbelastung wird die unterstützende Funktion der Flugführungsanzeige anhand der lateralen Abweichungen zu einem vorgegebenen Flugpfad erfasst. Räumliches Situationsbewusstsein wird operationalisiert über die Aufzeichnung des Blickverhaltens mittels Blickerfassung und der Aufmerksamkeit auf einer visuellen Zweitaufgabe. Piloten fliegen zweimal; einmal mit dem Lateral Guidance Line Format auf der Datenbrille und einmal mit konventionellen Instrumenten ohne Datenbrille an zwei unterschiedlichen Flugplätzen in balancierter Reihenfolge.

Die Effektivität des Anzeigeformats Lateral Guidance Line kann in der Hauptuntersuchung belegt werden. Die lateralen Abweichungen von der Solltrajektorie fallen in der Gruppe mit dem Lateral Guidance Line Format signifikant geringer aus als in der Gruppe mit konventionellen Instrumenten (ohne Datenbrille), bei ähnlich hoher Arbeitsbelastung. Eine signifikante Erhöhung der Blickzeit nach draußen sowie eine geringere Anzahl verpasster Signale in der visuellen Zweitaufgabe belegen das Potential, das räumliche Situationsbewusstsein des Anzeigeformats zu erhöhen.

Insgesamt zeigen die Untersuchungsergebnisse auf, dass Datenbrillen zukünftig das Potential haben, Piloten in der allgemeinen Luftfahrt zu unterstützten und Unfälle zu vermeiden. Einschränkungen seitens der Hardware müssen durch die Hersteller behoben werden. Die Arbeit leitet aus den Ergebnissen mehrere Empfehlungen für unterschiedliche Interessengruppen in der allgemeinen Luftfahrt ab und zeigt weitere Forschungsfragen auf.

Für meinen Großvater, der mein Interesse für Wissenschaft und Technik geweckt hat, den ich aber nie richtig kennenlernen konnte.

Danksagung

Ich möchte mich bei allen Personen, die zur Ermöglichung dieser Arbeit beigetragen haben herzlichst bedanken. Voran den Betreuern dieser Arbeit Prof. Dr-Ing. Uwe Klingauf und Prof. Dr. rer. medic. Josef Wiemeyer für die wissenschaftliche Begleitung und Unterstützung.

Mein aufrechter Dank gilt meinen Kollegen, die mich über die Zeit am *Institut für Flugsysteme und Regelungstechnik* sowie in der Graduiertenschule *Topologie der Technik* begleitet haben und damit den wissenschaftlichen Austausch ermöglichten. Das Teilen von Wissen zwischen den Disziplinen verdeutlichte mir den bekannten Spruch, dass Wissen das einzige Gut ist, was sich vermehrt, wenn man es teilt.

Bedanken möchte ich mich bei denjenigen, die diese Arbeit lektorierten, und die mit Rat und Tat zur Seite standen. Insbesondere meinem Bürokollegen und Freund Dr.-Ing. Christoph Santel, der mich ans Fliegen herangeführt und meine Arbeit herausgefordert hat. Namentlich möchte ich mich ebenso bei Prof. Dr. Joachim Vogt, Paul Gerber, Dr.-Ing. Jendrick Westphal, Martin Nowara, Dr. Patrick Liborius und Florian Müller für den wissenschaftlichen Austausch bedanken.

Mein Dank gilt ebenso meinen Studierenden, welche mich durch Ihre Arbeiten zu meinem Thema unterstützt haben. Besonders bedanken möchte ich mich bei Arvid Lange, Christian Spatz und Daniel Vancura.

Weiterhin gilt Dank denjenigen, die an den Versuchen und Befragungen dieser Arbeit teilgenommen haben und mir Einblicke in die Tätigkeiten eines Piloten und den Abläufen im Cockpit gegeben haben. Mein Dank gilt auch den Fluglehrern, die mich zum Privatpiloten ausgebildet haben, insbesondere Peter Eckhardt.

Nicht zuletzt möchte ich mich bei meiner Familie bedanken, die mir diesen Weg ermöglicht haben.

Der größte Dank, gilt meiner Frau Britta. Sie trug in ungeahntem Umfang zu dieser Arbeit bei. Seit Beginn meines Studiums in Darmstadt hielt sie meine Begeisterung für die Psychologie aufrecht und lenkte meinen Blick in die richtige Richtung. Ohne Ihre Motivation, ihren Rat und ihre Unterstützung hätte ich diese Arbeit nicht fertigstellen können. Ich möchte mich bei ihr von tiefstem Herzen bedanken.

Darmstadt, im Juli 2017 *Paul Haiduk*

Contents

Nomenclature

Symbols

Notation	Description
L	Luminance
T	Transmission
M	Mean of sample
N	Total number of participants in the sample under study
n	Number of participants in a subgroup
SD	Standard deviation of sample
D	distance
δ	angular distance
κ_i	Constant
lon	Longitude
lat	Latitude
p	Statistical significance
Φ	Bank angle
Ψ	Bearing, heading
R	Earth's radius
r^2	Coefficient of determination
Θ	Pitch angle
t	Time
w	Weight for calculation of TLX score
α	Predefined significance level of test
η^2	η^2 effect size measure

Subscripts

Notation	Description
$_{ct}$	Cross-track distance

Continued on next page

Continued from previous page

Notation	Description
ref	Landing reference speed
NE	Never exceed speed
1	Bearing from or to start point
2	Bearing from or to end point
3	Bearing from or to current aircraft position
$combiner$	Transmission of combiner
$shades$	Transmission of shades
$canopy$	Transmission of canopy
$ambient$	Ambient luminance
HMD	Luminance of HMD

Acronyms

AGL above ground level

AIP Aeronautical Information Publication

AIXM Aeronautical Information Exchange Model

ANOVA analysis of variance

AOI area of interest

AOPA Aircraft Owner and Pilot Association

AR augmented reality

ATC air traffic control

ATPL airline transport pilot license

AV augmented virtuality

COTS commercial off-the-shelf

CPL commercial pilot license

CS coordinate system

DFS Deutsche Flugsicherung

DOF degree of freedom

DV dependent variable

EASA European Aviation Safety Agency

EFB electronic flight bag

FAA Federal Aviation Administration

FTD flight training device

FTE flight technical error

FMS flight management system

FOV field of view

FSR	Institute of Flight Systems and Automatic Control
GA	general aviation
GPS	Global Positioning System
HUD	head-up display
HDD	head-down display
HMD	head-mounted display
HMI	human-machine interface
HSD	honest significant difference
HSI	horizontal situation indicator
HTS	head tracking system
IAE	information access effort
ICAO	International Civil Aviation Organization
IFR	instrument flight rules
IR	instrument rating
ILS	instrument landing system
IPD	interpupillary distance
IR	infrared
LAPL	light aircraft pilot license
LED	light-emitting diode
LGL	Lateral Guidance Line
LOC-I	loss of control in flight
LOSA	loss of situation awareness
MFD	multi-function display
MR	mixed reality
MTOM	maximum takeoff mass
NASA	National Aeronautics and Space Administration
OST	optical see-through
OTW	out-the-window
PED	personal electronic device
PFD	primary flight display
PPL	private pilot license
PTT	push to talk
RPM	revolutions per minute
SA	situation awareness
SUS	System Usability Scale
TLX	Task Load Index
UDP	User Datagram Protocol
UL	ultra-light
VC	virtuality continuum
VFR	visual flight rules
VMC	visual meteorological conditions
VR	virtual reality
VST	video see-through

1 Introduction

Most general aviation (GA) pilots fly under visual flight rules (VFR)[1]. They apply the so-called see-and-avoid principle; pilots must look outside to avoid other traffic and obstacles. Pilots need to continuously monitor flight information, traditionally displayed on conventional (*steam*) gauges, compare track deviations with printed approach charts, and manually maintain the correct airplane attitude all at the same time. This can be demanding particularly in high task load situations. There are several causes that are more likely to occur than others. Forty-seven percent of all GA accidents can be attributed to loss of control, the most common underlying cause in the last 10 years [34]. A proposed method for reducing the number of accidents is to reduce pilot workload [2] to an adequate level and to provide supporting assistance systems to the pilot.

Pilot assistance systems aim to provide easily accessible information for pilots in order to enhance situation awareness (SA) and to lower task load (see Section 2.3 for an explanation on situation awareness). "The problem with today's systems is not a lack of information, but finding what is needed when it is needed", states Endsley [39, p. 1]. In addition, it must be ensured that the information is provided in a way that is "useable cognitively as well as physically" [p. 1]. Popular pilot assistance systems for GA on the market are based on tablet applications, also known as electronic flight bags (EFBs) ([59**?**]. However, systems fail to keep the pilot *eyes-out*. These displays may drive the pilot's attention head-down at the expense of monitoring the scene outside [56], which has the potential to lead to breakdowns in task management [152]. Above all, when pilots switch their attention between inside the cockpit and the outside view, vision needs to constantly adjust to changing light as well as to different distances [35].

In order to establish a high level of SA, it is necessary to increase the time spent with *eyes-out*. It is assumed that smart glasses have the potential to support pilots while flying. Smart glasses project information directly into the field of view and

[1] Based on the issue of new pilot licenses in Germany between 2011 and 2015, it can be concluded that 4.1% of issued private pilots licenses on aircraft (PPL-A) included entitlement to fly under instrument flight rules (IFR) [97]. Therefore 95.9% of issued private pilots licenses permit to fly under VFR only.

[2] This thesis uses the term *task load* to indicate imposed work and the term *workload* to indicate the humans response.

thus, could help reduce the time spent monitoring conventional head-down instruments. What is needed is a display that combines all task relevant information, including warnings, and projects the information directly into the field of view.

A broad field of research on head-up displays for military [54; 117] as well as commercial applications [7; 93] exists in the literature. Until now, however, to the best of the author's knowledge, no plausible use cases have been addressed in academic research for light aircraft under VFR. There are no known publications within research that focus on commercially available smart glasses as a platform.

1.1 Aim of this Thesis

The aim of this thesis is to develop pilot assistance systems for general aviation, which facilitate access to information and enhance SA especially under high-task load situations. For this purpose, (see-through) head-mounted displays (HMDs), also called smart glasses, are considered as an enabling technology. Furthermore, the purpose of this thesis is to provide scientific insight and guidance for future developments of display formats on smart glasses. The methodological focus is on human factors research and the human-technology interaction.

In order to better understand the activity of pilots and their interaction with the cockpit and the instruments, techniques such as interviews, questionnaires and observations were applied. This results in a user-centered design process. Suitable use cases were defined by focusing on the user's needs. Technically feasible display formats and their application scenarios are identified and converted into prototypical systems. The potential of smart glasses for pilots under VFR is evaluated. In terms of human-system integration, a large number of questions arise from the human factor perspective. This thesis tries to answer the following questions:

- Which display formats are suitable to support the pilot with regard to the SA and to keep the task load at an appropriate level for the flight situation?
- How does the performance change while flying with smart glasses?
- What sort of information is most crucial to pilots?
- How can this information be displayed as effectively and intuitively as possible for the user?

These questions are answered by empirical experiments as well as analyses by experts, observations and surveys. In a simulator experiment the new technology was evaluated regarding to SA, workload and performance. Objective indicators and measurement methods are developed.

1.2 Structure of this Thesis

This thesis consists of seven chapters. The main structure is visualized in Figure 1.1. The introduction in Chapter 1 familiarizes the area of investigation, namely GA, and provides initial motivation. The chapter points towards the necessity for reducing task load and providing better task relevant information to reduce accidents in general aviation.

Chapter 2 provides insight into cockpit instrumentation and current display systems. The computerization of avionics is exemplified by emerging technologies. An overview on HMDs is given, with regards to technical characteristics and operation purpose. A taxonomy on smart glasses, as a subcategory of HMDs, is proposed and requirements for their integration in a simulator is presented. Available commercial off-the-shelf (COTS) smart glasses are compared. The smart glasses BT-200 by Epson are described. Furthermore, relevant models and methods from aviation psychology and human factors are presented, which will assist in understanding underlying principles and design choices.

Chapter 3 details the design and implementation process of three different display formats. Pilots helped to design the display formats by reviewing the concept at different stages, making it a user-centered design process. After a preliminary evaluation of three implemented display formats regarding subjective workload and subjective usability, the primary flight display (PFD) display format was selected for further development and final evaluation.

Chapter 4 presents the further development of the PFD display format with consideration of the user's feedback from the preliminary study. The format's new features are explained. A lateral guidance symbology was added, resulting in the so-called Lateral Guidance Line (LGL) display format. Four hypotheses were postulated that covered the performance, spatial awareness, subjective usability and subjective workload. Theses hypotheses will be tested in the evaluation.

Chapter 5 describes the methodology for the evaluation of the LGL display format in a flight simulator. The equipment used, test procedure and scenarios are explained. Pilots flew multiple traffic patterns at an unfamiliar airfield either with smart glasses or with conventional printed approach charts. Recorded measurements included flight technical error (FTE), *eyes-out* time and visual attention. Pilots rated the workload and perceived usability.

Chapter 6 presents the results of the evaluation. This chapter will start with the descriptive analysis of the recorded data followed by inferential statistics methods for testing each hypothesis. After each section the results will be discussed.

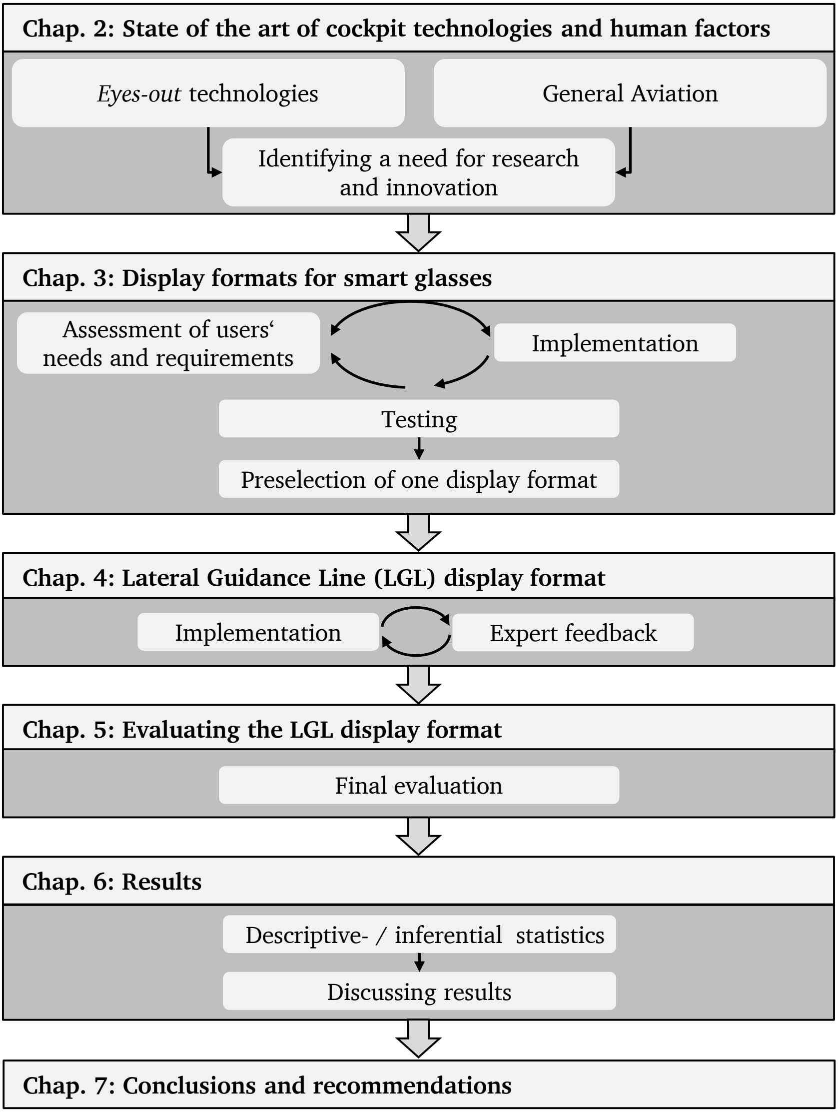


Figure 1.1.: Structure of this thesis.

Chapter 7 summarizes this thesis. Recommendations for relevant stakeholders are given. An outlook to further possible research on this topic shows which further developments should be considered.

1.3 Area of Investigation

This work mainly focuses on GA. The research results, in particular the design principles of the display formats, can be transferred to commercial aviation. In the past, developments in general aviation have emerged as important innovation drivers for other aviation sectors [4].

In addition, there are numerous overlaps between general- and commercial aviation. Improving processes through innovation for the pilots of general aviation could have a positive impact on the air transport system as a whole. This is particularly evident for airspace infringements of controlled airspace. Airspace infringements represent not only an increased task load for air traffic controllers, but a potential danger for all flight movements.

In some respects, the overall safety of the air transport system is determined by the proper functioning of the weakest link. This may often turn out to be the single-crew, low-technology, less rigorously trained private pilot than the multi-crew, high-technology extensively trained air commercial aviation pilot.

To increase flight safety and process efficiency, new technologies are constantly considered for their practical applicability. These may also include head-worn displays, such as smart glasses. Smart glasses could be applied to multiple other areas besides the pilot group. In literature, aviation related use cases comprise maintenance [107], ground / cabin personnel [106] or air traffic control [121].

Within the professional context, smart glasses could be put into use whenever the user needs to work hands-free or requires additional information, especially in high task load situations. Moreover, situations are reasonable in which the overlapping information could benefit the users mental image of the situation.

The need for technical solutions to support the general aviation pilot is comparatively larger than in commercial aviation. Smart glasses, which project the most important task relevant information directly into the field of view, therefore, enable a high innovation leap in general aviation. The different characteristics between general and commercial aviation are listed in Table 1.1.

Table 1.1 exemplifies the increased need for assistance systems within GA in order to reduce task load, enhance SA and eventually reduce accidents. Mainly, the need is higher within GA, because the view outside the windows is exceedingly important for maneuvering the airplane and avoiding traffic and obstacles. Within

Table 1.1. Comparison of Aviation Sectors

General aviation	Commercial aviation
∘ mostly operated under VFR ∘ low stage of technical development ∘ pilots fly occasionally ∘ view to the outside is essential	∘ mostly in controlled airspace under IFR ∘ high degree of technical development ∘ regular trainings ∘ view outside the window often not necessary

GA, most flights are operated under VFR; pilots apply the so-called see-and-avoid principle.

Flying under VFR and applying the see-and-avoid principle is in conflict with the common head-down instrumentation.

Within commercial aviation, most flights are operated within controlled airspace under instrument flight rules (IFR). The cockpit has a high degree of technical development, including an instrument landing system (ILS) which allows flying under low-visibility conditions. Simplified, outside monitoring is less relevant within commercial aviation.

1.4 General Aviation

GA refers to all civil flights other than scheduled air transport services and non-scheduled air services for remuneration or hire. [81]. GA can further be categorized into instructional flying, pleasure flying and aerial work. Flights in GA can be private or, to some extend commercially motivated. The taxonomy adapted from International Civil Aviation Organization (ICAO) is shown in Figure 1.2.

In Germany, approximately 20,000 GA aircraft are registered [96]. The categories of aircraft in general aviation are diverse, ranging from helicopters over ultra-light aircraft to turbines-driven business jets. Fixed-wing planes below 2,000 kg maximum takeoff mass (MTOM) generally reflect the bulk of general aviation (6,596 in year 2015). The biggest increase in the last thirty years came from more affordable, smaller aircraft, such as ultra-light aircraft, amateur built aircraft and smaller helicopters [80, p. 13].

Motivations for flying are varied. Besides practical transportation, most pilots fly for sheer recreational purposes while others engage in sport-like competitions [14, p.216]. Even if fully automatized flying would be possible at some point in the future, it may be assumed that some private pilots will still want to engage in manual flying [101].

Only little research has yet been conducted on the processes inside the GA cockpit. While the task of pilots in commercial aviation has been analyzed thoroughly [40], research on human performance of GA pilots lags behind [24].

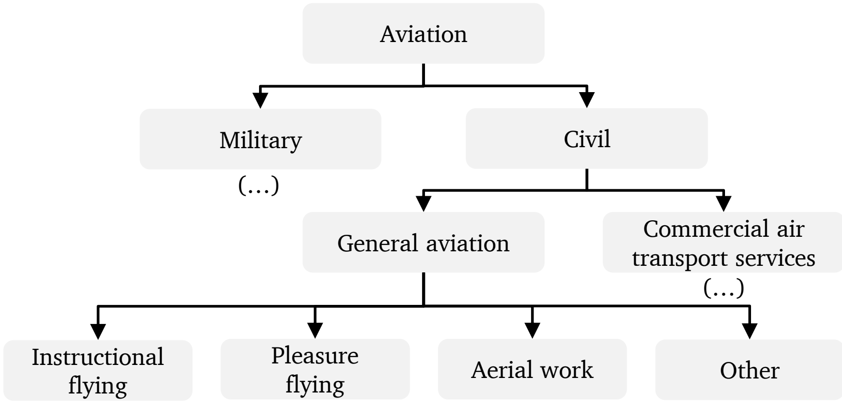


Figure 1.2.: ICAO classification of civil aviation activities. Adapted from [81].

1.4.1 Pilot's Tasks

Flying is a complex task that requires the pilot to switch continuously between competing tasks and to apply many of the available body senses and skills. Private pilots go through a training that enables pilots to fly aircraft on their own, and supposedly, most flights are executed as so-called single-pilot operations. However, if pilots wish to join each other on a flight, it was observed that they often agree on splitting some tasks as it is usually practiced on commercially operated multi-crew flights. With certain exceptions, the pilot-in-command, who is required to sit on the left seat, is ultimately responsible for the flight and its safe operation (LuftVO, segment 1, §2). The pilot-not-flying often assists with navigation, traffic monitoring and flight logging.

Especially on single-pilot operations, task management becomes important, referred to as single-pilot resource management. The concept covers the resources (both on-board the aircraft and from outside) available to a single pilot (prior and during flight) to ensure the successful outcome of the flight [112, p.17-4]. Nonetheless, this task management may fail. The Federal Aviation Administration (FAA) describes current issues with distraction by less essential tasking and the subsequent loss of control of the aircraft [52]. In reaction to this, the FAA published a safety briefing that reminds pilots to maintain aircraft control at all times. It uses a workload management principle, called aviate - navigate - communicate (A-N-C), which is a widely used phrase by pilots. It is a reminder of the pilot-in-command priorities during all flying situations, especially emergency conditions. This phrase is used here to illustrate task prioritization.

Aviate - Navigate - Communicate [53]:

Aviate Maintain control of the aircraft.

Navigate Know where you are and where you intend to go.

Communicate Let someone know your plans and needs.

The top priority is to aviate. That means flying the airplane by using the flight controls and flight instruments to direct the airplane's attitude, airspeed and altitude. Flight instruments from inside the cockpit provide important information about the flight situation. The pilot receives information on airspeed, attitude with relation to the horizon, altitude, vertical speed and rate, magnetic heading, and turns and coordination.

Rounding out those priorities are navigational tasks (navigate), and, as appropriate, talking to air traffic control (ATC) or someone outside the airplane (communicate). This may lead to delays in responding to ATC communications and passenger requests, or not responding at all unless positive aircraft control can be maintained throughout the flight.

While in the air, pilots rely heavily on their visual senses to avoid obstacles and to keep the aircraft in the intended attitude. The vestibular system is a sensory system that provides the leading contribution to the sense of balance and spatial orientation. Together with the feeling of seat cushion pressure changes from their somatosensory system, pilots create their own mental model of the aircraft's attitude, which can, however, be misleading in low-visibility situations [12].

Visual air navigation creates a unique set of cognitive demands as the pilot or navigator repeatedly compares features of the map with the outside view and quickly determines whether or not the two are congruent [6]. Pilots continuously compare track deviations with chart material, and manually maintain the correct airplane's direction. Especially in high task load situations this can be challenging. Typical chart material will be presented in the following sections.

1.4.1.1 Aeronautical Charts

Pilots needs to carry the latest and appropriate map for the area they want to fly in, as well as a map of the area of possible evasive routes. Commonly, these charts exist in printed form, but the regulation does not exclude charts on electronic devices. EU regulation 800/2013 [49] governs the non-commercial operation of non-complex aircraft (see EASA OPS Part NCO, Material Acceptable Means of Compliance / Guidance Material). The regulation states that "current and suitable aeronautical charts" are required. Charts for VFR flights need to show airspaces, airports, radio frequencies, distinctive landmarks, settlements, topology and isogonic lines. Aeronautical charts utilize the Lambert conformal conic projection (LCC), which guarantees equality of angles and minimal distortions. Updates are provided every 28 days, via the Aeronautical Information Publication (AIP). The most common types of aeronautical charts use a 1:500,000 scale, other used scales are 1:250,000, 1:300,000 and 1:1,000,000 for printed charts [89]. Most of them are foldable and need to be switched, when passing into the area of another card. For obvious reasons, this is neither comfortable nor practical, however, independent from electrical devices.

With the advent of modern computerized navigational systems and precision guidance, such as the Global Positioning System (GPS), pilots have been replacing traditional paper aeronautical charts with sophisticated moving map displays [78]. Such digital charts exist for integrated avionics as well as retrofittable devices. Digital charts are either vector based or simple scans of conventional maps. Vector based electrical maps allow zooming without distortions. Knobs or gestures allow fluent interaction, e.g. for zooming, which enables coverage of large areas without switching to other charts, as you would need to do with printed charts. Existing methods have the disadvantage that topographic information on two-dimensional maps must be transformed and integrated into the pilot's three-dimensional perspective. The underlying processing mechanisms are cognitively demanding and prone to error [78].

1.4.1.2 VFR Traffic Patterns

A VFR traffic pattern, also called traffic circuit, is a standardized path, followed by aircraft when taking off or landing. Traffic patterns are published as printed charts, called approach plates. Just as aeronautical charts, they may be used in electronic versions as well (see Section 2.1.2 on regulatory aspects). Deutsche Flugsicherung (DFS) [25] describes its importance for safe approaches and departures, but also for the protection of noise sensitive areas around the airfield. In the pattern, aircraft remain close to the airport, providing visual contact with the airfield at all times. Traffic patterns are usually left-hand turns unless otherwise specified. The standard traffic pattern altitude for small GA aircraft is 1,000 ft above ground level, if not described otherwise. Patterns are typically rectangular in basic shape (as depicted in Figure 1.3), and depend on the wind direction, which defines the landing direction. There are however numerous examples of non-rectangular, more complex patterns (see Figure 5.12), for the reason of ground features (e.g. terrain) or because of noise abatement. Each leg of the pattern has a particular name [112]:

Departure The section extending 1.5 km from the runway center line.

Crosswind A short climbing flight leg, 1.5 km in length, perpendicular to the runway heading.

Downwind A long level leg parallel to the runway but in the opposite direction. Aircraft usually enter the pattern here in a 45° angle.

Base A short descending leg, 1.5 km in length, perpendicular to the runway heading.

Final A descending leg from the end of base leg to the start of the runway along the extended runway center line. The last section of the final approach is sometimes referred to as short final.

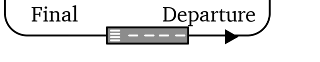


Figure 1.3.: VFR traffic pattern based on [112, p.13-12].

Non-rectangular and complex traffic patterns require the pilot to concentrate on charts, which is often described as demanding, especially at airfields the pilot is not familiar with. The Aircraft Owner and Pilot Association (AOPA) reminds that the VFR traffic pattern is not a binding regulation, but rather a recommendation.

However, municipal authoritatives have tried to control the adherence to published patterns and punish deviations [137]. Additionally, adjoining controlled airspaces will eventually require pilots to precisely adhere to published traffic patterns and complex approach procedures (compare Section 1.4.2).

1.4.2 Airspace Infringements

Air traffic control has varying executive control over aircraft in controlled airspace opposed to uncontrolled airspace [28]. Despite available modern technologies (such as GPS) and despite the prosecution of infringements by law, airspace violations are a common phenomenon. Infringements of controlled airspace (such as class C and D) are reported to the German supervisory authority of the German air traffic control, DFS, on a daily basis [4]. Airspace infringements, especially in high density areas, pose a threat to all airspace users. Additionally, airspace infringements have been associated with incidents, increased workload of air traffic controllers and delays [89]. In Germany, the legal consequences for infringing are regulated by LuftSiG, segment 5 “Bußgeld- und Strafvorschriften”, §18, §19, §20 and usually range from warnings to imprisonment.

To clarify what airspace infringements exactly are, the definition of Eurocontrol is used as stated in the Eurocontrol action plan [50]:

> Airspace infringement (also referred to as ’unauthorised penetration of airspace’) is generally defined as a flight into notified airspace without previously requesting and obtaining approval from the controlling authority of that airspace in accordance with international and national regulations. Notified airspace includes controlled airspace (ICAO airspace classes A to E, such as airways, TMAs [Terminal Areas], and CTRs [Control Zones]), restricted airspaces (e.g. Prohibited, Restricted and Danger Areas, Temporary Reserved Airspace or airspace notified by a restriction of flying in accordance with national requirements) and aerodrome traffic information zones or areas (ATZ [Aerodrome Traffic Zone] or TIZ [Traffic Information Zone]) implemented by a number of European states.[p.4]

Since 2004, the overall number of reported incidents is constantly increasing [50]. It is not clear whether infringements are really taking place more often or the increase of awareness through information distribution results in a larger amount of reports of airspace infringements. Nonetheless, Eurocontrol argues that in comparison with the evolution of the number of reported incidents assigned to other key risk areas (such as separation minima infringements), airspace infringements show a particularly marked trend.

Further analysis of the incident numbers indicated that the majority of infringements are committed by GA VFR flights. Eurocontrol inferred that lower levels of training and experience of pilots, flying only under VFR, are reasons for the increased share of infringements.

A supervised student research project, within the scope of this thesis, explored the anticipated evolution of future air traffic management systems under initiatives such as Single European Sky ATM Research (SESAR) with regard to GA airspace users [101]. Analysis suggested that airspaces will remain distinguishable in the future between managed (in other words controlled) airspace and unmanaged airspace (uncontrolled). For most private pilots, it will remain favorable to fly within the unmanaged airspace, due to its higher degree of freedom. Conflicts arise because managed airspace will, by all accounts, be more complex and relatively larger in space covered. Hence, it will be challenging for private pilots to maintain separation from managed airspace.

This poses the need for an easily accessible information of airspace information for private pilots.

1.4.3 Accidents in General Aviation

Reports from the European Aviation Safety Agency (EASA) [34] have shown that technical improvements (aircraft structure, engine) and automation (flight management system (FMS), autopilot, traffic alert and collision avoidance system) lead to a steady increase in safety of commercial aviation. Accidents in GA, on the other hand, stayed at the same high level over the last few years.

Ninety-two percent of recorded aviation fatalities can be assigned to GA [129]. Despite all improvements, the human will remain the most vulnerable factor in the chain of underlying causes. One of the main reasons for accidents is pilot error (84% in GA), also referred to as failure of the socio-technical system [21]. In four out of five cases the pilot misjudges the situation, or what is called a loss of situation awareness (LOSA). [41]

Exceeding the appropriate levels of task load is considered a critical factor for piloting errors that may result in loss of control over the aircraft. Underlying causes are grouped into categories. Forty-seven percent of all European accidents within light GA, i.e. aircraft up to 2,250 kg, may be grouped to the category loss of control in flight (LOC-I) [34], therefore being the most frequent underlying cause. Private pilots are most likely not trained to fly under IFR. Furthermore, training levels underlie strong variations. Some only fly for leisure and fun. This lack of routine may eventually lead to exceptionally high levels of workload, which in the case of an unexpected situation, eventually, resulting in mishaps [156]. Generally speaking, landing and take-off belong to the phases of flight with increased task load and higher rates of mishaps and failures. [91].

This chapter showed that accidents rates are particularly high within general aviation. Reducing the numbers of accidents can only succeed when pilot's task load is reduced to an adequate level and supporting information is provided to support the pilot's SA and spatial awareness. A proposed solution is the use of smart glasses, which project the most important information directly into the field of view. The following chapters will illustrate how HMDs displays may help the pilot to maintain an *eys-out* view.

2 State of the art of cockpit technologies and human factors

The intention of this chapter is to introduce the relevant technologies used in the cockpit, and in particular to describe the computerization of the cockpit. The focus is on EFBs as a technology that is related to smart glasses, as they form the group of personal electronic device (PED). Section 2.2 presents an overview of existing HMDs. Smart glasses as an emerging technology are presented and a taxonomy is developed. Section 2.3 will present the systematic approach that is used within this thesis. Relevant principles related to human factors and psychology are explained.

2.1 The General Aviation Cockpit

The instruments in a GA aircraft are typically located on an instrument panel and are all within physical reach in order to perform calibrations or manual alterations of the instruments. The minimum instrumentation of aircraft differs between VFR and IFR flights and is regulated in Germany by the "Dritte Durchführungsverordnung zur Betriebsordnung für Luftfahrtgerät" (3. DV LuftBO), segment 2, "Ausrüstung von Luftfahrzeugen". In general, most aircraft are equipped with six instruments, sometimes referred to as a *six pack* of instruments [112], to provide an overview of the aircraft's current situation. These instruments include the airspeed indicator, the attitude indicator, the altimeter, the vertical speed indicator, the horizontal situation indicator and the turn indicator. In modern aircraft, these instruments are combined into a digital PFD, referred to as a *glass cockpit*. A comparison of both conventional analogue instruments and glass cockpit instruments is illustrated in Figure 2.1. Conventionally, the PFD is placed on the left side of the cockpit while on the right side a so-called multi-function display (MFD) shows auxiliary information, such as navigational charts or engine parameters. Despite the glass cockpit's high capability of displaying flight-relevant information in an appealing manner, they have not shown to be a major safety improvement for GA [66; 67].

Instruments are arranged in a specific way, called a *T-pattern*, which follows the shape that would emerge by scanning the instruments in the way that is taught in flying schools [9]. Training for scanning the instruments is intended to reduce the

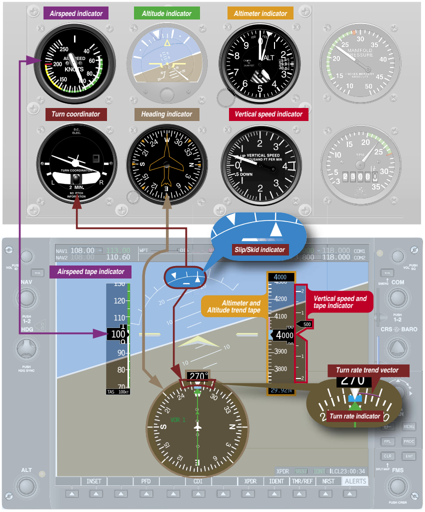


Figure 2.1.: Performance indicators. Source: [112].

time required for this task. Nevertheless, pilots must direct their view downwards inside the cockpit in order to read the instruments. When pilots switch their view between inside the cockpit and outside, their eyes must constantly adjust [35]. Both the adjustment for distance, or so-called accommodation, and adaptation for brightness - play roles in a delayed or constrained perception.

One of the greatest needs in aviation is for navigation and flight guidance. Besides using land marks and topographic charts for navigation, radio navigation can be used. Radio navigation enables aircraft with a receiving unit to determine their position by receiving radio signals, transmitted by a network of fixed ground radio beacons. To stay on course, instruments such as the horizontal situation indicator (HSI) (see bottom left of Figure 2.2) can display the deviation from a desired path.

2.1.1 Pilot Assistance Systems and Retrofittable Devices

Computerization of the cockpit has allowed more complex representations of information and has also shaped the way that information is perceived within the cockpit. In the beginning of aviation, classic *steam* gauges all looked similar (pointer on scale) and interaction was limited. Computerization of the cockpit has increased the complexity and availability of information. On conventional analogue instruments, information was always retrievable at the same location within the cockpit. Modern integrated avionic systems allow the digital presentation of information on multipage displays, thereby increasing automation.

Figure 2.2.: Electronic flight instrumentation with navigational information. PFD (left) with synthetic vision and a HSI on the bottom (Garmin G3X). Ipad with map (right) and vertical situation display (ForeFlight). Source: [57; 61].

Most modern navigation systems for GA rely on GPS. Waypoints are programmed into a flight plan and the pilot flies the aircraft from one waypoint to the next. Deviations from the desired course can be displayed by various means, ranging from representations similar to a horizontal situation indicator to so-called highway-in-the-sky on integrated synthetic vision displays. Examples of such systems are illustrated in Figure 2.2. Highway in the sky (sometimes known as tunnel-in-the-sky) symbologies allow three-dimensional flight guidance, and in early studies done with helicopter pilots in the 1980s [68] have been shown to effectively decrease workload. They also later proved beneficial for curved approaches with fixed-wing aircraft [58].

One problem that remains with tunnels-in-the-sky is the potential to cause an increased amount of clutter on the PFD [153]. In addition, the track recovery is often less effective when compared to other representations [141]. It may be doubted that the precision offered by such systems is necessary for the execution of a flight under VFR. For private pilots, navigational performance may often be of secondary importance, unless they are flying within controlled airspace.

For pilots, computers became viable for other tasks besides flying, e.g. flight planning, navigation and many others. In aviation, hand-held computers or tablet devices operated in the cockpit are often referred to as PEDs [**?**]. Mainly COTS devices such as the iPads are used, and applications (i.e. software) are distributed separately [59].

EFBs are used to present supplemental flight information (traditionally presented in printed format and carried by the pilot in flight bags). In GA, applications on PEDs are used for a multitude of operational purposes. Available applications range from those used in the calculation of weight and balance to multifunctional applications that assist with pre-flight planning, flight execution and post-flight operations [**?**]. An example for a multifunctional application is ForeFlight [57], depicted in Figure 2.2.

The use of EFBs, (or PEDs in general), provides the opportunity to equip aircraft with updated functions at a lower cost than an update to the aircraft's integrated avionics. Financial expenses for updates to the integrated avionics certainly play a big role when considering the average life span of a GA aircraft compared to the life span of most computerized devices.

GA aircraft can be characterized by their long product life span. In 2008, the average age of U.S registered light GA aircraft was 38 years [63]. Considering that electronic consumer products have a much shorter life span, it is clear that built-in computerized avionics quickly become outdated in terms of processing power.

Another benefit of EFBs is the mobility of the devices. This becomes particularly obvious, considering that aircraft chartering is often practiced. Each chartered aircraft may be equipped differently, and pilots must become accustomed to the installed avionic system, which may require different methods of accessing functions. On the other hand, pilots who charter aircraft profit tremendously from EFBs. Pilots may bring their own PED and the functions and the accustomed use stays the same across different chartered aircraft. Therefore, fewer mistakes may arise due to incorrect use.

2.1.2 Regulatory Aspects related to Personal Electronic Devices

In Europe, the EASA is the institution in charge of the regulatory and executive tasks in the field of civil aviation. A tedious certification process and high prices for integrated avionic devices deter aircraft owners from upgrading their aircraft. EFBs have gained in popularity because they are affordable and easy to retrofit [23]. However, the use of PEDs is not explicitly regulated. EU regulation 800/2013 governs the non-commercial operations of non-complex airplanes. It lists a number of documents, manuals and information that must be carried [49, p.48]. A supplementary document to ED Decision 2013/022/R clarifies: "The documents, manuals and information may be available in a form other than on printed paper.

An electronic storage medium is acceptable if accessibility, usability and reliability can be assured." [46, p. 20]. Under the current versions of the issued guidelines, EFBs are regulated carefully in the transport and passenger categories, but very little in GA. It is not further specified, which types of devices are suitable for the storage and depiction and how they may be used throughout the flight.

Guidelines applicable to commercial air transport operators (ED Decision 2014/001/R AMC 20-25) [47] classify EFBs either as *portable* or *installed* EFB, based on their mounting and connectivity with the aircraft.

Portable "A portable EFB is a portable EFB host platform, used on the flight deck, which is not part of the certified aircraft configuration."[p. 5]

Installed "An EFB host platform installed in the aircraft and considered as an aircraft part, covered, thus, by the aircraft airworthiness approval."[p. 6]

2.1.3 Ubiquitous Computing

Ubiquitous computing refers to the use of computers in any form that are available anywhere at any time [145]. This trend is reflected in the proliferation of wearable devices and mobile devices. PEDs, such as smart phones and smart watches are a companion and pioneer of this development due to their increasingly smaller size. M. Weiser describes his visionary idea of ubiquitous computing: "The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it." [146, p.1]. Want, Borriello, Pering, and Farkas remark that such technology "brings us one step closer to a world where we can access personally relevant information quickly and conveniently, without relying on bulky, fragile display systems" [144, p.41]. There are three problems which have been challenging for ubiquitous computing hardware and which will probably continue to be challenging in the future: "size and weight, energy, and the user interface" [144, p.42]. It is anticipated that the main goal for the user of assistance systems will remain the fulfillment of a specific task in the most efficient manner, with as few interactions as possible using one device.

2.2 *Eyes-out* Technologies

For many flight operations, the visual sensing of the outside view and elements around the aircraft is highly important. To increase flight safety as well as efficiency of procedures in aviation, new technologies are considered for practical applications. Head-worn displays are regarded as a technology that could possibly have a fundamental impact.

Eyes-out technologies offer a possibility to project information in such a way that less time must be spent monitoring head-down instruments. While remaining the ability to keep the eyes fixed on the outside world, users are more likely to detect important changes within the FOV [72; 98].

The following sections describe HMDs (Section 2.2.1) and smart glasses (Section 2.2.3), which are a subcategory of HMDs. A taxonomy for smart glasses is proposed. Existing technologies are explained and their applicability to GA cockpits described. In this thesis, head-up display (HUD) will not be covered [1].

2.2.1 Head-mounted Displays

HMDs are attached to the user's head and allow the most valuable information to be projected directly into the field of view (FOV). Potential areas of application could be those where users can benefit from visualized information that is either impossible or difficult to obtain due to specific task constraints. HMDs have been discussed for application in multiple aviation related groups, including maintenance [107], cabin- / ground personnel [106] and ATC [121]. In general, head-worn displays are applied to tasks that involve multiple simultaneous processes and tasks where it is necessary for the operator to work hands-free.

The idea of a head-worn display is not new. In the mid-1960s Ivan E. Sutherland from the Massachusetts Institute of Technology developed a closed-circuit television link between an HMD and multiple remote cameras [135]. The system generated graphics for the augmentation of the user's view. Due to the heavy weight of the system, it was suspended from the ceiling, and it was later know as "The Sword of Damocles" [130]. The device is depicted in Figure 2.3.

The first practical historical application of this to aviation was for military pilots. Early research began with night vision systems. In the 1980s, the U.S Air Force introduced a system, known as the Aviator's Night Vision Imaging System (ANVIS). For the later fielded AH-64 Apache helicopter, an augmented vision system, called Integrated Helmet and Display Sighting System (IHADSS) was applied. It featured the projection of augmented information to only one eye so that it would not obscure the view of the outside world. [117]. Both are illustrated in Figure 2.4. Because these systems are usually attached to a helmet, they are referred to as helmet-mounted displays.

[1] The benefits and limitations of HUD technologies are well-known [138]. However, HUDs are regarded as less applicable to the GA cockpit, because of their difficulty to retrofit.

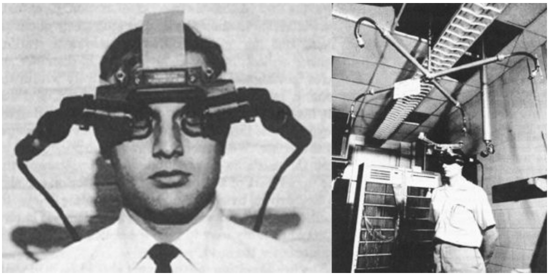


Figure 2.3.: Early HMD from 1965 made by Ivan E. Sutherland, called The Sword of Damocles. Source: [130].

Figure 2.4.: Military use of head-mounted displays. ANVIS (left) and IHADDS (right). Source: [117].

HMDs have been studied in past at the Institute of Flight Systems and Automatic Control (FSR). The *AddVisor 100* from Ericsson-Saab Avionics was used, among other things, for the projection of flight guidance information [93]. It featured a 30° FOV with a resolution of 1280x1024 pixels, which can stand up to today's consumer products. However, the projection to both eyes was only monochromatic and the weight of the system was 650 g.

HMDs can be categorized by their practical application in augmented reality (AR) and virtual reality (VR). Furthermore, different working principles can be used, namely optical see-through (OST) and video see-through (VST). A comparison of OST and VST is illustrated in Figure 2.5 and will be explained in the following sections.

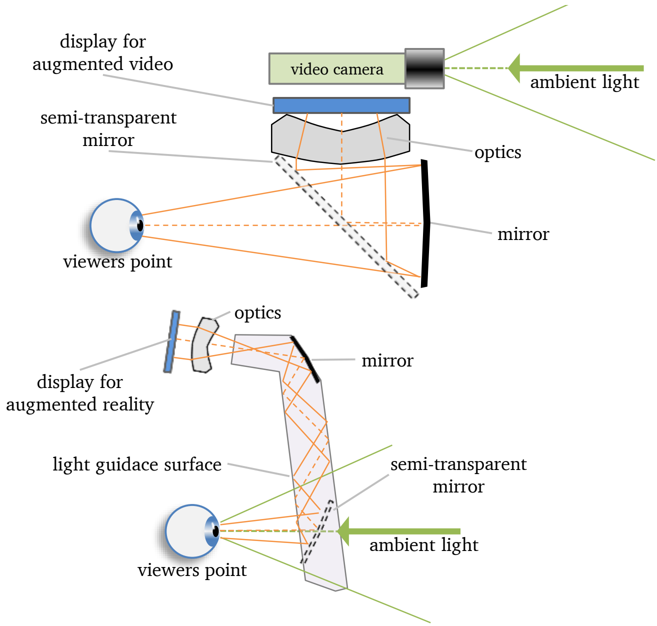


Figure 2.5.: Comparison between video based (top) after Dörner et al. [32, p.273] and optical see-through system (bottom) based on the Epson Moverio BT-200 [43].

2.2.1.1 Video See-through

One or more cameras record the environment and display the video images with optional virtual elements in the user's FOV. Because the resolution of the video camera is limited and image dynamic is restricted, a relative loss of information is evident. Besides the inferior video quality, a latency is unavoidable. Increased latency has an impact on the user's sense of balance [31]. The VST principle should only be applied when the user's perception of the environment is not crucial or safety-relevant.

2.2.1.2 Optical See-through

With the OST working principle, the virtual image is displayed in the user's FOV while the user is still able to directly perceive the environment. Most often a combiner, usually a half transparent mirror, is used for displaying the image. Another, not yet fully employed technology includes a direct projection of the image on the user's retina, which allows image projection to a limited position within the user's FOV.

To reduce the form factor, some smart glasses employ so-called wave-guide technology [124]. Light rays are reflected within the glass by a light guidance surface until they reach the combiner. All of these principles have in common the fact that some sort of image generator is necessary. However, this can be perceived by the user as being disturbing. In addition, the construction itself as well as the combiner does not allow all possible light to pass through. The transmission rate of most models is less than 80% [83]. To increase contrast on the screen, transmitted light may even be reduced by applying shades to the outer front.

2.2.1.3 The Reality-Virtuality Continuum

Display systems are distinguished by type or by degree of reality of the information presented. Milgram and Kishino [103] define the entire range of the representations as the virtuality continuum (VC). Figure 2.6 shows the course between reality and virtuality. VR display systems show a purely virtual, computer-generated image without any real external influence. The area between reality and VR is described as mixed reality (MR). The MR range is characterized by the fact that a mixture of reality and VR is presented. On the basis of the degree of reality, this area is again divided into AR and augmented virtuality (AV). If a real image is augmented by a virtual overlay, this is considered to be AR. AV means that the virtual image is enhanced by real content.

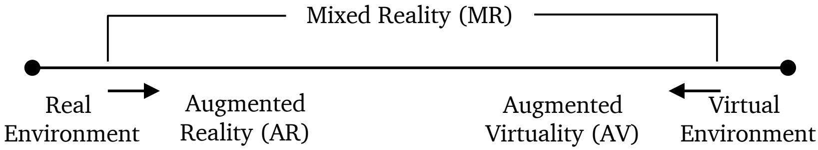


Figure 2.6.: Simplified representation of a Reality-Virtuality Continuum. Illustration by author after [103].

2.2.2 Collimation

Both virtual and real world objects must be at the same optical distance for them to be perceived as one unit. Virtual images are set to a particular virtual collimation distance [93]. A collimation, (meaning a parallelization of light rays), is achieved by the collimator which may consist of a curved mirror or lens. This can be used to project objects at a prescribed focus distance. Looking through the combiner, it may seem as if one was looking at a screen a few meters away [161] or in the case of a collimation distance of more than 6 meters [111] virtual objects appear to be focused at infinity with no or little parallax.

2.2.3 Smart Glasses

In contrast to conventional HUDs, so-called smart glasses are affordable, lightweight, and can possibly be retrofitted in most cockpits, which makes them attractive for light GA aircraft [69].

Miniaturization and declining prices have led smart glasses to be used in consumer applications (e.g. video games). With the introduction of Google Glass, numerous start-ups have taken a similar approach. Nowadays, smart glasses are available at reasonable prices, comparable to premium smart phones. However, no product has managed to persuade consumers with a compelling use case [119]. In general, once the cost benefit issues have been evaluated and found to be acceptable, one of the remaining main barriers to commercial applications of HMDs is user acceptance [118, p.68]. Remaining challenges for smart glasses developers are the limited FOV, occlusions by the combiner and the glasses' frame as well as inferior attitude and position sensors. This implies that display formats must prevent cluttering and cannot use augmented reality to its full extent.

Smart glasses can be traced back to a patent registered in 1991 by Benjamin Wells. Ten years later, patent US 5003300 "head mounted display for miniature video display system" [147] inspired Google's R&D. After Google patented *Project Glass* [76] in 2011, the company launched their first smart glasses *Google Glass* intended for the consumer market in 2014. The product was temporarily available to developers and for academic research, but was then discontinued before the official release to consumers. Besides technological shortcomings such as low battery capacity and unreliable speech recognition [42], the technology lacked adequate use cases for a broad go-to-market strategy. Business Insiders' BI Intelligence was predicting a sale of 21 million smart glasses a year for 2018, only by Google [16]. This optimistic prognosis led other companies, (e.g. Epson and Canon, as well as numerous start-ups), to develop similar products.

Other analysts, such as the head of the *technology and research* department at Gartner Institute, Brian Blau, are more sceptical about the technology. In the Gartner's Hype Cycles report, technologies are rated based on their expected impact. In a way, this resembles the technology-readiness level or you could say market-readiness of a product. Figure 2.7 shows that HMDs are positioned in the "trough of disillusionment". Furthermore, they argue that it will need worthy use cases for consumers [10].

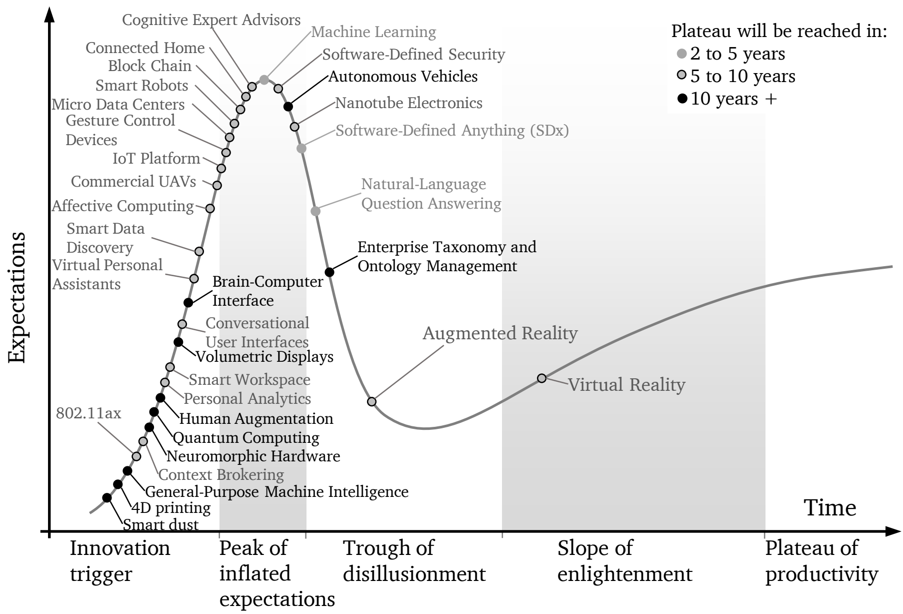


Figure 2.7.: Hype Cycle for emerging technologies 2016 by Gartner Institute. Adapted and modified from [62].

2.2.3.1 Taxonomy of Smart Glasses

A taxonomy for different types of smart glasses is suggested. They can be grouped based on their ocularity (monocular vs. bi(n)ocular) and on the area of projection (central vs. peripheral). In this context, the term *AR mode* - derived from augmented reality[2] - for central picture placement, and *glance mode* for peripheral projections is introduced. The concepts used for classification will be explained more thoroughly in the following sections. The classification is shown in Figure 2.8.

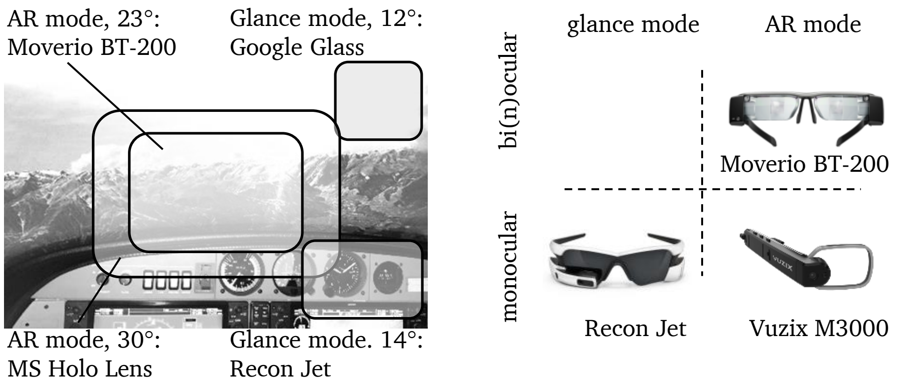


Figure 2.8.: Area of projection (left). Taxonomy of smart glasses based on ocularity (left). Source:[45; 120; 143].

2.2.3.2 Area of Projection

The picture placement is relevant for later intended purposes. Smart glasses that use a glance mode picture placement, are generally more adequate in use cases that cover information, that is to be retrieved only from time to time. On the other hand, AR mode glasses are suited for projecting information that requires immediate attention [3], or for superimposing objects with virtual information (compare

2 Smart glasses that employ augmented reality usually have the display positioned centrally. Nevertheless, a centrally placed image is not technically required.

3 Studies have shown that a deviation of 15° between the fixation point and the displayed information may already result in impaired perception [125]. Only intense flashing lights will draw the user's attention if placed in the periphery. Information that requires immediate attention should therefore be placed in the center (AR-mode) and not the periphery (glance-mode)

augmented reality). AR mode smart glasses project the information in the center of the FOV and may therefore occlude real world objects that lie behind it. The user might be somewhat limited in visual tasking. A solution for this could be masking or temporary deactivation of the display. Within a cockpit scenario, masking the virtual information whenever the pilot is looking down at the inside of the cockpit has been discussed in previous research [93]. Even if the glasses don't show a picture, the visual perception will in some way be limited in a way, because combiners are not totally clear.

A direct recommendation for or against the glance- or AR mode depends on the intended purpose. For the visualization of auxiliary information, that does not need be retrieved regularly, (e.g. radio frequencies or checklist), glance mode smart glasses would be more applicable.

2.2.3.3 Ocularity

Another classification scheme suggested by Rash et al. [118] uses the terms monocular, biocular, and binocular. These terms refer to the presentation mode of the symbology. In this context, monocular means that the virtual image is viewed by a single eye; biocular means that the smart glasses provide two visual images, from one or two display units, but each eye sees exactly the same image from the same perspective; binocular means that the smart glasses provide two visual images, one for each eye and each eye sees a different image. The visual images may be manipulated in order to provide perspective. All known current smart glasses that provide a visual image to both eyes use two display units, one on each side. Most of the time, unless in 3D mode, they will show exactly the same image and therefore operate as in the biocular mode. Table 2.1 lists possible advantages and disadvantages. For simplification, biocular and binocular will be grouped together.

As stated in literature, one major problem with monocular projections is the visual rivalries that may occur when different images are presented to each eye. Normally both eyes see the same picture of the environment (apart from slight differences resulting from disparity). When using monocular HMDs, only one eye sees the reality and the other sees the superimposed virtual image. The brain eventually responds by suppressing either image. As reported by Laramee and Ware [90], the characteristics of visual rivalry include the following:

Table 2.1. Human performance considerations of biocular/binocular optical design approaches. Adapted from [102; 118].

	Advantages	Disadvantages
Monocular (one image source viewed by one eye)	◦ lowest weight ◦ simple hardware design ◦ eye with no display remains unimpaired ◦ easy adjustment ◦ low price	◦ small FOV ◦ symmetric center of mass ◦ no stereoscopic depth-information ◦ possible visual rivalry problems, such as image suppression (involuntary)
Bi(n)ocular (one / two image source(s) viewed by both eyes)	◦ wider FOV ◦ no visual rivalry ◦ better depth perception in movements ◦ stereoscopic depth information ◦ symmetrical center of mass	◦ higher weight ◦ higher price ◦ complex alignment and adjustment ◦ higher occlusion by form factor ◦ symmetric center of mass

- The duration of the suppression of an image cannot be foreseen
- Moving images and brightness attract dominance
- Aspects of both images may be mixed and may change over time
- There is no deliberate influence on visual rivalry

Visual rivalry occurs more often when both images are similar (e.g. in color). A suppression of the virtual image may be inhibited by using eye-catching symbologies. Despite the potential problems reported in the literature, they may only be triggered in certain scenarios. Valimont et al. [138] showed that whether pilots used monocular or binocular HMDs did not affect their performance and pilots reportedly did not become consciously aware of any effects.

Yeh and Wickens [159] summarized literature results on the suitability of either monocular, biocular or binocular HMD. The results of ten studies were considered and meta-analyzed with regard to depth perception, target detection, task management, orientation and subjective usability. Binocular display systems were superior in direct comparison.

A recommendation for or against monocular or binocular displays depends on the operation purpose. For the visualization of information such as checklist, monocular smart glasses may be preferred, because the eye with no display remains unimpaired of any occlusion. For augmentation of real word objects binocular displays are in favor.

2.2.3.4 Applications for Smart Glasses

Only a few enterprises have developed aviation display format concepts on smart glasses. Among them are *Aero Glass* [1], *Aerocross Systems* [2] and *Headapp* [75]. The concepts are shown in Figure 2.9. None of the presented concepts have been evaluated in an academic environment, but all concepts suffer serious shortcomings. These will be presented and discussed as the personal opinions of the author of this thesis.

The Italian based company Headapp used the Recon Jet smart glasses for their prototype *Eye 4 flight* (Figure 2.9a), projecting information monocularly to the right bottom corner of the FOV. These therefore belong to the category of glance mode smart glasses that require the pilot to glance to the periphery in order to perceive the presented information. The display format consisted of flight state information, an artificial horizon and an arrow on a compass rose, pointing out the course to the destination. The use of digital numerical values for only altitude and airspeed is believed to be unfavorable. Additionally, the display format featured non-relevant flight data, cluttering the display.

(a) Headapp. Eye 4 Flight [75] **(b)** Aerocross Systems. Brilliant Eyes [2] **(c)** AeroGlass [1]

Figure 2.9.: Companies and their prototypes of display formats for smart glasses.

Aerocross Systems (Figure 2.9b) used the ORA smart glasses from Optinvent. The image positioning of the glasses could be switched between the center (AR mode) and the top (glance mode). The demonstrator showed a PFD with an artificial horizon. The horizon was not aligned with the real horizon as the concept was not intended for operational flight but rather as a proof of concept. Aerocross systems did not implement a tracking, therefore the symbologies were head-referenced instead of real world-referenced (which would have been necessary in order for the artificial horizon to be aligned with the real horizon). The concepts suffered from high clutter and inappropriate symbologies for a head-referenced projection. Development was discontinued after 2013.

The company Aero Glass (Figure 2.9c) presented computer generated images for a concept that was futuristic, but probably unrealistic with current means. The company was funded with 1.1 million euros by the European Union. In December 2014 Aeroglass was selected for a Phase 2 grant under the small and medium-sized enterprise instrument of Horizon 2020 [48]. The company chose the Moverio BT-200 smart glasses for their hardware integration, and the stated aim was to tackle symbologies for terrain, navigation, traffic, flight state information, weather and airspace information [1]. The company attracted a great deal of attention, despite the non-existence of any sort of demonstration. No research on the feasibility of conceived symbologies was published [4]. On which timeline such a display format could be realized and which hardware requirements need to be established has yet to be ascertained.

2.2.3.5 Head Tracking System

The orientation and position of the glasses is determined by a head tracking system (HTS). The HTS is necessary for a contact-analogue projection, which means a position-accurate placement of real objects and overlapping virtual objects. Two relevant approaches for determining the position are explained: via internal sensors (inertial sensor systems) and camera-based methods [20, p.18]. Internal HTSs employ gyroscopic, magnetic and acceleration sensors. The disadvantage of internal HTSs is the necessity for calibration before each use. Camera-based methods can be further differentiated into inside-out and outside-in approaches. Outside-in methods use at least one external camera and generally markers are attached to the object to be tracked. Image processing software then calculates the orientation and position in space [131]. Inside-out approaches are most promising because they can utilize the front-facing camera, that is a component of most smart glasses. Meers, Ward, and Piper [100] proposed a method that included only one camera that records (infrared) dots in space, with their positions known. Within the scope of this thesis, multiple algorithms were compared [see 36]. LED's of different color were positioned on the glare shield of the DA 40-180 flight simulator. Implemented algorithms were able to determine the smart glasses' orientation with an error $< 1°$. The method was limited by the camera's recording frequency and resolution and the limited computing speed. The method did not prove feasible for implementation in current smart glasses but could prove beneficial on future platforms, when

[4] A. Maróy, personal communication, August 2, 2016.

limitations are overcome. Nevertheless, the method could be used to augment existing HTSs, e.g. internal HTSs that as trials have shown are agile but tend to drift after a period of time and are susceptible to electromagnetic radiation [36].

2.2.3.6 Requirements for Smart Glasses in a Simulator Setup

Different requirements apply to the use of smart glasses in a real aircraft environment compared to the use within a simulator. Relevant differences are brightness and collimation distance. In the simulator environment, a lower brightness may be desired. In addition, the collimation distance should be the same distance as the projection screen. The required collimation distance for the DA 40-180 simulator was 3 m. In real aircraft environments, the collimation distance should be near infinity to allow for fusion of both background and virtual symbologies.

In order for the pilot to perceive the display of the HMD in different light conditions, the display should have an optimal luminance. Ambient light in the simulator differs from most conditions in real aircraft. What is relevant is the perceived contrast ratio CR, determined by the difference in color and brightness of virtual objects on the display and the background. Adapted from Melzer [102], contrast ratio depends on the luminance L of virtual objects on the display and the background as well as the transmission T of the combiner, shades and the canopy of the aircraft.

$$CR = \frac{L_{HMD}}{T_{combiner} \cdot T_{shades} \cdot T_{canopy} \cdot L_{ambient}} \tag{2.1}$$

Generally, a contrast ratio of 1:3 is recommended [86; 116]. Measured ambient luminance from the projection screen inside the canopy of the FSR research flight simulator was less than 50 $\mathrm{cd\,m^{-2}}$. Obviously, the application in a simulator does not require high levels of luminance. However, a sunlit cloud may reach luminance levels of up to 34,000 $\mathrm{cd\,m^{-2}}$ [158].

2.2.3.7 Hardware Selection

Subsequent to the release of Google Glass in 2013 [104], multiple companies began developing similar products for the consumer market. This study decided to focus on the consumer market rather than on the multi-thousand euro systems that were already in military use because it was believed that consumer products will eventually bring the most innovation to this technology and will facilitate a purchase price that would be affordable for private users. Available products current as of August 2014 were reviewed and resulted in a comparison as seen in Table 2.2.

Table 2.2. Considered Hardware. Status of August 2014. [83]

Model	Price	FOV	Resolution	Bat.	Ocularity	Mode
Optinvent ORA	699€	24°	640x580px	4h	monocular	both
Epson Moverio BT-100	499€	23°	960x540px	6h	binocular	AR
Epson Moverio BT-200	699€	23°	960x540px	6h	binocular	AR
Vuzix M2000AR	5,300€	30°	1280x720px	2-3h	monocular	AR
Vuzix M100	750€	15°	428x240px	2h	monocular	both
Recon Jet	450€	16°	400x240px	4-6h	monocular	Glance

Not all products were immediately obtainable at that time. Based on known specifications and the fact that the display could be switched between AR mode and glance mode, the Optinvent ORA smart glasses were preferred. Unfortunately, the product could not be obtained within the time constraints of this project because the assembly of the product was delayed. The alternative product, the Epson Moverio BT-200, depicted in Figure 2.10, was therefore obtained.

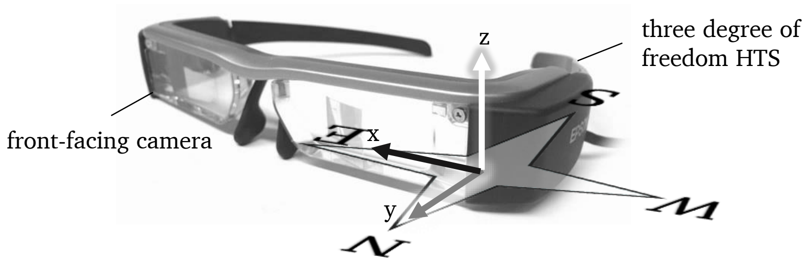


Figure 2.10.: Epson Moverio BT-200 smart glasses and coordinate system used by internal HTS.

The Moverio BT-200 is a see-through model with a binocular display and a resolution of 960x540 pixels in 2D, covering about 23° of the user's horizontal FOV. The luminance can reach a maximum of 3,000 $\mathrm{cd\,m^{-2}}$ and the weight is 88 g [44]. The brightness is adjustable to meet different requirements with changing ambient light. Display formats were planned to be implemented both monocularly and binocularly. Unfortunately, early attempts with the Moverio BT-200 indicated that the device did not allow the deactivation of one of the two display units, but forced the use of both display units, unless operated in 3D mode. In 3D mode, however, the resolution decreases to half the width. The collimation distance of the Moverio BT-200 was 3 m, which was in order with the set requirements for operation within the simulator. The product did not allow hardware-based adjustment forinterpupillary distance (IPD).

2.3 Utilized Models and Methods from Human Factors

The present study employs several concepts and theories from human factors. These theories will be explained in the following sections. Additionally, common methods are briefly explained. Evaluation methods and design principles for the development of the prototype used in this thesis, will be comprehensively presented in the particular methods chapter later in this thesis.

2.3.1 Situation Awareness

One of the underlying factors involved in many aircraft accidents is the misjudgment of the situation as a result of a loss of SA. Endsley [37] describes the SA that is necessary for a safe process and aircraft operations in three levels. The model is depicted in Figure 2.11.

The three levels of Endsley's model of Situation Awareness are:

1. The first level describes the perception of cues in the environment. Pilots must first perceive cues in their environment for later memorization and processing. Without basic perception of relevant information, the likelihood of forming an incorrect perception of the situation increases dramatically. Jones and Endsley [82] found that 76% of SA errors could be attributed to deficient perception of information. Reasons for problems in perception may be failures or shortcomings on the system side as well as problems with cognitive processes.

2. The second level includes the understanding of a situation. As Endsley points out [39], SA as a construct encompasses how pilots combine, interpret, store and retain information. Pilots need to know the meaning of a cue. This implies that pilots require a basic understanding of an object's properties and abilities. Flach [55] mentions that "the construct of SA demands that the problem of meaning be tackled head-on. Meaning must be considered

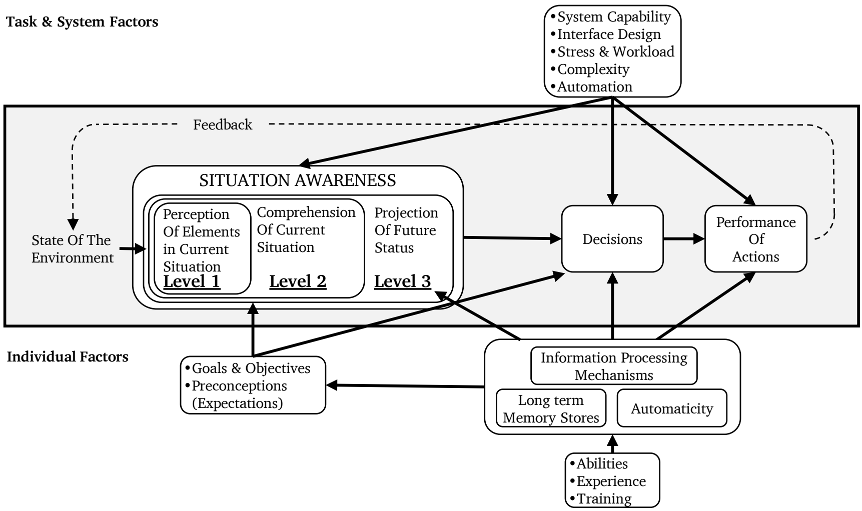


Figure 2.11.: Model of situation awareness in dynamic decision making. Adapted and modified from [37].

both in the sense of subjective interpretation (awareness) and in the sense of objective significance or importance (situation)" [p.3]. At level 2 SA, data perceived was interpreted in terms of operationally relevant meaning and significance.

3. The third level describes the ability to project from the current dynamics and events of the situation. To allow for timely decision making, pilots must plan ahead and anticipate future events and their implications. Such predictions, (mainly of a spatial nature), involve perceptual anticipation, which is defined as the acquisition of a moving target and subsequent prediction of its future positions [114].

Maintaining SA requires a thorough understanding of the relative significance of all factors related to flight and their future impact on the execution of the flight. When a pilot understands what is going on and has an overview of the total operation, he will not be fixated on a single perceived factor. Not only is it important for a pilot to know the aircraft's geographical location, it is also important to understand what is happening.

From a design perspective, we want to develop system designs that support the pilot's ability to receive the required information under dynamic operational constraints [39]. SA may be influenced by certain *task and system factors*. Systems should have a clear interface design, help to maintain an adequate level of stress and workload, have appropriate complexity and should provide automation that reduces unnecessary manual work and data integration [38].

This study employs the concept of spatial awareness, as defined by Wickens [150]. Wickens describes spatial awareness as the degree to which pilots perceive three-dimensional space and the objects in the surrounding environment (Level 1), the degree to which pilots build an understanding of their relative location to the ownship (level 2) and an understanding of their relative location to ownship in the future (level 3) [150]. Wickens describes connected concepts with relevance to spatial awareness: Understanding of airspeed and attitude (attitude awareness), position and track (track awareness), and the understanding of weather phenomena (weather awareness). Pilots tend to rapidly misinterpret their aircraft's attitude. A spatial disorientation will be noticed too late or events can even occur unnoticed, resulting in a loss of control and fatal accidents [113].

2.3.2 Performance and Task Load

Fatigue, stress, and work overload can cause a pilot to focus on a single isolated task and reduce overall SA of the flight. A contributing factor in many accidents is a distraction that diverts the pilot's attention from scanning for traffic or monitoring cockpit instruments. Phases of flight with high task load, for example, when multiple tasks compete for attention, prove to be particularly dangerous [128]. The so-called Yerkes-Dodson [160] law describes the relationship between stress and performance. Yerkes and Dodson found that within simple discrimination tasks, performance increases linearly with increases in arousal. However, in more difficult tasks, arousal and performance are related by an inverted U-shaped curve: With moderate levels of arousal the performance increases, but at the highest levels of arousal performance is impaired. The performance of each task therefore depends on the difficulty of the task and the current level of arousal. In high complexity tasks, too much arousal is just as performance-constraining as too little. In the sense of flight safety, an optimal level of arousal can be found.

At low levels of arousal, certain tasks may not be performed effectively. For example, in more monotonous flight phases, the pilot's level of arousal may be low. The pilot is then likely to miss visual information, make bad decisions, scan insufficiently and have longer reaction times [3]. High levels of arousal are more likely to be experienced than low levels of arousal, and this can also lead to a deterioration of performance. Each individual has a personal stress limit, and if this is exceeded stress overload occurs, which can result in an inability to handle even a moderate task load. This personal stress limit varies with different people, as it is affected by the physiological and psychological characteristics of each individual [17, p.164]. The characteristics of over-arousal are a tendency to narrow attention and focus, incapability of risk assessment and priority allocation and an eventual complete failure to function.

2.3.3 SEEV Model

Flying under VFR is a highly visually demanding task. Subtasks can include scanning for traffic or perceiving warnings on instruments as well as reading information on smart glasses that competes for other visual attention from real world objects behind.

Research on the visual attention that is required for such tasks has identified four main factors that determine where pilots look: salience, effort, expectancy and value (SEEV). In this SEEV model of scanning behavior [155], the probability that a given area of interest (AOI) will attract attention is modeled. The SEEV model has been found to be reliable in predicting visual scanning patterns of pilots [154]. The SEEV model tries to explain how and why different AOIs receive different amounts of attention.

SEEV model

Salience defines the extent to which an AOI stands out from the background due to its size, color, intensity or contrast. Thus, more salient objects attract higher attention while less salient objects can easily be overlooked.

Effort refers to the cost of shifting attention from one AOI to another. In the event that the relevant AOI lies within the field of foveal vision, no eye movement is necessary, and the effort to access information in this AOI is low. The information access effort (IAE) rises when scanning requires eye movements. Whenever head or body movement is required, the IAE is particularly high. Considering the human tendency to avoid effort, pilots will less often scan AOIs that require head or body movement.

Expectancy describes the tendency to look more at AOIs where changes are expected to occur. This can be biased, based on individual experience. Thus, pilots will tend to more often check instruments and displays that have frequently shown new and relevant information in the past.

Value refers to the usefulness of information in one AOI to a task. This means that the relative importance of the task guides where the eye will look. While scanning for traffic is crucial in order to see and avoid other air traffic, there are other critical tasks that require attention in different phases of flight.

Salience is directly dependent on the design characteristics of presented information, e.g. color, contrast and size. This plays a crucial role for gradated levels of relevance. Urgent information must immediately grab the pilot's attention. On the other hand, less relevant information should not distract the pilot from his primary task, namely flying the aircraft. Effort is relevant for the pilot's change of gaze to head-down information. *Eyes-out* displays clearly reduce the effort required for perceiving information. Expectancy and value vary with the experience and subjective assessment of the pilot. In situations where no traffic is expected, pilots may direct their visual attention away from the out-the-window (OTW) view towards displays in order to perceive information. Value is linked to the relevance of presented information. Display formats that present non-relevant information will not draw the user's attention.

2.3.4 Dual-task Paradigm

Pilots often execute multiple tasks simultaneously, which can be described as dual-tasking. In addition to the first task (flying the aircraft) a second task (for example, operating the radio) is often successfully coordinated. Wickens [149] proposed the multiple resource model. An underlying assumption of this model is the availability of several different pools of resources that can be tapped simultaneously. However, tasks may also interfere with each other because they compete for the same class of information processing [84]. If the task characteristics exceed the available processing resources, a reduction in processing is to be expected. The pilot is consequently distracted.

2.3.5 Human Factors Evaluation Methods

The appropriateness of particular human factors methods depends on a number of factors, including time and resources available, access to a test sample, the skills of the analyst and the type of data that is required [134]. Evaluation procedures can be divided into subjective and objective evaluation methods. Both types can be used to measure SA and workload. The Situation Awareness Rating Technique (SART) is a subjective procedure for the measurement of SA. SA can also be measured objectively, for example with the Situation Awareness Global Assessment Technique (SAGAT) or with in situ tests, which are mostly used within simulator environment testing conditions. Both methods are rather time costly and have shown inferior practicability in research within the scope of this thesis [42]. Spatial awareness and traffic awareness may be examined with eye-tracking methods. Subjective workload measurements include the National Aeronautics and Space Administration (NASA) Task Load Index (TLX), which is a widely-used and thoroughly validated instrument [73]. Psycho-physiological indicators such as heart rate or pupil diameter were found to be very valid indicators for measuring workload [126], but have the disadvantages of not being sensitive enough to smaller changes and imposing higher levels of intrusion on the test subject [157]. Other methods for measuring workload are tests showing reaction time to a secondary task. Reaction times may be an indicator of available resources and the eventual deterioration of resources due to higher levels of task load. There are also standardized methods for measuring usability. The System Usability Scale (SUS) is a widely-used tool for the subjective assessment of usability and user friendliness. In Table 2.3, the methods that are relevant to this project are listed.

Table 2.3. Selected Human Factors Evaluation Methods.

Construct	Objective	Subjective
Usability	◦ observation	◦ System Usability Scale (see Section 5.4.3)
Situation Awareness	◦ Situation Awareness Global Asessmet Technique (SAGAT) ◦ Eye-tracking	◦ Situation Awareness Rating Technique (SART)
Workload	◦ reaction times on a secondary task ◦ heart rate	◦ NASA Task Load Index TLX (see Section 5.4.3)
Performance	◦ flight technical error (FTE)	◦ questionnaire

2.3.6 User-Centered Design

Products that require the interaction of the user will evoke higher levels of product satisfaction when the functions that the product offers will function as expected and the interface is comprehensible for its users. High levels of usability can be achieved by including the user in some or all phases of development. König [87] describes different approaches to product development, among which the degree of user feedback is of relevance. The importance of involving users in the design process is highlighted, and this may be relevant to the final success of the product. The degree of user participation may vary from no participation at all to participation throughout every stage of the development process. It may even include co-decision or veto rights for commenting users.

In this thesis, every iterative step of the implemented prototype was presented to a panel of expert pilots. The participating were associated with the Technische Universität Darmstadt. Prototypes included mockups that illustrated the function and behavior of envisioned display formats, videos or sketches.

3 Preliminary Tests: Display Formats for Smart Glasses

This chapter presents the development and implementation of three different display formats. The design of the display formats is user-centered. During the concept development and implementation phases, interviews with expert pilots were conducted. These influenced the design of the display formats and helped to understand the users and their requirements for the applications. Concepts are based on use cases derived from interviews with private pilots and are presented in Section 3.1. Section 3.2 explains the implementation for each display format separately. After implementation, the display formats were compared against each other within a preliminary evaluation, presented in Section 3.3. Results led to the selection of the prototype that was further developed and researched.

3.1 Needs Assessment

This thesis applied needs assessment of the pilot's tasks. The following sections describe the user-centered process. In order to better understand the pilot's tasks and the information flow in the cockpit, various approaches were adopted. Structured interviews were used to determine the user and several use cases were identified. To acquire in depth understanding each use case was then presented to expert pilots prior to implementation.

Furthermore, eye-tracking was used to analyze AOIs in different phases of flight [105]. Building on this, an online survey with $N = 170$ participants was conducted to identify the most relevant instruments in different phases of flight. Furthermore, participants were asked for their design preferences on the selected instruments.

3.1.1 Interview Campaign

In order to identify users' needs, an interview campaign to obtain qualitative data on the range of information available in the cockpit was conducted. The aim of the campaign was to identify non-available information that is necessary for safe and enjoyable flights. With the help of structured interviews, questions were asked about information required during the flight. As a result, the gap between required information and available information was used to develop use cases for display formats for smart glasses. The interview campaign took place between March and April 2014 at a local GA airfield, Frankfurt-Egelsbach (EDFE). In total 24 male pilots participated. Their age varied between 21 and 65 years ($M = 48.3, SD = 14.4$). The average total flight experience under VFR was 1,176 h ($SD = 1357.4$). Participants were specifically guaranteed complete anonymity at each stage.

Considering that the interview targeted tacit [1] knowledge, it had to be first made verbal by using an interview technique that is similar to a cognitive walkthrough. The interview began by asking the subjects to describe a successful flight. Since the pilots were invited to participate in the interview on their way from their aircraft to the airfield tower, just after they had landed, all pilots described the flight they had just made as a successful flight. Answers to the questions were counted and then grouped. The results, given in relative percentages, are presented in Figure 3.1. Most frequently (43%), participants stated that a positive flight would be defined by an adequate level of perceived workload. This statement is rather imprecise because high levels of workload could be the result of a chain of errors. Apart from this, participants mentioned predictable weather, staying on track, planning ahead (no unforeseeable events) as necessities for a successful flight.

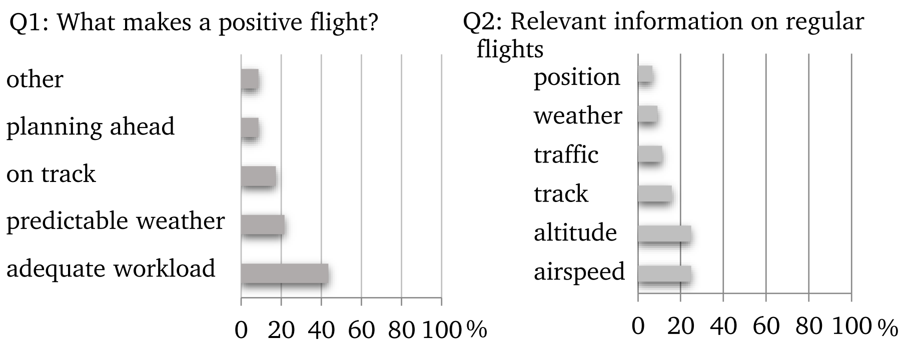


Figure 3.1.: Answers of $N = 24$ on what makes a successful flight (left). On the right side are the information needs (multiple responses allowed)[2].

[1] Tacit knowledge is described as knowledge that cannot be put on paper, formulated in sentences, or captured in drawings [142]. Tacit knowledge is difficult to verbalize because it is "tied to the senses, skills in bodily movement, individual perception, physical experiences, rules of thumb, and intuition" [p.6].

After familiarizing the participants familiar with the procedure and structure of the interview and enabling participants to verbalize tacit knowledge, pilots were asked to describe a flight that they would consider to be a negative flight. Most frequently, unexpected weather changes were mentioned followed by difficulties while landing, which is again rather unspecific and could be attributed to various underlying factors. As the interview continued this was refined by elaborating on the missing information. In addition, unintended airspace infringements, missed compulsory reporting points and technical malfunctions and user errors were mentioned. Results are shown in Figure 3.2.

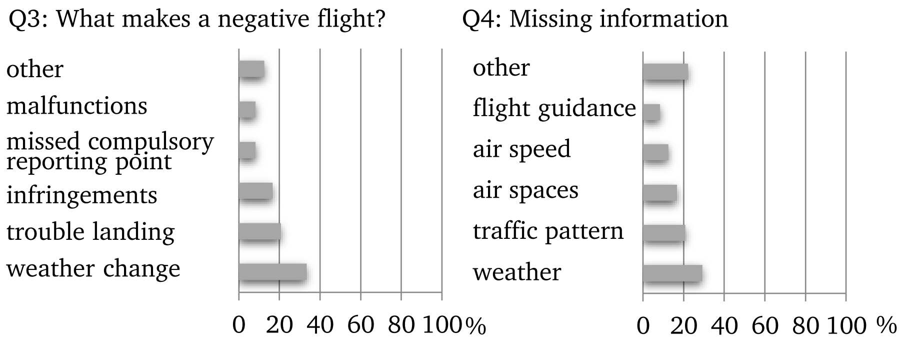


Figure 3.2.: Answers of $N = 24$ on what makes a negative flight (left). On the right side is the missing information (multiple responses allowed).

After identifying examples of regular flights and flights that pilots considered to be negative, participants were asked to identify the information that was missing during the negative flight. Most participants (25%) stated that in-flight weather information was missing, which led to a negative outcome for the outlined flight. It was followed by a request for more robust and intuitive information about the approach information related to the published VFR traffic pattern. Furthermore, participants stated that updated information about airspaces (including restricted and temporary airspaces), airspeed (including operating limits) or information on the intended flight direction was missing in the outlined flight, resulting in a negative outcome.

The gap between available and missing information was used to construct use cases for display formats on smart glasses. Even though weather information was

[1] The interview was conducted in German. The original answers were also given in German and were translated.

mentioned as missing information, it was not considered to be a use case for smart glasses because weather can be presented on more suitable platforms in the head-down area. Later consultations with the expert panel confirmed that weather is best displayed on head-down displays. Solely based on the missing information, four use cases were derived:

- Intuitive assistance for approach information (based on published traffic patterns)
- Airspace visualization
- Primary flight state information with an emphasis on operating limits (too slow, too high angle of attack)
- Flight guidance that shows the desired direction

Moreover, participants were asked for their expert opinion on what information they would select for displaying on smart glasses. Results are shown in Figure 3.3. Not surprisingly this partly matches with what has been derived from the information gap.

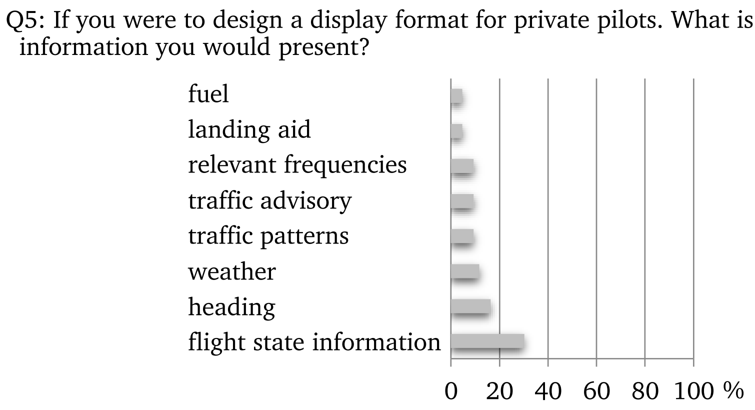


Figure 3.3.: Suggestions of $N = 24$ for selecting symbologies for a display format on smart glasses (multiple responses allowed).

3.1.2 Online Survey

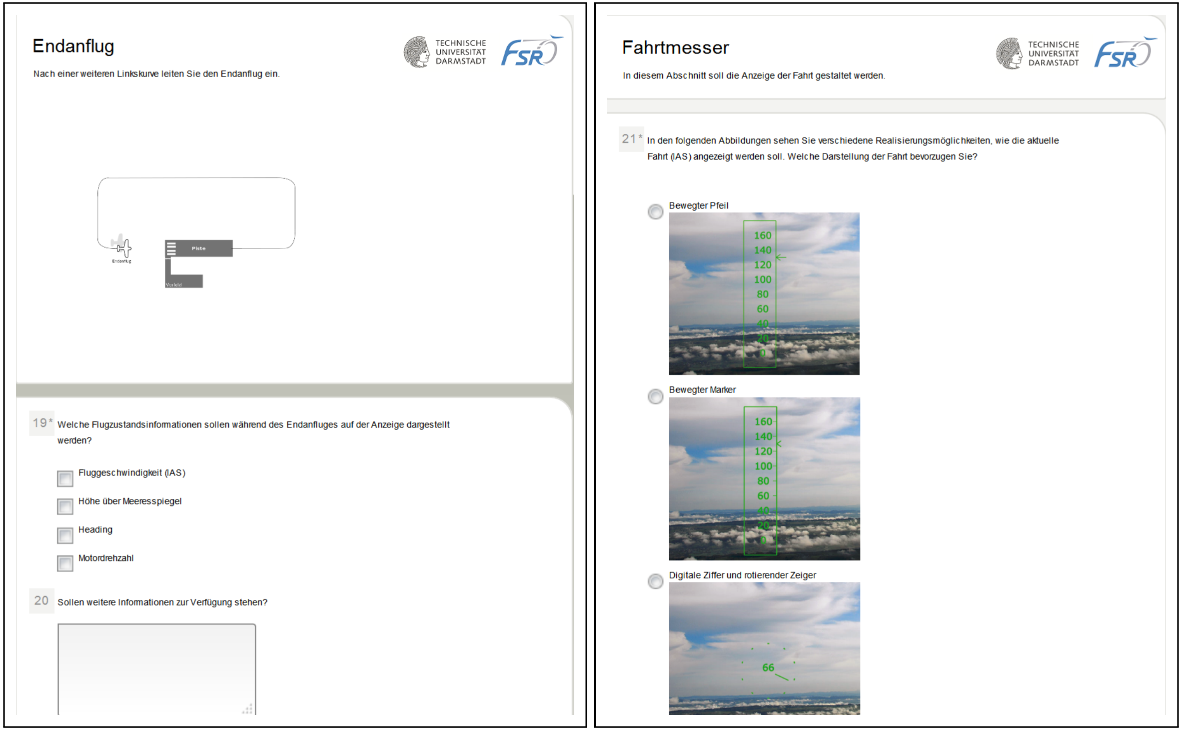
Within the scope of this thesis an online survey was conducted [133]. This was part of the design phase of the construction of the prototype of the PFD display format. The purpose of the study was to identify flight-phase-specific information needs. Because not all available information could be visualized at the same time without cluttering the view, we endeavored to reduce the information to that which was most relevant for specific flight phases. The online web survey system known as Kwiksurveys[3] was used to create the online questionnaire. Despite weaknesses known from the literature [51] such as declining response rates, unclear answering instructions and questions about sample selection, the advantages outweigh the disadvantages. The working group at the FSR has had positive experience in the past with the utilization of online surveys for data collection from pilots and maintains a database of pilots who frequently participate. $N = 170$ pilots, 2 female and 168 male, participated. Their age ranged from 18 to 73 years ($M = 46.2, SD = 14.6$). A link for the online survey was distributed via internet forums and mailing lists. Every participant held at least a valid license for engine powered aircraft or was a student pilot.

The online survey was divided into several sections. At the beginning of the survey, participants were asked to provide general personal information. Participants were then guided through each flight phase, beginning with taxiing, followed by run-up and initial climb and so on. An example from the online study is shown in Figure 3.4. Pictorial depictions and textual information was shown to clarify each flight phase.

After querying the relevant information, participants were asked for their design preferences (see Figure 3.4b). In the case of airspeed, participants could choose between four design proposals or could suggest something different. Design proposals were visualized with illustrations that were rendered in graphics interchange format (GIF) with animations in order to better clarify different motion possibilities, e.g. whether a pointer was moving or the scale itself was moving. Other versions included the round dial and the round pointer as depicted in Figure 3.4b. For each instrument, participants could choose from multiple design versions.

Results revealed that information could be limited to airspeed, altitude, heading and engine revolutions per minute (RPM). There are certain phases where airspeed plays a secondary role. Only 19% of the pilots stated that an airspeed indicator was required while taxiing. Further analysis of the comments indicated that pilots prefer to check the correct functionality of cockpit instruments while taxiing, but beyond this, it is assumed that the airspeed indicator is not required while on the ground. On the other hand, results suggest that airspeed is the most relevant information in final approach. For the design of the airspeed and altitude indicators, the two most preferred designs were the conventional tape-like moving scale. The second most preferred design was the *round dial*. Based on the par-

[3] http://www.kwiksurveys.com.



(a) Selection of information

(b) Selection of design preferences

Figure 3.4.: Untranslated examples from online study [133].

ticipant's preferences, the heading should be visualized with a moving scale. For engine RPM participants preferred a digital number only. Remarkably, participants generally preferred representations similar to the ones on existing glass cockpits. The user's habituation could have been reasons for their tendency.

3.2 Implementation

Display formats were implemented within student projects under the author's guidance and were evaluated in the FSR research flight simulator DA40-180 (covered in Section 5.4). Implemented display formats included a tunnel-in-the-sky [139], an airspace viewer [89] and a primary flight display [70; 133]. What all three display formats had in common was their aim was to reduce the amount of time a pilot requires to look at the instruments inside the cockpit, to increase SA and to reduce task load. The simulator's specification is thoroughly described in Section 5.4. Simulation data such as the orientation and position of the simulated aircraft as well as all relevant data from within the cockpit avionics was accessible via a User Datagram Protocol (UDP). The UDP port is described in [83].

The primary flight display was implemented using VAPS XT [92]. This is a human-machine interface (HMI) design tool with a graphical editor and an extensible architecture. The tool allowed the definition of the visual appearance of a display format and was also used to specify the behavior and data ranges of all elements that were rendered on-screen. Moreover, it was used to specify the behavior and data ranges of all elements which were rendered on-screen. The program allowed the user to define the connectivity between the components and their respective data sources (internal or external) and to program an advanced logic (e.g. to enable multiple pages). Data from the simulator was included in the program via a receiver module which interfaced with a written program that read the UDP port. VAPS XT uses an automatic code generator which analyzes the contents of the application, generates C++ code for all user-created content and compiles the results into a stand-alone executable file. The code was compiled to run on Windows platforms. In order to display the image from the Windows machine on the smart glasses, a virtual network computing tool called Real VNC was used. The tunnel-in-the-sky display format and the Airspace Viewer display format were contact analogue, meaning that objects match the shapes displayed on the smart glasses. This requires a fast and precise tracking of the position and orientation of the smart glasses and the internal HTS was used. Both contact-analogue display formats were written as Android applications using Open GL ES and run on all compatible Android devices of version 4.0.4 and later. Even though the Moverio BT-200 supports a three-dimensional representation of objects using stereoscopy, it was decided not to use this. Firstly, this was because of reduced resolution: The smart glasses resolution reduced to only half the display size when used in stereoscopic 3D. Secondly, because the virtual placement of symbologies in the nearby (within short collimation distance) would only be useful for the augmentation of instruments inside the cockpit. The final reason for not using stereoscopic 3D was usability. However, research [93] suggests that not everyone is able to perceive stereoscopic 3D as utilized with smart glasses. Designs that do not use this should therefore be explored.

3.2.1 Contact Analogy

Using OpenGL for visualization of objects enables the use of matrix operations to correctly position these objects and to eventually transform their real-world coordinates into screen coordinates in order to show the OpenGL scenery on a two-dimensional display [139].

The impression of contact analogue picture placement is achieved by multiple coordinate system transformations. The stepwise process is illustrated in Figure 3.5. At first, virtual objects consisting of polygons with corner points described by x, y and z coordinates are positioned in the OpenGL coordinate system (CS). In the second step the transformation in space can be described as a combination of scaling, translation and rotation. To display the same object relative to the aircraft (also called aircraft-fixed) the object was first transformed to account for altitude and relative position of the aircraft. After translating objects, the whole OpenGL scenery was rotated to account for the current orientation of the aircraft. Coordinates in the OpenGL CS were transformed using a single matrix operation. The object is now in the aircraft-fixed CS. In the third step the user's head orientation was taken into account. The position determined from sensors of the smart glasses was used to calculate the movement of the head. The CS of the internal HTS can be seen in Figure 2.10. After another transformation of the CS, the virtual object is within the head-fixed CS.

In this process, the user's head position equaled the position of the aircraft. It was assumed that for larger distances, such as for OTW scenery, smaller changes in head position would not make a difference. The visualization of virtual objects should therefore not be affected if the pilot is sitting either in left or right seat. This is certainly true if the system is implemented within a real aircraft, but this was only partly true for the simulator's projection. Due to perspective distortion of the simulated outside view on the screen, especially towards the periphery, it was not possible to visualize airspace in the correct position. Because of the close distance

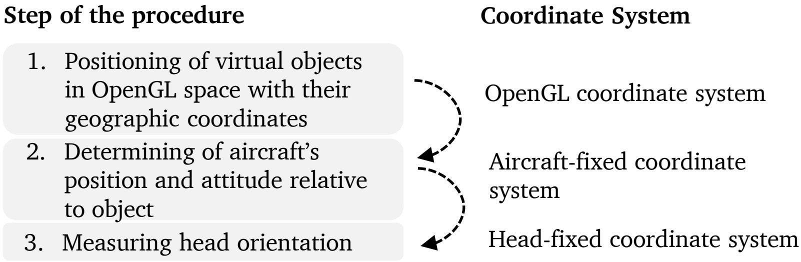


Figure 3.5.: Creating contact analogue objects as a series of CS transformations.

to the projection screen, the non-central point of view of the pilot complicated matters considerably.

The following sections will present the display formats separately.

3.2.2 Airspace Viewer

The aim of this concept was the three-dimensional visualization of airspaces. These are conventionally visualized on aeronautical charts, and the use of these charts is explained in Section 1.4.1.1. The display format has an emphasis on airspace infringements and therefore on airspaces requiring ATC clearances or restricted airspaces. Data for the georeferenced positions of airspaces were imported from a file from Deutscher Aeroclub [29] in the so-called *Open Air* format. In order to avoid clutter, only those airspaces in the vicinity were considered and transparency was used so that the OTW view would not be occluded. Airspaces were shown on the smart glasses in correct size and location, which required contact-analogue projection. Different patterns were tested to enable the pilots to estimate the distance. The visualization of a restricted airspace ED-R 134 is shown in Figure 3.6.

3.2.3 Tunnel-in-the-sky

The concept of tunnel-in-the-sky evolved from the idea of displaying the published VFR traffic patterns in an easy to follow manner. It includes a three-dimensional tunnel consisting of virtual rectangles marking the traffic pattern. Multiple concepts based on Parrish et al. [109] were implemented and checked with experts to estimate feasibility of application with smart glasses. A conventional wire frame design is shown in Figure 3.7a. The so-called crow's foot design visualizes only

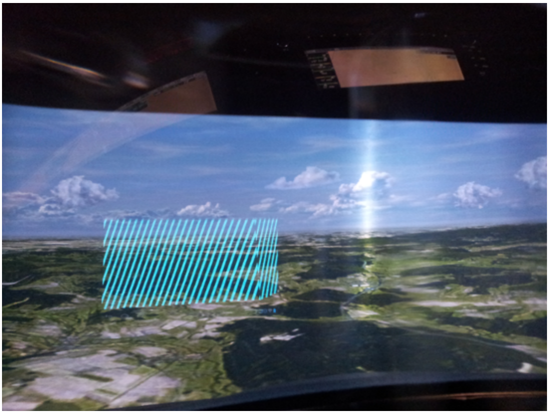
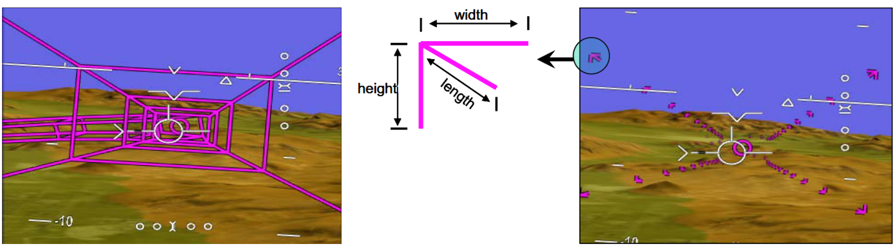


Figure 3.6.: Airspace Viewer display format as seen by pilots. Restricted airspace ED-R 134. Picture taken through smart glasses within DA 40-180 research flight simulator [89].

edges of tunnel segments (see Figure 3.7b) in order to declutter the display. Several versions of the crow's foot design were implemented. The tunnel is static, meaning that it will not move with the aircraft. Once pilots move outside the tunnel, they must recover and re-enter. Dynamic tunnels [132] were not considered for this thesis.



(a) Connected rectangle tunnel

(b) 'Crow's foot' tunnel

Figure 3.7.: Visualizations of different tunnel designs [109]-

In order to reduce the number of unnecessary objects on the display, which can eventually lead to clutter, tunnel segments at larger distances were faded out. As suggested by Sindlinger [132], segments close to the aircraft were also faded out.

With initial climb and final approach only a lateral guidance was given. The flight tunnel consisted only of segments to each side but was open to the bottom and the top respectively. The intention was to allow fast climbing and a steeper glide slope according to the pilot's choice. Pilots of small aircraft generally try to climb as quickly as possible and retain a steeper glide slope than the 3° standard slope for IFR [115]. Limiting the climb rate would clearly be unfavorable, as would be a limitation in glide slope. The implemented concept for the descent tunnel is closed at the bottom (based on a 3° glideslope) but open to the top, allowing for steeper approaches.

A first draft implemented on Laminar Research X-Plane 9.70 can be seen in Figure 3.8. The left lateral boundaries near the crow's feet were colored in red while the right boundaries were green. This color coding was chosen for its correspondence to aeronautical and nautical position lights. Each tunnel segment had a span of 150 m and 45 m in height. The distance between each consecutive segment was 150 m. Values were iteratively defined in flight tests.

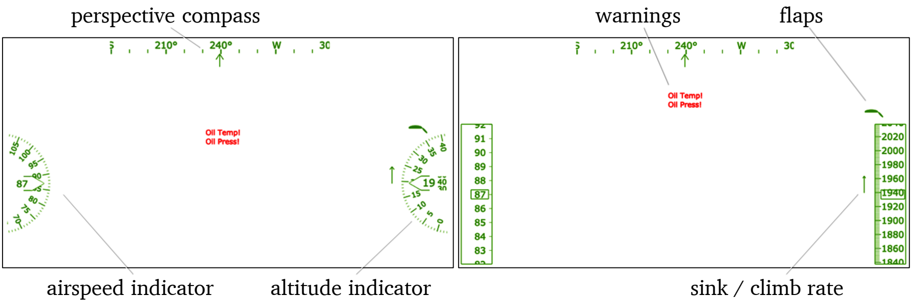
(a) Climb tunnel   **(b)** Descent tunnel

Figure 3.8.: Flight Tunnel display format prototype. Early implementation in Laminar Research 'X-plane 9.70' for demonstration. [139]

3.2.4 Primary Flight Display

The aim of the PFD format was to project the most important flight state information into the pilot's FOV in order to increase the *eyes-out* time. It was designed to ensure intuitive and, above all, fast readability without cluttering real-world objects. Two versions of the PFD were realized after conducting a study of user's preferences and flight phase specific information needs (see Section 3.1.2). The two realized versions were modularly designed based on the selections made in the online questionnaire, and they represent the most preferred design versions: A vertical *tape* representation, as also used on other head-down displays (HDDs), and a representation which is similar to round analogue instruments (see Figure 3.9).

As suggested by Kaiser [85], bright green was used as a basic color for all information. Airspeed was displayed on the left side while the altitude indicator was placed to the right. The airspeed indicator used standardized markings, which are colored bands on the face of the instrument. In accordance with the flight manual, the yellow band indicated the range, in which the aircraft may be operated in smooth air, and then only with caution in order to avoid abrupt control movements. Red marks indicated V_{NE} or the velocity that should never be exceeded. Engine RPM was left out for implementation because the simulated aircraft constant speed engine. The current heading was shown on the center top of the display. This location was chosen so that it would not cover up any head-down instruments inside the cockpit. In addition, a symbology visualized the setting of the flaps at the right bottom corner, analogous to its real position within the cockpit. Possible warnings were shown at the center of the display (e.g. excessive oil temperature etc.)



(a) Round dial concept

(b) Vertical tapes

Figure 3.9.: Primary flight display format. Black translucent background was set to white in this exemplary screenshot. Round dial concept (left) and vertical tapes (right). [133]

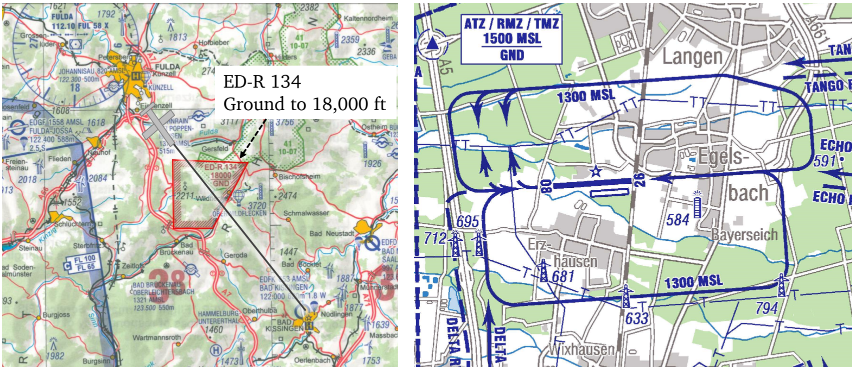
Compared to symbologies on PFDs, which use tapes, the round dial display version showed an arched scale that was somewhat similar to conventional analogue instruments. The design was based on a prototype by Korn, Schmerwitz, Lorenz, and Döhler [88] from DLR (German Aerospace Center). It was assumed that the similarity of the round dial to conventional round instruments would positively impact usability. It was believed that changes in speed or altitude could be more easily perceived, which could be attributed to the display's similarity to analogue instruments. Research points towards inferior readability for tape-like digital instruments [148].

3.3 Preliminary Tests

The preliminary evaluation was conducted in order to compare subjective usability and subjective workload of the implemented display formats and to uncover design limitations and advantages. Display formats were evaluated in two separate user trials using the FSR research flight simulator. A total number of $N = 14$ pilots participated in the simulator studies. The Airspace Viewer and the Flight Tunnel display format were evaluated by the same group of participants ($n = 7$), resulting in a within-subject design. The two versions of the PFD display format (tapes vs. round dial) as well as the recording of a baseline were evaluated with a separate group of participants ($n = 7$). The study design was chosen because it allowed a time efficient method.

The pilot age varied between 19 and 63 years ($M = 37.7, SD = 14.68$), and pilots' total flight time under VFR ranged from 140 h to 1,100 h ($M = 415.4, SD = 359.5$). Participants were recruited through mailing lists and notices posted at local airfields. Pilots with prescription glasses were restricted from participating unless they wore contact lenses. Participants were not compensated monetarily, but were offered the opportunity to fly in the simulator for their own enjoyment. For the evaluation of subjective workload and subjective usability, the NASA TLX [73] and the System Usability Scale (SUS) [15] were used. The questionnaires are thoroughly described in Section 5.4.3.

Following familiarization, the participants were briefed on their assigned scenario and were given all required materials for VFR flight execution. The recording of the baseline was performed before either version of the PFD display format was used. The baseline scenario included flying with the built-in Garmin G1000 moving map display and also with only printed approach charts. The test scenario for each display format on the smart glasses, (except the Airspace Viewer), included flying the VFR traffic pattern of local GA airfield Frankfurt-Egelsbach (EDFE) (see Figure 3.10a). The scenarios for the Airspace Viewer differed, and included a flight towards the restricted airspace ED-R 134, ranging from ground to 18,000 ft, starting at waypoint "SOLVU" near Fulda, Germany at 5,000 ft altitude to Airfield Bad Kissingen (EDFK) (see Figure 3.10b). Pilots were asked to avoid the airspace using the visual means provided on the smart glasses and return to heading 144° after passing the restricted airspace.



(a) Scenario for Airspace Viewer. Symbolic aircraft, track and conflicting ED-R 134 highlighted

(b) Scenario for PFD format and flight tunnel display format. Runway 26, right handed traffic pattern

Figure 3.10.: Provided material for flight execution of the selected scenarios. Left: Aeronautical charts [26]. Right: Approach charts [27].

After each scenario, participants were asked for verbal feedback and were given the written questionnaires regarding the TLX and SUS. Data processing of both scales is described in Section 5.7.2.

3.3.1 Results and Discussion

Even though the chosen scenarios differed, the test procedure and large parts of the methodology were identical, making it possible to compare the results. TLX and SUS scores were computed as described in Section 5.7.2. The score on the SUS was computed as general usability only and the results are displayed in Figure 3.11. These are shown in a way that compares ratings for workload and usability. Both scales range from 0 to 100. Higher values on the workload scale represent higher workload levels, which is generally not favored [4]. Higher values on the SUS represent better usability, which is favorable.

In one scenario participants flew with conventional printed charts and the support of the built-in Garmin G1000. This scenario can be considered as a baseline.

[4] Extreme low levels of workload are also unfavorable. According to the Yerkes-Dodson Law [160] (see Section 2.3.2) low levels of arousal are associated with low performance. Monotonous task may result in a variety of impaired decision making [3].

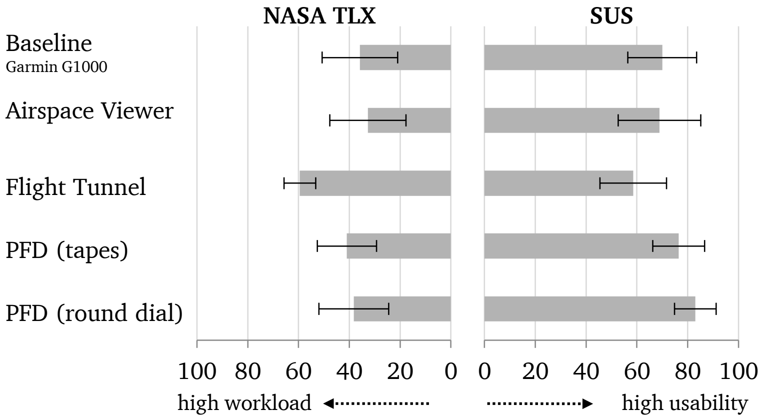


Figure 3.11.: Results on workload and usability (means and 95% confidence intervals, $n = 7$).

Values for baseline on the TLX were $M = 35.85$ ($SD = 20.01$) and for the SUS $M = 70$ ($SD = 18.26$).

To test whether the levels on the scales for workload and usability differed significantly between the display formats, separate analyses of variance (ANOVAs) were calculated. First, the results that compare the scales are given. Furthermore, display-format-specific results will be presented and discussed in the following sections, separately for each format.

There was a significant effect of the type of display format on levels of workload ($F(3{,}27) = 2.26, p = .023, \bar{r}^2 = .21$). This suggests that mean values are generally different. The Tukey's honest significant difference (HSD) post-hoc test revealed that levels of workload in the Flight Tunnel group ($M = 59.48, SD = 8.49$) were significantly different from all other groups $p < .01$. The effect size, calculated using eta squared, was .270. Display formats, other than the Flight Tunnel display format, were not significantly different from each other in terms of mean workload ratings.

There was a significant effect of the type of display format on levels of usability ($F(3{,}27) = 2.78, p = .005, \bar{r}^2 = .33$). The Tukey's HSD post-hoc test revealed that levels of usability in the Flight Tunnel group ($M = 58.57, SD = 17.73$) were significantly different from all other groups. The effect size, calculated using eta squared was .343. Display formats, other than the Flight Tunnel display format, were not significantly different from each other in terms of mean workload ratings.

Results suggested that subjective ratings on both the TLX and the SUS scale were not different from each other, except for one display format. The Flight Tunnel display format showed significantly different subjective ratings, pointing towards insufficient usability and increased task load for the pilot.

3.3.1.1 Airspace Viewer

All pilots successfully maneuvered around the restricted zone and the relative position of the restriction was correctly estimated by all pilots. In addition, pilots were asked to estimate the distance between their current position and the restricted airspace. Absolute estimation was rather difficult for participants. When prompted, the average distance to the airspace was 6600 m ($SD = 2767$ m). On average, pilot's estimations were off by 4250 m ($SD = 2449$ m).

However, it is not known whether the deviations were the result because of missing spatial context with ground, or the general difficulty of estimating distances with the given texture used in the simulated outside view. Using textures as cues for distance was not supporting pilots in estimating the absolute distance. This was certainly the case when airspaces were not reaching down to ground level, but began at a certain height. It was assumed that the virtual airspace was missing a linkage to fixed ground objects that would give the virtual objects context in the real world, which in turn would make it easier to estimate the airspace absolute position and distance. Such a concept is presented in Section 7.2.

3.3.1.2 Flight Tunnel

Subjective ratings of workload and usability turned out to be less positive for the Flight Tunnel than ratings of other implemented display formats. The implemented prototype of the Flight Tunnel display format was not suited to support pilots because of the high levels of induced task load. This can partly be attributed to the irritating effect of the image lagging behind the head movements by the pilots. Due to the application requiring high amounts of computing effort, images were rendered slowly, which resulted in a lagging effect. Participants reported that this lag induced reduced head movements. The absolute lag was not determined.

In addition, the subjective workload appears to have been negatively affected by the troublesome interpretation of the symbology for climb and descent. Pilots stated that it was problematic to track turns in the traffic pattern, as the minimal symbology - which was intended to reduce clutter on the display - overlapped with the path ahead [139]. The color coding in left and right segments did not effectively help to anticipate segments in turns.

Recorded flight tracks suggested that both lateral and vertical FTE was reduced. Results are compared to the control group that flew with printed approach charts. The average lateral FTE was smaller when using the Flight Tunnel display format on smart glasses ($M = 76.57\,\text{m}, SD = 91.20\,\text{m}$) compared to the control group ($M = 142.48\,\text{m}, SD = 150.17\,\text{m}$). However, differences were non-significant. Similarly, the average vertical FTE was not significantly smaller ($M = 20.95\,\text{m}, SD = 17.45\,\text{m}$) compared to the control group ($M = 28.01\,\text{m}, SD = 24.94\,\text{m}$).

Figure 3.12 shows the flown traffic patterns at EDFE (Frankfurt-Egelsbach) as a baseline (see Figure 3.12a) and with the support of the tunnel-in-the-sky display format on smart glasses (see Figure 3.12b).

It is clear from the recorded tracks that with the use of a tunnel-in-the-sky symbology, downwind and base segments showed fewer deviations from the published traffic patterns. However, some participants showed greater deviations in the turn from crosswind to downwind. It appears that the turn was initiated too late, resulting in a wider loop. Participants reported to have difficulty correctly interpreting the display format, which was resolved when flying the downwind. The interpretation issues became noticeable only in the second turn. The first turn (from upwind to crosswind), was located just before the highway and therefore allowed good visibility, which caused pilots to turn intuitively even without full understanding of the working principle. After unintentionally leaving the tunnel, participants had difficulty reentering, which can be seen in the recorded tracks.

(a) With conventional means **(b)** Tunnel-in-the-sky support

Figure 3.12.: Recorded flight tracks during preliminary testing of Flight Tunnel at EDFE (Frankfurt-Egelsbach) [139].

3.3.1.3 Primary Flight Display

The two versions of the PFD were not statistically different. Nevertheless, the round dial display format was slightly better and qualitative feedback tended towards a preference for the round dial display format.

Results showed that participants could fly effectively without extensive training. Low levels of subjective workload indicate an efficient manner of information acquisition. Participants reported that they spent less time monitoring the head-down instruments. Suggestions for improvements included extended color markings for operating speeds, precise vertical speed information and dynamic adaptability for different phases of flight. All suggestions for improvement were implemented in the successor display format: the LGL display format, which is covered in the following section.

3.3.2 Tracking

Contact-analogue display formats require an agile and precise tracking. In the course of the evaluation it became clear that the internal HTS (see Section 2.2.3.5 on tracking methods) was insufficient for a precise contact-analogue projection within the research simulator. Motion sensors turned out to be susceptible to drifting and signal distortion. Since the measurement of the glasses' orientation is strongly dependent on sensor data from the magnetometer, drifting was induced by the strong electromagnetic field within the simulator caused by the electrical servomotors. To improve the precision of the tracking capabilities, alternative visual tracking methods based on the built-in camera were taken in consideration [36]. Unfortunately, the implemented algorithms did not perform fast enough on the system to be considered feasible. Graphical data processing extracted visual markers in the cockpit. The potential of visual tracking methods was narrowed by the recording of images. The internal camera could record images at a highest frequency of 25 images per second, which supposedly is not fast enough to supply a flowing, uninterrupted superimposed image. However, the graphical data processing was the limiting bottleneck. The processing of one frame took around one second.

3.4 Pre-selection of Display Format

Opportunities for improvements were identified in the preliminary evaluation and user feedback provided valuable insights. The PFD display format received the best subjective ratings for usability. At the same time, becoming accustomed to the PFD format took very little time and induced no extra task load for the pilot.

The flight tunnel display format did effectively lead to smaller FTE. However, low subjective usability indicated that the chosen symbology was not user friendly, which simultaneously evoked high levels of workload. Low usability could be attributed to the lagging of the visualization as well as to clutter and ambiguity.

The Airspace Viewer display format was rated positively but would require larger displays to fully deploy its potential.

Since tracking challenges continued, it was decided not to further develop contact analogue AR display formats but to focus on those display formats that are realizable using existing technology or products that will be available in a shorter time frame. Improved tracking capabilities are needed to fully deploy the potential of AR display formats. It was decided to not further elaborate on contact analogue formats, but focus instead on head-referenced visualization.

The PFD display format was chosen for further development. User feedback in the evaluation was used to improve the visual design as well as the instrument behavior. Instead of using a flight tunnel for flight guidance within the VFR traffic pattern, a more minimalist approach was taken, resulting in the LGL display format, which will be covered in the following chapters.

4 Lateral Guidance Line (LGL) Display Format

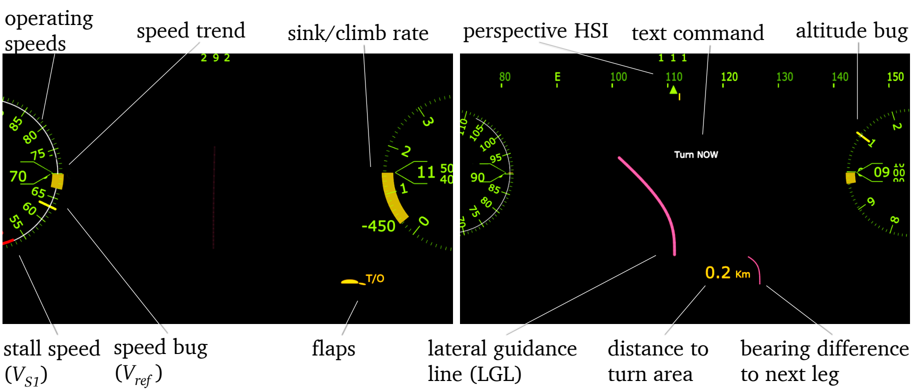
In the previous chapter, multiple display formats were implemented and preliminary tests were conducted. The PFD format resulted in the highest levels of usability with the lowest subjective workload. The display to be presented in the following chapter, the LGL display format, was built upon the existing PFD format (see Section 3.2.4) and further developed in a user-centered design approach. The appearance of the display format was redesigned according to suggestions of participants in the preliminary tests.

The functionality for a lateral flight guidance was added to the format. Instead of using a flight tunnel symbology within the VFR traffic pattern, a more minimalistic approach was taken: this resulted in a lateral guidance symbology. In the following sections, the implementation is described. Hereafter, hypotheses for investigation in the final evaluation are formulated.

4.1 Implementation

In order to reduce cluttering, the display changed its appearance based on the flown flight phase or position within the pattern. Figure 4.1 shows the display format during final approach and in the crosswind before turning downwind. Only task or content relevant information was displayed while less relevant information was faded out. For example, in final-approach, the airspeed and altitude indicator increased in size to allow better readability, while the HSI was faded out. An algorithm based on the cross-track error effectively detected the current position within the pattern.

A magenta colored lateral guidance line in the center of the format supported the pilot in maintaining track. The lateral guidance symbology was designed to show when to initiate a turn and when the turn should be completed in order to minimize deviations from a desired track. This minimalist working principle was based on feedback obtained from the preliminary evaluation. Feedback from the panel of expert pilots suggested that the process involved in navigation, especially in VFR traffic patterns, could be broken to two essential tasks: Firstly, pilots need to know when the turn should be initiated and secondly, when the turn should



(a) Final approach **(b)** Crosswind shortly before turn

Figure 4.1.: Features of LGL display format on smart glasses while in final approach (left) and before a turn area (right). Image modified for better readability. Black background color is transparent on smart glasses.

be completed. Vertical guidance has only a secondary role. Vertical guidance was achieved by an altitude bug, displaying the traffic pattern's altitude.

The symbology that provided guidance back to the desired ground track became visible when deviations from a desired heading or prescribed ground track exceeded a defined threshold. The larger the lateral deviation became, the stronger the lateral guidance line bent in the required direction.

The bending of the lateral guidance line to either side was achieved via a simple control law. The commanded course change ψ_{change} was calculated with a rhumb line control law (see Equation 4.1) and created either a positive (steer right) or negative (steer left) result

$$\psi_{change} = \kappa_1 \cdot D_{ct} + \kappa_2 \cdot (\psi_{desired} - \psi_{true}) \tag{4.1}$$

where

κ_i $\quad$ $i = 1{,}2$ is a constant (as a scaling factor),

D_{ct} $\quad$ is the cross-track distance,

ψ $\quad$ is the bearing.

The working principle is visualized in a pictorial schematic in Figure 4.2. Waypoints were considered to be so-called fly-by waypoints. In order to intercept from one track to the next, it is necessary to penetrate the circle around the way point, here called turn area. Once inside the turn area, the flight guidance seamlessly switched to the next leg. Turn areas for 90° turns were larger than turn areas for 45°. They were calibrated with the mean turn rates of test pilots who flew turns at 75 knots in the simulator. Furthermore, turn areas adjusted in radius, depending on the airspeed flown, in order to minimize FTE on the next leg. With a constant bank angle, the flown radius of a curve is proportional to airspeed. It was assumed that it took pilots a constant time to react to turn commands, therefore faster aircraft required the command to display earlier. With greater airspeed, the turn areas slightly increased in size, therefore the transition from one leg to the next was initiated at a farther distance before reaching the switch point.

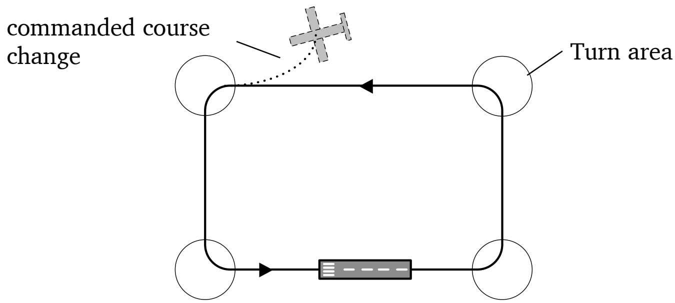


Figure 4.2.: Pictorial schematic of the flight guidance working principle within VFR traffic pattern.

The algorithm that predicted the position within the traffic pattern was based on the detection of the smallest cross-track error to the different legs. The cross-track error was calculated using Equation 4.2 [140]. The equation makes the assumption of a spherical earth.

$$D_{ct} = \arcsin(\sin(\delta_{13} \cdot \sin(\psi_{13} - \psi_{12})) \cdot R \quad (4.2)$$

where

δ_{13} is the angular distance from start point to current aircraft position,

ψ_{13} is initial bearing from start point to current aircraft position,

ψ_{12} is initial bearing from start point to end point,

R is the earth's radius in kilometers.

The bearing ψ_{13} and ψ_{12} used in Equation 4.2 were calculated with the atan2 [99] function. The function calculates the arc tangent of $\frac{y}{x}$ followed by a computed determination of six cases as described in Appendix D. The function and its arguments are

$$\psi_{1i} = \mathrm{atan2}(y,x) \quad \text{with} \quad i = 2,3 \tag{4.3}$$

$$y = \sin(\Delta\,\mathrm{lon}) \cdot \cos(\mathrm{lat}_2) \tag{4.4}$$

$$x = \cos(\mathrm{lat}_1) \cdot \sin(\mathrm{lat}_2) - \sin(\mathrm{lat}_1) \cdot \cos(\mathrm{lat}_2) \cdot \Delta\,\mathrm{lon} \tag{4.5}$$

where

$\mathrm{lat}_1, \mathrm{lon}_1$ is the start point (latitude,longitude),

$\mathrm{lat}_2, \mathrm{lon}_2$ is the end point (latitude,longitude),

$\Delta\,\mathrm{lon}$ is the difference in longitude.

The area at the bottom contained information about the direction of the next turn (right or left), the bearing difference to next leg (90° vs. 45°) and the distance until reaching the turn area. Expert pilots suggested absolute distance in kilometers[1], but it would also be reasonable to display the time until reaching the turn area. Once inside a turn area, a textual command instructed “turn now”. The HSI on the top of the format visualized the course over ground

Furthermore, the appearance and behavior of the instruments was adapted to expert suggestions. An additional marking for the operating speeds with regard to the flaps was added. The white band on the airspeed indicator resembled the normal range of operating speed for the aircraft (in accordance with the flight manual) with the flaps extended as for landing or takeoff. A yellow speed bug helped to maintain V_{ref}, which is the landing reference speed. V_{S1}, the stall speed, was marked red. The vertical speed indicator was redesigned. The arc shaped marking included the precise climb and sink rate next to it. Analogous to this, a speed trend was designed. This displayed the projected airspeed in 10 seconds.

[1] Pilots suggested kilometers instead of nautical miles because they would have a more intuitive feeling for distances in meters rather than in nautical miles.

4.2 Hypothesis

Using a head-mounted display could help the pilot maintain an *eyes-out* view. Because the most important information is presented in the FOV, pilots do not necessarily need to move their gaze downwards to conventional head-down displays. Spatial awareness could increase due to the fact that fewer shifts in gaze between inside and outside the cockpit are necessary.

> **Hypothesis 1:** Spatial awareness will be higher for the LGL display format than for the conventional moving map display on an HDD. *Eyes-out* time, as measured with an eye-tracking device, will be higher. Awareness for signals on the outside view will thus increase. Reaction times to signals on the OTW projection in turn will be shorter and more signals will be detected.

Pilots should maintain a certain track, especially within the VFR traffic pattern, for example. The lateral guidance symbology might assist the pilot in maintaining a prescribed track and therefore would help reduce the cross-track error to it.

> **Hypothesis 2:** Lateral deviations will be lower for the LGL display format than for the conventional moving map display on an HDD. The flight precision within the VFR traffic pattern will increase and thus the cross-track error will decrease.

An adequate level of task load is crucial to safely manage all tasks in the cockpit. The LGL display format could possibly lower the task load because it presents the most crucial information directly into the FOV. Consequently, information can be perceived effortlessly.

> **Hypothesis 3:** Subjective perceived mental workload will be lower for the LGL display format than for the conventional moving map display on an HDD.

Multiple display formats were developed within the scope of this thesis. The LGL display format is based on the PFD display format. Feedback from the evaluation of the PFD display format was integrated into the design of the LGL display format. Even though the primary information is similar for many aspects, additional features and improvements will presumably have a positive influence on user satisfaction.

Hypothesis 4: Subjective user satisfaction is higher for the LGL display format than for the precursor display format, the PFD display format (see Section 3.2.4). This applies to the general usability as a scale.

5 Method: Evaluating the LGL display format

This chapter describes the methodology for the evaluation of the LGL display format. To test the effects on spatial awareness, flight technical error, subjective usability and subjective workload, the LGL format on the Moverio BT-200 was experimentally compared to a control group, using a flight simulator study. The control group scenario included flying with a conventional head-down display (HDD) moving map symbology on the Garmin G1000. Subjective usability and subjective workload were queried. For evaluation, a total of $N = 20$ pilots participated. Pilots flew multiple traffic patterns at an unfamiliar airfield either with smart glasses or with conventional means. Flight technical error, *eyes-out* time and reaction times to a secondary task for visual attention in the outside view were recorded. Pilots rated the workload and perceived usability.

5.1 Participants

Between February and March 2016, a total of $N = 20$ pilots participated in the simulator study. One pilot was female, the remainder male. The ages varied between 19 and 55 years ($M = 37.3$ years, $SD = 9.2$ years). The pilots' total flight time under VFR ranged from 60 h to 2,275 h ($M = 444.2\,\text{h}, SD = 554.5\,\text{h}$). All pilots held at least a light aircraft pilot license (LAPL), private pilot license (PPL) or ultra-light (UL) license and had valid medical certificates. Their qualification varied: Three pilots held a commercial pilot license (CPL), one pilot held an airline transport pilot license (ATPL). 47% of the pilots were flying most regularly at Frankfurt-Egelsbach (EDFE), 13% in Babenhausen (EDEF) and 13% in Aschaffenburg (EDFC). None of the pilots had prior experience with smart glasses. Participants were recruited through printed advertisements on notice boards and by electronic mailings to former study participants. Pilots with prescription glasses were restricted from participating unless they wore contact lenses. There was no monetary reimbursement offered. Participation was promoted by offering the opportunity to fly the DA 40-180 simulator after the session.

Table 5.1. Demographics and Flight Experience of Participants.

	M	SD	min	max
Age in Years	37.3	9.2	19	55
Hours of Flight under VFR	444.2	554.5	60	2,275
Landings under VFR	654.1	385.5	55	1,500

5.2 Research Design

In this simulator study a 2 x 2 mixed factorial research design was applied in order to evaluate the influence of the independent variables *smart glasses* on multiple dependent variables. Each participant flew with the LGL display (experimental condition: *smart glasses*) and also in a control condition *moving map HDD* in a balanced order, thus making it partly a within-subjects design. In the control condition, pilots flew with a conventional moving map symbology on an HDD and with approach charts. This means that in the session M1, participants used the smart glasses and subsequently in M2 the head-down display or vice versa.

To prevent unwanted familiarization between the two sessions, two different airfields were used, creating the factor *airfield*. The *airfield* was balanced between participants, basically creating four different groups of subjects. The research design is illustrated in Figure 5.1.

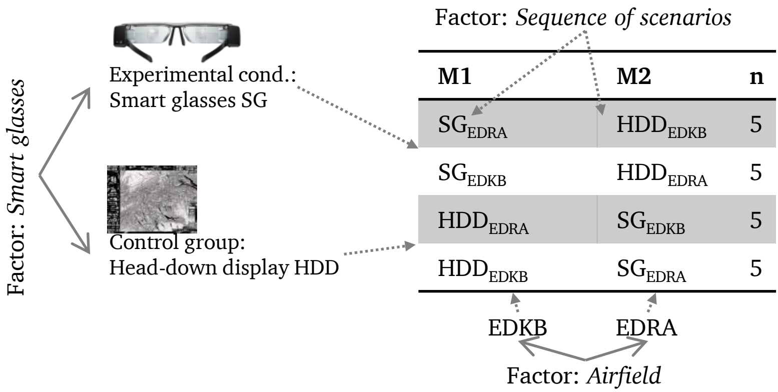


Figure 5.1.: Schematized research design and forming of the groups.

The *sequence of scenarios*, or in other words whether the participants first flew with or without the smart glasses was checked for influence on the dependent variable and is therefore another factor that varied between subjects.

5.3 LGL Display Format for Smart Glasses

The display format on the smart glasses used in this experiment was the so-called LGL format, shown in Figure 5.2. This format provides the pilot with dynamic flight state information and in addition lateral navigational information. See Chapter 4 for a detailed description of the display format.

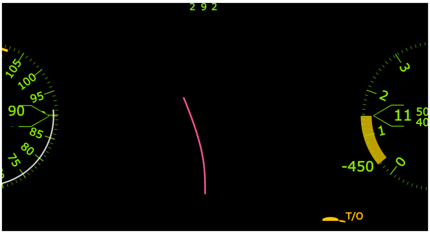


Figure 5.2.: Evaluated LGL display format.

5.4 Equipment and Measurements

Hypothesis 1 postulated that the LGL display format for smart glasses would lead to an improved spatial awareness compared to a quasi-control group that flew the scenario with a conventional moving map display format on an HDD and with printed approach charts. Two measurement approaches were selected to test this hypothesis. Firstly, a camera-based eye tracking device was chosen to measure the time spent scanning for traffic or obstacles (in other words *eyes-out* time). Additionally, a second approach to measure spatial awareness was selected. Reaction times on a secondary visual task were used.

The Diamond DA 40-180 flight simulator [79] was used for the study. Its cockpit was equipped with a Garmin G1000 avionics suite [60]. The simulator resembled a single engine DA 40-180 Diamond Star aircraft, a widely used and mostly IFR avionics equipped aircraft. The OTW view was generated using a modified version of the Diamond Global Canvas image generator software [30]. These OTW images were projected onto a cylindrical projection screen having a 200° × 30° FOV (see Figure 5.3). An instructor station allowed the experimenter to control the simulator and set parameters such as the simulated settings of the aircraft (e.g. mass) and the visual system (e.g. weather phenomena). The simulator was equipped with a force-feedback on the stick, resulting in a realistic impression of the dynamic flight control forces.

(a) Opened **(b)** 180° Outside Projection

Figure 5.3.: Diamond DA 40-180 flight training device.

5.4.1 Secondary Reaction Times Experiment

Response times to a secondary visual task were recorded using self-constructed measurement equipment based on an *Arduino Uno* micro controller [22]. The task was designed to resemble a traffic monitoring task. The schematic setup is depicted in Figure 5.5. A red dot, possibly other traffic, was projected to the OTW screen. The two degree of freedom (DOF) movable red laser beam was directed to predefined changing locations on the horizontal line of the OTW view. The system was placed above the instructor station (see Figure 5.4). The red laser diode was attached to two servo engines in a kinematic chain to allow movements. The kinematic chain of servo engines is displayed in Figure 5.6.

To prevent anticipation, the position of signals on the OTW view were initially randomized and then kept constant across trials. Pauses between signals were fully randomized, varying between 5 s and 20 s to allow for temporal uncertainty. The number of recorded signals therefore varied (M=18.05, SD=1.48). In retrospect, the variation in the number of presented signals between subjects was not desired. Instead, the number of presented signals should be the same between groups.[1]

[1] A Kruskal-Wallis test confirmed that there were no systematic differences in number of presented signals between groups $\chi^2(1, n = 40) = .16, p = .68$. The number of presented sig-

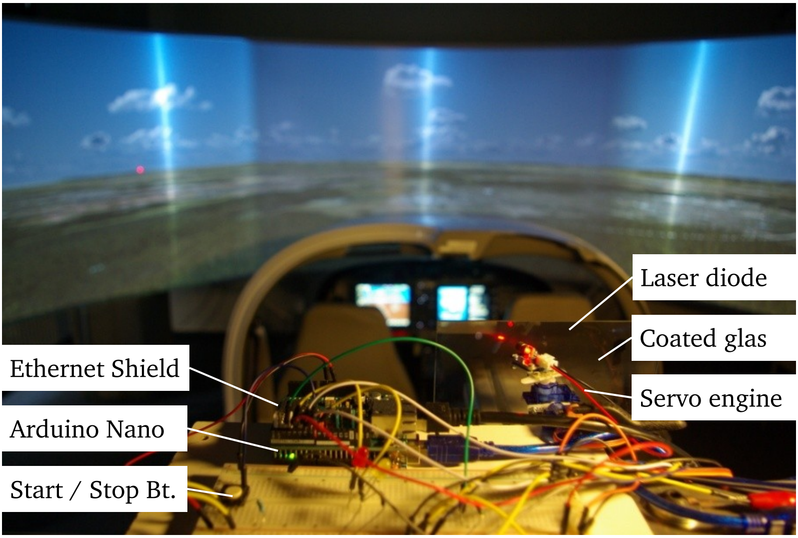


Figure 5.4.: Flight simulator cockpit mock-up and projection screen [back] with equipment for measurement of reaction times on a visual secondary task [front].

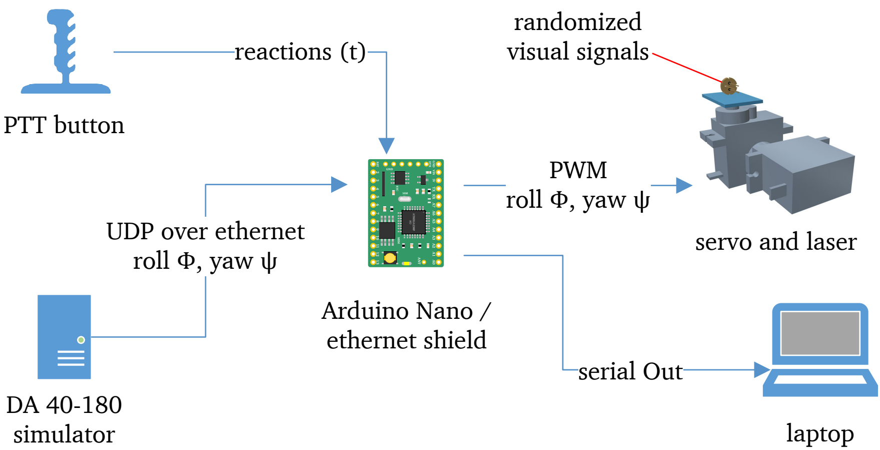


Figure 5.5.: Schematic setup of measurement equipment for recording of reaction times on a secondary visual task.

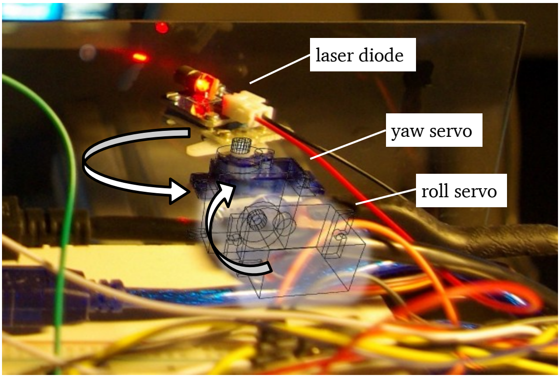


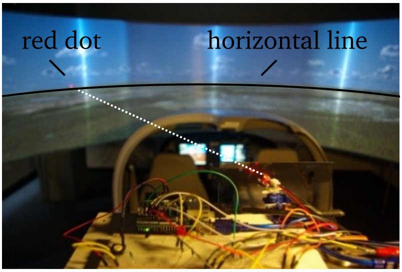
Figure 5.6.: Kinematic chain of servo engines. One servo for roll and for one for heading.

Each signal was displayed for a maximum of 10 s. The system used the push to talk (PTT) button as an input device. As soon as a signal was shown, the participant had to press the PTT button to indicate that the signal had been seen. The signal then terminated until the next signal was displayed. When the PTT button was pressed without a signal being shown, a false positive was recorded. If the participant did not react after 10 s, the signal disappeared and a missing value was recorded. Hence the recorded reaction times could possibly vary between 0 s and 10 s. The secondary task was paused between final approach and initial departure.

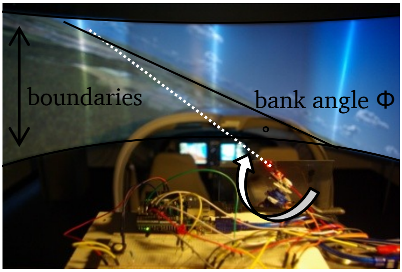
The position was automatically adjusted. If there was no adjustment for the flown bank angle the red dot of the laser would have appeared somewhere above or below the horizontal line, which is an unusual position for distant traffic. Pilots are accustomed to monitoring potentially conflicting traffic close to the horizontal plane [19]. Traffic is in conflict when at the same altitude and will then appear on the horizon line. As long as other traffic is distant, it will appear close to the horizon even if flying at a different altitude. In order to stabilize the laser beam on the horizon line in the OTW view, and thus giving the appearance of another aircraft flying, the bank angle of the simulator was mapped to the servo controlling the movement around the x-axis (roll axis). In Figure 5.7 the movement of the roll servo engine is illustrated. Any bank angle that would cause the red dot on the projection screen to appear displaced was corrected by the roll servo. In this way, the red dot was held close to the horizon line, appearing in an area where a pilot

nals in the smart glasses group ($M = 18.1, SD = 1.51$) did not differ from the control group ($M = 18.0, SD = 1.49$), $d = 0.06$.

would suspect traffic, making it a more realistic measurement procedure for spatial awareness.



(a) at level flight



(b) with bank angle

Figure 5.7.: Measurement equipment for reaction times.

The Arduino Ethernet Shield II [5] was attached to the Arduino Uno micro controller [22] to receive the simulator's UDP data. The servo's position was updated in accordance with the simulators bank angle. Due to the limited vertical FOV of 30° large parts of the horizon may not be visible to the pilots when flying with a large bank angle Φ. The swing of the roll servo was therefore limited with coded rules for the laser beam to remain within the boundaries of the projection screen. With only moderate turns the laser beam would stay close to the horizontal line and would thus give the impression of another aircraft flying. There was no stabilization for pitch angle θ or heading ψ.

5.4.2 Eye Tracking

Spatial awareness was assessed using an eye tracking system. Head and eye movements were recorded using a camera-based eye tracking system: FaceLab [127] from Seeing Machines. This is a non-intrusive system measuring head pose (orientation, position), gaze (direction, position, pupil diameter, vergence distance and saccade events), eyelid data (duration and frequency of blink events, aperture and fatigue) along with facial features for emotional and effort related interpretation. Features are detected with a frequency of 60 Hz and an accuracy of ±1% translational and ±1° rotational. Depending on the selected field of interest the gaze rotational precision is between $0.5 - 1°$. The system's operational range allows head rotations of up till ±90° (around the vertical axis) and a tilt (around the lateral axis) of up to ±45° [127]. For a placement inside the DA 40-180 simulator a two camera / infrared (IR) pod setup was selected to allow capture in the so-called precision mode, i.e. an IR source is reflected in the subject's eyes and the relative position is then tracked. The working principle is illustrated in Figure 5.8.

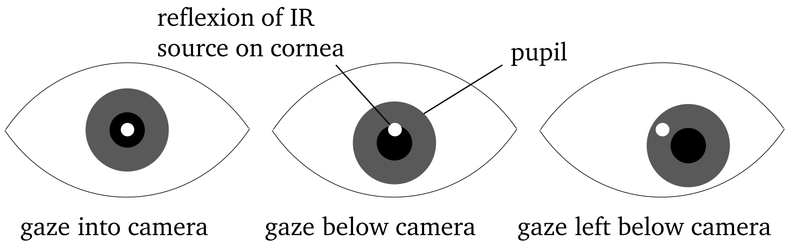


Figure 5.8.: Working principle of camera-based eye tracking systems. Illustration by author based on [33, p.58].

The cameras, with 12 mm focal distance lenses, were mounted onto both top corners of the pilot's Garmin G1000 display. The IR pod was placed 9 cm above the midpoint between the cameras. Figure 5.9 shows the position of the cameras' IR pod. The adjustable brackets enabled the cameras to rotate around the vertical axis and tilt around the lateral axis. This was necessary to adjust the picture frame for each participant.

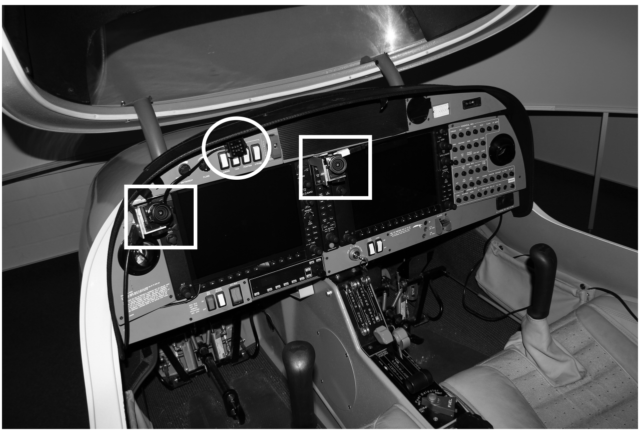


Figure 5.9.: camera-based eye tracking and mountings inside the DA 40-180 simulator cockpit. cameras (framed) and IR pod (circled).

Within the provided software, a world model was created in previous studies [77] by measuring the cockpit dimensions. Within the world model (see Figure 5.10), AOIs were specified: inside the cockpit, OTW projection (left, front, right), MFD.

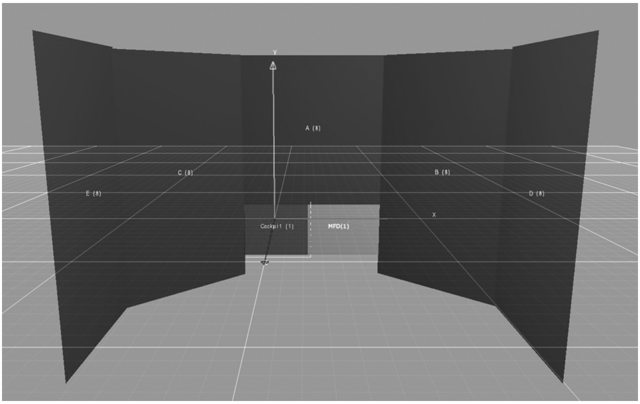


Figure 5.10.: FaceLab's world model of the DA 40-180 simulator used in the evaluation for capturing relative time on AOIs. [77]

A head model was created for each participant. This included taking five photographs of the participants head while oriented in different directions. Reference points and facial features were then marked on the photographs in order to allow the head pose estimation.

5.4.3 Questionnaires

To assess the usability of the system, a German translation of the widely used SUS by Brooke [15], translated to German, was used. Besides the perceived ease of use as a single dimension, research has shown that the usability and learnability sub-scales can be interpreted independently [94]. Items 4 and 10 provide the learnability dimension and the other 8 items provide the usability dimension. Ratings on the SUS are usually positioned relative to other systems, thus allowing relative comparison. Apart from contrasting ratings between systems, an extension to SUS by Bangor, Kortum, and Miller [8] provides an interpretation using adjectives.

The perceived workload was assessed using a German translation of the NASA TLX by Hart [73]. This is a multi-dimensional rating scale used to assess task load. It uses six dimensions of workload to provide information about the nature and relative contribution of each dimension in influencing overall task load. Test subjects rated the contribution made by each of six dimensions of workload to identify the intensity of the perceived workload. The NASA TLX provides an overall score based on a weighted average of ratings on the six sub-scales. The first required the test subject to evaluate the contribution of each factor (its weight) to the workload of the specific task by choosing between 15 possible pair-wise comparisons of words (semantic differential). Examples of questions are phrases such as "Which of the two contributed the most to mental workload during the task", "Mental or physical demand?", "Physical demand or performance?" and so forth. On the second page of the test, the test subject rates the magnitude of the dimension on a 21-level scale. The overall workload score for each participant is then calculated by multiplying each rating by the weight given to that dimension.

5.5 Procedure

A schematic of the experimental procedure is given in Figure 5.11. After welcoming, the participants were made familiar with the flight simulator until they reported feeling comfortable flying traffic patterns with the smart glasses. The familiarization lasted 30 minutes on average and included flying the VFR traffic pattern at a well-known local airfield, Frankfurt-Egelsbach (EDFE) and systems operation of the DA 40-180 flight simulator as well as the smart glasses. During parts of the familiarization, participants were asked to react to a secondary visual task (red dot) for familiarization with the reaction time experiment. Participants were asked to confirm that they had seen the projected red dots by pushing the PTT button. The response time task was paused during final approach and initial climb and continued when traffic pattern altitude was reached.

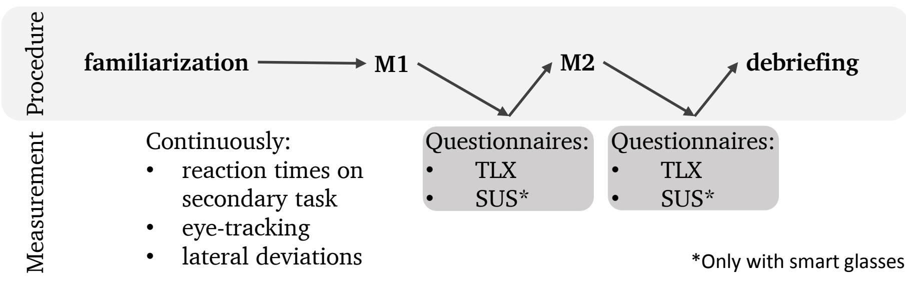


Figure 5.11.: Schematized procedure and measurements.

Following the familiarization, two sessions were flown. Before each scenario, the participants were briefed on their assigned scenario and were given all required materials for VFR flight execution. For better comparison, especially of the eye tracking data, smart glasses were worn throughout the experiment. In the control condition, the smart glasses remained black. After each scenario, participants were asked for general impressions and were given the written questionnaires of the TLX [73]. In addition, the SUS [15] was administered after completing the experimental condition. Depending on their assigned condition, pilots first flew with or without the support of the LGL display format on the smart glasses. Similarly, in each trial one of two airfields was selected, in accordance with the assigned condition. After completing the simulator session, participants were debriefed.

5.6 Experimental task

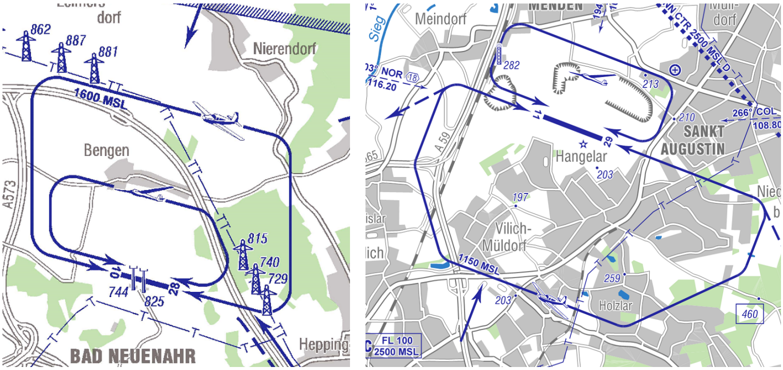
Participants were instructed to fly along the briefed traffic pattern of their scenario, applying their usual tolerances. At the beginning of each session, they were reminded to acknowledge all perceived red dots on the OTW view by pressing the PTT button.

5.6.1 Scenarios

Two German airfields were selected because their VFR traffic patterns feature an unconventional geometry, i.e. they are non-rectangular. Both traffic patterns are shown in Figure 5.12. Each trial consisted of two complete circuits in the pattern, with touch-and-gos in between, beginning with take-off, crosswind, downwind, base and final. In total, two participants (one in each condition) stated they had previously flown at EDKB. One participant had flown at EDRA.

Compared to conventional traffic patterns, the chosen patterns are non-rectangular and therefore more complex and demand greater navigational performance and attention from the pilots (see Section 1.4.1.2).

At the beginning of each scenario the aircraft's position was set to take-off position. Weather condition in the outside view were set to meet VFR conditions - at least 1.5 km flight visibility and clear of clouds. Also, the ground textures of the OTW view were of lower resolution compared to the familiarization airfield. The intention of this was to increase the level of difficulty, similar to a situation where a pilot's workload is slightly increased when flying at an unfamiliar airfield.



(a) EDRA Bad Neuenahr-Ahrweiler **(b)** EDKB Bonn Hangelar

Figure 5.12.: VFR traffic pattern used in the evaluation. Source: Deutsche Flugsicherung [27].

5.6.2 Validity of the Secondary Reaction Task

The equipment necessary for the secondary task was set up in such a way as to provide discernible indications of attention. Three major factors influence response speed in this paradigm: stimulus intensity, temporal uncertainty and expectancy. These factors were considered in order to meet validity. For example, if the stimulus intensity is too high, the signal will be easily detected even with low attention directed to the area. Thus, the test will not be able to discriminate between high levels and low levels of attention. In previous tests with research personnel the detection difficulty was tested and influencing factors were adjusted accordingly. The stimulus intensity, i.e. the laser brightness, was lowered to reduce contrast with the surrounding visual of the OTW view. In other words, the stimulus brightness was decreased. To prevent anticipation of the next signal, pauses between signals were randomized. The system was positioned well behind the cockpit so that it could easily be operated without affecting the expectancy of the upcoming signals.

5.7 Recording and Treatment of Data

Flight data was recorded with a frequency of 1 Hz and later processed with Mathworks Matlab 2016. Data was recorded to an SD memory card using the built-in function of the Garmin G1000. Reaction times on a secondary task were recorded on a self-built *Arduino* micro controller and transmitted via the serial port to a laptop computer. This data was manually saved as comma separated values after each session and later processed with IBM SPSS 17. Reaction times were averaged to calculate a score. Relevant flight data was filtered, and consisted of the position (latitude, longitude) and the altitude of the aircraft. The (lateral) flight technical error, the so called cross-track error, was calculated for each data entry using Equation 4.2. The cross-track error for each data entry was then summed to make up one value for each circuit in the traffic pattern for each participant.

5.7.1 Treatment of Reaction Times

Reaction times for the secondary task were averaged and the standard deviation was calculated for each scenario of the test subject. This was to compensate for different amounts of recorded data points. The secondary task was only run while the pilot was at a comfortable height —approximately 1,000 ft above ground level (AGL) in the VFR traffic pattern. No signals were given when the pilot was in the final approach or upwind. The signals were shown with variations between test subjects: A new signal was shown every 5 s to 20 s after a reaction by the pilot or a time out. This variability led to greater differences in the amount of data points, which made an averaging necessary in the subsequent statistical analysis.

5.7.2 Treatment of SUS and TLX ratings

Scores on both tests were computed from the ticks marked on the written questionnaires. Scores on the SUS were computed according to Brooke [15]. Inverted items were reversed accordingly. The sum of the score was then scaled to return a range of 0 to 100. Even though the SUS was intended as a one-dimensional measurement, the global score may be split up into two dimensions, (learnability and usability), as suggested by Lewis and Sauro [94]

Scores on the TLX were computed using Equation 5.1. The weights w is the total number of times that each dimension was selected. This ranged from 0 (not relevant) to 5 (more important than any other attribute). The global mental workload score was computed as a weighted average, considering the subjective rating of each attribute d_i (for the 6 dimensions) and the correspondent weights w_i.

$$\text{computed score} = \left(\sum_{i=1}^{6} d_i \cdot w_i \right) \frac{1}{15} \tag{5.1}$$

The value was converted in $[0; 100] \in \mathbb{R}$ for better comparison.

5.7.3 Treatment of Eye Tracking Data

The FaceLab system recorded data with a frequency of 60 Hz, meaning that approximately every 16.67 ms it computed and saved relevant head and gaze information. This was exported into a text format and further processed. Each data set for each participant consisted of two traffic patterns. For each scenario, a value was calculated that represents spatial awareness. It can be described as the relation between time focused on the OTW projection or in general cockpit outside and cockpit inside. The formula is shown in Equation 5.2. The expression CO:COI will be used further to refer to *eyes-out* time, which is an indicator for potential established spatial awareness, as measured within this study. A higher value of CO:COI represents a greater amount of time spent observing the OTW projection and therefore potentially more occurrences noticed, which is necessary for a high level of spatial awareness. On the other hand, a decrease of the computed variable would imply less time spent *eyes-out* because the visual attention was directed inside the cockpit.

$$\text{Time} \quad \textit{eyes-out} = \text{CO} : \text{COI} = \frac{T_{outside}}{T_{outside} + T_{inside}} \tag{5.2}$$

where

$T_{outside}$ is the total time spent by the pilot looking outside the cockpit, as registered by the eye tracking equipment,

T_{inside} is the total time spent by the pilot looking inside the cockpit, as registered by the eye tracking equipment.

Recorded data suffered a high rate of missing frames (as an analyzed frame within a (video-)stream). Missing frames reflect a situation in which the system's algorithm was not successfully tracking the virtual facial markers and/or the IR pod's reflection on the cornea. The system's algorithm was in certain cases unsuccessful with computing the gaze and head orientation, supposedly due to the wearing of the smart glasses. Other reasons that sometimes led to a complete loss of data were a faulty setup of the cameras, insufficient calibration, disruption of recording due to software crashes or accidental bumps against the cameras by the participants.

In order for the subject's gaze to be analyzed, it was determined that it was necessary that at least 40% of the recorded frames were successfully analyzed. In case more than 60% of the frames were missing, the complete data set was ignored, with the intention of eliminating those data sets with insufficient and non-useful data. The listwise deletion seemed to be the most appropriate method of dealing with the samples. It was assumed that data were missing at random and that there were no systematic differences of elimination between groups. Imputation of values by replacing missing values with substitute values (which is usually done with missing data) is not possible with raw eye tracking data because of the unpredictability of a participant's gaze in raw data.

Data sets of $n = 9$ participants were ignored due to high levels of missing data (see Section 5.7.3). It was checked whether there were systematic differences between groups before elimination of selected participants in terms of missing frames. An analysis revealed that there was no significant differences in the percentage of missing frames between the experimental ($M = 44.7$,$SD = 13.3$) and the control condition ($M = 45.9$,$SD = 13.03$);$t(38) = -.083$,$p = 0.775$. The actual difference in mean scores between groups, calculated using Cohen's d, was 0.09.

6 Results

$N = 20$ pilots participated in the study. Every participant completed the study. Subjective as well as objective measurements were recorded to test the hypotheses. Objective measurements were recorded continuously throughout the experiment, and two subjective scales were administered after each scenario. Results were first analyzed descriptively with common mathematical tools. Mathworks Matlab 2016 was used for data preparation and some visualization. Inferential statistics were calculated with IBM SPSS 17.0. The transformation of data and further linear modeling was performed using R Statistics. This chapter will begin with the descriptive analysis of the recorded data followed by inferential statistics methods for testing each hypothesis. After each section the results will be discussed.

6.1 Spatial Awareness (Hypothesis 1)

It was expected from Hypothesis 1 that the *eyes-out* time would increase by using a display format on the smart glasses, presumably because information is projected into the FOV and therefore reorientation towards inside the cockpit is not necessary. As a second approach, reaction times to a secondary task were recorded.

Results on Eye Tracking

To compare the time eyes-out (CO:COI) in conditions with and without the support of the LGL display format on the smart glasses, a paired-samples t-test was conducted.

As visible from Table 6.1, there was a significant difference in the scores of the condition with smart glasses ($M = 0.87, SD = 0.08$) and without smart glasses ($M = 0.69, SD = 0.07$); $t(10) = -7.147, p < .001$. These results suggest that using smart glasses did result in differences in the calculated parameter for *eyes-out* time. This would imply that the *eyes-out* time would increase significantly from 69% to 87%. With limited certainty, it can be concluded that using smart glasses for visualization of navigational information in highly dynamic situations may increase the time available for monitoring outside view.

Table 6.1. Paired Samples Test of *Eyes-out* Time (OC:OCI).

			95% Confidence Interval				
M	SD	SEM	Lower	Upper	t	df	Sig.
-.18	.08	.02	-.25	-.23	-7.15	10	.000

6.1.1 Results on Reaction Times

Reaction times varied between 612 ms and 9988 ms. The visual stimulus was presented on the OTW screen at changing locations on the horizon line. This can be interpreted as a changing bearing ψ towards the stimulus. Reaction times to signals in front of the pilot were shorter compared to signals in the periphery. This is illustrated in Figure 6.1.

A curve was fitted onto the data plot using polynomial regression. A second-degree polynomial shown in Equation 6.1 fits the nonlinear model to the data.

$$\text{ms} = 1.54 \cdot \psi^2 + 7.41 \cdot \psi + 3280 \tag{6.1}$$

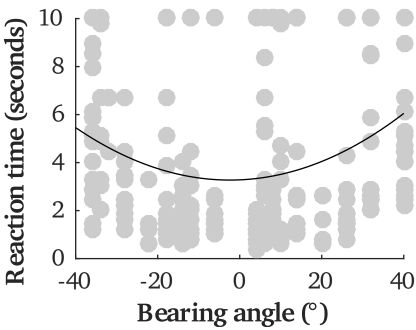


Figure 6.1.: Reaction times on a secondary task: familiarization at EDFE with a fitted line to the data. Missing values were computed in the data as a 10 s punishment.

Unlike the prediction, there were no differences in reaction times between the smart glasses group (M = 2,815.28 ms, SD = 565.07 ms) and the control group ($M = 2{,}718.42\,\text{ms}, SD = 498.87\,\text{ms}$). However, a Kruskal-Wallis Test revealed a statistically significant difference in numbers of missed signals between two groups $\chi^2(1, n = 40) = 7.83, p < .01$. The smart glasses group recorded fewer missed signals ($M = 1.45, SD = 1.31$) than the moving map HDD group ($M = 2.65, SD = 1.34$). Cramér's $V = .44$ points towards a moderate to large effect (see Appendix C). Mean values of the reaction times and the number of missed signals are presented in Figure 6.2.

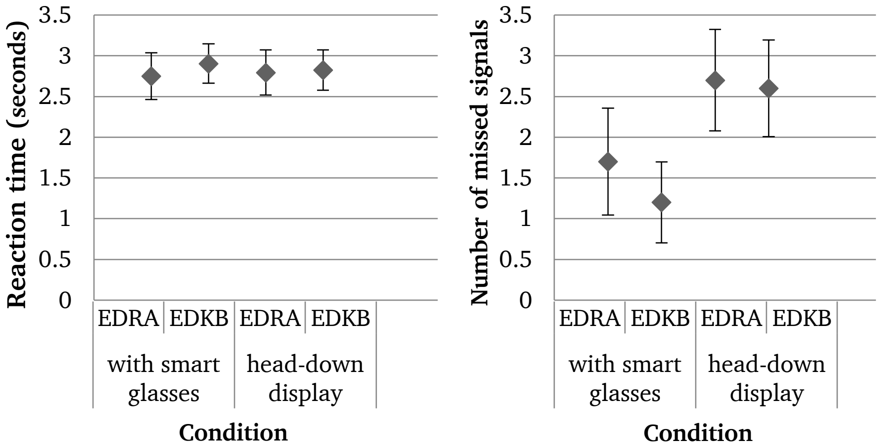


Figure 6.2.: Reaction times (left) and number of missed signals (right) on a secondary visual task (means and 95% confidence intervals).

As expected other independent variables did not show a statistically significant effect, meaning that neither the *airfield* nor the *order of scenarios* had an impact on reaction times or number of missed signals. Therefore, the selected experimental design was appropriate.

6.1.2 Discussion on spatial awareness

This study evaluated whether pilot's spatial awareness can be enhanced by a LGL display format for smart glasses. Spatial awareness was operationalized by two measures: The change in duration attended to the OTW projection (OC:OCI) and in addition reaction times on a secondary visual task.

There was a significant increase in the calculated parameter (OC:OCI) or in other words *eyes-out* time. It was increased on average from 69% to 87%, an increase of 18%. This suggests that the pilot has more time available for observing the outside view, which is crucial for establishing a high level of spatial awareness. The missing values or in other words missing frames in the recorded data sets that were included in the analysis ($N = 11$) was exceptionally high ($M = 43\%, SD = 9.01$) compared to prior studies using the research flight simulator and the same eye tracking device. With this measurement approach, it was not possible to verify with any certainty, whether or not objects were really perceived and cognitively processed. Therefore, a second measurement approach (reaction times on a secondary task) was chosen that analyzed the visual attention towards the outside view.

Even though mean reaction times did not significantly change, the number of missed signals (red dots on the OTW screen) were significantly less than in the moving map HDD group. It was assumed that due to the increase in *eyes-out* time, fewer signals were overlooked.

In summary, analysis of both variables provided evidence for an increased spatial awareness during the simulated flight. Hypothesis 1 was retained. Monitoring the OTW view is crucial for establishing a high level of spatial awareness during VFR flight. It is believed that smart glasses can prove beneficial in increasing the *eyes-out* time.

6.2 Flight Precision (Hypothesis 2)

This study evaluated the ability of smart glasses to reduce lateral flight technical error. It was hypothesized that a flight guidance system for smart glasses will lead to a higher precision in flying the pre-designated route. The displayed lateral guidance line, indicating deviations from the intended flight path, increases flight precision and therefore reduces the cross-track error compared to flying with conventional HDD using approach charts. Especially when flying in highly dynamic situations, such as a VFR traffic pattern, the flight technical errors would possibly decrease. Recorded flight paths are displayed in Figure 6.3. As a measure of the flight precision, the so called cross-track error was chosen. This is the lateral deviation of the predesignated track course. The cross-track error was calculated from the recorded flight data (see Section 5.7 for further information about the treatment of data).

6.2.1 Results on Flight precision

The average cross-track error, is shown separately for every participant in Figure 6.4. The cross-track error is the lateral deviation from the desired track, in

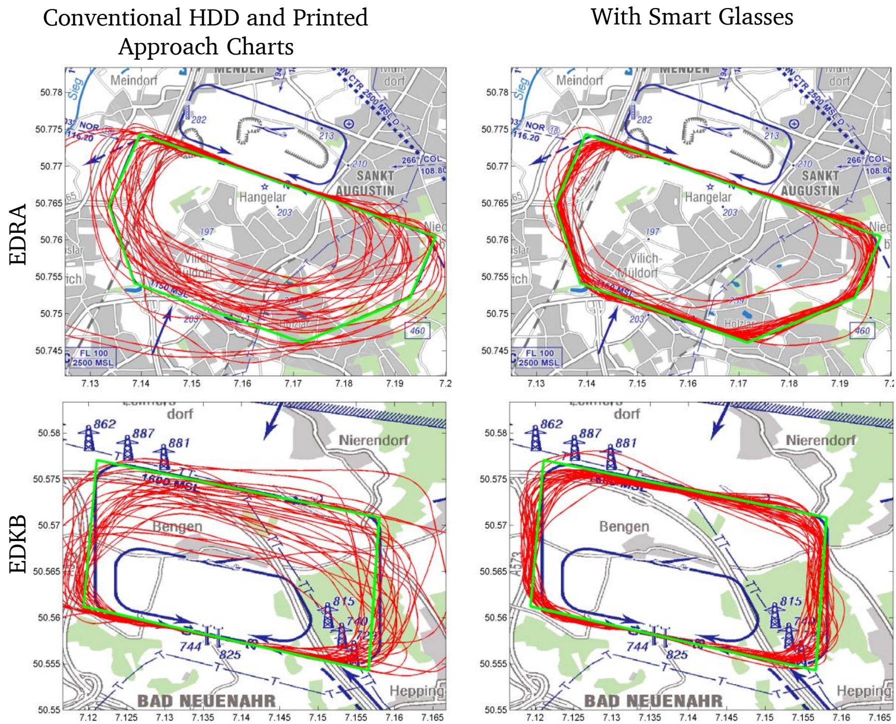


Figure 6.3.: Recorded flight paths of $N = 20$. The green line resembles the intended flight path, the red line the actual flown flight track. Background image source: Deutsche Flugsicherung [27].

this case, the deviation from the VFR traffic pattern. Data from $N = 20$ participants was analyzed. Data ranged from min $= 27.59$ m to max $= 373.89$ m.

From Figure 6.4, a difference between the condition with and without the LGL format on the smart glasses is observable. For every participant except one, the mean cross-track error decreased. In the next step the differences are analyzed using inferential statistical methods.

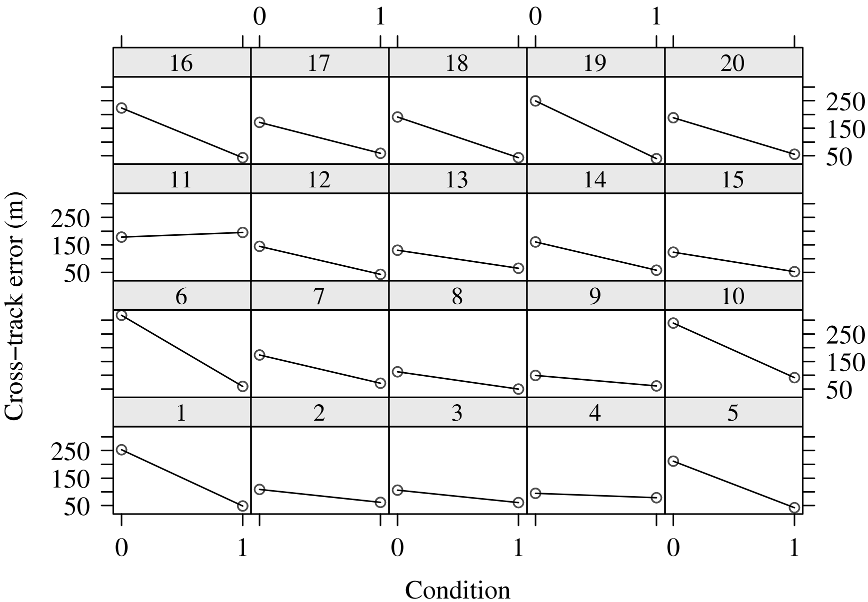


Figure 6.4.: Mean cross-track error, separately for every participant. Condition 1 = with the lateral flight guidance on the smart glasses, 0 = without.

The assumption of normality was violated for some participants during some experimental conditions. Please see Appendix B for a thorough analysis of the assumptions. In order to meet the requirements for inferential statistics the abnormally distributed data was transformed using the Box-Cox transformation. The transformation is described in Appendix B. After transformation, an ANOVA was calculated (see Table 6.2) resulting in a significant effect of the smart glasses condition ($F(1,76) = 117.92, p = .000$) with $r^2 = 0.60$, which points toward a small to moderate amount of explained variance (see Appendix C). The average cross-track error was smaller in the smart glasses group ($M = 63.95, SD = 37.09$) than in the conventional HDD group ($M = 176.52, SD = 77.6$). Means are displayed in Figure 6.5. Pilots could reduce their lateral flight technical error by 64% on average compared to flying with conventional HDD and approach charts.

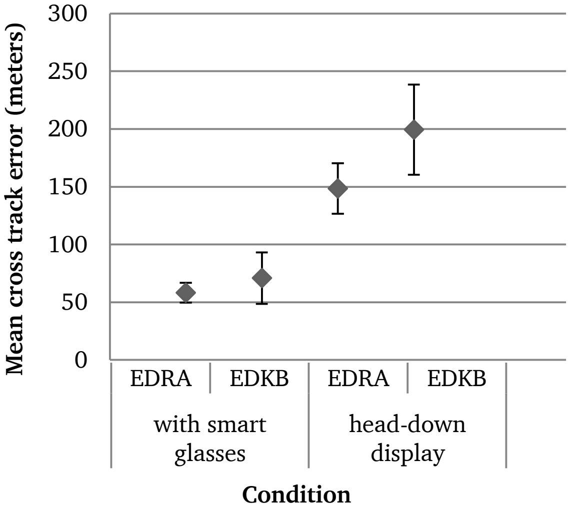


Figure 6.5.: Cross-track error. $N = 20$ (means and 95% confidence intervals).

6.2.2 Discussion on Flight Precision

There was a statistically significant difference: Pilots could reduce their lateral flight technical error. Hypothesis 2 was retained. It is believed that the LGL display has the potential to be used as an auxiliary display for navigation systems. Especially in highly dynamic situations, e.g. within approach and departure procedures, a flight guidance display format presented on smart glasses could be advantageous. As seen in Figure 6.4 a decrease in cross-track error was not observable for every participant. Participant number 11 has a slight increase in mean cross-track error. It was suspected that in this special case the participant exceeded the advised airspeed, which caused the algorithm to give turning instructions unusually early, eventually leading to missed turn areas subsequently.

Table 6.2. Summary of ANOVA on Transformed Values

	df	Sum Sq	Mean Sq	F	p
Smart glasses	1	47.29	47.29	117.92	.000
Airfield	1	1.23	1.23	3.08	.08
Order of scenarios	1	0.00	0.00	0.00	.99
Residuals	76	30.48	0.40		

6.3 Workload (Hypothesis 3)

One aim of providing pilots with assistance systems is to lower workload. Hypothesis 3 tested the effect of the display format on the pilot's subjective workload. It was hypothesized that the subjective workload will be smaller when the pilot flies with the LGL display format on the smart glasses compared to the control condition. In the control condition the pilot used a moving map symbology on the conventional HDD. Workload was measured after each scenario with the NASA TLX, a subjective questionnaire.

6.3.1 Results on Workload

Weighted ratings on the NASA TLX questionnaire as well as their underlying components (i.e. weights) are displayed in Figure 6.6. The importance weights result from ratings on a semantic differential. The six dimensions provide diagnostic information about the nature and relative contribution of each dimension influencing the overall workload. The component ratings are usually not analyzed separately [73], yet, they can help designers pinpoint the source of a workload or performance problem.

A paired-samples t-test was conducted to compare the overall TLX workload score in conditions with and without the support of the LGL display format on the smart glasses. As can be seen in Table 6.3 there was no significantly difference in the scores of the condition with smart glasses ($M = 53.26, SD = 14.09$) and without smart glasses ($M = 55.74, SD = 16.03$); $t(19) = -.475, p = 0.640$. The actual difference in mean scores between groups, calculated using Cohen's d, was 0.16. These results suggest that using the smart glasses did not result in differences in subjective workload. There was no significantly difference between flying con-

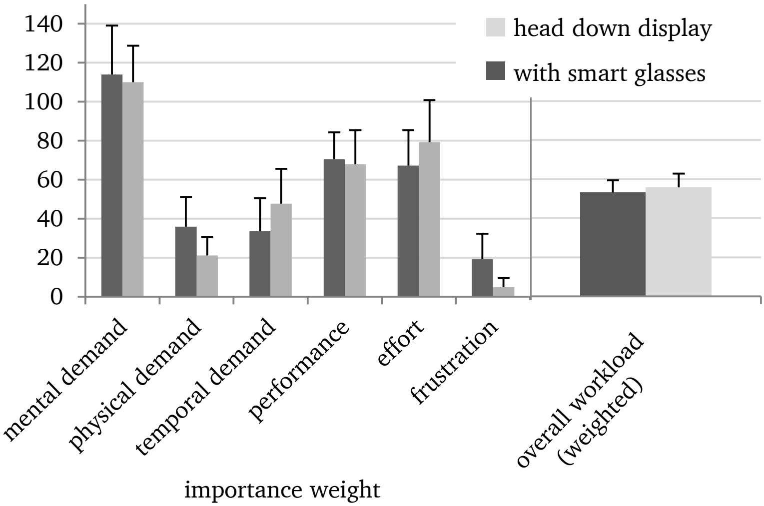


Figure 6.6.: NASA TLX scores (Means and 95% confidence intervals). Overall workload = mean of weighted ratings.

ventionally (paper map and digital moving map displays) and flying with the LGL display format on the smart glasses.

Table 6.3. Paired Samples Test on NASA TLX

			95% Confidence Interval				
M	*SD*	*SEM*	Lower	Upper	*t*	df	sig.
-2.49	23.42	5.24	-13.44	8.47	-.475	19	.640

6.3.2 Discussion on Workload

Results showed that there were no significant differences in workload levels between the groups with the LGL display format and the group with conventional HDD and paper maps. Hypothesis 3 cannot be retained. In general, subjective workload ratings in this study were relatively high (compared to other pilot assistance systems previously evaluated at our laboratory). Supposedly this has to do with the characteristics of the experimental task that induced high levels of task load. This can be attributed to an insufficient familiarity with the simulator and the smart glasses as well as the low-resolution scenery within the scenarios. Participants received a relatively short period of familiarization. After only three rounds both with and without smart glasses in the traffic pattern, most participants reported that they felt sufficiently familiarized to continue with the actual task. However, after the experimental task half of the participants stated that they underestimated the task and that more training would be necessary to comfortably use the system with minimal effort. The decreased pixel density on the OTW projection may also be responsible for an unusual high subjective workload.

6.4 Usability (Hypothesis 4)

Subjective perceived system satisfaction (also called ease of use) was assessed after each trial with the smart glasses.

6.4.1 Results on Usability

The average score, (*general usability*), for the LGL display format was 73 ($SD = 14.18$). It ranged from min = 45 to max = 92.5.

An independent-samples one way t-test was conducted to compare usability ratings for the LGL display format and the precursor, the PFD format (see Section 3.2.4). There was no significant difference in the scores for the LGL ($M = 73, SD = 14.18$) and the PFD ($M = 81.25, SD = 10.28$) format; t (25)=1.645, p = .121, d=.66. These results suggest that the mean SUS scores of both display formats are not significantly different from each other.

The sub-scales of the SUS were analyzed further, as suggested by Lewis and Sauro [94]. The mean values on each sub scale are illustrated in Figure 6.7. The two sub-scales were correlated as expected from the literature [94] (Pearson's bivariate correlation coefficient $r = .42$, $p = .04$). To test whether the sequence of scenarios had an influence on the ratings, a one-way ANOVA was calculated. There was no significant effect of the sequence of presentation on the overall SUS score ($F(1{,}19) = .006, p = .493, \eta^2 = 0.02$). Whether the participant first experienced the scenario with a LGL display format on smart glasses or flew with a conventional moving map display format had no significant influence on the ratings. Thus, the chosen experimental design proved useful.

6.4.2 Discussion on Usability

Usually, ratings on the SUS are positioned relative to other systems. Bangor et al. [8] developed an additional adjective rating scale for the SUS based on a review of hundreds of usability studies in order to help interpret individual scores and aid in explaining the results. This extension to the SUS is shown in Figure 6.8. Based on the adjective rating scale the general usability of the LGL display format for the given task would translate to *good*. Hypothesis 4 suspected that further improvements in the precursor of the LGL display format would result in even better ratings. In fact, ratings on the SUS tend to be lower for the LGL display format than for the precursor, even though it is not significant on a 5% level. Hypothesis 4 cannot be retained. The usability of the LGL display is not significantly different from its precursor.

The unexpected lower usability rating for the LGL display format could be explained by deviating levels of task load in the experimental task. The experimental task within the evaluation of the LGL display format was to fly two consecutive traffic patterns at an unknown airfield with prior familiarization at a different airfield. However, the experimental scenario for the precursor, the PFD format, took

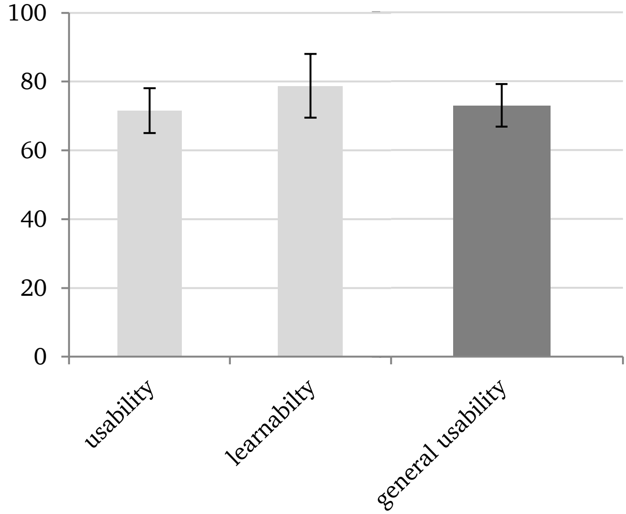


Figure 6.7.: System Usability Scale: Usability and learnability subscale ratings (Means and 95% confidence interval).

place at the same airfield as the familiarization. Longo and Dondio [95] showed a linkage between task load and perceived usability. Increased levels of workload tend to lead to lower ratings on usability. Presumably the higher task load within the evaluation of the LGL display format could explain the lower usability ratings.

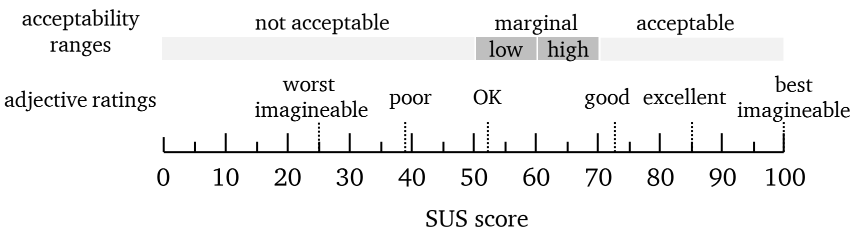


Figure 6.8.: Interpretation of System Usability Scale adapted from Bangor et al. [8].

6.5 Verbal feedback and participant responses

Usability of the evaluated display format was not undisputed. Circumstantial evidence from unstructured interviews suggested that 2 of 20 participants were at least concerned with the practicability. Participants expressed discomfort and distress, which could be resolved by adjusting the hardware. Readability problems that could not be resolved by readjusting the nose pads of the glasses were reported by 5 of 20 participants. Further feedback was interpreted qualitatively. Foremost participants remarked that the system distracted from the information in the background; information presented on the smart glasses was described as being in front, and participants stated that it was hard to combine an image of both foreground and background. They also found it difficult to have a sharp image on the combiner of the smart glasses. Participants stated that the HSI was noted but most ignored this information in favor of the lateral guidance line, which was more intuitive, as believed.

6.6 Hardware constraints

This study did not focus on the hardware itself, but instead on the display formats and the way that information can be visualized in order to enhance SA. For this reason, this thesis will not discuss hardware usability and constraints in depth. Even though the Moverio BT-200 was selected because they are versatile smart glasses, they face many drawbacks. Some of the constraints in conjunction with the execution of the simulator studies are discussed. The main hardware constraints are:

- Occlusion from combiner and glasses' frame
- Small FOV that is clearly visible to the user
- Missing adaptability for IPD
- Poor tracking capabilities
- Uncomfortable cable to input device
- Stereoscopic visualization reduces the resolution
- Not possible to operate monocularly

It was assumed that the quality of merging of both virtual information and real world imagery could depend on certain form factors of the hardware as well as the selected reference frame of the projected information. Foremost, the focal point of the smart glasses must match the distance to real world imagery. For optical reasons, the focal point varies for different IPDs. A person with a greater IPD may see the virtual information at a shorter focal point while a person with a smaller IPD may see the virtual information at a greater focal point. Smart glasses should be adjustable for each user in order to project information at the right focal point.

Within the evaluation, it was observed that head movements decreased when wearing smart glasses with the designed LGL display format. Comparing the total head movements, measured with the eye tracking device, a difference is noticeable. While wearing the smart glasses without anything displayed (within the familiarization scenario), head movements were more frequent and covered a broader area compared to the scenarios where the LGL display format was shown. It could be speculated that the pilot perceived an uncomfortable motion and tried to eliminate this by not moving the head. This perceived motion could be attributed to the fact that the display format was head-referenced and consequently the projected information remained at the same position relative to the head. It is suggested that the future display format will use cockpit-referenced projection for flight state information. This would reduce the relative perceived motion, that is possibly responsible for the reduced head movements.

It was assumed that relative motion between the projected and real-world imagery could also negatively affect merging of both images. There is no research on whether a specific visualization technique could improve the balancing of perception between virtual information and real world imagery. However, it was assumed that this may play a role whenever the focal points of the virtual information and real world imagery do not match. In accordance with the SEEV model [151], relative motion draws visual attention. In case of a non-matching focal point and thus a rivalry in dominance of either virtual or real imagery, relative motion could be negatively influencing the merging. Relative motion could be decreased with contact-analogue information placement, given sufficient speed and precision of the tracking of the glasses.

Other factors believed to have an influence on the merging of virtual information and real world imagery are:

- Movement
- Color
- Frame of reference

6.7 Limitations of selected research design and potential for future work

The selected research design proved to be valid. The balanced assignment of the participants ($N = 20$) to four conditions and the control of influencing variables seemed in good order with common research methods. Results showed that there were no systematic differences between the sequence of presentation, which makes the partly within-subject design also a good choice. The flight technical error varied between the selected experimental airfields due to different geometric characteristics, which is acceptable because of the balanced design.

Measuring spatial awareness with a reaction time experiment is believed to be a precise way to measure this phenomenon. Because it resembled a traffic monitoring task it has a high level of ecological validity. Using a reaction time experiment as a secondary visual task is recommended for future research. However, the difficulty should be increased. Mean reaction times within this thesis were around 3 s. The difficulty (in other words the brightness of the signal) was calibrated with participants who were non-pilots. It is recommended to decrease the brightness and therefore increase the difficulty for detection in order to achieve higher test sensitivity.

The measurement techniques showed some lack of validity. The eye tracking device had difficulties tracking the participant's gaze and the number of missing frames in the recorded data sets was exceptionally high. The reason for the high level of missing values was the software's poor tracking performance when smart glasses were worn. Even though the issue was anticipated and the system's manufacturer stated that the system may not work properly when the person to be tracked is wearing glasses, it was decided to apply this measurement method. Especially with wider head movements the data tended to be missing. In other words, when pilots turned their head to either side the tracking was frequently interrupted. In order to gain a better understanding of the changes in visual attention of pilots, more appropriate eye-tracking methods must be used. Camera-based systems may still have problems with the reflections on the glasses, but improvements of the underlying processing algorithms might improve tracking capabilities.

When analyzing the data of the reaction time task, repeating data points with the same value were noticeable. This was supposedly due to the microcontroller's tendency to run at a certain frequency. Even though a call-back function was implemented on an analogue input pin, many recorded values are alike. A certain measurement error is therefore suspected. It is assumed that some variance was neglected, but since the values were averaged for further statistical analysis this was most likely not impairing the results.

The selected airfield scenery textures were of lower resolution. This was because the simulators OTW view featured high resolution ground texture for only a few selected airfields. Some participants argued that the provided OTW view was of a too low fidelity to navigate proficiently. The airfields were selected because of their unconventional geometry. Using scenarios with a high fidelity of the OTW view is recommended. Recommendations for future research are provided in Section 7.2.

7 Summary and Final Discussion

This thesis considered smart glasses as an assistance system for general aviation pilots that fly under visual flight rules (VFR). The display formats, which were developed within the scope of this thesis, aimed at reducing pilot's task load and increasing their spatial awareness, especially under high task load situations.

Chapter 1 started by motivating the conducted research and defined the area of investigation. Accidents in general aviation (GA) were used to illustrate the need to develop systems that reduce task load and enhance situation awareness (SA). The pilot's tasks were explained.

In Chapter 2, available cockpit instrumentation was described. The view out-the-window (OTW) was considered essential for pilots to maintain the aircraft's attitude and avoid obstacles as well as other aircraft. Most instruments are placed in the head-down area, which requires the pilots to switch focus between inside of the cockpit and the OTW view. It was concluded that flying under VFR and applying the see-and-avoid principle conflicts with the common head-down instrumentation. As a possible technological solution, smart glasses were described, and a taxonomy of available technologies was proposed. Today's commercial off-the-shelf (COTS) see-through smart glasses, a subcategory of head-mounted displays (HMDs), are lightweight, retrofittable and affordable and could be used to assist pilots to maintain an *eyes-out* view. Furthermore, utilized models and methods from aviation psychology and human factors were presented that would assist in understanding underlying principles and design choices made.

Chapter 3 described the design and implementation process of three different display formats. Pilots helped to design the display format by reviewing the concept in different stages, making it a user-centered design process. Moreover, preferences of over 170 participants were gathered beforehand in an online survey. Structured interviews were conducted. Display formats included the so-called Airspace Viewer (for displaying airspace information in an augmenting way), the flight tunnel display format (for guidance within the traffic pattern) and a primary flight display (PFD) that displayed the most relevant flight state information. The prototypes were integrated in the Institute of Flight Systems and Automatic Control (FSR) research flight simulator, Diamond DA 40-180. After a preliminary evaluation of three implemented display formats regarding subjective workload and subjective usability, the PFD display format was selected for further development and final evaluation.

Chapter 4 covered the further development of the selected display format. The display format was modified based on the participant's feedback and a lateral guidance symbology was added, resulting in the so-called Lateral Guidance Line (LGL) display format. It displayed the most relevant flight state information and a line that supported the pilot to maintain track. The format's behavior was developed for guiding the pilot within a VFR traffic pattern and is context sensitive: it showed different information based on the position within the traffic pattern. The lateral guidance symbology signaled when to initiate a turn and when the turn should be completed in order to minimize deviations from a desired track. The symbology became visible when the aircraft went off-track and consisted of a curved arrow that bent to either side and dynamically straightened when the turn should be finished. Whenever the aircraft was on track, the LGL symbology faded out.

Chapter 5 described the methodology for the evaluation of the LGL display format. In the evaluation, a total of $N = 20$ pilots participated in a simulator study. Pilots flew multiple circuits at an unfamiliar airfield either with smart glasses or with conventional means that included printed approach charts and the multi-function display (MFD). Flight technical errors, *eyes-out* time and reaction times to a secondary task for visual attention in the outside view were recorded. Pilots rated the workload and perceived usability.

Chapter 6 presented the results of the evaluation. An increase in time *eyes-out* and more detected signals on a secondary task demonstrated the potential of the LGL display format to increase spatial awareness. Time *eye-out* was measured with an eye-tracking system. Less time was spent monitoring the instruments from inside of the cockpit. The analysis of the reaction times showed that the number of missed signals were significantly less when pilots used the LGL display format compared to navigating with conventional maps, even though mean reaction times did not change. It was assumed that due to the increase in time *eyes-out*, fewer signals were overlooked. It is believed that smart glasses can prove beneficial in increasing the time *eyes-out* and spatial awareness.

Pilots could reduce their lateral flight technical error within the VFR traffic pattern when flying with the smart glasses compared to flying with a head-down moving map display and conventional approach charts. It is believed that the LGL display format has the potential to be used as an auxiliary display for navigation systems.

Unexpectedly, results on subjective mental workload, measured with the NASA Task Load Index (TLX), showed that there were no significant differences in workload levels between the LGL display format and conventional navigation. This could also be interpreted in a positive way: With only little familiarization, participants were able to use the display format on smart glasses without an extra workload. Subjective usability, measured with the System Usability Scale (SUS) was *good*. However, the perceived usability of the LGL format did not significantly differ from its previous stage, the PFD format (see Section 3.3.1.3).

In summary, smart glasses are a promising technology. Low-cost consumer smart glasses will have the potential in the future to support pilots in GA especially when flying under VFR. Smart glasses may help to clean up the cockpit and get rid of the remaining paper-based checklists, maps and other printed information. Unlike tablet devices, smart glasses will eventually allow to retrieve information hands-free, so pilots can focus on flying. The suggested display format could prove beneficial even during highly dynamic phases of flight. The proposed lateral guidance line symbology is a simplistic but still powerful symbology that possibly performs lateral guidance as well as the tunnel-in-the-sky display format [139] but with lower occlusion of the background. However, tracking restrictions of the internal head tracking system (HTS) still make it difficult to deploy full benefit of this technology.

The following sections will present practical recommendations. The thesis will finish with an outlook and recommendations for future research.

7.1 Recommendations

Smart glasses are still a futuristic technology. It is not predictable whether smart glasses will make their way into general aviation cockpits. In this vision of the future, the thesis' author assumes that pilots will use smart glasses from the consumer market. Recommendations for their requirements, introduction and their utilization are given for selected stake holders in the following sections.

The recommendations presume that commercially available smart glasses will use the same operating systems as current personal electronic devices (PEDs) or eventually just work as an auxiliary display for smart phones or tablets. Furthermore, it is presumed that the software will be developed and distributed similar to today's apps.

Manufacturers of smart glasses

In order for smart glasses to be beneficial for pilot's current challenges of the hardware have to be overcome. Necessary future developments are mainly:

- Foremost occlusions by the combiner and glasses' frame must be reduced to allow real see-through and as little occlusions of the periphery as possible. Research suggest that see-and-avoid has its limitations [71; 123]. The effort in scanning the peripheral field of view (FOV) is high, and occlusions by the combiner may result in failing detection of traffic.

- Besides the limited FOV, the tracking sensors that were susceptible to e.g. electromagnetic radiation limit the range of realizable display formats. Tracking capabilities of the HTS must be improved and should include optical calibration and stabilization.

- For operation within a real aircraft, the collimation distance should be near optical infinity to allow for an easy fusion with the background. Furthermore, at near optical infinity, virtual images theoretically allow the eye to relax (reducing visual fatigue) and provide easier accommodation for older users. [118, p.49].

- It is recommended to make the distance between the two display combiners adjustable. This does not necessary have to be an adjustment by hardware. It also could mean a software-based adjustment, or in other words horizontal adjustment of the right and left projected image. Nevertheless, the adjustment should be hardware-based in order to be able to perceive the fullest possible image.

- It is recommended that smart glasses product be specifically designed for use in general aviation cockpits. It should be interfaceable with today's existing PEDs, allowing them to act as image generators, data sources and textual data input devices. Some interactions (e.g. pre-flight planning) require the pilot to enter textual information. The current input methods on smart glasses would make pre-flight planning very difficult. PEDs are advantageous for operating the smart glasses especially because they offer an already available text-entry method. Therefore, it is recommended that developers of smart glasses think about their products as companions of other PEDs (smart phones or tablets).

Software Developers

- It is recommended to apply a user-centered design process to all phases of development. The focus on the user's need can assure that the final product will meet the user's expectations. The interface design should have adequate complexity and should support the pilot's decision. Modern displays allow to display information on multiple pages and apply advanced logic to the automation of the symbologies shown. This automation needs to be understood by the user.

- To increase the precision and calibrate the HTS, it is recommended to use optical marker based methods. The development of augmented reality (AR) display formats requires a precise tracking. As this study showed, the precision of inertial sensors within the smart glasses' HTS are insufficient (see Section 3.3.2). It is recommended to consider optical methods for the calibration of the HTS within the cockpit (see Section 2.2.3.5 on HTS). For flight in night-time, optical markers could be replaced with infrared light-emitting diodes (LEDs) as suggested by Meers et al. [100].

- Assumed that future smart glasses will use the same operating systems as contemporary one's, developers should include a mode to run apps in a way that inhibit all other applications from interfering. The architecture of today's operating systems for PEDs does not support the operation of applications in a *dedicated* way, meaning that one application inhibits other applications from operating. Ideally, the operating system would allow to give one application priorities, if wished so by the user. The user could be prompted if he wishes to run the application in a *dedicated mode*.

- Developers of applications could signal to the user that it is running in a *dedicated mode*, and that no interference must be anticipated. If the utilization of the proposed *dedicated mode* is not available, the application could indicate if other applications are active. Alternatively, the application should instruct the user on how to make sure the application runs safe. This could include reminders in a checklist fashion to ensure other applications are disengaged or prompting the user to configure the device correctly prior to use.

Manufacturers of avionics

- It is recommended that flight data shall be shared between built-in avionic systems and PEDs to enable the visualization of flight data on either PEDs or smart glasses in the future. Most relevant flight data for systems such as the one presented in this thesis are airspeed, altitude, position and heading. Avionic systems that transmit their data allow modular interoperability. Such systems would eventually be more appealing to customers. Avionics should include an open data protocol under which the information is transmitted wirelessly. Bluetooth, for example, allows multiple connections between devices, uses low energy and has a sufficient range of functioning [65].

Regulatory Authorities

- Based on the gathered experience on smart glasses it cannot yet be advised how regulatory authorities should govern the use of smart glasses. It is most likely that smart glasses should be handled the same way as portable electronic flight bag (EFB), namely without a certification as long as devices are not interfaced with other cockpit avionics. Prior to any ruling, further ergonomic limitations need to be identified. It is assumed that in order for EASA to allow smart glasses to be worn throughout the flight, they may not significantly diminish the pilot's ability to perceive the surroundings and perform see-and-avoid actions.

- It is advised to include general lessons on PEDs, the computerization of the cockpit and the avoidance of user errors within basic pilot training. Retrofittable devices, such as EFBs and eventually smart glasses in the future do not only enhance pilot's SA but also do have a downside to it. As studies on the usability of EFB applications have shown [59], the computerization is often misleading. Users feel safe while using assistance features, however, their underlying working principles may often be misunderstood. Pilots get trained to fly conventionally equipped aircraft. Therefore, it is advised to included lectures on trust in automation in the curricular.

- For enabling the development of future display formats in GA, georeferenced data needs to be made available. Unfortunately, today's situation in Europe can be described by a multitude of different authoritative data feeds and national organizations that do not make geospatial data openly accessible. Machine-readable data should be defined in its structure, simple, human-readable (such as the extensible markup language) and exchangeable. The Aeronautical Information Exchange Model (AIXM) is, to some extent, an example for such a system, however, only focusing on the execution of commercial aviation. Among others, the data base features geospatial information on airspace, navaids and procedures. It is advised to extend the scope of AIXM to GA in an open accessible way.

Pilots

- Thorough training on how to use retrofitted electronic devices is necessary. Studies on PEDs have shown that automation has a downside to it [59], namely a misconception of underlying automation principles. A mismatch between the expected degree of automation and the actual automation is noticeable. The actual automation may be non-functioning or not to the degree to which the pilot expects it to be. Pilots assume that whatever an assistance system suggests must be correct. In other words, pilots show an over trust in automated systems, which may eventually fail. Besides electronic failures (i.e. battery failure) software also has its flaws (bugs, inter-application conflicts).

- It is recommended to use PED for flying exclusively and not to utilize devices for non-flight related tasks. Unlike one-purpose devices (e.g. navigation system), PEDs can be used for multiple purposes, both task related but also task unrelated. Pilots are reminded that it is the user's responsibility to configure applications and the operating system in such a way that other task-unrelated apps are disabled from interfering. In an unfavored scenario, other interfering apps send notifications to the device's screen or the pilot simply receives a call because the pilot forgot to set the PED to 'do not disturb' mode.
- Pilots need to be careful with additional technology in the cockpit and consider the cost of new technologies, which is mostly a higher degree of complexity.
- Pilots are reminded to bring all necessary material in a redundant way. An assistance system is there to assist, but the pilot must be able to retain full SA even if the system fails. This requires that pilots have a backup plan and think ahead.
- It is recommended to establish checklists to ensure safe operation of PEDs including smart glasses. Checklists should include items such as the reminder to charge batteries before departure or closing non-flight related apps.

7.2 Future work and outlook

Future research could focus on elaborating different types of display formats on smart glasses. Some of the use cases, as the one discussed in this thesis, are feasible with existing smart glasses while others will eventually be feasible with future generation smart glasses. The success of smart glasses as pilot assistance systems heavily relies on the technology to overcome restrictions. Foremost is the need for more precise and rapid HTS allowing for true contact analogue display formats. With a precise tracking, display formats that are either aircraft- or world referenced display formats are feasible. Another restriction relates to the FOV. The existing smart glasses constrain developers from employing the full potential of the technology.

Further research on attention tunneling needs to be conducted. Presumably, the projection characteristics of the smart glasses are mainly responsible for the occurrence of attention tunneling. In the context of developing display formats on smart glasses it would be beneficial to investigate the effect of the design itself, the reference frame and further contributing factors on attention tunneling.

Participants mentioned that the head-referenced information was distracting. They stated that head movements were kept to a minimum in order to prevent nausea. Therefore, it is suggested that the projected information on smart glasses should rather be aircraft-referenced. Further empirically validation is needed. Aircraft-referenced placement would imply that the information will stay at the same spot in the surrounding cockpit, despite any head movements. Information attached to certain places on the windshield give the impression of a virtual head up display.

Lenhart [93] suggested the usage of stereoscopic cues for the representation of symbologies or, in other words, to display symbologies at different focal distances. To minimize adaptations of the eyes, virtual imagery should be set to the same focal distance as the background [110]. Most of the in-flight time, the pilot focuses on objects in the far distance. Nonetheless, the change of the virtual distance of the symbology, as suggested by Lenhart [93], would be an interesting approach, in case it was the intended to superimpose cockpit instrumentation in the near field with additional information. On the other hand, pilots may find this disrupting, particularly if instruments are cluttered. Further research could cover the effect on perception of objects.

Not every participant performed as expected with the LGL display format. A few participants missed the so called 'turn areas', causing interruptions in flight guidance. This situation is visible from the visualized flight paths in Figure 6.3. From these figures, it is also visible that the turn rates differed between participants. Some participants reacted immediately to commands with steep bank angles while others showed a delayed reaction with slower turn rates. The LGL display format did not tie the pilot to predefined turn rates (and connected to this predefined roll angles), as did the flight tunnel display format. Pilots generally maintain their individual roll angles they apply to almost every turn, resulting in some more steeper or flatter turns. Instead, the developed flight guidance instructs the pilot when to initiate the turn. Depending on the pilot's preferences this results in some variation at the end of the turn. To reduce variation, future developments could include information on the pilot's turn rates in turns lagging and tailor the timing for turn instructions accordingly. The ideal moment to give a command could then be anticipated accordingly.

A minor limitation of the designed lateral guidance line was the missing vertical guidance. A three-dimensional guidance is feasible by showing a three-dimensional arrow-like symbology. In the implemented design, vertical guidance was realized with altitude bugs that indicated the desired altitude. This design could be supported with arrow-like symbols on the altimeter or vertical speed indicator.

The participant's feedback on the LGL display and the flight tunnel format suggested that under visual meteorological conditions (VMC), pilots navigate by landmarks on the ground. Within the approach under VFR, the ground must be visible, and thus flight guidance information could be placed as AR referenced symbology on the earth's surface. Another way of representing flight guidance are flag-like representations on the ground to guide the pilot, similar to a 'Runways Lead-in Lighting System'. The symbology could be projected on ground landmarks. A concept for this display format is shown in Figure 7.1b. Other pilots suggested a concept of 'gates' through which a pilot would need to fly, as during training lessons in the Microsoft 'Flight Simulator X' (see Figure 7.1a).

(a) Gates. Source: Flight Simulator X [122]

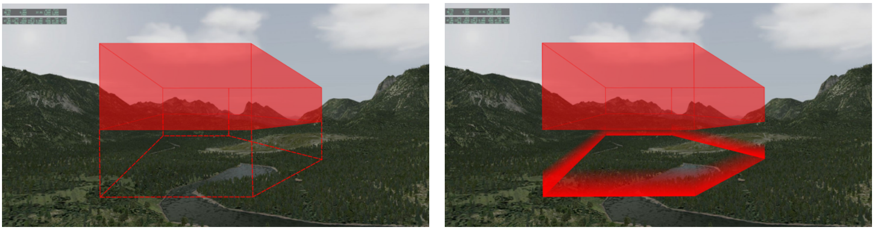
(b) Ground-based flags. Background image source: [122]

Figure 7.1.: Future Work: AR-symbologies for navigation.

Instead of consecutive segments spaced not far apart, this design uses single gates at more distant locations, which eventually mark a turn. These gates have no spatial context and thus are more suitable for en-route guidance (in higher altitude).

In the evaluation of the Airspace Viewer display format, it became clear that pilots were missing spatial context that would help estimate the absolute distance to airspaces, which reach down to ground. Two designs were suggested to create spatial context. Both concepts relied on the airspace edges to be projected on the ground to span a wire frame beneath the airspace (see Figure 7.2a). In another version, shadows supported the visibility (see Figure 7.2b). The pillars on each corner reaching down to the ground required a surface model of the earth, which was not available.

Nevertheless, AR concepts cannot be implemented until available smart glasses and their HTS meet all requirements, which is needed for a practical implementation. Once the overlapping of contact analogue information can be handled, further display formats can be considered.



(a) Wireframe below airspace **(b)** Shadow projected to the ground

Figure 7.2.: Future work: Concepts of airspace representation to create spatial context. [89]

Future research should focus on improvements of the display formats and its effect on occlusion and usability. Even though a broad spectrum of research has previously focused on head-up display formats, results cannot be applied to smart glasses one-on-one. This thesis could not study the eligibility of monocular or binocular projections. Future research must compare their suitability for different display formats.

Overall, smart glasses are in fact a very promising technology to support pilots in general aviation. This technology could work favorably with PEDs, facilitating text input and interaction with the system. Instead of presenting additional information on a tablet device placed on the pilot's lap, smart glasses could allow pilots to fly safer and more relaxed in the future.

References

1. Aero Glass. (2016). *Augmented Reality Aerial Navigation.* https://glass.aero/.
2. Aerocross Systems. (2013). *Brilliant Eyes Head-mounted Display.* http://www.aerocross.com/projects.htm.
3. Airbus Industries (Ed.). (1995). *Coping with long-range flying. Recommendations for Crew Rest and Alertness : A Flight Operations View*. Blagnac, France: Airbus Industries.
4. AOPA Germany e.V. (2012). *Luftraumverletzungen* (Tech. Rep.). Egelsbach.
5. Arduino. (2016). *Arduino Ethernet Shield 2.* http://www.arduino.org/products/shields/arduino-ethernet-shield-2.
6. Aretz, A. J. (1991). The design of electronic map displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *33*(1), 85–101.
7. Bailey, R. E., Arthur III, J. J., Prinzel III, L., & Kramer, L. J. (2007). Evaluation of head-worn display concepts for commercial aircraft taxi operations. In *Defense and Security Symposium* (pp. 65570Y–65570Y). International Society for Optics and Photonics.
8. Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies*, *4*(3), 114–123.
9. Bellenkes, A. H., Wickens, C. D., & Kramer, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine*.
10. Blau, B. (2013). *Hype Cycle for Human-Computer Interaction, 2013* (Tech. Rep.). Gartner Institute.
11. Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
12. Bos, J. E., Bles, W., Hosman, R. J., & Groen, E. L. (2003). *The cause of spatial disorientation* (Tech. Rep.). DTIC Document.
13. Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 211–252.
14. Bridges, D., & Neal-Smith, J. (2016). *Absent Aviators: Gender Issues in Aviation*. Routledge. (Google-Books-ID: 51XACwAAQBAJ)
15. Brooke, J. (1996). SUS: A quick and dirty usability scale. In P. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability Evaluation in Industry* (pp. 189–194). London: CRC Press.

16. Business Insider. (2014). *Google Glass $11 Billion Market By 2018 - Business Insider.* http://www.businessinsider.com/google-glass-11-billion-market-by-2018-2013-5.
17. Campbell, R. D., & Bagshaw, M. (2002). *Human Performance and Limitations in Aviation* (3rd ed.). Oxford, United Kingdom: Blackwell Science.
18. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, N.J: Lawrence Erlbaum Associates.
19. Colvin, K., Dodhia, R., & Dismukes, R. K. (2005). Is pilots' visual scanning adequate to avoid mid-air collisions. In *Proceedings of the 13th International Symposium on Aviation Psychology* (pp. 104–109). Citeseer.
20. Craig, A. B., Sherman, W. R., & Will, J. D. (2009). *Developing Virtual Reality Applications: Foundations of Effective Design*. Morgan Kaufmann.
21. Dambier, M., & Hinkelbein, J. (2006). Analysis of 2004 German general aviation aircraft accidents according to the HFACS model. *Air Medical Journal, 25*(6), 265–269. doi: 10.1016/j.amj.2006.03.003
22. D'Ausilio, A. (2011). Arduino: A low-cost multipurpose lab equipment. *Behavior Research Methods*, *44*(2), 305–313. doi: 10.3758/s13428-011-0163-z
23. Dave Hirschman. (2011). *Avionics: The iPad-centric cockpit behold the power of the tablet.* https://www.aopa.org/news-and-media/all-news/2011/august/01/avionics-the-ipad-centric-cockpit.
24. David O'Hare. (1999). Introduction to Human Performance in General Aviation. In *Human Performance in General Aviation.* Hants, England: Ashgate.
25. Deutsche Flugsicherung. (2000). *Nachrichten für Luftfahrer Teil II.*
26. Deutsche Flugsicherung. (2015a). *ICAO Luftfahrtkarte 2015 Frankfurt (NO 49/6)* [Aeronautical chart]. Langen, Germany.
27. Deutsche Flugsicherung. (2015b). *Luftfahrthandbuch Band II VFR.* Langen.
28. Deutsche Flugsicherung. (2015c). *Sicherer Sichtflug* (Brochure). Langen, Germany: Deutsche Flugsicherung.
29. Deutscher Aero Club. (2016). *Luftraumdaten Deutschland.* http://www.daec.de/fachbereiche/luftraum-flugbetrieb/luftraumdaten/.
30. Dilo, J. T. (2013). *Visualization of External Air Traffic in the Diamond Global Canvas Visual System* (Unpublished diploma thesis). Technische Universität Darmstadt, Institute of Flight Systems and Automatic Control, Darmstadt, Germany.
31. Di Luca, M. (2010). New method to measure end-to-end delay of virtual reality. *Presence: Teleoperators and Virtual Environments*, *19*(6), 569–584.
32. Dörner, R., Broll, W., Grimm, P., & Jung, B. (Eds.). (2013). *Virtual und Augmented Reality (VR / AR)*. Berlin, Heidelberg: Springer Berlin Heidelberg.
33. Duchowski, A. T. (2003). *Eye Tracking Methodology: Theory and Practice*.

Secaucus, NJ, USA: Springer-Verlag New York, Inc.
34. EASA. (2016). *Annual Safety Review 2016* (Tech. Rep.). Cologne, Germany: European Aviation Safety Agency.
35. Edgar, G. K. (2007). Accommodation, cognition, and virtual image displays: A review of the literature. *Displays*, *28*(2), 45–59. doi: 10.1016/j.displa.2007.04.009
36. Eichler, J. (2015). *Implementation and evaluation of current data glasses to support pilots in the general aviation.* (unpublished master's thesis). Technische Universität Darmstadt, Telecooperation Group, Darmstadt, Germany.
37. Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32–64. doi: 10.1518/001872095779049543
38. Endsley, M. R. (1999). Situation awareness in aviation systems. *Handbook of aviation human factors*, 257–276.
39. Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. *Situation awareness analysis and measurement*, 3–32.
40. Endsley, M. R., Farley, T. C., Jones, W. M., Midkiff, A. H., & Hansman, R. J. (1998). *Situation awareness information requirements for commercial airline pilots* (Tech. Rep.). International Center for Air Transportation.
41. Endsley, M. R., & Garland, D. J. (2000). *Situation Awareness: Analysis and Measurement*. Routledge.
42. Engeroff, M. (2014). *Analysis, Implementation and Evaluation of Use Cases of Current Data Glasses to Support Different Users in Commercial Aviation* (Masterarbeit). Technische Universität Darmstadt, Darmstadt.
43. Epson. (2014a). *Merging Epsons's Core Technologies. Moverio BT-200 Smart Glasses.* http://global.epson.com/innovation/engineer/moverio.html.
44. Epson. (2014b). *Moverio BT-200 datasheet.*
45. Epson. (2017). *Moverio BT-200 - Epson.* https://www.epson.de/products/see-through-mobile-viewer/moverio-bt-200.
46. European Aviation Safety Agency. (2013). *Acceptable Means of Compliance (AMC) and Guidance Material (GM) to Part-NCO* (Supplementary document to ED Decision 2013/022/R). Cologne, Germany: European Aviation Safety Agency.
47. European Aviation Safety Agency. (2014). *AMC-20 - Amendment 12 - Airworthiness and operational consideration for Electronic Flight Bags (EFBs)* (Supplementary document to ED Decision 2014/001/R). Cologne, Germany: European Aviation Safety Agency.
48. European Commission. (2016a). *Augmented Reality for more Aviation Safety.* /easme/en/news/augmented-reality-more-aviation-safety.
49. European Commission. (2016b). *Commision Regulation (EU) No. 800/2013.*

50. European Organisation for the Safety of Air Navigation. (2010). *European Action Plan for Airspace Infringement Risk Reduction.*
51. Evans, J. R., & Mathur, A. (2005). The value of online surveys. *Internet Research*, *15*(2), 195–219. doi: 10.1108/10662240510590360
52. Federal Aviation Administration. (2010). *Press Release – FAA Calls on Airlines to Limit Cockpit Distractions.* https://www.faa.gov/news/press_releases/news_story.cfm?newsId=11338.
53. Federal Aviation Administration. (2015). *Fly the Aircraft First.*
54. Ferrin, F. J. (1998). Current issues in helmet-mounted display systems for military applications. In *Helmet- and Head-Mounted Displays III* (Vol. 3362, pp. 71–79). doi: 10.1117/12.317457
55. Flach, J. M. (1995). Situation awareness: Proceed with caution. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 149–157.
56. Flemisch, F. O., & Onken, R. (2000). *Detecting usability problems with eye tracking in airborne battle management support* (Tech. Rep.). DTIC Document.
57. ForeFlight. (2014). *ForeFlight Mobile 6.0.* https://blog.foreflight.com/2014/04/10/foreflight-mobile-6-0/.
58. Funabiki, K., Muraoka, K., Terui, Y., Harigae, M., & Ono, T. (1999). In-flight evaluation of tunnel-in-the-sky display and curved approach pattern. In *Guidance, Navigation, and Control Conference and Exhibit* (p. 3966).
59. Funk, M. (2015). *Learnability and Usability of Mobile Applications for Supporting Pilots in General Aviation under Visual Flight Rules.* (unpublished master's thesis). Technische Universität Darmstadt, Institute of Flight Systems and Automatic Control, Darmstadt, Germany.
60. Garmin (Ed.). (2006). *G1000 Integrated Flight Deck Pilot's Guide.*
61. Garmin. (n.d). *G3X.* http://www.garmin.com/us/products/intheair/sport-aviation/g3x.
62. Gartner Institute. (2016). *Gartner's 2016 Hype Cycle for Emerging Technologies Identifies Three Key Trends That Organizations Must Track to Gain Competitive Advantage.* http://www.gartner.com/newsroom/id/3412017.
63. General Aviation Manufacturers Association. (2015). *2014 General Aviation Statistical Databook & 2015 Industry Outlook.*
64. Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of educational research*, *42*(3), 237–288.
65. Gomez, C., Oller, J., & Paradells, J. (2012). Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology. *Sensors*, *12*(9), 11734–11753.

66. Groff, L. (2010a). *Introduction of Glass Cockpit Avionics into Light Aircraft - Qualitative Analysis Results* [Presentation at the National Transportation Safety Board]. Washington, D.C..

67. Groff, L. (2010b). *Introduction of Glass Cockpit Avionics into Light Aircraft - Quantitative Analysis Results* [Presentation at the National Transportation Safety Board]. Washington, D.C..

68. Grunwald, A. J., Robertson, J. B., & Hatfield, J. J. (1980). *Evaluation of a computer-generated perspective tunnel display for flight path following* (Tech. Rep.).

69. Haiduk, P. M. (2013). *Head-mounted-displays for accident avoidance in general aviation.* [Posterpresentation at the 7th International Summer School on Aviation Psychology]. Graz, Austria: European Association for Aviation Psychology.

70. Haiduk, P. M. (2015). *Data glasses for GA Pilots: Primary Flight Display (PFD) with Low Clutter* [Poster presented at the 7th International Summer School on Aviation Psychology.]. Graz, Austria: European Association for Aviation Psychology.

71. Haiduk, P. M., Santel, C. G., Fuchs, K.-U., & Klingauf, U. (2014). Sehen und Gesehen werden: Simulatorstudie mit Fremdverkehr zur Untersuchung von Erfassungswahrscheinlichkeiten [See and avoid: A simulator study investigating the detection probabilities of conflicting traffic]. *Der Mensch zwischen Automatisierung, Kompetenz und Verantwortung [Man between automatization, compentency and responsibility]*, 77–91.

72. Harris, D., & Muir, H. C. (2005). *Contemporary issues in human factors and aviation safety*. Gower Publishing, Ltd.

73. Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 years lateral. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*, 904–908. doi: 10.1177/154193120605000909

74. Harwell, M. R., Rubinstein, E. N., Hayes, W. S., & Olds, C. C. (1992). Summarizing Monte Carlo results in methodological research: The one-and two-factor fixed effects ANOVA cases. *Journal of Educational and Behavioral Statistics*, *17*(4), 315–339.

75. HeadApp. (2016). *Eye 4 Flight.* http://www.headapp.eu/.

76. Heinrich, M. J., & Olsson, M. I. (2012, May). *Wearable display device* (No. USD659741 S1). (U.S. Classification D16/309, D16/335, D16/326)

77. Herzog, M. (2016). *Effects of Voice Guidance in Mobile Navigation Applications on Spatial Awareness, Flight Precision and Subjective Workload of Pilots in General Aviation* (unpublished master's thesis). Justus-Liebig Universität Gießen, Gießen.

78. Hickox, J. C., & Wickens, C. D. (1999). Effects of elevation angle disparity, complexity, and feature type on relating out-of-cockpit field of view to an

electronic cartographic map. *Journal of Experimental Psychology: Applied*, *5*(3), 284–301. doi: 10.1037/1076-898X.5.3.284

79. Institute of Flight Systems and Automatic Control. (2014). *Diamond DA 40-180 simulator.* http://www.fsr.tu-darmstadt.de/forschung_und_dienstleistung/ausstattung/dsimulator.de.jsp.
80. International Business Publications. (2006). *International Civil Aircrafts Registration Procedures Handbook: Strategic Information and Procedures*. Washington, DC: Int'l Business Publications.
81. International Civil Aviation Organization. (2009). *Review of the Classification and Definitions used for Civil Aviation Activities* (Working Paper). Montréal, Québec.
82. Jones, D. G., & Endsley, M. R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space, and Environmental Medicine*.
83. Junghenn, S., Lux, M., Lingenau, A., Müller, C., & Rödelsberger, M. (2014). *Head Mounted Displays – Entwurf eines Displays für Cockpits in der allgemeinen Luftfahrt* (unpublished advanced design project). Technische Universität Darmstadt, Institute of Flight Systems and Automatic Control, Darmstadt.
84. Kahneman, D. (1973). Attention and effort.
85. Kaiser, J. (2004). *Verwendung stereoskopischer Informationsdarstellung in durchsichtfähigen Anzeigen am Beispiel eines Head-Up Displays*. Stuttgart, Germany: Ergonomia Verlag.
86. Kocian, D. F. (1990). *Design considerations for virtual panoramic display (VPD) helmet systems* (Tech. Rep.). DTIC Document.
87. König, C. (2012). *Analyse und Anwendung eines menschzentrierten Gestaltungsprozesses zur Entwicklung von Human-Machine-Interfaces im Arbeitskontext am Beispiel Flugsicherung* (Unpublished doctoral dissertation). Technische Universität.
88. Korn, B., Schmerwitz, S., Lorenz, B., & Döhler, H.-U. (2009). Combining enhanced and synthetic vision for autonomous all-weather approach and landing. *The International Journal of Aviation Psychology*, *19(1)*, 49–72.
89. Lange, A. (2015). *Implementation and Evaluation of a Dataglass Display Format for Visualizing Airspaces to Support Pilots in the General Aviation* (unpublished master's thesis). Technische Universität Darmstadt, Telecooperation Group, Darmstadt, Germany.
90. Laramee, R. S., & Ware, C. (2002). Rivalry and interference with a head-mounted display. *ACM Transactions on Computer-Human Interaction*, *9*(3), 238–251.
91. Lee, Y.-H., & Liu, B.-S. (2003). Inflight workload assessment: Comparison of subjective and physiological measurements. *Aviation, space, and environmental*

medicine, *74*(10), 1078–1084.
92. Lefebvre, Y. (2008). A flexible solution to deploy avionics displays to multiple Embedded Platforms. In *2008 IEEE/AIAA 27th Digital Avionics Systems Conference* (pp. 5.A.3–1–5.A.3–9). doi: 10.1109/DASC.2008.4702843
93. Lenhart, P. M. (2006). *Räumliche Darstellung von Flugführungsinformationen in Head-Mounted Displays* (1st ed.). Ergonomia Verlag.
94. Lewis, J. R., & Sauro, J. (2009). The factor structure of the system usability scale. In *International Conference on Human Centered Design* (pp. 94–103). Heidelberg: Springer.
95. Longo, L., & Dondio, P. (2015). On the Relationship between Perception of Usability and Subjective Mental Workload of Web Interfaces. In (pp. 345–352). IEEE. doi: 10.1109/WI-IAT.2015.157
96. Luftfahrt Bundesamt. (2016a). *Bestand an Luftfahrzeugen in der Bundesrepublik Deutschland.* https://www.lba.de/DE/Presse/Statistiken/Statistik_Luftfahrzeuge.html.
97. Luftfahrt Bundesamt. (2016b). *Statistik der vom LBA erteilten Lizenzen für fliegendes Personal.* https://www.lba.de/DE/Presse/Statistiken/Statistik_Lizenzen.html.
98. Manning, S. D., & Rash, C. E. (2007). A moveable view. *J. Flight Safety Foundation*, *2*(8), 41–44.
99. MathWorks. (2016). *Four-quadrant inverse tangent - atan2.* https://de.mathworks.com/help/matlab/ref/atan2.html.
100. Meers, S., Ward, K., & Piper, I. (2008). Simple, Robust and Accurate Head-Pose Tracking Using a Single Camera. In *Mechatronics and Machine Vision in Practice* (pp. 111–122).
101. Mehringskötter, S. (2014). *Analyse von zukünftigen (ATM-) Einsatzmöglichkeiten aktueller Datenbrillen zur Unterstützung von Piloten in der allgemeinen Luftfahrt* (unpublished student research project). Technische Universität Darmstadt, Institute of Flight Systems and Automatic Control, Darmstadt, Germany.
102. Melzer, J. (2000). Head-Mounted Displays. In C. Spitzer (Ed.), *The Avionics Handbook* (1edition ed.). Boca Raton: CRC Press.
103. Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, *77*(12), 1321–1329.
104. Miller, C. C. (2013). Google Looks to Make Its Computer Glasses Stylish. *The New York Times*.
105. Minkner, O. (2014). *Identifizierung von relevanten Instrumenten in unterschiedlichen Flugphasen bei Sichtflug. Eine Eye-Tracking Studie* (unpublished master's thesis). Technische Universität Darmstadt, Institute of Flight Systems and

Automatic Control, Darmstadt, Germany.
106. Moores, V. (2014). Into the looking glass: Virgin to roll out Google Glass, but SITA has its concerns. *AIR TRANSPORT WORLD*, *51*(8).
107. Nicolai, T., Sindt, T., Kenn, H., & Witt, H. (2005). Case study of wearable computing for aircraft maintenance. In *IFAWC-International Forum on Applied Wearable Computing*. VDE VERLAG GmbH.
108. Pallant, J. (2010). *SPSS Survival Manual: A step by step guide to data analysis using SPSS*. Open University Press.
109. Parrish, R. V., Williams, S. P., Arthur, J. J., Kramer, L. J., Bailey, R. E., Prinzel, L. J., & Norman, R. M. (2006). *A Description of the "Crow's Foot" Tunnel Concept* (Tech. Rep. NASA/TM-2006-214311). Hampton, VA: NASA Langley Research Center.
110. Patterson, R., Winterbottom, M. D., & Pierce, B. J. (2006). Perceptual issues in the use of head-mounted visual displays. *Human factors*, *48*(3), 555–573.
111. Peterson, S., & Pinska, E. (2006). Human performance with simulated collimation in transparent projection screens. In *Proceedings of the Second International Conference on Research in Air Transportation (ICRAT)* (pp. 231–237).
112. *Pilot's Handbook of Aeronautical Knowledge*. (2008). Oklahoma City, OK: United States Department of Transportation, Federal Aviation Administration.
113. Pongratz, H. (2006). *Kompendium der Flugmedizin*. H. Pongratz.
114. Poulton, E. C. (1957). On prediction in skilled movements. *Psychological bulletin*, *54*(6), 467.
115. Puranik, T. G., Harrison, E., Min, S., Jimenez, H., & Mavris, D. N. (2016). General Aviation Approach and Landing Analysis using Flight Data Records. In *16th AIAA Aviation Technology, Integration, and Operations Conference*.
116. Rash, C. E. (1999). *Helmet mounted displays: Design issues for rotary-wing aircraft* (Vol. 93). SPIE Press.
117. Rash, C. E., & Kalich, M. E. (2009). A Limited Rotary-Wing Flight Investigation of Hyperstereo in Helmet-Mounted Display Designs. , 111.
118. Rash, C. E., Russo, M. B., Letowski, T. R., & Schmeisser, E. T. (2009). *Helmet-mounted displays: Sensation, perception and cognition issues*. ARMY AEROMEDICAL RESEARCH LAB FORT RUCKER AL.
119. Rauschnabel, P. A., & Ro, Y. K. (2016). Augmented reality smart glasses: An investigation of technology acceptance drivers. *International Journal of Technology Marketing*, *11*(2), 123–148.
120. Recon Instruments. (2017). *Recon Jet*. http://store.reconinstruments.com/Recon-Jet.
121. Reisman, R. J., & Brown, D. M. (2006). Design of augmented reality tools for air traffic control towers. In *6th AIAA Aviation Technology, Integration and*

Operations (ATIO) Conference, Wichita, KS.

122. Sagamir. (2011). *FSX - Flugstunde 2 - Grundlagen des Fliegens.*
123. Santel, C. G., Haiduk, P. M., Fuchs, K.-U., & Klingauf, U. (2014). The potential for analytical modeling and flight simulator studies of midair collisions to support accident investigation. *Proceedings of the European Association for Aviation Psychology Conference*, *31*, 395–407.
124. Sarayeddine, K., & Mirza, K. (2013). Key challenges to affordable see-through wearable displays: The missing link for mobile AR mass deployment. In (Vol. 8720, pp. 87200D–87200D–5). doi: 10.1117/12.2018184
125. Schattenberg, K. (2002). *Fahrzeugführung und gleichzeitige Nutzung von Fahrerassistenz-und Fahrerinformationssystemen* (Unpublished doctoral dissertation). Bibliothek der RWTH Aachen.
126. Schulz, C. M., Schneider, E., Fritz, L., Vockeroth, J., Hapfelmeier, A., Wasmaier, M., ... Schneider, G. (2011). Eye tracking for assessment of workload: A pilot study in an anaesthesia simulator environment. *British journal of anaesthesia*, *106*(1), 44–50.
127. Seeing Machines. (2011). *FaceLab 5 user manual* (User Manual). Canberra, Australia.
128. Shappell, S. A., & Wiegman, D. A. (2003). *A human error analysis of general aviation controlled flight into terrain accidents occurring between 1990-1998* (Tech. Rep.). DTIC Document.
129. Shook, R. W., Bandiero, M., Coello, J. P., Garland, D. J., & Endsley, M. R. (2000). Situation awareness problems in general aviation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 44, pp. 185–188).
130. Shore, J. (2012). *Where Did Augmented Reality Come From?* http://mashable.com/2012/09/24/augmented-reality/.
131. SIMI Motion. (2008). *Simi Motion: Kalibrierung, Videoaufnahme, Tracking, Standbildvermessung, DLT-Berechnung, Präsentation, HTML-Report.* http://www.simi.de/de/products/motion/overview/analysis/index.html.
132. Sindlinger, A. (2008). *Ein Beitrag zur dreidimensionalen Darstellung von Nominal-Trajektorien in perspektivischen Flugführungsanzeigen* (Unpublished doctoral dissertation). TU Darmstadt.
133. Spatz, C. (2015). *Implementation and evaluation of a dynamic HMI for displaying the primary flight information on data glasses to support pilots in the general aviation* (unpublished master's thesis). Technische Universität Darmstadt, Institute of Flight Systems and Automatic Control, Darmstadt, Germany.
134. Stanton, N. A., & Young, M. S. (1999). What price ergonomics? *Nature*, *399*(6733), 197–198.

135. Sutherland, I. E. (1965). The ultimate display. *Multimedia: From Wagner to virtual reality*.
136. Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Boston, MA: Pearson.
137. *Thema Platzrunde* (Tech. Rep.). (2003).
138. Valimont, B., Wise, J. A., Nichols, T., Best, C., Suddreth, J., & Cupero, F. (2009). When the wheels touch Earth and the flight is through, pilots find one eye is better than two. In *Proceedings of SPIE* (Vol. 7326, p. 732603).
139. Vancura, D. (2015). *Implementation and evaluation of current data glasses to support pilots in the general aviation* (unpublished master's thesis). Technische Universität Darmstadt, Telecooperation Group, Darmstadt, Germany.
140. Veness, C. (2016). *Calculate distance and bearing between two Latitude/Longitude points using haversine formula in JavaScript.* http://www.movable-type.co.uk/scripts/latlong.html.
141. Verschragen, E., Mulder, M., & van Paassen, M. (2005). Track-Recovery Support for Tunnel-in-the-Sky Displays. In *AIAA Guidance, Navigation, and Control Conference and Exhibit* (p. 5960).
142. von Krogh, G., Ichijo, K., & Nonaka, I. (2000). *Enabling Knowledge Creation: How to Unlock the Mystery of Tacit Knowledge and Release the Power of Innovation*. Oxford University Press, USA. (Google-Books-ID: JVESDAAAQBAJ)
143. Vuzix. (2017). *M3000 Smart Glasses.* https://www.vuzix.com/Products/m3000-smart-glasses.
144. Want, R., Borriello, G., Pering, T., & Farkas, K. I. (2002). Disappearing Hardware. *IEEE pervasive computing*, *1*(1), 36–47.
145. Weiser, J. (1995). The coming age of calm technology. In P. J. Denning (Ed.), *Beyond Calculation: The Next Fifty Years of Computing* (1st ed.). Copernicus Books.
146. Weiser, M. (1999). The Computer for the 21st Century. *SIGMOBILE Mob. Comput. Commun. Rev.*, *3*(3), 3–11. doi: 10.1145/329124.329126
147. Wells, B. A. (1991). *Head mounted display for miniature video display system* (No. US5003300 A). (U.S. Classification 345/8, 348/115, 340/980, 348/838, 348/E03.009; International Classification H04N5/66, H04N5/64, G02B27/00, G02B13/16, H04N3/08, G02B27/01, G02B27/22; Cooperative Classification G02B2027/0132, G02B13/16, G02B27/017, G02B27/0176, G02B27/0172, H04N3/08, G02B2027/014, G02B2027/0143, G02B2027/0159, G02B2027/0156, G02B27/2228, G02B2027/0178; European Classification G02B27/01C2, H04N3/08, G02B27/22S, G02B13/16, G02B27/01C1, G02B27/01C)

148. Wesslen, E. P., & Young, J. P. (2011). Pilot Performance: Round Dial and Vertical Tape Altimeters. *Aviation Technology Graduate Student Publications. Paper 16*.
149. Wickens, C. (1991). Processing resources and attention. *Multiple-task performance*, 3–34.
150. Wickens, C. (2002). Situation awareness and workload in aviation. *Current directions in psychological science*, *11*(4), 128–133.
151. Wickens, C., McCarley, J., & Steelman-Allen, K. (2009). NT-SEEV: A model of attention capture and noticing on the Flight Deck. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *53*(12), 769–773. doi: 10.1177/154193120905301202
152. Wickens, C. D., & Alexander, A. L. (2009). Attentional Tunneling and Task Management in Synthetic Vision Displays. *The International Journal of Aviation Psychology*, *19*(2), 182–199. doi: 10.1080/10508410902766549
153. Wickens, C. D., Alexander, A. L., Horrey, W. J., Nunes, A., & Hardy, T. J. (2004). Traffic and flight guidance depiction on a synthetic vision system display: The effects of clutter on performance and visual attention allocation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 48, pp. 218–222). SAGE Publications.
154. Wickens, C. D., Goh, J., Helleberg, J., Horrey, W. J., & Talleur, D. A. (2003). Attentional Models of Multitask Pilot Performance Using Advanced Display Technology. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(3), 360–380. doi: 10.1518/hfes.45.3.360.27250
155. Wickens, C. D., Helleberg, J., Goh, J., Xu, X., & Horrey, W. J. (2001). Pilot task management: Testing an attentional expected value model of visual scanning. *Savoy, IL, UIUC Institute of Aviation Technical Report*.
156. Wickens, C. D., McCarley, J. S., Alexander, A. L., Thomas, L. C., Ambinder, M., & Zheng, S. (2008). Attention-situation awareness (A-SA) model of pilot error. *Human performance modeling in aviation*, 213–239.
157. Wierwille, W. W., & Eggemeier, F. T. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *35*(2), 263–281.
158. Wood, R. B., & Howells, P. J. (2000). Head-Up Display. In C. Spitzer (Ed.), *The Avionics Handbook* (1edition ed.). Boca Raton: CRC Press.
159. Yeh, M., & Wickens, C. D. (1997). Performance issues in helmet-mounted displays. *University of Illinois Institute of Aviation Technical Report (ARL-97-9IARMy-FEDLAB-97-l). Savoy, IL: Aviation Research Laboratory*.
160. Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of comparative neurology and psychology*,

18(5), 459–482.

161. Ziefle, M., Oehme, O., & Luczak, H. (2010). *Visuelle Information und Leistung bei Head-Mounted Displays mit erweiterter Realität*. GRIN Verlag.

List of figures

List of tables

A Checking the Assumptions for Applying Multiple Regression to Reaction Times

Multiple regression makes a number of assumptions about the data. Normality, linearity and homoscedasticity refer to various aspects of the distribution of data points and the underlying relationship between the variables. The following list was taken from from Pallant's manual on SPSS [108]. For further reading please refer to the statistics texbooks of Tabachnick and Fields [136]

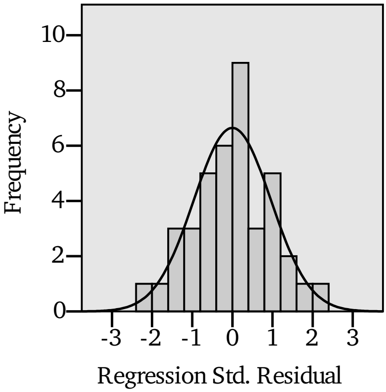
1. *normality:* the residuals should be normally distributed about the predicted dependent variable (DV)

2. *linearity:* the residuals should have a straight-line relationship with predicted DV scores

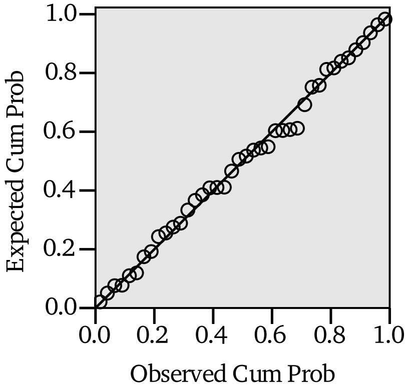
3. *homoscedasticity:* the variance of the residuals about predicted DV scores should be the same for all predicted scores

The histogram of standardized residuals (see Figure A.1a) indicated that the data contained approximately normally distributed errors, as did the normal P-P plot of standardized residuals (see Figure A.1b), which showed points that were not completely on the line, but close.

The scatterplot of standardized (see Figure A.2) residuals showed that the data met the assumptions of homogeneity of variance and linearity. An further analysis of standardized residuals was carried out, which showed that the data contained no outliers (Std. Residual Min = −2.05, Std. Residual Max = 2.13). Tests to see if the data met the assumption of collinearity indicated that multicollinearity was not a concern (tolerance = .96, VIF = 1.0).



(a) histogram of standardized residuals



(b) normal P-P plot of regression standardized residual

Figure A.1.: Plots for assessing the distribution of errors

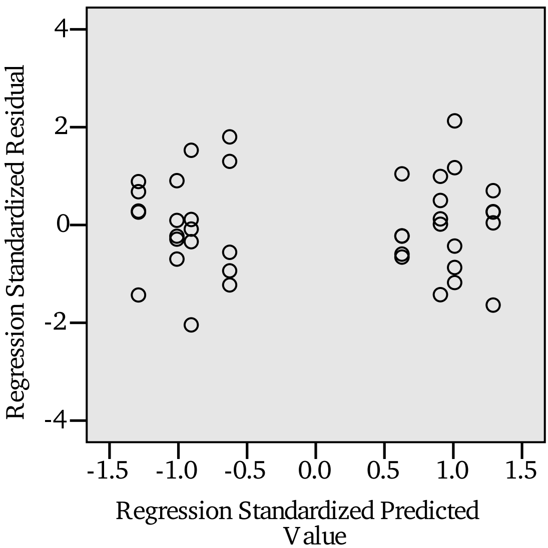


Figure A.2.: Scatterplot of standardized residuals for assessing homoscedasticity and linearity

B Checking the Assumptions for Inferential Statistics on the Flight Technical Error

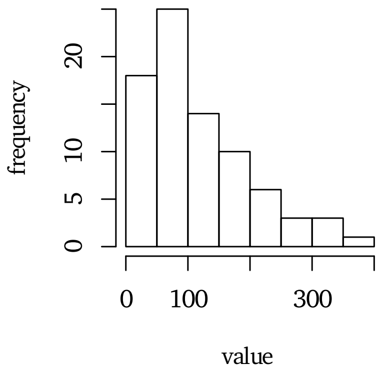
To conduct an analysis of variance (ANOVA) or a linear regression the data needs the to meet several requirements. The data should not contain significant outliers and be normally distributed. In addition, it requires a homogeneity of variances.

To test for normal distribution of the data a Shapiro-Wilk normality test was conducted. It turned out significant ($W = 0.865, p < 0.01$). The skewness is illustrated in Figure B.1a. It is clearly visible from the bar chart that the data is heavily skewed. An averaging of absolute values was necessary in prior data treatment. Hence, values tend towards zero.

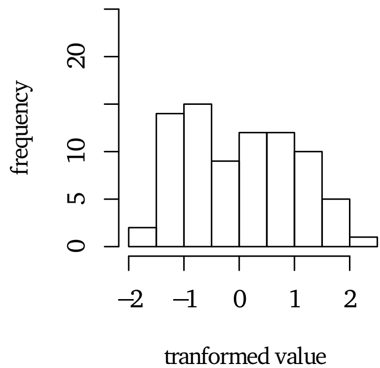
Data was transformed to meet the requirements of normal distribution. A Box-Cox power transformation [13] was applied to the data. The one-parameter Box-Cox power transformations are defined as:

$$y_i^{(\lambda)} = \begin{cases} \ln(y_i) & \text{when } \lambda = 0 \\ \dfrac{y_i^{\lambda} - 1}{\lambda} & \text{otherwise} \end{cases} \tag{B.1}$$

After transformation the Shapiro-Wild test for normal distribution still turned out significant ($W = 0.957, p = 0.01$). Nevertheless, it was decided to continue with inferential statistical analysis. The robustness of the ANOVA method has been shown in several publications like in the ones by Glass et al.[64; 74].



(a) non-normally distributed data



(b) data after Box-Cox transformation

Figure B.1.: Frequency distribution of mean deviations from predesignated track (cross-track error)

C Interpreting Effect Sizes

To assess the substantive significance of a result, effect sizes may be interpreted. Effect sizes indicate the strength of an effect. However, the interpretation of effect sizes is a subjective process. In literature there are some hints how to interpret effect sizes.

C.1 η^2 and r^2 Coefficient of determination

The coefficient of determination is a value that indicates the amount of the variance in the dependent variable that is predictable from the independent variable. Cohen [18, pp. 82] also recommended interpretation intervals for the r^2 effect size, which is generally used during linear models. η^2 is interpreted the same way.

$$
\begin{array}{llll}
\eta^2 \text{ and } r^2 & \leq & \pm.2 & \text{negligible explained variation}, \\
\eta^2 \text{ and } r^2 & \leq & \pm.5 & \text{small explained variation}, \\
\eta^2 \text{ and } r^2 & \leq & \pm.8 & \text{moderate explained variation}, \\
\eta^2 \text{ and } r^2 & \geq & \pm.8 & \text{large explained variation},
\end{array}
$$

C.2 Pearson's $\bar{r}^2$

Bortz stated that Pearson's multivariate regression coefficient $\bar{r}^2$ (*adjusted* r^2) should be interpreted similarly to the bivariate coefficient [11, pp. 449–451]. Therefore, it is interpreted using the following intervals.

$$
\begin{array}{lllllr}
0 & \leq & \bar{r}^2 & < & .01, & \text{negligible effect} \\
.01 & \leq & \bar{r}^2 & < & .09, & \text{small effect} \\
.09 & \leq & \bar{r}^2 & < & .25, & \text{moderate effect} \\
.25 & \leq & \bar{r}^2 & & , & \text{large effect}
\end{array}
$$

C.3 Cramér's V and Pearson's Bivariate r

Cohen [18, pp.25-27, 79-80] provides the following examples for interpreting the effect sizes r and Cramérs V.

$$
\begin{array}{lllllr}
0 & \leq & V & < & .1, & \text{negligible effect} \\
.1 & \leq & V & < & .3, & \text{small effect} \\
.3 & \leq & V & < & .5, & \text{moderate effect} \\
.5 & \leq & V & & , & \text{large effect}
\end{array}
$$

D Atan2 Function

The Atan2 function, as it was used in this thesis, is the arc tangent of $\frac{y}{x}$ followed by computed determination of six cases. Functions implemented in Mathworks Matlab and C++ were used.

$$\text{atan2}\colon\ \mathbb{R}^2 \backslash \{(0,0)\} \rightarrow (-\pi, \pi],$$

$$(x,y) \mapsto \begin{cases} \arctan(\frac{y}{x}) & \text{für } x > 0, \\ \arctan(\frac{y}{x}) + \pi & \text{für } x < 0, y > 0, \\ \pm\pi & \text{für } x < 0, y = 0, \\ \arctan(\frac{y}{x}) - \pi & \text{für } x < 0, y < 0, \\ +\frac{\pi}{2} & \text{für } x = 0, y > 0, \\ -\frac{\pi}{2} & \text{für } x = 0, y < 0. \end{cases}$$