Navigating Complexity: An Eye-Tracking Study on the Redesign of the Swiss VFR Aeronautical Chart

GEO 511 Master's Thesis

Author: Simeon Lüthi, 16-923-237

Supervised by: Dr. Armand Kapaj, Dr. Tumasch Reichenbacher

Faculty representative: Prof. Dr. Sara Irina Fabrikant

27.04.2025

Department of Geography, University of Zurich
Master of Science in Geography
Specialization: Geographic Information Systems and Science

Navigating Complexity:
An Eye-Tracking Study on the Redesign
of the Swiss VFR Aeronautical Chart

GEO 511 - Master's Thesis

Submitted by:
Simeon Lüthi (16-923-237)
Zürich, April 25th, 2025

Supervisors:
Dr. Armand Kapaj, University of Zurich
Dr. Tumasch Reichenbacher, University of Zurich

Faculty Representative:
Prof. Dr. Sara Irina Fabrikant

Abstract

This thesis investigates the effectiveness of a proposed design update to the Swiss 1:500,000 VFR area chart using eye-tracking methods. Drawing from ICAO Annex 4, which outlines worldwide design standards, the updated chart features reduced visual clutter and enhanced visual saliency. Both expert and non-expert participants, representing student pilots, completed visual search tasks on both charts in two separate regions, each exhibiting different levels of visual complexity. A within-subject, counterbalancing study-design was implemented to strengthen the internal validity. Results show that both task completion time and time to first AOI hit were significantly reduced with the updated chart design, while answer accuracy remained stable. Non-expert participants benefited more noticeably from the improved layout, and no systematic bias toward the current chart could be observed for expert participants. Findings from previous literature regarding differences between experts and non-experts could not be recreated for saccade and fixation metrics. However, heat maps and scan paths indicated more focused attention under the updated design. The study demonstrates that a more effective chart design is achievable. In light of the ongoing transition from paper charts to electronic flight bags, such updates to existing charts should be considered.

Keywords: Cartography, aeronautical chart design, visual clutter, eye movement analysis, aviation, map reading

Acknowledgments

First of all, I would like to thank all participants who took the time to complete the eye-tracking experiment here on campus. Without you, none of this would have been possible.

I am also thankful for the support and guidance of my two supervisors, Dr. Armand Kapaj and Dr. Tumasch Reichenbacher. Their availability, suggestions and helpful feedback brought me through this project.

On a personal note, I would also like to thank my parents and my girlfriend for their unconditional love and support throughout my studies and during my Master's thesis here at the University of Zurich.

Finally, I am grateful for having gotten to know my fellow peers, whose friendships have enriched my academic journey throughout the past years.

Contents

List of Figures

List of Tables

Abbreviations

Abbreviation	Definition
EFB	Electronic Flight Bag
FAA	Federal Aviation Administration
ICAO	International Civil Aviation Organization
IFR	Instrument Flight Rules
MFD	Multifunction Display
VFR	Visual Flight Rules
Navaid	Navigational Aids
DME	Distance measuring equipment
NDB	Non-directional beacon
VOR	VHF omnidirectional range
AWY	Airway
CTA	Control Area
CTR	Control Zone
FIR	Flight Information Region
FIZ	Flight Information Zone
HX	Not Permanently Active
RMZ	Radio Mandatory Zone
TMA	Terminal Control Area
TMZ	Transponder Mandatory Zone
ATCO	Air Traffic Control Officer
ATPL	Airline Transport Pilot License
CPL	Commercial Pilot License
PPL	Private Pilot License

Table 1: List of abbreviations used in this study.

Chapter 1

Introduction

1.1 Background

From the very beginning, pilots are introduced to the fundamental principle of "Aviate, Navigate, Communicate." While maintaining control of the aircraft is the top priority, in-flight navigation is the second most critical task. To achieve that, pilots rely on a combination of onboard instruments and a variety of maps, usually referred to as charts, either in electronic form, or traditional paper maps, to determine their position and further routing.

Early Days

Aviation cartography arguably traces its origins to the late 19th century when Prussian artillery balloonist Lieutenant Colonel Hermann Moedebeck recognized the necessity of maps that depicted suitable landing sites for balloons, making him the "spiritual godfather" of aviation charts (Steele, 1998; Svatek, 2014). In 1911, an international commission agreed on the implementation of standardized mapping practices under the supervision of a global governing body, a principle that has been upheld ever since. While Europe initially led charting advancements, the impact of World War I propelled the United States ahead in aeronautical development (Steele, 1998).

Throughout the early decades of the 20th century, the foundations of air traffic control began to take shape (Gilbert, 1973). As airline traffic increased, area controls were introduced to prevent mid-air collisions. A milestone came in 1938 with the first formal draft of flight rules, which established controlled airspaces around certain airports and made compliance with air traffic control instructions mandatory. This also marked the introduction of the two distinct rule sets, instrument flight rules (IFR), and visual flight rules (VFR), both of which remain in use to this day. VFR pilots navigate using visual

references in the environment, whereas IFR pilots rely solely on their flight instruments. Since IFR operations hence do not require external visual cues, flights can operate in all weather conditions, including at night (International Civil Aviation Organization, 2024).

During the same period, the first dedicated airway charts for IFR operations emerged. Recognizing the need for reliable navigation aids, Captain E. B. Jeppesen began compiling notes and charting air routes based on his own flight experiences, enabling him to conduct precise instrument flights when other peers remained grounded. His efforts laid the foundation for Jeppesen, a company that remains a leading provider of aeronautical charts today (Steele, 1998). A few years later, in November 1944, the establishment of the International Civil Aviation Organization (ICAO) in Chicago, Illinois, marked a determining moment in aviation history (Mackenzie, 2010). It serves as an international governing organization, with the aim of unifying procedures and regulations worldwide. Already four years later, in 1948, the newly founded ICAO introduced standardized symbology for various aeronautical charts types among its member states, forming the basis of what is now known as ICAO Annex 4. Still, the individual national aviation authorities remain responsible for the maintenance and publication of these charts (International Civil Aviation Organization, 2022).

Technological advancements

Over the past decades, technological advancements have significantly reshaped aviation, particularly for commercial IFR operations. One of the most fundamental changes has been the shift away from conventional radio-based navigational aids toward satellite-based navigation (GNSS), which has rendered many traditional ground-based navigation systems obsolete (ICAO, 2005). The introduction of glass cockpits has modernized aircraft instrumentation by replacing analogue gauges with multifunction displays (MFD). By consolidating critical flight information into a single interface, thereby saving space and improving situational awareness. The information displayed is dynamically adapted to the phase of flight, allowing pilots to selectively display relevant data and hide non-essential information (Young et al., 2006).

Similarly, at an individual level, the adoption of electronic flight bags (EFB), usually in the form of tablet devices, has revolutionized how pilots access and manage personal flight documentation and information. Previously, physical flight bags were filled with paper charts, flight plans, regulatory documents, and aircraft manuals, requiring painstaking manual updates, cross-checking multiple charts, and physical route plotting (Fitzsimmons, 2002). Today, EFBs have replaced nearly all paper-based documentation, on the one hand reducing payload, on the other hand providing up to date documentation and on other hand real-time integration of critical flight related data such as weather conditions or airspace restrictions directly into navigational documents. Newest generation aircraft

even allow for the communication between EFBs and cockpit avionics, enabling pilots to even more intuitively operate flight management systems (Skaves, 2011).

While commercial flight enjoys such innovations, VFR flying continues to rely on more rudimentary technologies. This is also due to the nature of VFR flight, which is not performed commercially, follows simpler principles, and is commonly performed with older aircraft. It is not uncommon for the age of single-engine piston aircraft usually used for leisure VFR flights to be well over thirty to forty years (ASFG Ausserschwyzer Fluggemeinschaft Wangen, 2025), meaning many pilots still rely on conventional flight instruments, unless their aircraft has been retrofitted with modern avionics. Paper charts remain a widely used, cost-effective method, as they do not require expensive subscription models or adequate devices to be displayed on (Techau, 2018). Nonetheless, EFBs are gradually gaining traction among VFR pilots as well. Moving maps, which provide real-time positional updates, can enhance situational awareness in flight when compared to stationary paper charts. Nonetheless, regulatory requirements still mandate that pilots carry paper maps as backups (Babb, 2016).

Swiss VFR Chart

In Switzerland, the VFR aeronautical chart is published annually in mid-March by *Skyguide*, the organization responsible for managing Swiss airspace, in cooperation with *swisstopo*, the federal office for topography. The Swiss VFR chart covers the entire country and its immediate surroundings at a scale of 1:500'000. Vertically, it spans from surface level upwards to Flight Level 195 (FL195), corresponding to approximately six thousand meters. The aeronautical chart is based on the CH1903+ coordinate system, which utilizes the Bessel ellipsoid and an oblique conformal cylindrical projection (Mercator projection) (Skyguide, 2025b).

Due to the contrast in navigational needs, IFR and VFR charts are designed with fundamentally different priorities. IFR charts mostly rely on abstract, fictional GPS or radio-based way points, often with minimal geographic reference. As they are intended for instrument-based navigation, topographic information is limited to essential elements such as major water bodies and terrain elevations, primarily to guide pilot attention and maintain spatial awareness, resulting in more uncluttered designs. In contrast, VFR charts are designed to support visual navigation and must incorporate both aeronautical data and detailed topographic features. This dual requirement significantly increases the visual complexity and feature density of VFR charts (Federal Aviation Administration (FAA), 2025a).

Like other VFR aeronautical charts around the world, the Swiss VFR area chart is composed of three distinct layers. The base layer consists of a topographical map

featuring land cover such as lakes, glaciers, and infrastructure such as highways, major roads, railways, and settlement boundaries. Smaller towns are represented by circular point symbols, while prominent landmarks such as castles, churches, monasteries, power plants, and tank farms are specifically marked. Additionally, mountain passes are displayed, as the performance limitations of piston-engine aircraft often require pilots to navigate through lower-lying valleys rather than flying in a direct line. Relief color grading and hillshading are used to enhance terrain comprehension (Skyguide, 2025c; Swisstopo, 2025).

Overlaying the base topographic layer is the aeronautical information essential for safe VFR flight. Airfields are depicted with their runway orientations, with their point symbols varying based on runway surface type, ownership (private or public), and usage (military, civil, or mixed operation). Controlled airspace boundaries (Class C and D airspaces) as well as restricted and danger areas, though these may not always be active, are shown to avoid infringements. Lastly, ground-based navigational aids (navaids) used for in-flight navigation are represented with the standardized ICAO symbol set (Skyguide, 2025c; Swisstopo, 2025).

Finally, a third layer contains all aviation obstacles represented using both line and point symbols. The most prominent line symbols is high-voltage power line network as well as spanned cables, particularly in mountainous regions. Point symbols indicate single object obstacles such as poles, masts, transmission towers, and wind turbines. Additionally, areas with increased aerial activity, like paragliding, parachuting, and glider operations, are marked to alert pilots to potential enhanced flight activity (Skyguide, 2025c; Swisstopo, 2025).

Map elements from all three layers are labeled to provide pilots with the information necessary for safe navigation. Mountains and hilltops are marked with spot elevation labels to assist pilots in maintaining terrain clearance. Airspace boundaries include annotations of both upper and lower altitude limits, frequencies for air traffic services, and the ICAO airspace class. Navaids are labeled with their identifiers, frequencies, and Morse code representations for signal verification. Airfields are shown with their most important details such as ICAO identifiers, elevation, runway lengths, and communication frequencies. Additionally, minimum sector altitudes (MSA) are indicated for every 30-arc-minute quadrant in both latitude and longitude directions (Skyguide, 2025c; Swisstopo, 2025).

1.2   Motivation

Since VFR aeronautical charts are notoriously dense and visually cluttered, efficient map reading becomes a challenge, especially with added time pressure. This issue poses a real challenge, especially for student pilots. Based on personal experience during flight training, it can be confirmed that students struggle to interpret the charts confidently and must invest significant time and effort to master the efficient reading of aeronautical charts.

Despite the central role of VFR charts in flight planning and navigation, the design of the official Swiss VFR aeronautical chart has remained largely unchanged for decades. A second set of maps at a larger 1:250,000 scale was introduced in 2010 to address clutter in the two busiest areas around Geneva and Zurich. But, these charts only cover those two regions, and the smaller map extent increases the need for in-flight folding. In this context, improving chart readability and overall efficiency remains a key concern, best if achieved directly on the 1:500'000 area chart (Swisstopo, 2025).

Piloting an aircraft is a multifaceted task that involves managing high cognitive loads. Unlike driving, where pulling over is an option, pilots cannot simply stop to reassess a situation mid-flight. They must always mentally stay "ahead" of the aircraft, especially if any changes in plans should occur. This continuous situational awareness must be maintained while simultaneously monitoring and adjusting flight parameters such as speed, heading, and altitude, and, when applicable, communicating with air traffic control, executing procedures, or running through checklists.

This need to juggle multiple tasks during flight makes piloting an aircraft inherently demanding and time-critical. As a result, efficient time management is essential. As noted by Aretz (1991), "navigation competes with flight control for limited spatial processing resources." Recent advancements, as discussed in Section 1.1, therefore aim to reduce this workload and improve pilot performance, yet VFR flight has not progressed at the same pace.

The official Swiss VFR chart is still distributed in raster format, effectively a digital reproduction of the paper map. It remains non-interactive, lacks georeferencing, and does not support different zoom levels, making it unsuitable for integration into modern digital tools. Already in 1993, Leary (1993) recognized the need to digitize and standardize aeronautical charting across nations. Yet, over three decades later, VFR pilots often continue to rely on fragmented sources that must be manually checked and cross-referenced, especially during pre-flight planning.

Popular EFB applications provide hybrid solutions by overlaying aeronautical data onto

open-source base maps such as *OpenStreetMap* or satellite imagery. These platforms offer a wide range of functions and a high degree of customization. However, they are not specifically optimized for aeronautical information communication. Therefore, there is a growing need for a modernized VFR chart that is compatible with digital platforms. As EFB use continues to rise and reliance on paper maps declines, it is a perfect moment to rethink how aeronautical information is visually communicated (ForeFlight, 2025) and possibly adapt the current map design.

To better understand how effectively pilots interpret visual information on VFR aeronautical charts, this study employs eye-tracking technology. Given that approximately 80% of human cognitive input is estimated to stem from visual perception (Hubel, 1988), eye-tracking offers a direct and unfiltered method to observe the perceptual and cognitive processes involved when interacting with a given stimuli. Since its early adoption in cartographic research, eye-tracking has gained widespread popularity due to its capacity to reveal user attention and behavior, or vice versa, the effectiveness of different presented designs (Fairbairn and Hepburn, 2023).

Human eye movements are characterized by two primary components: fixations, where the eye momentarily pauses to process information, and saccades, the rapid movements between fixations. Analyzing these elements enables both quantitative assessments through metrics such as fixation duration, saccade length, and saccade velocity, as well as qualitative insights, such as the visual scan paths and attention "hot spots" at individual task levels. While qualitative analyses of Swiss VFR charts and adaptations have been conducted, no prior studies have extracted and compared quantitative eye-tracking metrics in this context (Holmqvist et al., 2011).

The findings of this study aim to improve the understanding of how a redesigned VFR chart impacts visual behavior. These insights may inform future iterations of the Swiss VFR chart, especially in light of potential vectorization and integration into existing EFB systems.

1.3   Research Objectives

This study investigates the impact of an updated map design on participants' performance compared to the current map design. The redesigned chart aims to address existing challenges by reducing visual clutter and enhancing saliency. To evaluate and quantify the effectiveness of the new design, both expert and non-expert participants are tasked with solving a series of visual search tasks. By solving these tasks, expert and non-expert participants' eye-tracking data will give insights into the performance of the map designs.

From this framework, the following research questions are derived:

- **RQ 1:** How is an improvement in readability reflected in eye-tracking results?

- **RQ 2:** How is a potential bias towards the current map influencing map reading skills between non-expert and expert groups?

- **RQ 3:** What are differences in map reading patterns between the existing and proposed map designs?

1.4   Thesis Structure

This thesis is structured into six chapters. Chapter 1 introduces the research topic, motivation, and objectives. Chapter 2 reviews the current state of research on visual perception, aeronautical cartography, and eye-tracking. Chapter 3 outlines the methodological framework, including participant recruitment, experimental design, and the development of the updated map. Chapter 4 presents the results of the eye-tracking experiment, combining quantitative metrics and visual analysis. Chapter 5 discusses the findings in relation to the research questions and situates them within the broader context of eye-tracking. Chapter 6 concludes the thesis and outlines implications for future work.

1.5   Terminology

Throughout this thesis, the terms "map" and "chart" will be used interchangeably to refer to the 1:500'000 Swiss VFR aeronautical chart. Likewise, "new" and "updated" both refer to the chart version created for this project.

Chapter 2

Related Work

2.1 Eye-tracking to understand Map Use

Eye-tracking as an experimental method, originally rooted in fields such as psychology and neuroscience, has been adopted in cartography, based on the assumption that maps function as visual stimuli (Ciołkosz-Styk, 2012). It is a powerful method for obtaining an unfiltered view into participants' thinking. According to the eye-mind hypothesis (Chekaluk and Llewellyn, 1992), eye movement recordings can serve as a direct proxy for the current contents of conscious processing. Early examples of eye tracking in cartography date back to the 1970s, with Williams (1971) exploring symbol selectivity and Jenks (1973) analyzing visual attention on dot maps. Steinke (1987) was among the first to provide an in-depth review of eye movement studies in cartography, and in other disciplines such as psychology. Similarly, early research by Çöltekin et al. (2010) contributed to a deeper understanding of how people engage in visual interaction strategies and decision-making in pattern recognition.

Since then, the application of eye tracking has become increasingly influential in cartographic research. According to Krassanakis and Cybulski (2019), out of the existing experimental techniques, eye-tracking represents one of the most valuable methods towards the evaluation of different aspects related to map reading process. Krassanakis and Cybulski (2019) also highlights the growing integration of eye tracking technology within spatial research, including domains of spatial cognition, geographic information science, and cartography.

Cartographic symbolization plays a critical role in the readability and usability of maps. Based on Bertin (1983)'s visual variables, symbol priorities such as size, color, shape, and orientation guide how users perceive and interpret spatial data. Eye tracking studies

have increasingly been used to evaluate the effectiveness of these variables. A study of Çöltekin et al. (2009) assessed visual variables in change detection tasks and found that symbol size was the most effective and efficient user performance, while orientation was least effective. These findings align with the concept of visual selectivity - the ability of the visual system to isolate relevant symbols from a complex background Çöltekin et al. (2009); Krassanakis and Cybulski (2019). Several eye tracking studies have further examined how visual search behavior is influenced by symbol location and complexity. Krassanakis et al. (2011b,a, 2013) showed that symbols placed in the periphery resulted in more complex scan paths compared to centrally located ones. Moreover, the visual search in subsequent tasks was influences be the target's previous location, indicating memory effects and learning patterns in visual navigation. Fixation clusters also tended to concentrate on symbol-rich areas of the map, particularly where point symbols were used.

Eye tracking research also plays a significant role in exploring the differences between expert and novice map users in cartography. Fairbairn and Hepburn (2023) emphasize that grouping participants into experts and novices prior to the data collection is essential for meaningful comparison. Their study highlights how eye tracking can reveal differences in visual attention, search strategies, and interpretation between user different groups, helping improve map usability through targeted design adjustments.

Recent research also explores the use of entropy in eye tracking heatmaps as a quantitative measure of interface complexity. According to Lee et al. (2023), heat map entropy can serve as a predictor of judgment time, making it a valuable indicator for evaluating the suitability and usability of visual designs.

2.2   Visual Clutter

Visual clutter has been extensively studied due to its direct impact across various research domains. It is commonly defined as an excess of visual information that slows information retrieval and negatively affects search performance. In cartography, visual clutter plays an important role in map design and the effective communication of spatial information. However, there is no universally correct metric for measuring visual clutter. In their work, Rosenholtz et al. (2007) aimed to quantify visual clutter across different images. They demonstrated that three distinct clutter metrics—feature congestion, sub-band entropy, and edge density—can effectively predict search times and other task-related performance measures. Lohrenz et al. (2009) developed a refined model incorporating color density and saliency, which returned improved predictive results over the approach of Rosenholtz et al. (2007), which can be used to determine visual clutter levels.

The well-known memory model developed by Atkinson and Shiffrin (1968) introduces a three-layer structure, demonstrating that visual and auditory cues of a presented stimuli are initially processed by the sensory register, passed on to the short-term memory, often referred to as "working memory", before being permanently saved in the long-term memory. Items stored in the working memory can then be recalled within 30 seconds of the initial exposure to be processed. Intuitively, an increasing map complexity will therefore negatively affect memory performance, as demonstrated by Bestgen et al. (2017). Their study indicates that certain cartographic elements, such as grid overlays, facilitate the recollection of spatial relationships, enabling users to recall information more accurately from memory.

Zhang et al. (2022) further support the role of visual structuring in cognitive efficiency, showing that color-coding icons on mobile app displays leads to significantly faster visual search times. This finding highlights the importance of thoughtful styling in cartographic design to enhance information retrieval. Nonetheless, ICAO Annex 4 imposes certain limitations, for example explicitly restricting airspace boundaries to two colors and mandating a single color hue for all other aeronautical map features (International Civil Aviation Organization, 2022). Consistent with Zhang's findings, Cybulski and Ledermann (2024) report that as point symbol similarity increases, visual search times become longer, further emphasizing the necessity of differentiation in cartographic symbology.

In their overview, Doyon-Poulin et al. (2012) examined the effects of varying levels of visual clutter on pilot performance. Their findings indicate that a moderate level of clutter is generally preferable for optimizing task completion times. Still, they emphasize that visual clutter alone is not a sufficient parameter for evaluating the effectiveness of a display. Instead, they argue that "the optimization of clutter ought to be linked with the particular task, context, and settings the display has to support" (Doyon-Poulin et al., 2012, p. 7), stating the need for dynamic display styles that adapt to task-specific demands. Additionally, (Doyon-Poulin et al., 2012) highlight that piloting an aircraft involves considerably greater cognitive complexity than the controlled visual search tasks commonly studied in traditional research, thereby mitigating the expected impact of increased search times due to increased clutter. Nonetheless, Liu et al. (2023) found that increased levels of clutter on head-up displays make fixations longer, but saccade lengths stayed the same.

2.3   Aviation Cartography

Most research to date has focused on IFR (Instrument Flight Rules) charts, largely because the majority of commercial flights operate under IFR conditions. Multer et al. (1991) explored potential design improvements for Instrument Approach Procedure (IAP) charts

developed for the Federal Aviation Administration (FAA) of the US Department of Transportation. The authors recommend that 40% of the chart area should be reserved as white space. Unfortunately, this not applicable to VFR charts, given their denser information requirements. Still, their experiments with various highlighting techniques revealed that font size had a significant impact on readability and performance, which can be considered in VFR chart design. Around the same period, Osborne et al. (1992) investigated the use of icon-based representations for missed approach instructions in place of textual descriptions. The study found that participants performed significantly better with the visual icons. To this day, US charts feature this method in their IFR approach charts.

Butchibabu et al. (2012) investigated the impact of de-cluttering IFR approach and arrival charts by splitting up different arrival procedures across separate sheets. Their findings showed that pilot performance significantly improved across all airports included in the study. Similarly, Yeh et al. (2021) conducted a questionnaire-based study with pilots to assess which elements on IFR approach charts should remain permanently visible and which should be toggled on or off. Of 364 determined chart elements, only 173 (47%) were considered essential for permanent display. The studies highlights a direct impact on pilot performance through adjustments that are easily implementable within the scope of modern EFBs, even enabling personalized configurations without omitting safety-relevant information.

The use of EFBs in general is well studied, especially in context of pilot performances compared to paper charts. For example, Winter et al. (2018) found that pilots reported significantly lower workload when using electronic charts compared to traditional paper maps. The study also noted a negative correlation of response times and number of flight hours when using EFBs, whereas performance with paper maps remained constant across flight experience. A very recent study of Sarbach et al. (2025) analyzed the effects of three dimensionally displaying live weather into flight plan charts and its effect on situational awareness. Although a planar representation yielded highest situational awareness levels, these developments show the vast possibilities and potential of EFB technologies.

2.4   Swiss VFR charts

In the context of this thesis, only little research directly aligns with its specific focus of comparing chart design efficiencies. In particular, studies addressing the Swiss VFR aeronautical chart are very scarce. Two exceptions include a research paper by Sarbach et al. (2023) and a Master's thesis by Juliette Marx (2015) at the Department of Geography at the University of Zurich.

Sarbach et al. (2023) conducted a qualitative analysis comparing the Swiss VFR chart to three other national charts, namely the United States (FAA), Austria (Austro Control), and Australia (CASA). They developed a cartographic evaluation framework based on general design principles such as color hues, font sizes, and map symbolism. The Swiss chart consistently ranked first or second across all metrics according to the framework. In a second phase, the authors assessed user perception through a combination of visual search tasks and qualitative questionnaire among 27 participants. The results showed that the Swiss chart was the overall favorite, receiving the highest scores. Nevertheless, the rankings were of qualitative nature.

In her Master's thesis, Juliette Marx (2015) explored the redesign of selected elements on the visual approach charts (VAC). These charts are intended for use during the departure and approach phase of flights, unlike the 1:500,000 VFR area chart examined in this thesis, which is used for preflight planning and in-flight area navigation. Marx's modifications included eliminating township names, repositioning airspace vertical limit labels to the edges of airspace boundaries, and adjusting the color hues of lakes, forests, and settlement areas. She then evaluated these changes using visual search tasks, performed under increased cognitive load through visual and auditory stimuli designed to simulate an in-flight environment. Importantly, her study included both experts (pilots) and non-experts (geography students) as participants. While non-experts generally appreciated the modified designs, pilots tended to prefer the original charts, suggesting a possible bias toward familiar layouts. A similar trend emerged in Sarbach's study, where Swiss pilots showed a strong preference for their national chart, also likely due to familiarity.

Chapter 3

Methods

This chapter outlines the methodology used to conduct the eye-tracking study. It details the steps taken from participant recruitment to map preparation and the evaluation of results. The chapter is divided into four main sections:

- **Section 3.1** describes the characteristics of the study participants, including demographic information, flight experience, and how participants were grouped.

- **Section 3.2** explains the process of creating the updated map design used in the study. It covers everything from data acquisition to post-processing for the final visual stimuli.

- **Section 3.3** details the structure and procedure of the eye-tracking experiment, including the tasks given to participants, the setup of the experimental environment.

- **Section 3.4** outlines the statistical analysis methods used to evaluate performance and eye-tracking metrics collected in the study.

3.1 Participants

Participant recruitment took place through the personal network on a voluntary basis. There were no health requirements in place to participate in this study. Colorblindness, as with the pilot medical certificate, not allowed.

The study included 31 participants (M: 25, F: 7) with an average age of 33 years ($M = 33.55, SD = 13.09$). Of those 31 participants, 11 own a pilot license (7 PPL, 1 CPL, 3 ATPL), 2 are air traffic controllers, and 3 are in pilot school. On average, the pilots had roughly 250 hours of VFR flight experience. 9 of the pilots are based in the German speaking part
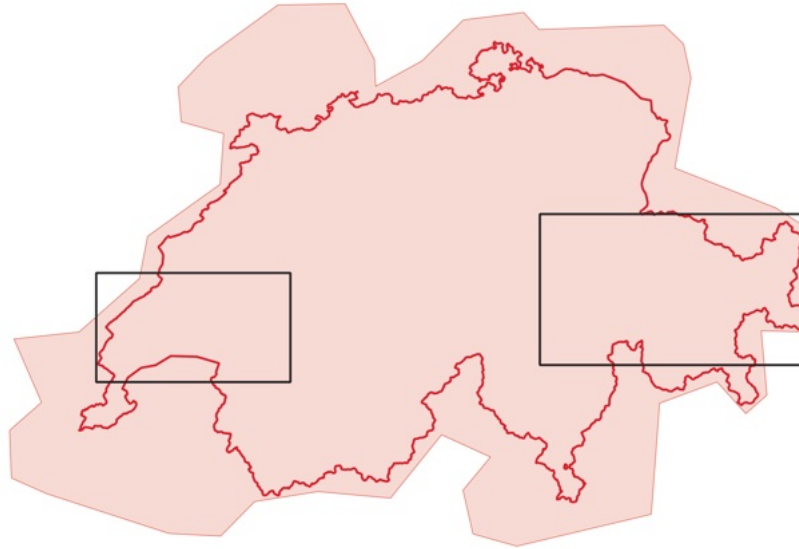
Figure 3.1: Extent of the updated Map (pink) and two Task Areas (black)

of Switzerland, 1 in Austria close to the Swiss border, and 1 in Ticino. All of the pilots stated using EFBs.

3.2   Map Design

This section outlines the process of creating the updated Swiss VFR map design developed for this thesis. The design integrates multiple cartographic elements in alignment with ICAO Annex 4 specifications, with adjustments to the current map aimed at improving visual clarity and usability for pilots. Although the current study focuses on static map outputs, the map was designed in vector format to support future adaptation for EFB use.

3.2.1   Data Collection

Base Map

The basemap geodata for the official Swiss VFR area chart is based on the 1:500,000 national topographic map. Vector features such as roads, railways, lakes, forested areas, and permanent snow cover are imported from the Swiss Map Vector 500 geodatabase[1], or the Swiss TLM Regio dataset[2]. The following datasets and selection criteria were applied:

- **Roads:** Imported from `DKM500_STRASSEN` dataset. Only freeways (`OBJEKTART=100`), expressways (`OBJEKTART=200`), large roads (`OBJEKTART=1200`), and small roads up to 4m in width (`OBJEKTART=1400`) are included. Smaller roads and pathways are

---

[1]swisstopo   Swiss   Map   Vector   500:   `https://www.swisstopo.admin.ch/de/landeskarte-swiss-map-vector-500`

[2]swisstopo SwissTLMRegio: `https://www.swisstopo.admin.ch/en/landscape-model-swisstlmregio`

excluded. Only above-ground segments are retained, tunnels are omitted. Freeways are extracted into a separate layer.

- **Railways:** Imported in the same manner as roads, without additional filtering. Underground segments are omitted.

- **Flowing Waters:** Sourced from the SwissTLMRegio[2] dataset, with values higher than 9 in BREITE excluded.

- **Forest cover:** Imported from the DKM500_BODENBEDECKUNG dataset, with patches smaller than 500'000m^2 (0.5km^2) excluded.

- **Lakes:** Included from the Swiss Map Vector 500 dataset without further filtering.

- **Settlement boundaries:** Imported from SwissTLMRegio[2].

- **Permanent snow cover:** Imported from SwissTLMRegio[2].

- **Hillshade:** The digital elevation model used for the hillshade was retrieved from the DHM25 / 200m dataset of swisstopo[3]

The udpated map design does not include any data from neighboring countries, except for areas below Swiss airspaces that stretch abroad.

---

[3]swisstopo DHM25 / 200m https://www.swisstopo.admin.ch/en/height-model-dhm25-200m

Aeronautical Data

The aeronautical data used in to create the new map consists of airspace boundaries, airfields, navaids, and certain airports' reporting points:

- **Airspaces:** Official airspace data is only available in raster format, which is not suitable for the purposes of this study. As an alternative, the airspace database of the Swiss Hang Gliding and Paragliding Association[4] was used. Their data is parsed from official documentation and made available for download in the OpenAir[5] format, which can be converted into an ESRI-Shapefile format for further geo-processing and.

- **Airfields:** Airfields and heliports were manually compiled into a `.csv` file containing the ICAO code, airfield type, runway orientation, and coordinates as attributes. All attributes were derived from official Aerodrome Information (AD INFO) charts[6]. In Switzerland, there is a total of 55 airfields and 22 heliports.

- **Navaids:** Like the airfields, navaids were manually compiled into a `.csv` file. Each entry included the name, identifier, type, and geographic coordinates. Until March 20th, 2025, there were 12 navaids in Switzerland: 7 VOR-DME, 4 DME, and 1 NDB (Skyguide, 2025a) (see Table 1).

- **Reporting Points:** Larger airports operate with predefined arrival and departure routes, requiring pilots to report their position to air traffic control upon reaching specific reporting points. These points were obtained from the openAIP database, which provided coordinate data. Altitudes were cross-referenced and extracted from official aeronautical charts. Both altitude constraints and recommended altitudes were included. For Basel Mulhouse airport, different altitudes are used for departing and arriving traffic, therefore the values were left out.

Obstacle Data

In the scope of this thesis, only point obstacle data and transmission lines were included in the map.

- **Obstacle point data:** Swisstopo provides aeronautical obstacle point data including all obstacle types, such as towers, masts, chimneys, and other structures. The obstacles were filtered to only include elements higher than 300ft (91.44m) above ground level, and divided into the three types present on the current VFR map.

---

[4]Swiss Hang Gliding and Paragliding Association SHPA `https://www.shv-fsvl.ch/en/`
[5]Open Air User Guide `http://www.winpilot.com/UsersGuide/UserAirspace.asp`
[6]Skyguide eVFR Manual `https://www.skybriefing.com/de/evfr-manual`

- **Transmission lines:** High-voltage transmission line data was obtained via the online service portal *geodienste.ch*. The dataset includes all installations operating at voltages greater than 36kV, represented as line geometries.

- **Wind turbines:** As the official obstacle dataset does not appear to include all wind turbines, an additional shapefile containing the locations of all wind energy facilities was downloaded from Swisstopo to ensure completeness. Identical to the point obstacle data, only windmills higher than 300ft were kept.

3.2.2 Design Guidelines & Map Colors

The design guidelines from ICAO Annex 4 (International Civil Aviation Organization, 2022) are found in Appendices 2 and 3, which define the recommended symbology and color usage for different aeronautical charts. Particularly important is the specification of two primary colors for representing aeronautical features. Although Appendix 3 explicitly states that for area charts "different colors may be required" (International Civil Aviation Organization, 2022, APP 3-1), the updated map design follows a two-colored scheme, similar to the FAA sectional charts (Federal Aviation Administration (FAA), 2025b). On those, blue and magenta are used in parallel to distinguish between different classes of controlled airspace.

Dark Blue	**#2E2F74**
Magenta	#732C24

Table 3.1: ICAO Aeronautical Data Colors

Other color choices were kept mostly in accordance with the current Swiss VFR 1:250'000 Area Chart Geneva / Zurich Skyguide (2025c).

Roads	#FFFFFF
Railways	#000000
Highways	#FFDB11
Obstacles	#DB1E2A
Settlements	#F7F208
Forest	#C2DEC9
Lakes	#BBE0F2
Streams	#487BB6
Permanent Snow	#DFEDF2

Table 3.2: Updated Map Colors

3.2.3   Base Map

The base map design is inspired by the 1:250,000 Swiss VFR area chart Skyguide (2025c), but deviates from the graduated color-based relief approach proposed by ICAO, which is also adopted by the current 1:500'000 area chart Skyguide (2025b). Instead of using a graduated color scale to represent elevation, the updated design omits background colors in favor of improved contrast for map elements. In practice, colored elevation bands offer almost no navigational value due to the complexity introduced by the numerous subclasses. The current Swiss area chart, for instance, features ten distinct elevation color classes. Cartographic literature recommends using no more than five classes to enable users to clearly distinguish between them (Brewer, 1994).

To retain essential topographic context, the hillshade layer was preserved. This element remains particularly valuable in mountainous terrain as found in Switzerland, where visualizing topography is critical for situational awareness and especially route planning. The hillshade was generated using commonly used parameters, with a sun azimuth of 315° and an altitude of 45°. It was rendered with increased transparency to ensure a visually subtle yet cartographically effective background layer.

3.2.4   Point Symbols

Airfields

The icons for airfield point symbols were derived from the current VFR aeronautical chart. It uses an efficient method to depict different airport types that is well established in Europe. According to International Civil Aviation Organization (2022), only runway surface type must be visually declared, all other information is extracted from the labeling or AD INFO charts.

In accordance with the airspace boundary coloring, military airfields are rendered in ICAO magenta, as specified in Table 3.1. This is consistent with the use of red color to indicate military airfields on the current Swiss VFR aeronautical charts, visually highlighting their closure to civilian traffic.

Navaids and Reporting Points

The symbology for navaids follows long-established international conventions. No changes were made to their representation in the updated map design. VOR stations are depicted with a compass rose, oriented to magnetic north. Reporting points, typically used for position reporting near controlled airports, are shown as filled triangles. Both types of features use the ICAO blue, maintaining consistency with other aeronautical symbols.

Table 3.3: Point Symbols for Airfields

Airfield Type (count)	Symbol
Private paved	
Private unpaved	
Private mixed use (3)	
Public paved (12)	
Public mixed use (1)	
Military (4)	
Heliport	

Other Point Symbols

Point symbols for obstacles follow ICAO specifications and distinguish between different obstacle classes, such as, lit obstacles, high obstacles, and wind turbines. Mountain passes are also marked using established symbology, as they are often used for navigational purposes.

3.2.5 Line Symbols

The updated map features seven types of line symbols. Infrastructure line symbols include highways, larger and smaller roads (as in the current chart), and railways. Streams are represented with a uniform line width. High-voltage power lines, which are a common obstacle, are shown in red, as defined in Table 3.2. Recommended VFR routes, as proposed in Appendix 2, are included using a dotted line style, which is significantly thinner than on the current Swiss chart.

3.2.6 Area Symbols

Basemap

Settlement extents are shown using filled areas, adapted from the 1:250,000 map rather than the 1:500,000 chart, which uses varying point symbol sizes for smaller townships. Forested areas are displayed with slight transparency to allow integration with the underlying hillshade. Permanent snow cover and glaciers are rendered in a subtle, semi-transparent grayish blue to visually distinguish them.

Airspace Boundaries

ICAO Annex 4 Appendix 2 recommends drawing airspace boundaries in blue, with the exception of Class A airspace. Following a comparable approach, the updated map design applies ICAO blue to all aeronautical features, except for military airfields and airspaces, which are depicted in ICAO magenta. This solution is suited best in the Swiss context, as there are no Class A airspaces, and airfields already are held in a different color as civil airfields.

Since not all airspaces are operational 24 hours a day, a distinction was made between 24/7 and different operating hour models (HX/H0/TEMPO). Dashed lines for non continuous use. FAA chart currently uses dashes as design feature for Class D airspace boundaries. The ICAO symbol for uncontrolled class G airspaces was used for class E airspace, since that's the lowest class controlled airspace in Switzerland.

In addition, the green boundary for sanctuaries as well as danger areas and restricted areas symbology was taken over from both existing Swiss VFR charts.

Table 3.4: Boundaries for Airspace Classes

Airspace Type (Count)	Symbol	Airspace Type (Count)	Symbol
CTA (4) & AWY (2)		Class D TEMPO (3)	
CTR (3)		Military Class D TEMPO (2)	
CTR HX (11)		Class E (3)	
Military CTR HX (5)		Class E HX (1)	
Class C (30) & D (8)		D- & R-Area (35)	
Class D HX (20)		Sanctuary (5)	
Military Class D HX (14)		Country border	

3.2.7 Labeling

The most significant change in the updated map design lies in the labeling system. A consistent typographic philosophy was applied across the layers, aiming to reduce unnecessary visual clutter. The font used throughout the map is *Helvetica Neue*, a sans-serif typeface selected for its good readability. Base map labels are rendered with standard character spacing, while all aeronautical features and their associated labels are condensed to conserve space. Labels are placed with the goal of avoiding overlap with

other map features. If overlap cannot be avoided, preference is given to obscuring vowels rather than consonants, as vowels carry less weight in word recognition and allow for faster visual processing (Berent and Perfetti, 1995). Boxed backgrounds are avoided, with the exception of the highest spot elevation on the map extent.

The current chart uses a combination of yellow settlement areas and round point symbols for smaller towns, often omitting labels for less prominent locations, leaving point symbols without a label. Yet, "Cities, towns, and villages shall be selected and shown according to their relative importance to visual air navigation." (International Civil Aviation Organization, 2022, Chapter 17.7.1.1) This allows for some flexibility in label placement. Particular attention was given to towns located at valley junctions or along airspace boundaries, as these are often used as navigational or reporting points during VFR operations.

Minimum Sector Altitude (MSA) labels are included per quadrant and do not interfere with other map elements. Spot elevations remain identical to the current chart and are displayed in italics. Major city names are presented in bold and capital letters, following the FAA chart visual hierarchy, with an added white outline to improve contrast against the background. Labels for mountain passes were retained, although the prominent boxes indicating recommended altitudes along VFR routes were removed and replaced with a single, bold label, in line with ICAO standards.

Airports and heliports are labeled on two lines, with the first line containing the airfield name and ICAO code grouped together to emphasize their association. The second line includes additional operational data as per ICAO conventions. Additionally, ATIS frequencies are now included where available,eliminating the need for pilots to cross-reference multiple charts. Military airfields are distinguished with magenta labels, consistent with their treatment of point symbols and surrounding military airspace boundaries.

Since 3D space is mapped on 2D planar map on aeronautical charts, airspace labels feature the semantically most complex labelling. Unlike the current Swiss chart, where frequencies and identifiers are separated, and vertical limits are enclosed in separate boxes, the updated design consolidates all relevant information into a single, unified label, as ICAO standards propose. This also aligns with practices seen in other countries and significantly improves information readability.

These labels are included not only for controlled airspaces such as AWY, CTA, CTR, and TMA, but also for FIZ, RMZ, and TMZ zones (see Table 1). Additionally, a new hybrid label format was developed to combine the information of multiple sequenced TMA into a single, long-format label. The information for these TMA, although not currently

visible on the official chart, are fully integrated on the updated design of the 1:500,000 map.

3.2.8 Post-processing

The post-processing of the QGIS export was carried out in Sketch.[7] The map was exported from QGIS in SVG vector format and imported into Sketch, where the vector data was organized into corresponding thematic layers for further editing.

The aeronautical labels mentioned above were also created within Sketch. In total, over 70 distinct airspace labels were created. Labels for navaids followed the same design principles, displaying the individual identifier, frequency, and Morse codes.

Township labels exported from QGIS were manually adjusted in Sketch to optimize legibility at the 1:500,000 scale. Label placement was refined to avoid overlaps and improve clarity in densely populated or topographically complex regions. The same applies to terrain spot elevations.

---

[7]Sketch, version 101.8, Sketch B.V., 2025, `https://www.sketch.com`.

3.3 Experimental Design

3.3.1 Pre-study Questionnaire

As described in Section 3.1, participants were asked to complete an online questionnaire prior to the experiment. The questionnaire was structured into three sections. The first section gathered general personal and demographic information. The second section contained aviation-related questions, giving insights into experience and preferred methods of navigation. If a participant did not hold a pilot license, the second section was automatically skipped. Lastly, the third section focused on participants' general map use habits and geographic knowledge of Switzerland.

Before participating in the study, all participants were informed about the nature and scope of the experiment. They were required to read and sign an informed consent form, which outlined that participation was voluntary, that the collected data would be used anonymized. Also, participants were given the opportunity to withdraw at any time without providing any justification.

3.3.2 Eye-tracking Study

The main part of the study consisted of a static eye-tracking experiment conducted in situ at the University of Zurich. Participants were asked to perform a series of map-based tasks while their eye movements were recorded using an eye-tracking system. The aim was to analyze visual behavior across different map designs and determine how changes in cartographic layout influenced task performance and visual search strategies.

Stimuli

The experiment included four visual stimuli, representing the two regions of Grisons (`T1`) and Romandie (`T2`), each shown in both the current and the updated map design. The stimuli are referred to as `T1C` and `T2C` for the current design, and `T1N` and `T2N` for the new design. All maps were displayed in full-screen resolution (1920 by 1080 pixels), with each region shown at identical scale and extent to allow for consistent comparison across map designs.

The two map extents were selected to represent different levels of visual clutter. The area in the Canton of Grisons is relatively sparsely populated, mainly due to its topography, with limited infrastructure and resulting minimal airspace complexity. Within the selected extent, only two airports are present, and no controlled airspace boundaries exist, aside from a flight information zone at Samedan airport. In contrast, the Romandie region exhibits a significantly higher degree of visual clutter, featuring several airspaces belonging to Geneva airport and Payerne airfield, extensive transportation infrastructure,

and large built-up areas.

The final map visuals were exported by cropping sections from the current map design (TIFF format) and from the updated map design project created in Sketch. The selected map extents covered different sized areas, as shown in Figures 3.2 and 3.3 below.

(2,706,300, 1,213,400)   (2,834,600, 1,213,400)

9,276 km^2

(2,706,300, 1,141,100)   (2,834,600, 1,141,100)
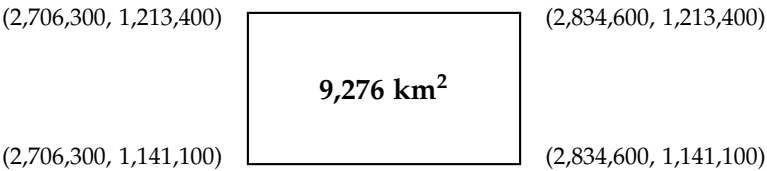
Figure 3.2: Map extent for Task 1 (Grisons), in CH1903+ coordinate format

(2,493,600, 1,185,200)   (2,586,700, 1,185,200)

4,868 km^2

(2,493,600, 1,132,900)   (2,586,700, 1,132,900)
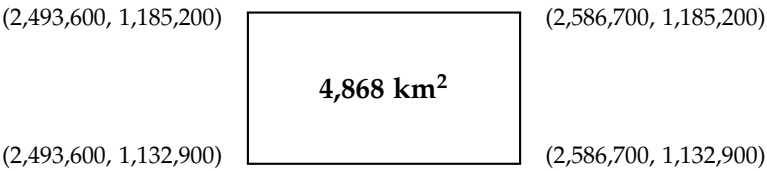
Figure 3.3: Map extent for Task 2 (Romandie), in CH1903+ coordinate format

Tobii Pro Labs Project Design

The structure of the study foresaw participants completing a visual search task, immediately followed by a multiple-choice question. Since Tobii Pro Lab (Tobii AB, 2024) does not support built-in question prompts, static slides (1920 by 1080 pixels) with answer options were created in Sketch and imported into the project. All remaining elements of the experiment were put together within the Tobii Pro Labs environment.

A sequential, non-randomized structure was used. The study followed a within-subject design with all subjects completing both tasks (scenarios) on both the current and updated map design. Each stimulus included one "flight scenario": Task 1 (Grisons) asked participants to imagine flying from Bad Ragaz to Samedan, Task 2 (Romandie) involved flying from Yverdon Les-Bains to La Côte. Task 1 included four sub-tasks, task 2 included five sub-tasks. This resulted in a total of 18 sub-tasks per participant. The sub-tasks were slightly varied between scenarios to avoid priming, but the underlying structure and AOIs remained the same, allowing for more or less direct comparison. It was expected that the second presentation of a similar task might be answered more quickly due to familiarity.

To avoid these systematic performance advantages throughout the second round of task completion, a counterbalancing scheme was implemented. Two contrasting stimuli sequences were used: One in the order `CNNC` (Current-New-New-Current), and the other `NCCN` (New-Current-Current-New). This ensured that both map designs were seen in varying orders across participants. Task 1 was always presented before Task 2 in both rounds to maximize the time between identical map extents. The order was also chosen based on the visual complexity of the stimuli, with Task 1 being less cluttered than Task 2, as described in sub-section 3.3.2.

The questions presented throughout the experiment can be categorized into three types, following the classification proposed by Polatsek et al. (2018): Retrieve Value (RV), Filter (F), and Find Extrema (FE). Retrieve Value tasks involve locating a specific element or reading a particular value from the map. Filter tasks require selecting information based on a defined set of criteria. Find Extrema tasks involve comparing multiple specific attributes to determine a maximum or minimum value. The majority of questions fell into the RV and FE categories. Since the task sequence was not structured by stimulus but rather by presentation order, each sub-task is grouped by stimulus and sorted by its sequence and iteration.

Table 3.5: Overview of sub-tasks, corresponding questions, and task types

Sub-task	Question	Task Type
Grisons (Task 1)		
Sub-task 1, Iteration 1	Which navigational aids are to be found in the map extent?	F
Sub-task 1, Iteration 2	What is the frequency of Corvatsch CVA (DME)?	FV
Sub-task 2, Iteration 1	Which of the three passes leading to the Engiadina Valley is the highest?	FE
Sub-task 2, Iteration 2	Which of the three passes leading to the Engiadina Valley is the lowest?	FE
Sub-task 3, Iteration 1	Are there any obstacles around Samedan?	F
Sub-task 3, Iteration 2	What is the top elevation of the antenna in St. Moritz?	FV
Sub-task 4, Iteration 1	What is the approximate elevation of Chur?	F (FE)
Sub-task 4, Iteration 2	What is the approximate elevation of Domat/Ems?	F (FE)
Romandie (Task 2)		
Sub-task 1, Iteration 1	Halfway through the flight you see an airfield with a grass runway. Which airfield is it?	F
Sub-task 1, Iteration 2	What is the airfield elevation of Montrîcher?	FV
Sub-task 2, Iteration 1	What airspace lies above Lausanne?	FV
Sub-task 2, Iteration 2	What is the frequency used for the TMA LSGG 5 airspace?	FV
Sub-task 3, Iteration 1	What is the nearest navaid to La Côte?	F
Sub-task 3, Iteration 2	What is the frequency of Gland GLA (NDB)?	FV
Sub-task 4, Iteration 1	Name a town to the west of the tank farm.	F
Sub-task 4, Iteration 2	Name a town to the east of the tank farm.	F
Sub-task 5, Iteration 1	What is the number of the restricted area above Bière?	FV
Sub-task 5, Iteration 2	What is the lower limit of the LSR-5 restricted area?	FV

The four folders containing the stimuli and corresponding questions and answers were duplicated and arranged to form both sequences. Since the sub-task questions differed slightly between versions, two separate sets of answer slides had to be created and assigned accordingly. The answer slides always listed three answers, one of which is correct. Participants navigated the study using a keyboard input, which was chosen to prevent accidental clicks during the map viewing phase, since backtracking is not possible.

The flight scenario was described at the beginning of the task. The sub-tasks were presented by a question slide featuring a target symbol, if applicable, to prime participants on what to search for on the subsequent map extent. After the question screen, a 1000 ms fixation cross was displayed to center the gaze for all participants. The stimulus was then shown without a time limit, and participants advanced with a key press once ready to answer the question. The cursor became visible and allowed participants to select their answer via mouse click on the multiple choice slide. They were not informed about the correctness of their answers.

After completing all sub-tasks, participants were shown a thank-you screen. Pressing the Escape key ended the study and stopped the recording.

Eye-tracking Recording

Participants were invited to the University of Zurich and guided to the Eye Movement Lab at the Department of Geography. Upon arrival, non-expert participants received a brief introduction to the map layout and symbology to familiarize themselves with basic visual elements. While understanding the semantics was not within the scope of this study, any questions regarding general map use were addressed. A PowerPoint presentation with a mock-up of the experiment was shown to explain the study procedure and structure to participants. They completed a set of warm-up tasks, using a stimulus that was geographically separate from those used in the main experiment to avoid contextual priming.

Eye movements were recorded using a Tobii Pro eye-tracking system operating at a frequency of 300 Hz. The system tracked both eyes individually, capturing gaze coordinates, fixation data, and pupil and eye openness. Participants were seated comfortably at an optimal viewing distance of approximately 50–70 centimeters from the screen on a chair without wheels. Once ready, a calibration procedure was conducted to align the gaze data with actual fixation points on the screen. Only after a successful calibration did participants begin the experiment at their own pace.

The experimenter remained present in the room to monitor the process but remained

seated behind the participant to avoid influencing their visual behavior. No audio or video recordings were made, only gaze data was collected. At the conclusion of the study, the recording session was saved within the Tobii Pro Lab environment. Input devices such as the mouse and keyboard were disinfected after each session.

3.3.3   Raw Data Handling

Tobii Pro Lab allows exporting eye-tracking recordings in various formats. For this study, the data was exported for each participant in the `.xlsx` Excel format to enable efficient handling using Python. Since the recordings were captured at 300 Hz, each second produces 300 timestamps as rows in the file. Due to the resulting file size, the data was trimmed using various Python scripts. To maintain the overview and store all the results, a dedicated folder was created for each participant.

First, the raw `.xlsx` dataset was filtered to only contain rows produced during exposure to one of the four stimuli `T1C`, `T1N`, `T2C`, `T2N`, i.e. while visually searching one of the four map extents. From there, only the columns containing relevant gaze data were retained for post-analysis. These contain the timestamp, fixation coordinates, eye movement type, gaze event duration, pupil diameter, eye openness, and AOI hits. Since jumps in the recording timestamp exist after filtering for exposure to the map extent stimuli only, the specific sub-task numbers (e.g. `T1C Task 1`, `T1C Task 2`) can be assigned. The trimmed participant dataset was then further split into individual `.csv` files per stimuli and saved in the folder.

Participant Results

From the different stimulus `.csv` files, all relevant metrics were extracted to generate a single task summary (`subtask_statistics.csv`) file for each participant. This file includes one row per sub-task containing key metrics such as its completion time, time to first AOI hit (for one or multiple AOIs), as well as the number of fixations and saccades. Task accuracy (correct/incorrect) was manually annotated after evaluating the experiment.

A folder named `for_copying` contains a reformatted version of the participant level sub-task statistics file. It reorders the entries into the correct chronological sequence of sub-task completion and splits `subtask_statistics.csv` into separate single-row files per metric. These individual single-row files were then used to compile a global dataset, aggregating all recorded data across participants per sub-task and metric.

Global Metrics

The four `.csv` stimulus files acted as the base for calculating the statistics regarding saccades, fixations, and AOI hits across the whole experiment. Three separate folders were created, in which all relevant extracted data was stored in.

A global fixations `.csv` contained a list of all fixations throughout the experiment with their X, Y screen coordinates and their respecitve durations. Similarly, a global saccade `.csv` stored all saccades. The start and end coordinates were determined using the preceding and subsequent fixation coordinates, and together with the resulting Euclidian distance the speed of the saccade was determined. Every fixation and saccade contained stimulus and participant data, enabling the calculation of normalized counts.

Apart from the statistics, the saccade and fixation data acts as the base for 2D visualizations. Heat maps are the most prominent ways to display fixation hot spots and durations (Raschke et al., 2013). For exploratory purposes linked to RQ3, four additional types of heat maps were generated. Differences were calculated using a log-cubic weighting function that preserved the directional sign of the values.

- Sub-task-level difference heat maps

- aggregate heat maps per stimulus

- stimulus-level difference heat maps

Similarly, scan paths are the easiest way to display saccadic eye movement (Raschke et al., 2013). The inclusion of multiple metrics as well as participant data in the global `.csv` data frames allows for additional visualizations. Analogue to the fixations, four different types of scan paths were created:

- sub-task-level disambiguation between map designs

- stimulus-level disambiguation between map designs

- sub-task-level disambiguation between experts and non-experts

- stimulus-level disambiguation between experts and non-experts

The AOI metrics were extracted using the binary AOI hit columns of the stimuli `.csv` files. From there average times to first hit, absolute number of hits, as well as total dwell time were extracted. The dwell time is calculated as a sum of all of the fixation durations for a given AOI or area (Burian et al., 2018).

3.4 Statistical Analysis

Since eye-tracking data typically do not follow a normal distribution, non-parametric statistical tests were employed. Specifically, Wilcoxon signed-rank tests and Aligned Rank Transform (ART) ANOVAs were used as robust alternatives to the paired t-test and standard ANOVA, respectively (Burian et al., 2018). All analyses were performed at a 95% confidence level. Participant expertise, task, and map design were treated as independent variables. The dependent variables included task completion time, time to first AOI hit, fixation metrics such as duration and count, and saccade metrics such as length and speed.

To assess overall variance and potential interactions between independent variables, ART ANOVAs were conducted. In addition, Wilcoxon signed-rank tests were used to examine pairwise differences between map designs, map extents, and expertise groups. Analyses were carried out at three levels: the sub-task level, the stimulus level, and the overall level. All statistically significant results are clearly marked in the corresponding tables and visualizations.

The measures for the 2D heat maps are calculated using the Shannon entropy and a weighted spatial dispersion. Entropy can be used to describe the switching between AOIs. (Krejtz et al., 2014) implement the Shannon entropy to denote the randomness. A higher entropy corresponds to more randomness and more frequent switching between AOIs.

$$H = -\sum_{i=1}^{n} p_i \log_2(p_i), \quad \text{with} \quad p_i = \frac{x_i}{\sum_{j=1}^{n} x_j} \tag{3.1}$$

Where: p_i is the normalized value at bin i, x_i is the raw fixation count at bin i, and n is the total number of bins.

The weighted spatial dispersion is calculated as follows:

$$I = \frac{\sum_{i=1}^{n} d_i^2 \, x_i}{\sum_{i=1}^{n} x_i} \tag{3.2}$$

Where: d_i is the Euclidean distance from bin i to the center of the heatmap, x_i is the fixation weight at bin i, and the center (x_c, y_c) is given by:

$$x_c = \frac{W}{2}, \quad y_c = \frac{H}{2}$$

where W and H are the width and height of the heatmap.

Chapter 4

Results

The results portrayed in this chapter were created using the methods laid out in the previous Chapter 3. The first part of the results gives an overview over various metrics across the two map designs and expertise in regards to the first two research questions, while the second part gathers exploratory results aimed at the third research question.

- The first section (4.1) looks at the answers given by the participants during the eye-tracking experiment.

- The second section (4.2) dives into the different results of the eye-tracking metrics discussed in sub-section 3.3.3.

- The last section (4.3) portrays the different 2D visualizations of the eye-tracking data. These help in the qualitative analysis answering RQ3.

Shapiro–Wilk tests confirmed that none of the dependent variables followed a normal distribution as laid out in Section 3.4: Completion Time ($W = 0.722$), Answer Time ($W = 0.542$), Average Time to AOI ($W = 0.547$), Fixations ($W = 0.728$), Fixations per Second ($W = 0.931$), Average Fixation Duration ($W = 0.703$), Saccades ($W = 0.706$), Saccades per Second ($W = 0.920$), Average Saccade Length ($W = 0.935$), Average Saccade Duration ($w = 0.816$), and Average Saccade Speed ($W = 0.970$). Accordingly, non-parametric methods were correctly applied for all statistical analyses.

4.1 Multiple Choice Results
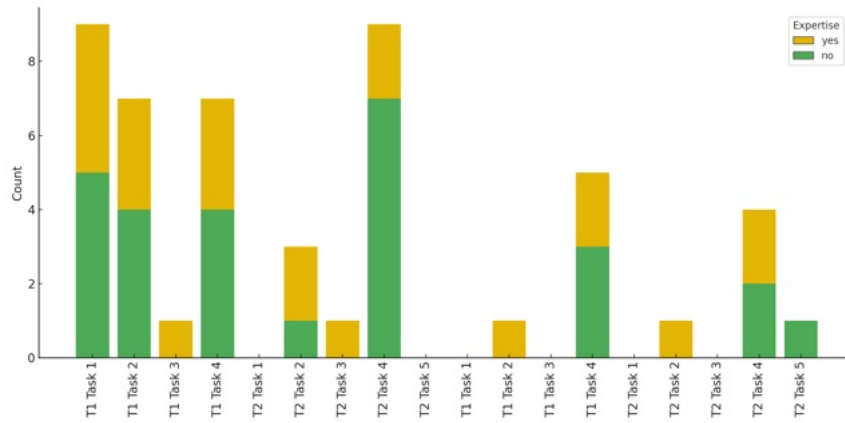
Figure 4.1: Mistake Count colored by Expertise
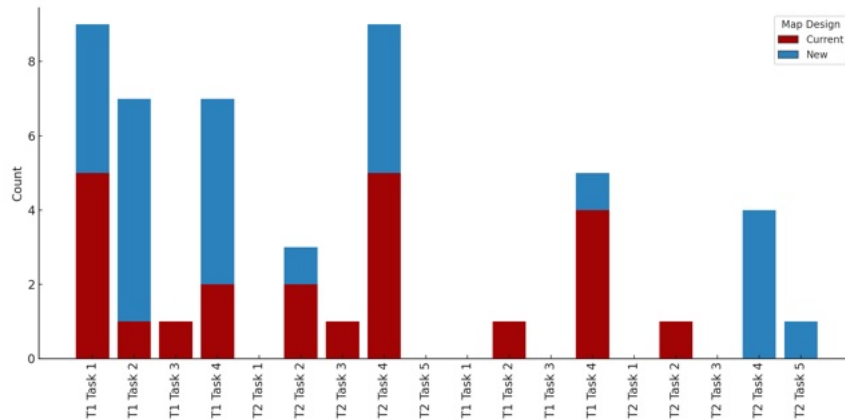
Figure 4.2: Mistake Count colored by Map Design

Across all participants and sub-tasks, 91.2% of answers were correct (509/558), with experts performing just slightly better at 92.7% (266/288). The updated map design scored 90.6% correct answers overall (253/279), 93.1% for experts (134/144) (see Figures 4.1 and 4.2).

Answer times did not differ significantly by expertise, $p = .068$, nevertheless, there is a significant interaction between Expertise and Map Design $p = 0.016$ (see Table A.1, Figure 4.3). Non-expert participants' answer time was 1.57 seconds lower on the updated map design, while expert times hovered around 3 seconds across both designs. The average answer time was 3.3 seconds, with experts responding faster on average (3.0 seconds).
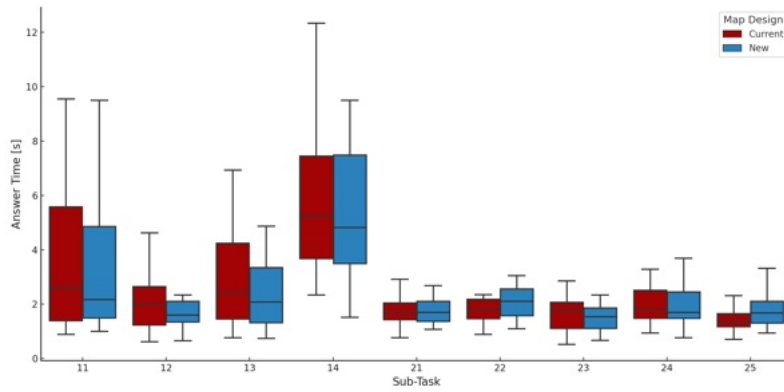
Figure 4.3: Answer Times (Seconds) per Sub-Task across Map Designs
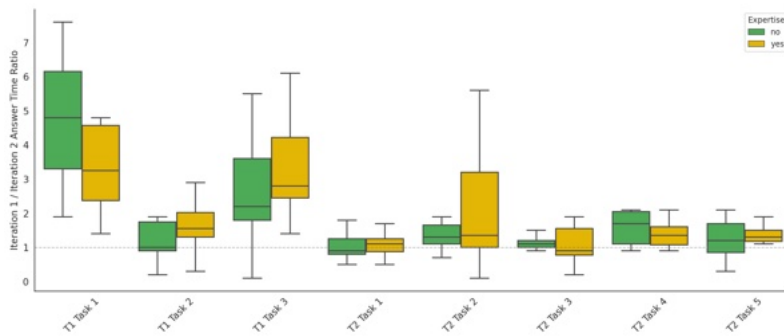
Figure 4.4: Answer Time Ratio ($\frac{AnswerTime_{Iteration1}}{AnswerTime_{Iteration2}}$) per Sub-Task, colored by Expertise (Grisons Task 4 omitted)

4.2 Eye-Tracking Metric Results

This section looks into the results across Task 1 (Grisons) and Task 2 (Romandie) combined, focusing on different metric results gathered throughout the eye-tracking experiment as outlined in Subsection 3.3.2. It aims to provide insight into the first two research questions on map reading efficiency and a bias toward the current map design.

4.2.1 Completion Time

Table 4.1: Linear Regression Predicting Total Completion Time

Predictor	Coef.	SE	t	p-value
Intercept	83.78	12.12	6.91	**<.001**
Grisons Knowledge	-4.63	2.95	-1.57	.129
Romandie Knowledge	1.97	3.27	0.60	.552
Map Use Frequency	-6.09	2.75	-2.22	**.036**
Expertise (binary)	-18.88	6.02	-3.14	**.004**
Model Fit: $R^2 = 0.489$, Adjusted $R^2 = 0.411$, $F(4, 26) = 6.23$, $p = .001$				

Evaluating the responses from the pre-study questionnaire (see Section 3.3) regarding participants' frequency of map use and geographic knowledge of the two regions, a linear regression predicting total task completion time revealed a significant effect of expertise

($p = .004$) and map use frequency ($p = .036$). While self-reported geographic knowledge of the Grisons region showed an expected negative association with completion time, this effect did not reach statistical significance ($p = .129$). In contrast, knowledge of the Romandie region showed no meaningful association with performance ($p = .552$).
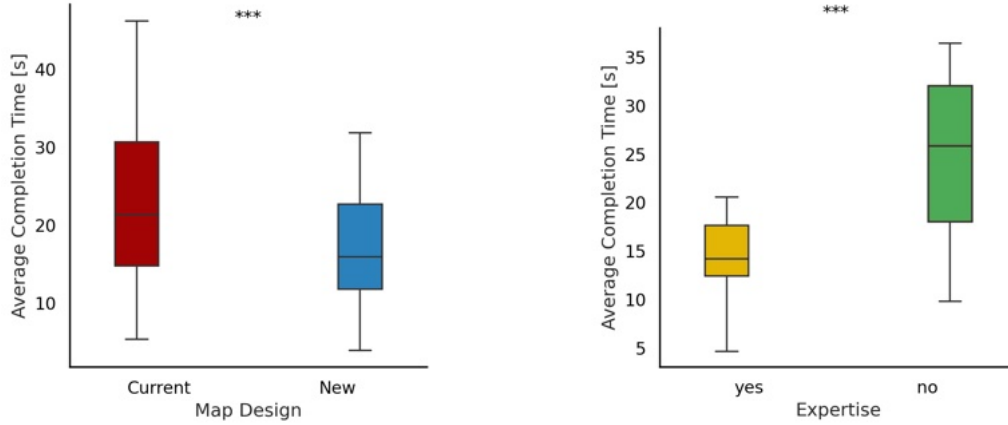
(a) Average Sub-Task Completion Time across Map Designs

(b) Average Sub-Task Completion Time by Expertise

Figure 4.5: Overall Average Sub-Task Completion Time, grouped by (a) Map Design and (b) Expertise Level.

Two separate 2×2×2 Aligned Rank Transform ANOVA (Tables 4.2 and 4.4) revealed that the map design had a highly significant effect on average sub-task completion time, $F(1, 504) = 20.03$, $p < .001$, as well as average total completion time, $F(1, 112) = 9.47$, $p = .003$, which combines all sub-tasks of a given stimulus. In addition, expertise significantly influenced task completion time, $F(1, 504) = 52.73$, $p < .001$ (Table 4.2).

There was no significant interaction between Task and Design, $F(8, 504) = 1.54$, $p = .139$, indicating that the benefit of the new map design on completion time remained relatively consistent across different tasks (Figure 4.7).

However, there was a significant interaction between Expertise and Design, $F(1, 504) = 4.23$, $p = .040$, indicating that the effectiveness of the design differed between experts and non-experts (Figure 4.8). No higher-order interactions reached significance.

Table 4.2: ART ANOVA on Average Total Completion Time.

Effect	df	Res. df	F	p-value
Stimulus	1	112	14.25	**<.001**
Design	1	112	9.47	**.003**
Expert	1	112	24.85	**<.001**
Stimulus:Design	1	112	1.82	.180
Stimulus:Expert	1	112	1.70	.195
Design:Expert	1	112	1.55	.216
Stimulus:Design:Expert	1	112	0.54	.465

A Wilcoxon signed-rank test (Table 4.3) showed that average total completion time was significantly lower with the new map design in both Grisons, $W = 98$, $p = .005$, and Romandie, $W = 126$, $p = .028$, in line with the findings of the Aligned Rank Transform ANOVA (Table 4.2). Figure 4.6 visually shows the large standard deviation ($SD = 72.06$) found across completion times of tasks on the Grisons extent using the current map design.

The effect of map design on completion time remained significant when analyzing individual sub-tasks. In the Grisons region, completion time was significantly lower with the new design, $p = .021$, as well as in Romandie, $p = .014$. Additionally, participants completed sub-tasks in the Romandie region significantly faster than those in Grisons (see Figures 4.7 and 4.8).

Table 4.3: Wilcoxon signed-rank test comparing Average Total Completion Time between Current and New map designs per region.

Region	Design	Mean (s)	SD (s)	W	p-value
Grisons	Current	122.75	72.06	98	**.005**
	New	84.75	42.55		
Romandie	Current	81.42	42.72	126	**.028**
	New	67.04	41.49		
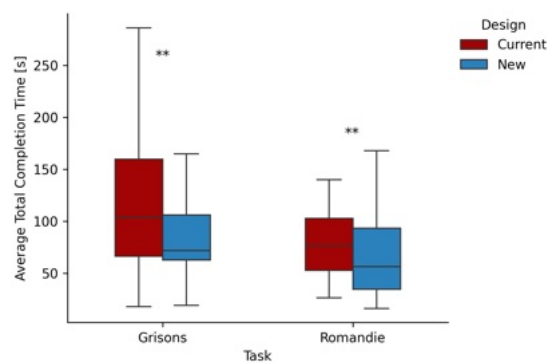
Figure 4.6: Average Total Completion Time per Map Design, grouped by Stimulus

Table 4.4: Aligned Rank Transform ANOVA on Average Completion Time per Sub-Task.

Effect	df	Res. df	F	p-value
Task	8	504	14.16	**<.001**
Design	1	504	20.03	**<.001**
Expert	1	504	52.73	**<.001**
Task:Design	8	504	1.54	.139
Task:Expert	8	504	2.76	**.005**
Design:Expert	1	504	4.23	**.040**
Task:Design:Expert	8	504	0.94	.487

Table 4.5: Wilcoxon signed-rank test comparing Completion Time across Map Designs.

Region	Design	Mean (s)	SD (s)	W	p-value	Effect Size (r)
Overall	Current	22.69	23.94	3454.0	**<.001**	0.338
	New	16.87	16.22			
Grisons	Current	30.69	30.47	2748.0	**.021**	0.211
	New	21.19	18.23			
Romandie	Current	16.28	14.16	4354.5	**.014**	0.200
	New	13.41	13.50			
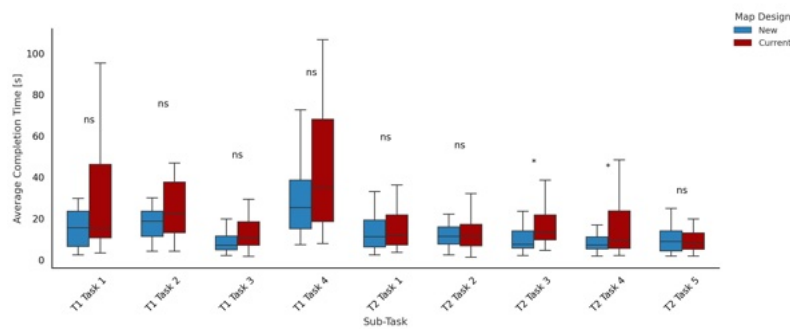
Figure 4.7: Average Completion Time (in Seconds) per Sub-Task across Map Designs
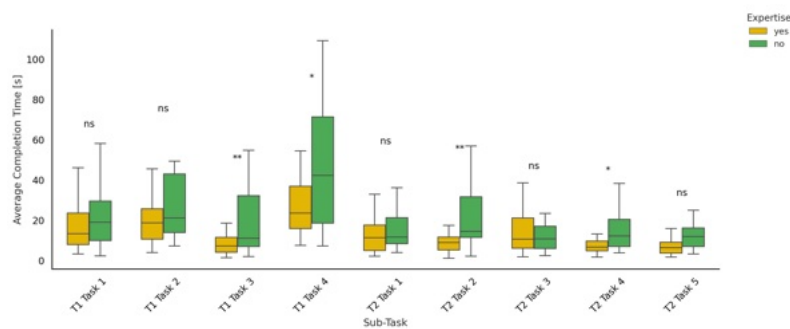
Figure 4.8: Average Completion Time (in Seconds) per Sub-Task by Expertise

4.2.2   Fixations

Beyond visual search completion time, fixation-based metrics were analyzed to assess visual processing demands. A 2×2×2 Aligned Rank Transform ANOVA on average fixation duration (Table 4.6) once again revealed a significant main effect of Task, $F(8, 504) = 2.51$, $p = .011$, as well as an expected significant effect of Expertise, $F(1, 504) = 38.53$, $p < .001$. No significant main effect of Design was found, $F(1, 504) = 0.05$, $p = .828$.

Table 4.6: Aligned Rank Transform ANOVA on Average Fixation Duration.

Effect	df	Res. df	F	p-value
Task	8	504	2.51	**.011**
Design	1	504	0.05	.828
Expert	1	504	38.53	**<.001**
Task:Design	8	504	2.95	**.003**
Task:Expert	8	504	0.35	.945
Design:Expert	1	504	0.08	.779
Task:Design:Expert	8	504	0.48	.870

There were no notable visual differences in fixation duration between the Grisons and Romandie regions, with most average values ranging between 250 and 350 milliseconds (see Figures 4.9 and 4.10). On average, fixation duration was 321 milliseconds for experts and 283 milliseconds for non-experts throughout the experiment.

Figure 4.11 shows the first sixty fixation durations for every participant of every sub-task. From there, the global temporal average of fixation durations could be taken. It starts at just over 250 milliseconds, increases slightly throughout the next 10 fixations, and hovers at around 300 milliseconds throughout the rest of the sub-task.
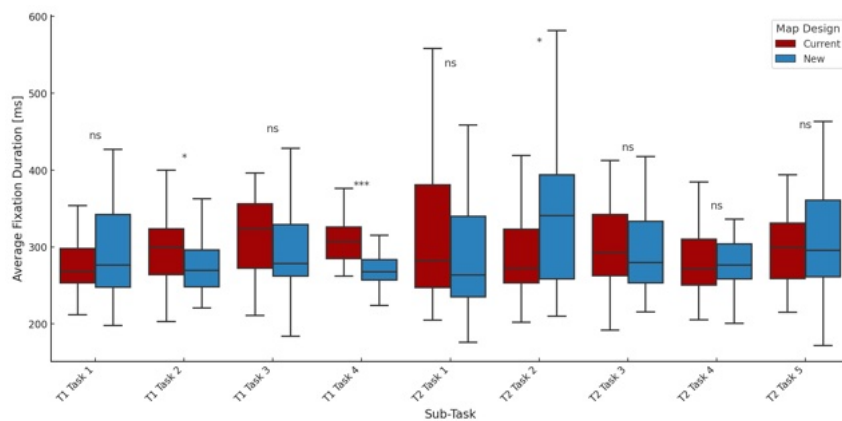
Figure 4.9: Fixation Duration per Sub-Task across Map Designs.

In terms of fixation frequency (fixations per second), there is no significant difference between the two map designs. A Wilcoxon signed-rank test yielded no significant difference overall, $W = 4724.0$, $p = .262$, $r = 0.092$. Yet, a significant difference was
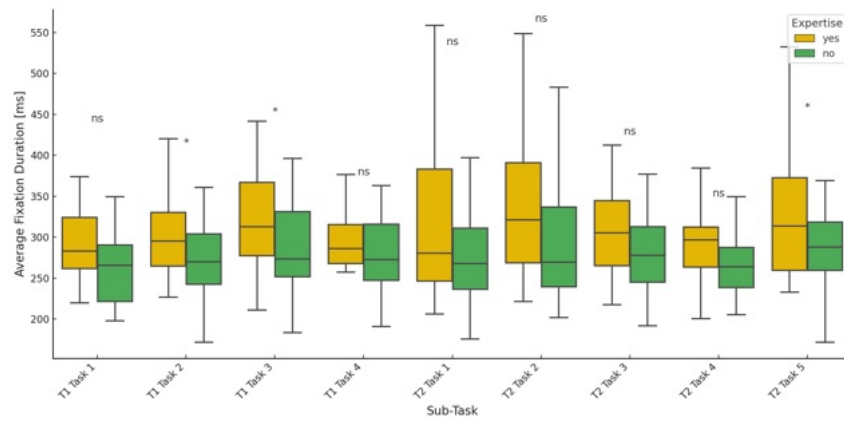
Figure 4.10: Fixation Duration per Sub-Task by Expertise.
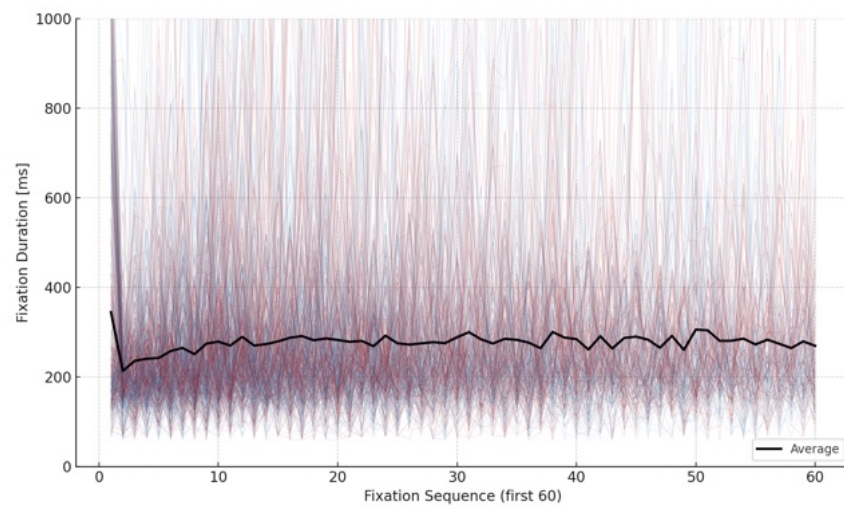
Figure 4.11: Fixation Durations throughout the first 60 fixations for every participant and sub-task. The average is highlighted as a black line.

found in the Grisons region, $W = 1698.0$, $p < .001$, $r = 0.369$. No significant difference was observed for Romandie, $W = 4612.0$, $p = .344$, $r = 0.077$ (Table 4.7). The individual frequencies per sub-task vary between 2.7 and 3.5 fixations per second (see Figures 4.12 and 4.13).

Table 4.7: Wilcoxon signed-rank test comparing Fixations per Second across map designs.

Region	Design	Mean	SD	W	p-value	Effect Size (r)
Overall	Current	3.04	0.46	4724.0	.262	0.092
	New	3.07	0.55			
Grisons	Current	2.98	0.53	1698.0	**<.001**	0.369
	New	3.19	0.59			
Romandie	Current	3.13	0.55	4612.0	.344	0.077
	New	3.02	0.76			
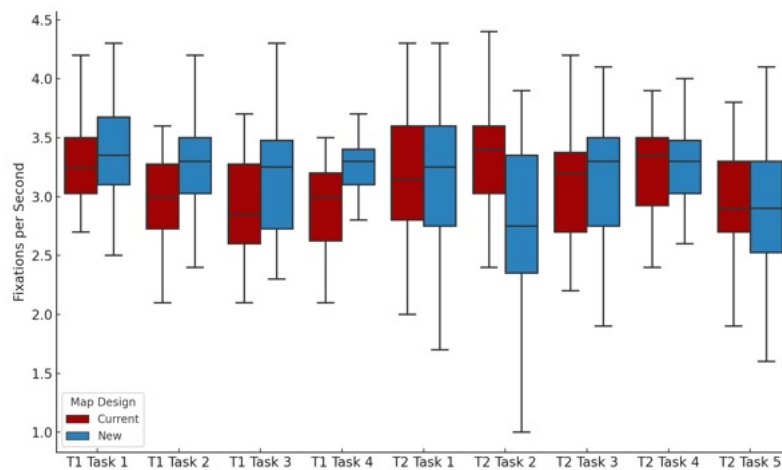
Figure 4.12: Average Fixation Duration per Sub-Task across Map Designs

4.2.3 Saccades

Alongside fixations, saccades represent the second core category of eye movements. A series of Wilcoxon signed-rank tests (Table 4.8) revealed significant differences between the current and updated map designs for saccade length, $p = .004$, as well as saccade speed, $p = .002$. In contrast, saccade duration did not differ significantly between designs, $p = .101$.

Table 4.8: Wilcoxon signed-rank tests comparing Saccade metrics between Current and New map designs.

Metric	Design	Mean	SD	W	p-value
Saccade Length	Current	136.95	51.31	14586	**.004**
	New	150.39	57.00		
Saccade Speed	Current	3.96	0.94	13053	**.002**
	New	4.15	0.95		
Saccade Duration	Current	31.32	6.45	15693	.101
	New	32.40	8.58		

A 2×2×2 ART ANOVA on saccade frequency (saccades per second) (Table 4.9) showed no main effect of map design, $F(1, 504) = 0.00$, $p = .971$, indicating a consistent number
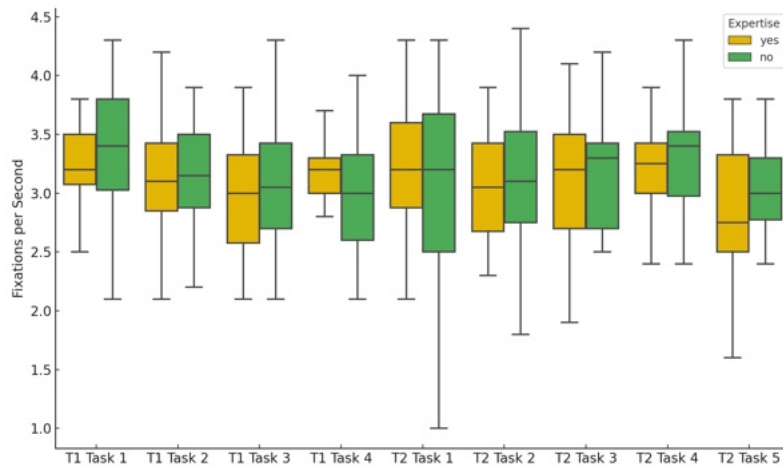
Figure 4.13: Fixation Frequency per Sub-Task and Expertise

of saccades across conditions (*Current* = 3.57 per second, *New* = 3.56 per second). Similar to fixations, significant main effects were found for Expertise, $F(1, 504) = 13.06$, $p < .001$, Task, $F(8, 504) = 4.02$, $p < .001$, and its interaction with map Design, $F(8, 504) = 2.00$, $p = .045$.

Table 4.9: Aligned Rank Transform ANOVA on Saccades per Second.

Effect	df	Res. df	F	p-value
Task	8	504	4.02	**<.001**
Design	1	504	0.00	.971
Expert	1	504	13.06	**<.001**
Task:Design	8	504	2.00	**.045**
Task:Expert	8	504	0.79	.611
Design:Expert	1	504	1.80	.180
Task:Design:Expert	8	504	0.52	.839

A series of 2 ×2× 2 ART ANOVAs on all other saccade metrics revealed consistent patterns (see Tables A.2, A.3, and A.4). The effect of *Task* yielded highly significant effects across all metrics, $p < .001$, and *Expertise* equally showed significant effects on all outcomes except saccade length, $p = .135$.

The accumulated saccade lengths allow for a better understanding of the intensity of the visual search on a stimulus, since lengthy scan paths indicate less efficient search behavior (Kotval and Goldberg, 1998). Figure 4.14 shows the average total saccade length per sub-task across all participants, while Figure 4.15 shows it for experts only. In both instances, only Task 4 of Romandie returned statistically significant values, with $p < .001$ for all participants, and $p = 0.025$ for experts only.
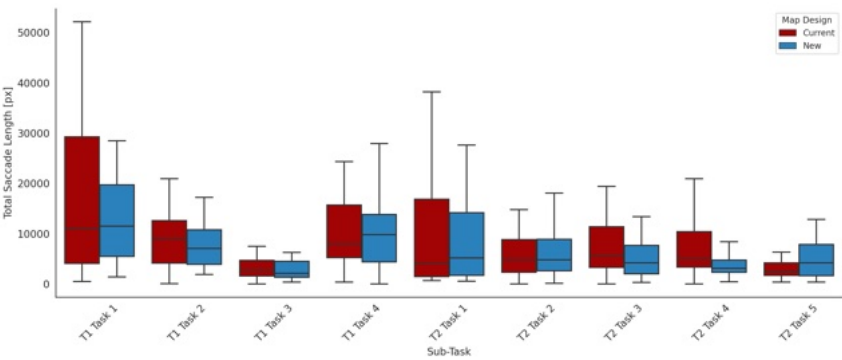
Figure 4.14: Average Total Saccade Length per Sub-Task across Map Designs
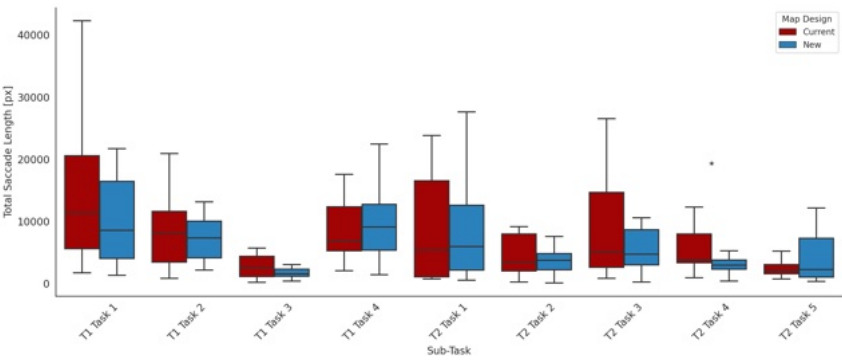
Figure 4.15: Average Total Saccade Length of Expert Participants per Sub-Task across Map Designs

4.2.4 Relation between Saccade Length and Fixation Duration by Map Design

The relationship between saccade length and fixation duration, as seen in Netzel et al. (2016), visualizes systematic differences in visual search behavior as a function of map design (Figure 4.16) and expertise (Figure 4.17). The current map design shows fixation durations that were slightly shorter and saccade lengths that remained largely comparable to those observed in the new design. In contrast, the new map design elicited longer fixation durations, although this did not translate into significantly greater saccade amplitudes.

Across both designs, experts consistently exhibited longer saccades and prolonged fixation durations, indicative of a more efficient and confident search strategy.
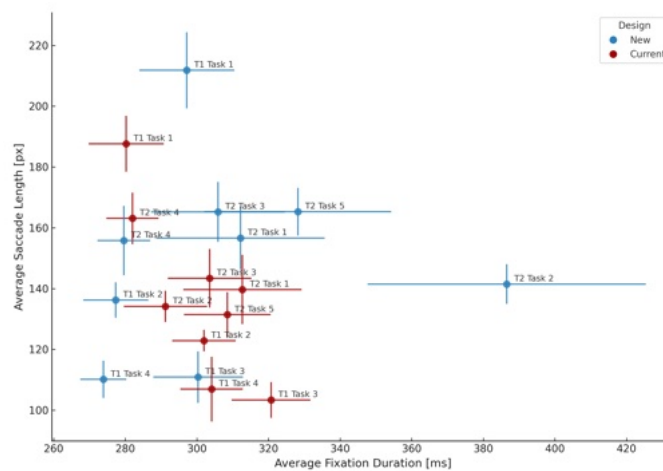
Figure 4.16: Saccade Length vs Fixation Duration Diagram.
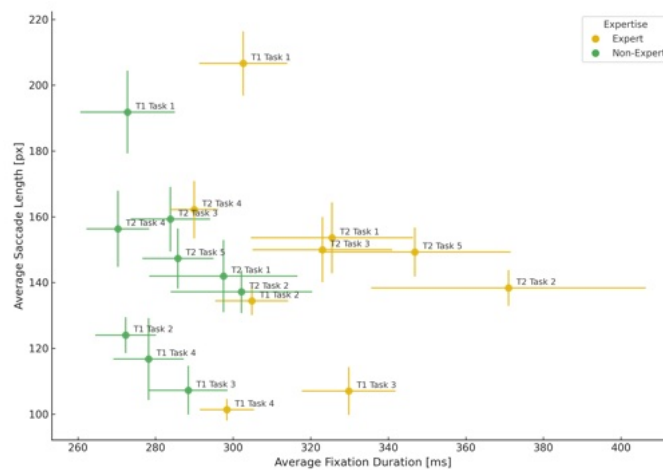
Figure 4.17: Relation between Saccade Length and Fixation Duration by Participant Expertise

4.2.5 AOI Hit Times

Same as with the completion time, the time to first AOI hit significantly varies between design, task and expertise, as can be seen with an ART ANOVA conducted on Average AOI hit times (Table 4.10). Similarly, the average time to first AOI hit across all sub-tasks was notably lower in the Romandie sub-tasks (mean = 3.63 seconds) compared to those in the Grisons region (mean = 8.26 seconds). As anticipated, expert participants reached AOIs more quickly than non-experts across the board. A summarized overview of the AOI hit times including the statistical measures can be found under Table 4.11, combining all AOI, as well as Figures 4.18 and 4.19 visualizing the hit times per AOI seperately.

Table 4.10: Aligned Rank Transform ANOVA on Time to First AOI Hit (in seconds)

Effect	df	F	p-value
Stimulus	1, 959.71	79.34	**<.001**
Design	1, 960.94	8.24	**.004**
Expertise	1, 28.57	10.74	**.003**
Stimulus:Design	1, 965.92	2.70	.101
Stimulus:Expertise	1, 960.41	7.72	**.006**
Design:Expertise	1, 960.88	3.18	.075
Stimulus:Design:Expertise	1, 965.75	0.01	.927

Table 4.11: Wilcoxon signed-rank test comparing time to first AOI hit between map designs, grouped by expertise

Group	U statistic	p-value
Non-Experts (Current vs. New)	30089.0	**.004**
Experts (Current vs. New)	38013.5	.146
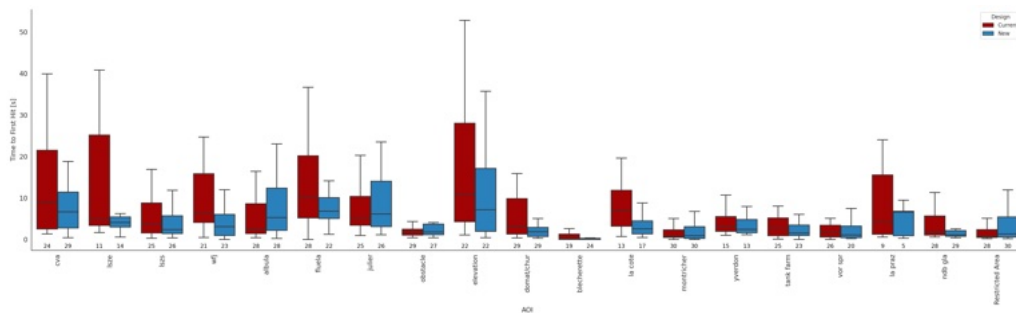
Figure 4.18: Time to first Hit for all AOI by Map Design
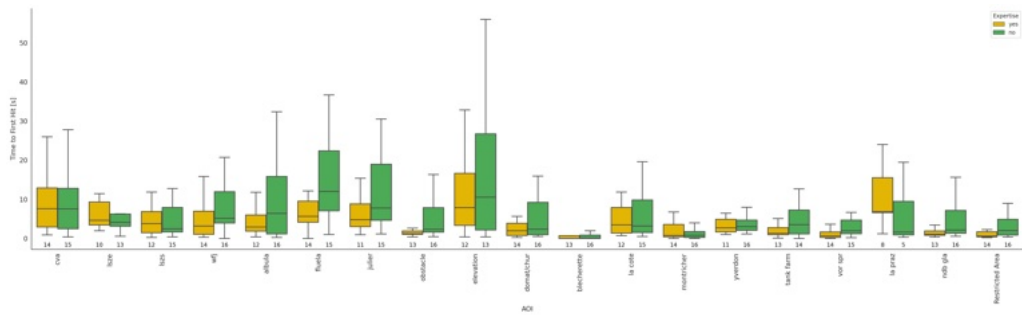
Figure 4.19: Time to first Hit for all AOI by Expertise
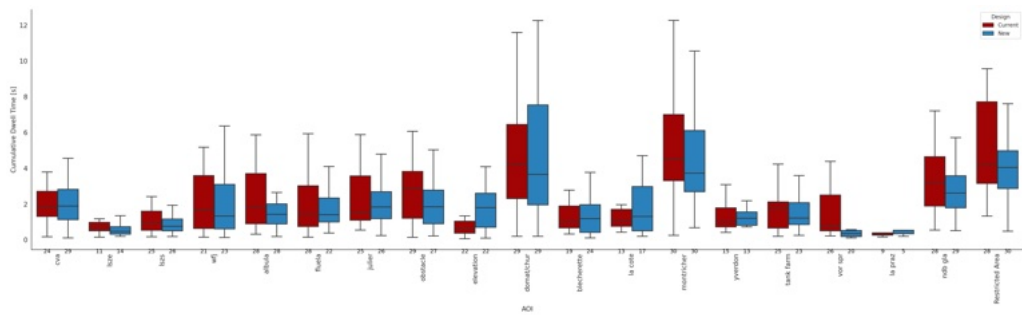
Figure 4.20: Average Cumulative Dwell Time per AOI by Design
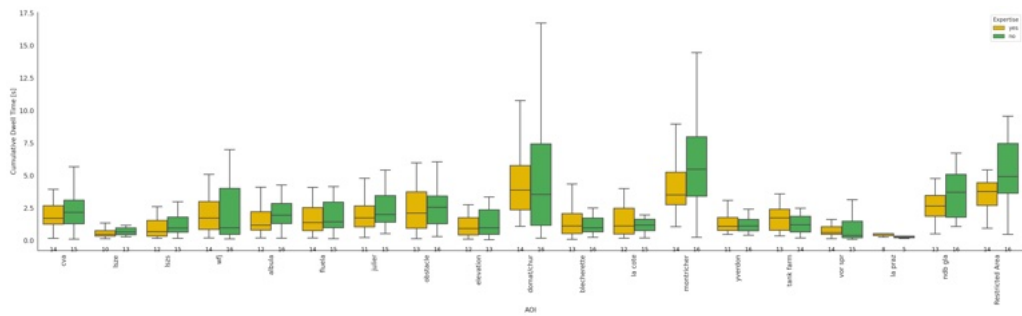
Figure 4.21: Average Cumulative Dwell Time per AOI by Expertise

4.3   Spatial Visualizations of Eye-Tracking Metrics

4.3.1   2D Fixation Heat Maps

Using the methods described in Section 3.4, 2D fixation heat maps show the spatial distribution of fixation points. The metrics of the different 2D fixation heat maps are listed in Table A.7.

Aggregate heat maps show the cumulative distribution of all fixation points throughout the sub-task for a given stimuli. Figure 4.22 shows the total fixations for the current map extent of Grisons and fixations on the updated map design of Romandie.

The two difference heat maps of the Grisons and Romandie extents can be seen in Figure 4.23.
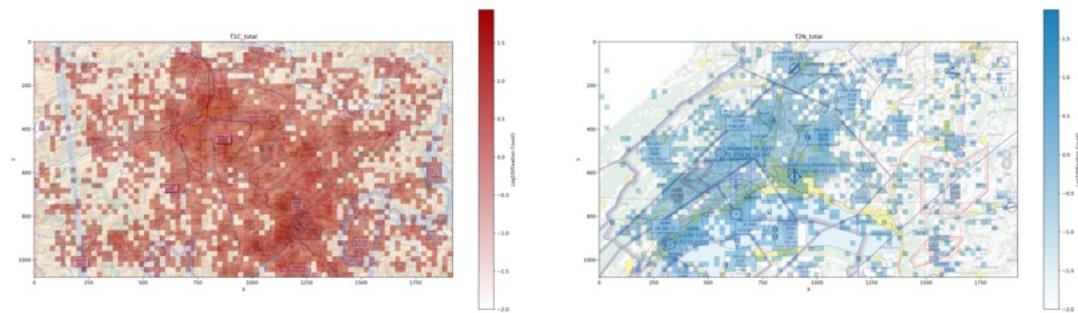
Figure 4.22: 2D Cumulative Fixation Heat Map of Grisons (Current Map Design) and Romandie (Updated Map Design)
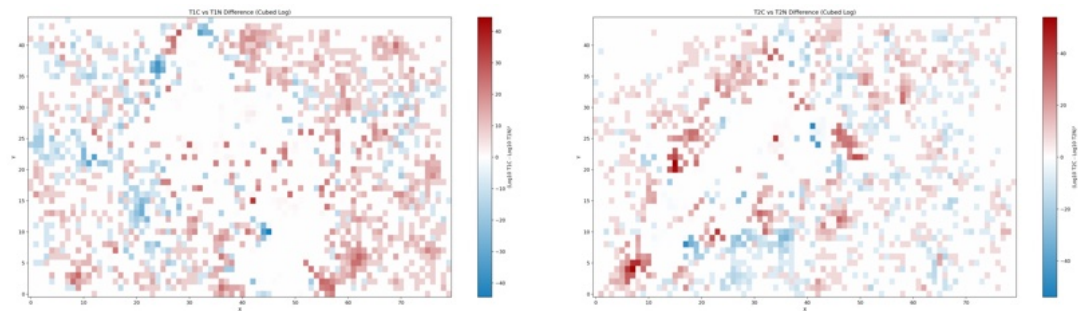
Figure 4.23: 2D Difference Heat Map of Fixations for Grisons and Romandie. Red areas were more looked at on the current map design, while blue ones were more looked at on the updated map design.

4.3.2   2D Saccade Scan Paths

Saccadic scan paths are the visual counterpart to the 2D fixation heat maps. In the figures below, saccades collected throughout a given sub-task collected and colored by map design for easier interpretation. The corresponding stimulus is lightly visible in the background. Not all tasks are included in this section.

Across all tasks, scan paths tend to cluster around prominent textual elements and labels, intuitively suggesting that these features attract early and frequent visual attention. This pattern is to be seen across all sub-tasks and map designs (see Figure 4.24).

Sub-tasks 2 and 4 of each stimulus give the best insight into saccadic eye movement due to the nature of the task. They are to be seen in Figure 4.25. Task 2 in Grisons shows the three mountain passes as hot spots, while Task 2 in Romandie clearly shows the labels for Geneva's TMA 5 airspace as hot spots.

(a) Grisons Current

(b) Grisons New

(c) Romandie Current

(d) Grisons New
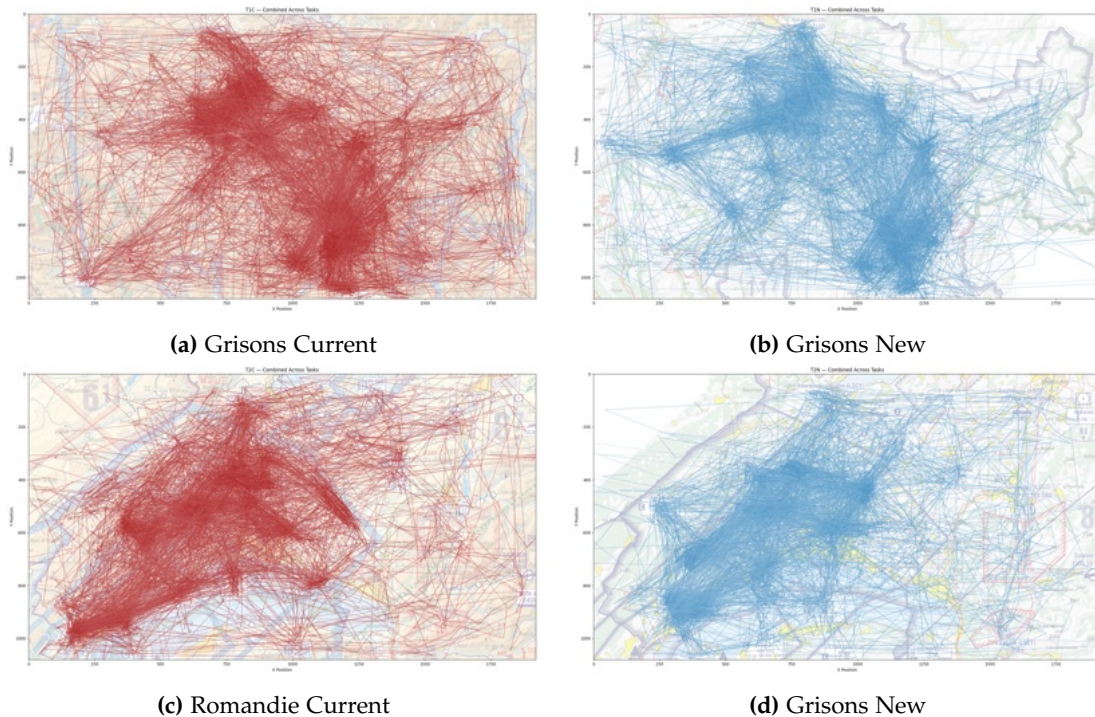
Figure 4.24: Cumulative Scan Paths across all four Stimuli

(a) Grisons Current Task 2

(b) Grisons New Task 2

(c) Grisons Current Task 4

(d) Grisons New Task 4

(e) Romandie Current Task 2

(f) Romandie New Task 2

(g) Romandie Current Task 4

(h) Romandie New Task 4
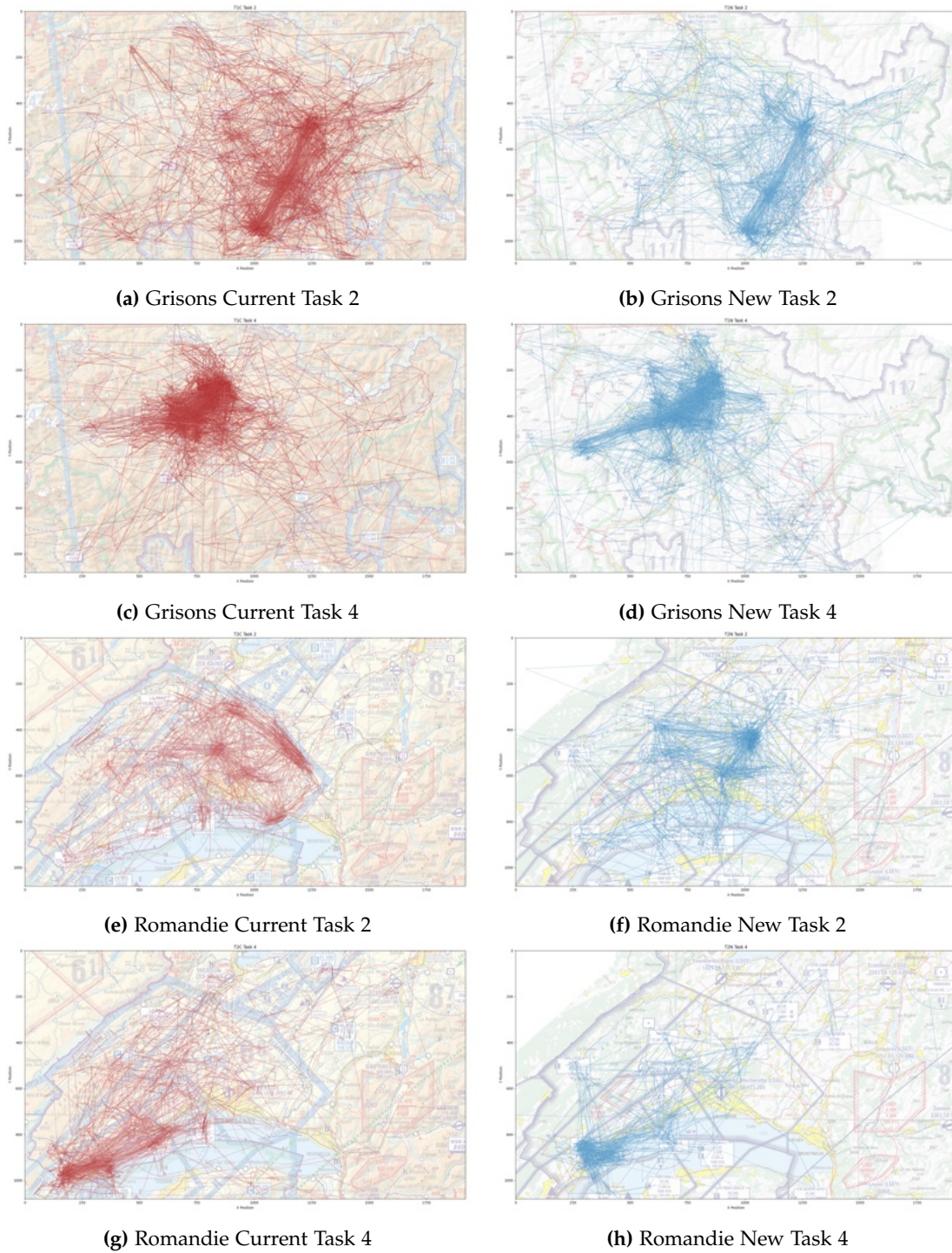
Figure 4.25: Scan paths across Task 2 and Task 4 for both regions and map designs

Chapter 5

Discussion

5.1 Task Questions

Firstly, looking into the correctness of the answered questions throughout the eye-tracking experiment, it becomes clear that most mistakes were made while exposed to the two first stimuli. This is indicative of a difficult start, as mistake counts rapidly declined after that. Visual priming is the most probable cause for these numbers, and makes sub-task performance highly individual. The answer time ratios are indicative of fast learning on the Grisons map extent.

The tasks for the Grisons area were slightly more difficult, with more mistakes happening while exposed to these stimuli. This can be tied to the nature of these tasks, as explained in Section 3.3). Experts, as expected, performed better than non-experts. Yet, the difference between correct answer percentages is surprisingly small, indicative of large discrepancies within the participant groups. Interestingly enough, similarly good results among novice participants were attained by Keskin et al. (2020) in their study. The highest individual mistake count for an expert was 4 (77% correct).

A central limitation of this study lies in the nature of the tasks themselves, which consistently yielded statistically significant effects across nearly all statistical tests. The Grisons tasks tended to be broader in scope, counterbalanced by visually less cluttered stimuli. Because flight planning constitutes a major part of pre-flight procedures, three of the four tasks were designed to reflect questions arising from such activities. These involved scanning larger spatial areas, which contributed to longer sub-task durations. Especially task 4 in Grisons stood out for its lack of a predefined target symbol or value, thus, participants had to make use of different visual cues. It was this particular task that let completion time shoot up across both expert and non-expert groups.

In contrast, the Romandie tasks were generally more straightforward, largely due to the higher complexity and density of the map, which allowed for more targeted questions. Answering times were significantly shorter in this region, averaging 2.8 seconds compared to over 4 seconds for the Grisons tasks. An exception was Romandie Task 2, where some expert participants mistakenly identified the airspace class instead of the airspace identifier. This resulted in prolonged dwell times on the answer screen.

5.2 General Eye-Tracking Results

5.2.1 Completion Time

The empirical results showed that completion times were generally shorter across the experiment when using the updated map design. This suggests that the redesign improved task efficiency. Still, the ART ANOVA did not reveal a significant interaction between expertise and design for this metric, indicating that the effect of the updated map design was consistent across both expert and non-expert groups.

Despite the absence of a statistical interaction, the descriptive results suggest that non-expert participants benefited more from the updated design. This is likely because experts, who already performed tasks more quickly, had less room for improvement. While expert participants also completed tasks slightly faster with the updated map, the relative gain was smaller.

Importantly, the ART ANOVA showed no indication of a bias toward the current map design for RQ 2 (see 1.3). There were no significant interactions between expertise and design across any of the evaluated metrics, except for average completion time, where the updated design actually performed significantly better. These results discard the idea of a strong bias towards the current map as seen with the Master's thesis of Juliette Marx (Marx, 2015) and reinforce the conclusion that performance gains outweighed participant familiarity with the current map design.

5.2.2 Saccades and Fixations

The collected fixation and saccade duration ranges are consistent with established findings in the literature (Holmqvist et al., 2011; Rayner, 2009). While these studies suggest a negative correlation between fixation duration and task difficulty, this relationship could not be recreated in the present study. Also, Task difficulty remains a subjective construct that varies between participants and was not taken into account in this eye-tracking experiment. Nevertheless, some patterns are visible, particularly in relation to the updated map design. Tasks from the Romandie region returned longer average fixation durations than those from the Grisons region, which would align with the higher error

rates recorded in the multiple-choice responses associated with Grisons tasks.

However, the findings of this study diverge from those reported by Ooms et al. (2013); Keskin et al. (2020); Netzel et al. (2016), who argue that novice map reading is typically characterized by longer fixation durations and a higher number of fixations per second. The present results of this experiment suggest the opposite. The recreated figure from Netzel et al. (2016) (see Figure 4.16) illustrates this discrepancy. Sub-tasks completed using the current map design are positioned in the lower-left quadrant, which is indicative of more efficient visual search behavior based on the findings of Netzel et al. (2016). Moreover, the influence of expertise in general may have been limited in this study, as none of the participants who owns a pilot license is based in Western Switzerland. For most, the Romandie region was unfamiliar, especially when tasked to find La Côte airfield.

Saccade and fixation frequencies remained stable throughout the experiment for both groups, indicative of consistent visual search performances. A decline in fixation rate over completion time, as suggested by McGregor and Stern (1996), was not observed. Fixation durations remained constant, although longer and more difficult tasks potentially have shown a similar effect. Due to faster completion times of experts, fewer overall fixations and saccades were needed, which aligns with the map use paradigm described by Keskin et al. (2020). As shown in Figure B.4, saccade speed may be lower in the Grisons region due to the larger map extent, and thus a smaller area dealt with throughout a sub-task.

5.2.3 Areas of Interest

Both time to first AOI hit and overall completion times were significantly faster with the updated map design. While the Vaud region presented higher visual complexity, the updated design still appeared to consistently score lower hit times. Similarly, non-expert participants benefited more from the updated map design, with significance levels at $p = .004$ when comparing AOI hit times between designs, as opposed to the insignificant $p = .146$ for experts.

Expertise also played a role in gathering the information in the Grisons tasks, as there was a highly significant interaction between *Task* and *Expertise*. The difference in hit times was closer in Romandie tasks, once again suggestive of the easier nature of the tasks. Additionally, dwell times on key AOIs were slightly longer, and saccades were faster during Romandie tasks, suggesting a more active visual search process. These findings align with Holmqvist et al. (2011).

When filtering by key AOIs, both the current and updated map designs yielded similar results overall. Peripheral AOIs were located more quickly in the updated design, likely

due to increased visual saliency of labels of the AOI, which also is visually reflected in the scan paths, discussed in Subsection 5.3.2.

Combining these eye-tracking metrics, we can answer RQ 1 (see 1.3). The significantly faster completion time for both participant groups shows less time required to confidently answer a given task, which is of great importance when flying. The faster completion time did not lower accuracy of the answers. Similarly to completion time, the AOI hits happened significantly earlier on the new map when looked at wholly per stimulus, indicative of a more salient design.

Most conventional eye-tracking metrics show no significant differences which correspond to improved readability. Especially the fixation durations and saccade length speak against improved legibility. Nonetheless, the improved AOI hit times most probably lead participants to an overall faster completion time in the end.

5.3   Qualitative Analysis

5.3.1   2D Fixation Heat Maps

The 2D heat maps are similar in nature, with most hot spots lying in central regions of the extent. Therefore, those central areas appear white on difference heat maps, since they are frequented a lot in both designs. In the peripheral regions, there is scattering of fixations between the two designs. It shall be noted that participants' gazes are not drawn to neighboring countries on the updated design, since the data are missing. Naturally, this effect is more prominent in the Grisons extent, as a large portion of the map is blank on the updated design.

Still, these heat maps also have drawbacks, and must be treated with caution. As Raschke et al. (2013) warns, the heat maps are an absolute representation of fixations, and do not reflect any temporal relationship between the hot spots. Furthermore, the different AOIs were not visited with equal frequency, resulting in some areas being underrepresented in heat map visualizations.

When looking at the difference heat maps on a sub-task level, qualitative analysis is rather difficult. Due to the overall smaller number of gazes per area when compared to a cumulative heat map, an individual fixation has more impact. Therefore, the heat maps show scattered gaze points originating from both map versions, which makes it difficult to draw clear conclusions.

Still, the statistical results for the heat maps show certain patterns. While the inertia is slightly higher for the new map design, the entropy scores lower. The higher inertia

corresponds with the increase in saccade lengths on the new map design, making attention more spread out over the stimulus.

5.3.2 Saccade Scan Paths

Saccade scan paths reveal that labels strongly attract visual attention due to their saliency. For instance, the minimum recommended altitude label for the VFR route to Julierpass is highly prominent on the current map in the Grisons region (see Figure B.5). However, this information is not particularly relevant for most navigational tasks.

Although the entropy values of the resulting heat maps are slightly lower for the updated design, the differences are not that big and should therefore be interpreted with caution. Nonetheless, the visualization suggests that saccadic scan paths can be effectively guided through strategic label placement. This effect is particularly evident in the Romandie region, where the updated map design features ICAO-standardized labels. A major drawback of the current map is its fragmented labeling of airspaces, which can be especially distracting in visually cluttered areas. In such contexts, standardized, salient labels appear to play a central role in guiding gaze behavior and improving visual efficiency.

Task 4 in the Grisons extent is exemplary for a typical task when being confronted with an emergency situation, such as an engine failure. Since most elevations other than mountain tops or passes are not labeled, the altitude of such large areas, which can be used for emergency landings, remain unknown and must be derived from as many sources as possible. Naturally, the heliport in Untervaz gives an exact number at 1786 feet, just 50 meters lower than the actual elevation of Chur (or Domat / Ems, depending on the iteration) of around 1950 feet.

For example, Tavanasa heliport in the upper Rhine valley gives second value which can be used to easily interpolate between the two. Since time is extremely limited, especially in emergency situations, fast information retrieval is vital. The fixation frequency of this task is significantly higher and completion time lower, yet not statistically significant on this particular task level. Nonetheless, the difference heat map of this task reveals a surplus in fixations of Tavanasa heliport, which is indicative of spotting the peripheral label better due to enhanced saliency.

In relation to RQ 3 (see 1.3), albeit not particularly pronounced, a few differences in visual patterns could be observed. These patterns suggest a more efficient and goal-oriented visual search. However, this interpretation should be approached with extreme caution, as the overall lower completion times also resulted in fewer saccades and fixations.

5.4   Limitations

As with any experimental study, several limitations must be acknowledged.

First, the redesigned map is not fully congruent with the official chart. Certain features, such as airspace class labels along airspace borders, were intentionally excluded due to a limited cost-benefit ratio. Their omission may have influenced especially expert participants' interpretation. Also, not all obstacle data been able to recreated, therefore less visual clutter especially in mountainous regions.

Additionally, some newly proposed features, such as dashed airspace boundaries or combined airspace labels, would first need separate testing and regulatory approval to be inserted on the chart. The two stimuli, however, did not feature any of those changes.

Another notable limitation involves the treatment of regions outside of Switzerland. As noted in Section 5.3, blank areas due to missing data have unintentionally guided gaze patterns, drawing attention more strongly to the central map content, and thus influencing heat map metrics.

Due to time constraints in place, a within-subject design had to be chosen for this study. With that come obvious drawbacks. Every participant is exposed to the same geographic extent twice. Although the tasks were posed that they were not identical, most participants still can recall the spatial relationships of AOI Bestgen et al. (2017), leading to improved strategies and lower completion times throughout the second iteration. This notably visible when looking at Figure 4.1 showing the number incorrect answers. Visual priming also occurred due to placement of the target symbol on the question slide before being exposed to the stimulus, since there is no space for legend next to map extent. Already knowing what to look for definitely helps non-expert participants and does not reflect a real-world scenario, where the legend is on the back of the chart.

With the fragmentation of the participants into smaller and smaller cohorts come more unreliable results due to the small sample size Fairbairn and Hepburn (2023). Although 31 participants consists of a not unusual size for eye-tracking experiment, the two divisions into experts and non-experts, as well as into one of two experiment sequences for counter-balancing drastically reduce the sample size of identical groups. The results of the 7 to 8 participants making up one group are very sensitive to outliers. This is reflected in the large standard deviations observed in certain metrics, which reduce the reliability of the results and complicate statistical testing (Cohen, 2013).

Regarding the choice of tasks it was particularly difficult to find a compromise between being too difficult for non-experts and too easy for expert participants. An initial plan to

include the Zurich region was ultimately abandoned due to its extremely high visual and semantic complexity, which makes it nearly impossible to navigate for non-expert participants. Therefore, the tasks generally were too easy, especially throughout the second iteration. As a result, the shorter completion time did not lead to participants becoming overly tired throughout the experiment, as was the case with the Marx's Master's thesis (Marx, 2015).

The eye-tracking equipment is also subject to variability, as some participants returned erratic eye movements that were not fully trustworthy. One participant's data was unusable and not taken into account in the post-experiment analysis. With one participant, the calibration was unsuccessful due to only one eye being recognized by tracker, reducing the sample size from 32 to 31. The calibration results are also participant specific, and therefore vary in accuracy. As the experiment took roughly seven to fifteen minutes to complete, it was impossible for participants to stay absolutely still throughout, further adding to the possibility of erroneous eye-tracking data.

Chapter 6

Conclusion

6.1 Main Findings

This study investigated whether an update to the current VFR aeronautical chart can improve the readability for private pilots, and if the effect is more pronounced among non-expert participants representing novice pilots. By removing current cartographic deficits, adhering more strictly to the well established ICAO design guidelines from Annex 4 (International Civil Aviation Organization, 2022), and implementing new design features, it aimed to enable lower visual search completion times.

The main findings support the potential of performance improvements when utilizing the updated map. Especially in the first map extent, spanning the canton of Grisons, the most prominent differences were found. In terms of completion time, a significant improvement was seen, both for experts and non-experts, showing potential benefits in a real-world scenario. While the fixation frequency, as well as saccade length and saccade speed significantly increased, times to first AOI hit significantly decreased on the updated map design. Still, the results obtained by the eye-tracking experiment must be treated with certain caution, as individual performance discrepancies paired with a small sample size ($N = 31$) account for high variances in some metrics.

Quantitatively analyzing the fixation heat map metrics revealed slightly lower entropies for the updated map design. Visually investigating the 2D plots showed labels influencing scan paths. By making the labels more salient and placing them at suitable positions on the map, gazes were more collected patterns on the updated map.

6.2   Contributions

This study compared two VFR aeronautical charts using eye-tracking technology, which has not been deployed in this niche of cartography. Additionally, it quantitatively put into perspective the effectiveness of the current Swiss VFR aeronautical chart, which previously only was looked at qualitatively. Its findings contribute to understanding how a de-cluttered design may increase map reading efficiency in the high visual clutter environment of a VFR aeronautical chart.

The distinction between expert and non-expert participants gives insight into different performances, which may especially benefit student pilots.

6.3   Future Research

To expand on the present work, a more extended eye-tracking study would offer deeper insights into the specific design elements that contribute to the effectiveness of VFR aeronautical charts, such as the sequence of AOI visits, or more thoroughly diving into gaze paths. Given the limited number of participants and the high degree of individual variability, some of the eye-tracking metrics observed in this study diverged from findings reported in literature. For future research, a between-subjects design may be more appropriate to reduce priming effects and isolate design-related differences more clearly.

Additionally, the inclusion of non-expert participants can be debated on, as the current findings already suggest that an updated map design provides greater benefits for this group. Future tasks should also be enhanced in complexity to better reflect real-world pilot decision-making processes. Ideally, the updated design could be evaluated in a mobile eye-tracking setup within a full flight simulator, allowing for more ecologically valid insights and higher fidelity to actual operational environments.

Bibliography

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.

ASFG Ausserschwyzer Fluggemeinschaft Wangen (2025). Flotte. Accessed: 2025-04-25.

Atkinson, R. C. and Shiffrin, R. M. (1968). Human Memory: A Proposed System and Its Control Processes. In Spence, K. W. and Spence, J. T., editors, *The Psychology of Learning and Motivation*, volume 2, pages 89–195. Academic Press, New York.

Babb, T. A. (2016). VFR Flight Planning with Paper Charts versus Electronic Charts. *International Journal of Aviation Sciences*, 1(2):1–8.

Berent, I. and Perfetti, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading English. *Psychological Review*, 102(1):146–184.

Bertin, J. (1983). *Semiology of Graphics: Diagrams, Networks, Maps.* University of Wisconsin Press, Madison, Wisconsin.

Bestgen, A.-K., Edler, D., Müller, C., Schulze, P., Dickmann, F., and Kuchinke, L. (2017). Where Is It (in the Map)? Recall and Recognition of Spatial Information. *Cartographica*, 52(1):80–97.

Brewer, C. A. (1994). *Color Use Guidelines for Mapping and Visualization*, page 123–147. Elsevier.

Burian, J., Popelka, S., and Beitlova, M. (2018). Evaluation of the Cartographical Quality of Urban Plans by Eye-Tracking. *ISPRS International Journal of Geo-Information*, 7(5):192.

Butchibabu, A., Grayhem, R., Hansman, R. J., and Chandra, D. (2012). Evaluating a de-cluttering technique for NextGen RNAV and RNP charts. In *2012 IEEE/AIAA 31st Digital Avionics Systems Conference (DASC)*, pages 2D2–1. IEEE.

Chekaluk, E. and Llewellyn, K. R., editors (1992). *The Role of Eye Movements in Perceptual Processes*, volume 88. North Holland, Amsterdam, 1st edition.

Ciołkosz-Styk, A. (2012). The Visual Search Method in Map Perception Research. Technical report, Instytut Geodezji i Kartografii, Warszawa.

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Routledge.

Cybulski, P. and Ledermann, F. (2024). The impact of point symbol similarity on visual search on maps. *Cartography and Geographic Information Science*, page 1–18.

Çöltekin, A., Fabrikant, S. I., and Lacayo, M. (2010). Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. *International Journal of Geographical Information Science*, 24(10):1559–1575.

Çöltekin, A., Heil, B., Garlandini, S., and Fabrikant, S. I. (2009). Evaluating the effectiveness of interactive map interface designs: A case study integrating usability metrics with eye-movement analysis. *Cartography and Geographic Information Science*, 36(1):5–17.

Doyon-Poulin, P., Ouellette, B., and Robert, J. (2012). Review of visual clutter and its effects on pilot performance: New look at past research. In *2012 IEEE/AIAA 31st Digital Avionics Systems Conference (DASC)*, page 1–36. IEEE.

Fairbairn, D. and Hepburn, J. (2023). Eye-tracking in map use, map user and map usability research: what are we looking for? *International Journal of Cartography*, 9(2):231–254.

Federal Aviation Administration (FAA) (2025a). *Aeronautical Information Manual (AIM), Chapter 9: Aeronautical Charts and Related Publications*, basic edition. Effective February 20, 2025.

Federal Aviation Administration (FAA) (2025b). San Francisco Terminal Area Chart. VFR Aeronautical Chart. Scale 1:250'000, Effective February 20, 2025 to April 17, 2025.

Fitzsimmons, F. S. (2002). The electronic flight bag: A multi-function tool for the modern cockpit. Technical report, Institute for Information Technology Applications, United States Air Force Academy, Colorado.

ForeFlight (2025). ForeFlight Mobile Electronic Flight Bag. `https://foreflight.com/products/foreflight-mobile/`. Accessed: January 18, 2025.

Gilbert, G. (1973). Historical Development of the Air Traffic Control System. *IEEE Transactions on Communications*, 21(5):364–375.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., and Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. oup Oxford.

Hubel, D. H. (1988). *Eye, Brain, and Vision*. Scientific American Library, New York.

ICAO (2005). *Global Navigation Satellite System (GNSS) Manual*. Montréal, Canada, 1st edition.

International Civil Aviation Organization (2022). *Annex 4 to the Convention on International Civil Aviation: Aeronautical Charts*. International Civil Aviation Organization, Montreal, Canada, 15th edition.

International Civil Aviation Organization (2024). *Annex 2 to the Convention on International Civil Aviation: Rules of the Air*. International Civil Aviation Organization, Montreal, Canada, 11th edition.

Jenks, G. F. (1973). Visual Integration in Thematic Mapping: Fact or Fiction? Manuscript, University of Kansas.

Keskin, M., Ooms, K., Dogru, A. O., and De Maeyer, P. (2020). Exploring the Cognitive Load of Expert and Novice Map Users Using EEG and Eye Tracking. *ISPRS International Journal of Geo-Information*, 9(7):429.

Kotval, X. P. and Goldberg, J. H. (1998). Eye Movements and Interface Component Grouping: An Evaluation Method. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 42(5):486–490.

Krassanakis, V. and Cybulski, P. (2019). A review on eye movement analysis in map reading process: the status of the last decade. *Geodesy and Cartography*, 68:191–209.

Krassanakis, V., Filippakopoulou, V., and Nakos, B. (2011a). An Application of Eye Tracking Methodology in Cartographic Research. In *Proceedings of EyeTrackBehavior 2011 (Tobii)*, Frankfurt, Germany.

Krassanakis, V., Filippakopoulou, V., and Nakos, B. (2011b). The Influence of Attributes of Shape in Map Reading Process.

Krassanakis, V., Lelli, A., Lokka, I. E., Filippakopoulou, V., and Nakos, B. (2013). Investigating Dynamic Variables with Eye Movement Analysis. In *Proceedings of the 26th International Cartographic Conference (ICC)*, Dresden, Germany.

Krejtz, K., Szmidt, T., Duchowski, A. T., and Krejtz, I. (2014). Entropy-based statistical analysis of eye movement transitions. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '14. ACM.

Leary, N. (1993). Maps and Charts for Visual Air Navigation. *Journal of Navigation*, 46(1):1–9.

Lee, Y., Jung, K., and Lee, H. (2023). Human Performance and Heat Map Entropy in System State Judgment Task using a Visual Interface Screen. *Engineered Science*.

Liu, Q., Wang, Y., Bai, Y., Yu, M., Cao, Z., Yu, X., and Ding, L. (2023). Development of a quantitative measurement on visual clutter in see through display. *Frontiers in Neuroscience*, 17.

Lohrenz, M. C., Trafton, J. G., Beck, M. R., and Gendron, M. L. (2009). A Model of Clutter for Complex, Multivariate Geospatial Displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 51(1):90–101.

Mackenzie, D. (2010). *ICAO: A History of the International Civil Aviation Organization*. University of Toronto Press, Toronto, ON, Canada.

Marx, J. (2015). Perzeption der VFR-Karten: Analyse der Effizienz und Effektivität von Visual Flight Charts. Master's thesis, Department of Geography, University of Zurich, Zurich, Switzerland.

McGregor, D. K. and Stern, J. A. (1996). Time on task and blink effects on saccade duration. *Ergonomics*, 39(4):649–660.

Multer, J., Warner, M., Disario, R., and Huntley, M. S. (1991). Design Considerations for IAP Charts: Approach Course Track and Communication Frequencies. Technical Report DOT/FAA/RD-91/19, U.S. Department of Transportation, Federal Aviation Administration.

Netzel, R., Ohlhausen, B., Kurzhals, K., Woods, R., Burch, M., and Weiskopf, D. (2016). User performance and reading strategies for metro maps: An eye tracking study. *Spatial Cognition & Computation*, 17(1–2):39–64.

Ooms, K., De Maeyer, P., and Fack, V. (2013). Study of the attentive behavior of novice and expert map users using eye tracking. *Cartography and Geographic Information Science*, 41(1):37–54.

Osborne, D. W., Huntley Jr, M. S., et al. (1992). Design of Instrument Approach Procedure Charts Comprehension Speed of Missed Approach Instructions Coded in Text or Icons. Technical report, United States. Federal Aviation Administration. Office of Aviation Research.

Polatsek, P., Waldner, M., Viola, I., Kapec, P., and Benesova, W. (2018). Exploring visual attention and saliency modeling for task-based visual analysis. *Computers & Graphics*, 72:26–38.

Raschke, M., Blascheck, T., and Burch, M. (2013). *Visual Analysis of Eye Tracking Data*, page 391–409. Springer New York.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8):1457–1506.

Rosenholtz, R., Li, Y., and Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, 7(2):17.

Sarbach, A., Kwok, T. C., Kiefer, P., and Raubal, M. (2025). 2D versus 3D aviation weather visualisations. *Ergonomics*, page 1–14.

Sarbach, A., Weber, T., Henggeler, K., Lutnyk, L., and Raubal, M. (2023). Evaluating and Comparing Airspace Structure Visualisation and Perception on Digital Aeronautical Charts. *AGILE: GIScience Series*, 4:1–13.

Skaves, P. (2011). Electronic Flight Bag (EFB) Policy and Guidance. In *2011 IEEE/AIAA 30th Digital Avionics Systems Conference*, pages 8D1–1–8D1–11.

Skyguide (2025a). *eVFR Manual Switzerland (GEN Section of the AIP)*. Federal Office of Civil Aviation (FOCA). Accessed: 2025-04-25.

Skyguide (2025b). Switzerland Liechtenstein Aeronautical Chart ICAO. Printed by Federal Office of Topography (swisstopo), CH-3084 Wabern. Scale 1:500,000.

Skyguide (2025c). Zurich Area and Geneva Area Area Chart ICAO. Printed by Federal Office of Topography (swisstopo), CH-3084 Wabern. Scale 1:250,000.

Steele, P. (1998). History of Air Maps and Charts. *Journal of Navigation*, 51(2):203–215.

Steinke, T. R. (1987). Eye Movement Studies In Cartography And Related Fields. *Cartographica*, 24(2):40–73.

Svatek, P. (2014). Civil Aviation Cartography in Austria 1908–1938. *Journal of Navigation*, 68(1):126–141.

Swisstopo (2025). Digital Aeronautical Chart ICAO Switzerland 1:500,000.

Techau, T. E. (2018). *General Aviation Pilot Acceptance and Adoption of Electronic Flight Bag Technology*. Ph.d. dissertation, Embry-Riddle Aeronautical University, Daytona Beach, FL.

Tobii AB (2024). Tobii Pro Lab (Version 1.241) [Computer software].

Williams, L. (1971). The Role Of The User In The Map Communication Process: Obtaining Information From Displays With Discrete Elements. *Cartographica*, 8(2):29–34.

Winter, S. R., Milner, M. N., Rice, S., Bush, D., Marte, D. A., Adkins, E., Roccasecca, A., Rosser, T. G., and Tamilselvan, G. (2018). Pilot performance comparison between electronic and paper instrument approach charts. *Safety Science*, 103:280–286.

Yeh, M., Jaworski, J., Swider, C., and Chase, S. (2021). Examining Minimum Information Requirements for Electronic Aeronautical Charts. *International Journal of Human–Computer Interaction*, 37(7):601–610.

Young, J. P., Fanjoy, R. O., and Suckow, M. W. (2006). Impact of Glass Cockpit Experience on Manual Flight Skills. *Journal of Aviation/Aerospace Education & Research*, 15(2).

Zhang, M., Gong, Y., Deng, R., and Zhang, S. (2022). The effect of color coding and layout coding on users' visual search on mobile map navigation icons. *Frontiers in Psychology*, 13.

Appendix A

Tables

Table A.1: Aligned Rank Transform ANOVA on Answer Time.

Effect	df	Res. df	F	p-value
Task	8	504	14.30	**<.001**
Design	1	504	0.56	.454
Expert	1	504	3.36	.068
Task:Design	8	504	0.64	.744
Task:Expert	8	504	1.04	.401
Design:Expert	1	504	5.85	**.016**
Task:Design:Expert	8	504	2.34	**.018**

Table A.2: Aligned Rank Transform ANOVA on Average Saccade Duration.

Effect	df	Res. df	F	p-value
Task	8	504	8.07	**<.001**
Design	1	504	1.87	.172
Expert	1	504	15.56	**<.001**
Task:Design	8	504	0.46	.881
Task:Expert	8	504	0.88	.531
Design:Expert	1	504	0.09	.759
Task:Design:Expert	8	504	1.01	.426

Table A.3: Aligned Rank Transform ANOVA on Average Saccade Length.

Effect	df	Res. df	F	p-value
Task	8	504	23.37	**<.001**
Design	1	504	11.88	**<.001**
Expert	1	504	2.24	.135
Task:Design	8	504	1.51	.150
Task:Expert	8	504	0.55	.818
Design:Expert	1	504	2.01	.157
Task:Design:Expert	8	504	1.34	.222

Table A.4: Aligned Rank Transform ANOVA on Average Saccade Speed.

Effect	df	Res. df	F	p-value
Task	8	504	18.38	**<.001**
Design	1	504	8.40	**.004**
Expert	1	504	16.61	**<.001**
Task:Design	8	504	0.75	.647
Task:Expert	8	504	0.42	.912
Design:Expert	1	504	1.91	.167
Task:Design:Expert	8	504	1.12	.351

Table A.5: ART ANOVA results for Time to First AOI Hit by Design ($_D$), Expertise ($_E$) and their Interaction ($_I$).

AOI	Sub-Task	df$_{res}$	F$_D$	p$_D$	F$_E$	p$_E$	F$_I$	p$_I$
Grisons								
CVA DME	1	49	1.129	.293	0.242	.625	1.583	.214
WFJ DME	1	40	5.759	**.021**	1.791	.188	0.450	.506
Samedan	1	47	1.536	.221	0.376	.543	1.854	.180
Bad Ragaz	1	21	1.862	.187	0.092	.764	0.282	.601
Julierpass	2	47	0.310	.581	3.527	.067	3.742	.059
Albulapass	2	52	1.678	.201	2.051	.158	1.141	.290
Flüelapass	2	46	3.354	.074	8.817	**.005**	2.012	.163
Obstacle	3	52	6.038	**.017**	9.492	**.003**	9.981	**.003**
Domat/Chur	4	40	0.581	.450	0.136	.714	0.281	.599
Heliport	4	54	7.352	**.009**	2.296	.136	3.078	.085
Romandie								
Montrîcher	1	56	0.195	.661	0.230	.634	0.427	.516
Yverdon	1	24	0.642	.431	0.002	.969	0.035	.854
La Côte	1	26	4.075	.054	0.336	.567	0.399	.533
Blécherette	1	39	0.850	.362	0.162	.689	0.119	.731
Tank Farm	3	44	1.218	.276	4.688	**.036**	2.235	.142
GLA NDB	4	53	8.126	**.006**	9.950	**.003**	2.577	.114
SPR VOR-DME	4	42	0.363	.550	8.761	**.005**	0.008	.930
LAP DME	4	10	0.643	.441	0.515	.489	0.062	.809
LSR-5	5	54	0.146	.704	2.028	.160	0.864	.357

Table A.6: ART ANOVA results for AOI Dwell Time by Design ($_D$), Expertise ($_E$) and their Interaction ($_I$).

AOI	Sub-Task	df$_{res}$	F_D	p_D	F_E	p_E	F_I	p_I
Grisons								
CVA DME	1	49	0.304	.584	0.336	.565	1.031	.315
WFJ DME	1	40	0.001	.973	0.208	.651	0.005	.943
Samedan	1	47	0.222	.639	1.684	.201	2.451	.124
Bad Ragaz	1	21	1.505	.234	0.623	.439	1.192	.287
Julierpass	2	47	0.001	.974	1.195	.280	0.013	.911
Albulapass	2	52	2.423	.126	3.905	.053	0.090	.766
Flüelapass	2	46	0.000	.997	0.423	.519	0.003	.960
Obstacle	3	52	1.549	.219	0.540	.466	1.335	.253
Domat/Chur	4	54	0.188	.666	0.103	.750	0.141	.709
Heliport	4	40	7.016	**.012**	0.467	.498	0.009	.925
Romandie								
Montrîcher	1	56	0.316	.576	3.888	.054	0.025	.874
Yverdon	1	24	0.271	.607	0.004	.952	0.010	.919
La Côte	1	26	0.059	.811	0.002	.963	0.083	.775
Blécherette	1	39	0.039	.845	0.009	.926	0.034	.855
Tank Farm	3	44	0.314	.578	0.939	.338	0.849	.362
GLA NDB	4	53	0.310	.580	1.847	.180	0.036	.850
SPR VOR-DME	4	42	9.799	**.003**	0.060	.807	3.991	.052
LAP DME	4	10	0.656	.437	1.216	.296	0.001	.976
LSR-5	5	54	3.187	.080	5.720	**.020**	0.716	.401

Task	Inertia	Std. Dev.	Mean	Entropy
Grisons - Current				
Task 1	769.98	6.84	1.60	8.74
Task 2	662.68	7.08	1.72	8.74
Task 3	822.44	6.93	1.24	7.53
Task 4	593.66	7.96	1.84	8.21
Grisons - New				
Task 1	805.82	6.84	1.37	8.10
Task 2	736.13	6.91	1.42	8.16
Task 3	843.05	6.82	1.20	7.56
Task 4	652.27	7.52	1.69	8.17
Romandie - Current				
Task 1	798.13	6.95	1.36	7.96
Task 2	762.37	6.83	1.30	7.85
Task 3	778.49	6.94	1.31	7.80
Task 4	940.38	6.89	1.29	7.81
Task 5	906.81	6.86	1.17	7.46
Romandie - New				
Task 1	799.51	6.90	1.30	7.79
Task 2	780.61	6.96	1.25	7.59
Task 3	837.02	6.80	1.17	7.52
Task 4	967.61	6.82	1.10	7.09
Task 5	889.22	6.82	1.21	7.62

Table A.7: Global fixation heatmap metrics grouped by stimulus and map design.

Task	Inertia	Std. Dev.	Mean	Entropy
Grisons - Current				
Task 1	876.89	6.72	1.20	7.81
Task 2	863.04	6.75	1.11	7.30
Task 3	943.24	6.73	1.00	6.82
Task 4	773.76	6.90	1.26	7.53
Grisons - New				
Task 1	927.10	6.71	1.06	7.16
Task 2	857.17	6.73	1.13	7.48
Task 3	976.80	6.71	0.95	6.53
Task 4	820.53	6.84	1.18	7.37
Romandie - Current				
Task 1	909.74	6.73	1.09	7.29
Task 2	915.92	6.71	1.02	6.96
Task 3	864.48	6.78	1.11	7.26
Task 4	1008.06	6.73	1.01	6.88
Task 5	999.28	6.72	0.94	6.51
Romandie - New				
Task 1	912.89	6.73	1.06	7.13
Task 2	931.08	6.72	0.99	6.79
Task 3	921.11	6.72	1.02	6.98
Task 4	1013.16	6.71	0.95	6.57
Task 5	962.51	6.71	1.01	6.94

Table A.8: Expert fixation heatmap metrics grouped by stimulus and map design.
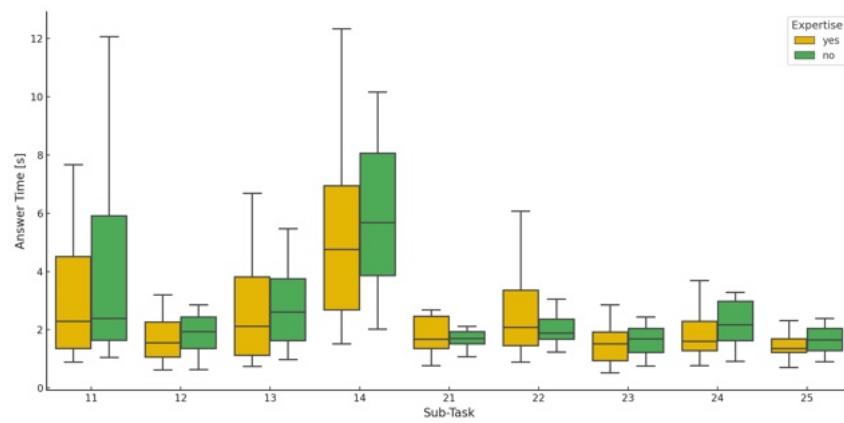
Appendix B

Figures

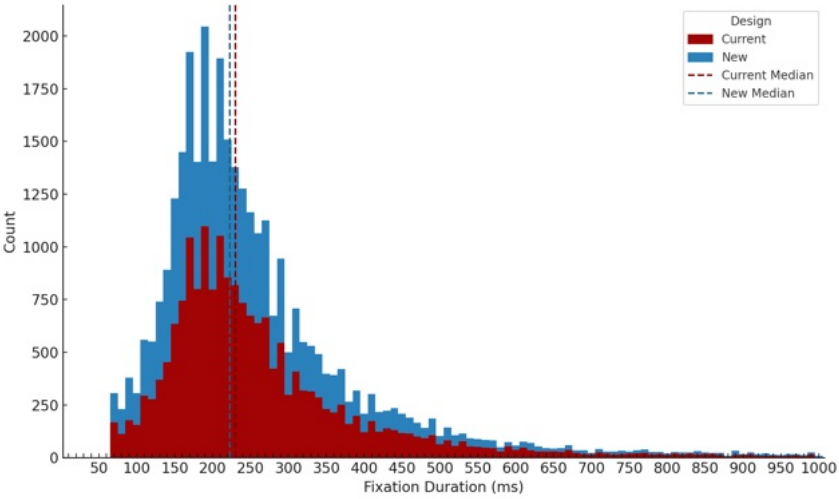**Figure B.1:** Answer Times per Sub-Task by Participant Expertise

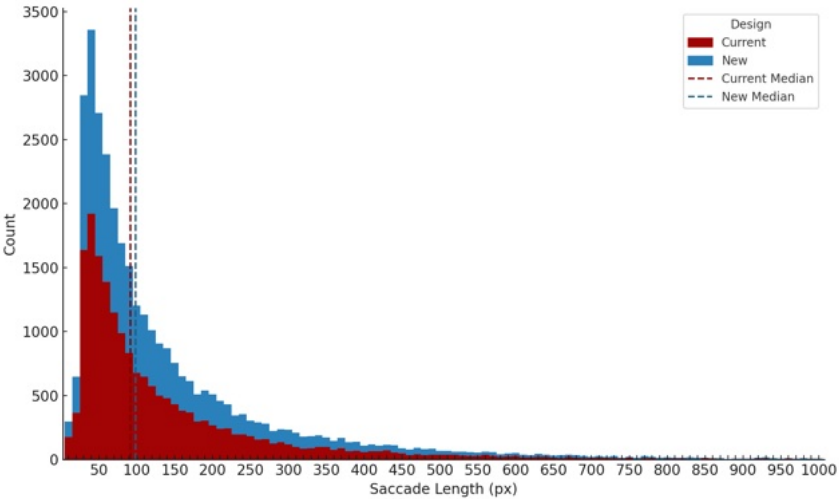**Figure B.2:** Fixation Duration Frequency Count
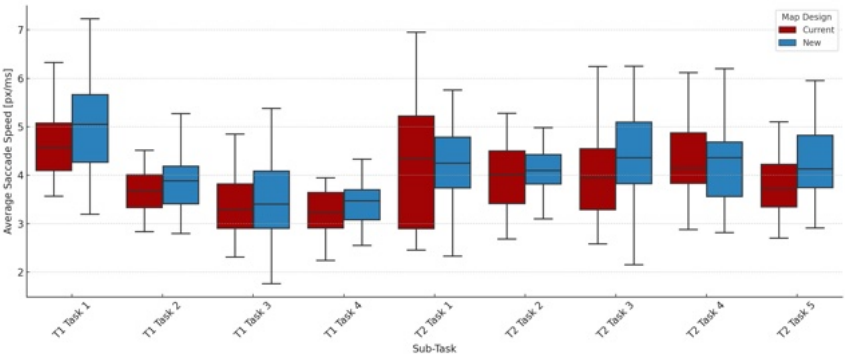

Figure B.3: Saccade Length Frequency Count

Figure B.4: Saccade Speed per Sub-Task across Map Designs

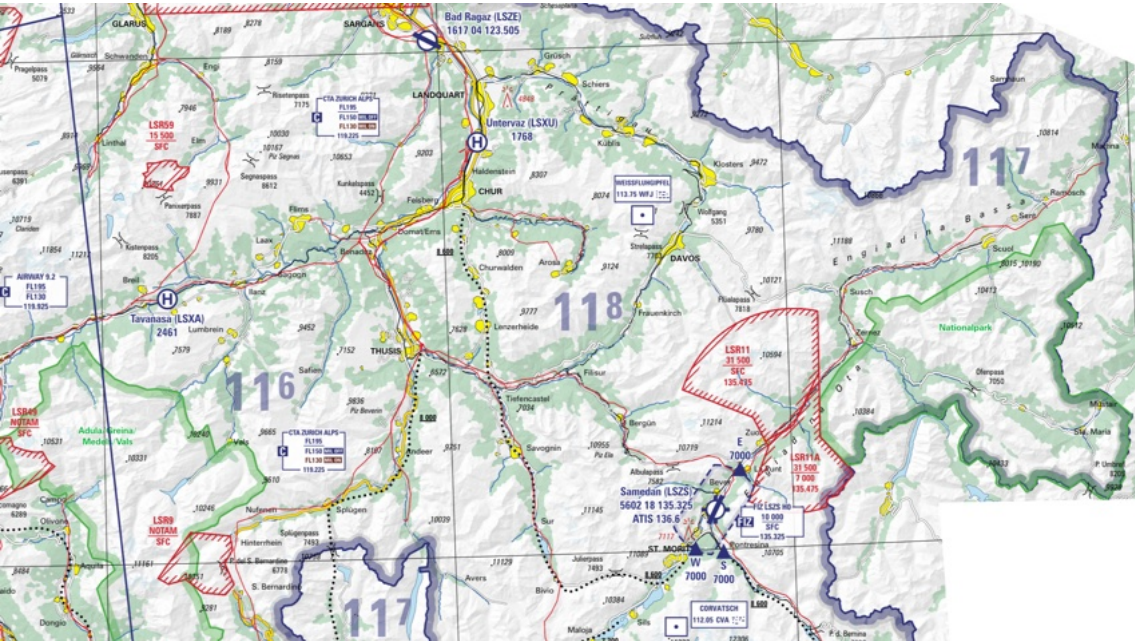**Figure B.5:** Grisons Current Map Extent (`T1C`)

Figure B.6: Grisons New Map Extent (`T1N`)

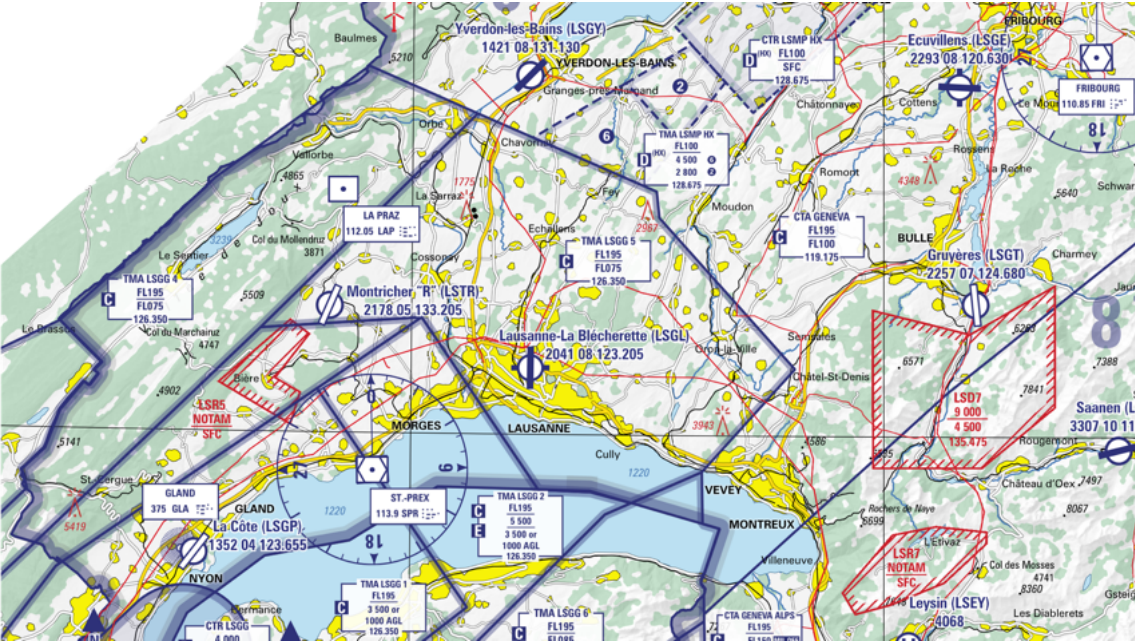**Figure B.7:** Romandie Current Map Extent (T2C)

Figure B.8: Romandie New Map Extent (T2N)

Appendix C

Used R Packages

Package	Version	Description
ARTool	0.11.1	Implements the Aligned Rank Transform for nonparametric ANOVA on factorial models.
dplyr	1.1.4	A grammar of data manipulation.
ggplot2	3.5.1	Create elegant data visualizations using the grammar of graphics.

Table C.1: Overview of R packages used in this analysis.

Appendix D

Used Python Packages

Package	Version	Description
collections	Built-in	Built-in module for specialized container datatypes (e.g., defaultdict, Counter).
matplotlib	3.10.1	Plotting library for creating static, animated, and interactive visualizations.
numpy	2.2.4	Core library for numerical computing with arrays and mathematical functions.
pandas	2.2.3	Data structures and data analysis tools (especially for tabular data).
pingouin	0.5.5	Statistical package for easy and robust parametric and non-parametric tests.
scipy	1.15.2	Scientific computing tools, including optimization, integration, and statistical functions.
seaborn	0.13.2	Statistical data visualization based on matplotlib.
statsmodels	0.14.4	Statistical modeling and testing (e.g., linear models, time series, ANOVA).

Table D.1: Overview of Python packages used in this analysis.

Appendix E

Additional Documents

Thank you for volunteering to participate in my Master's thesis study on the map use behavior of the VFR aeronautical area chart! This form collects background information about each participant, which will be used for statistical analysis.

All collected data will be used solely for research purposes in the context of this study. Personal information will not be shared with third parties, and all responses will be anonymized in the final thesis publication.

Background Information

- Email

- Name

- Age

- Gender

- Have you ever used the VFR Map of Switzerland? (Yes/No)

- Do you own a pilot license? (Yes/No)

Flying Experience

- Approximately how many flight hours have you logged until today?

- Where are you based? (ICAO code)

- What is your preferred method of enroute navigation? (Multiple choice + Other)

- What kind of base map does your electronic application use? (Multiple choice + Other)

- How likely would you use a digital, scalable version of the official Swiss VFR map on an EFB? (Likert scale: 1 = Unlikely to 5 = Very likely)

General Map Knowledge

- How often do you use map products? (Paper maps, Google Maps, etc.)

- How familiar are you with the overall geography of Switzerland? (Likert scale: 1 = Not at all familiar to 5 = Very familiar)

- How familiar are you with the geography of Western Switzerland (Romandie)? (Likert scale)

- How familiar are you with the geography of Southeastern Switzerland (Grisons)? (Likert scale)

Study Information and Declaration of Consent

This document is a declaration of consent for the study "Navigating Complexity: A Study of Map Use Behavior in Swiss VFR Aeronautical Charts" at the GIVA unit in the Department of Geography of the University of Zurich. The master's thesis is supervised by Dr. Tumasch Reichenbacher and Dr. Armand Kapaj. Should any questions arise at a later date, Simeon Lüthi (simeon.luethi@uzh.ch) will be happy to answer them.

Purpose of the Study
The purpose of the study is to empirically determine the efficiency of a newly proposed VFR aeronautical chart design against the current map design using stationary eye-tracking. The study site is located in the eye movement lab at UZH Irchel Campus (Y25-L9). The study is being carried out as part of Simeon Lüthi's master's thesis.

Procedure of the Study
During the study, you will be presented with static map extents. The extents cover two regions in Switzerland and will be displayed both using the current, as well as the newly proposed map design. Throughout the experiment, you will be given simple visual search tasks followed by multiple-choice questions. The study takes about 10-20 minutes to complete. You can stop the experiment at any time and without giving a reason.

Confidentiality of the Data
The data collected during the study will be treated confidentially and will not be passed on to third parties. With your signature, you give your consent for us to publish the anonymized and aggregated results of the study. The data published in this way will not allow any conclusions to be drawn about you personally. The eye-tracking device does not collect any audio or video material.

Harms
This study is classified as low-risk. However, please note that we do not provide specific insurance coverage for any potential harm during the experiment.

Declaration of Consent
By signing this form, you consent to the anonymized and aggregated results of the study being published. No personally identifiable data will be shared.

_____   _____
Full Name   Signature

_____
Date, Location

Appendix F

Personal declaration

I hereby declare that the submitted thesis results from my own independent work. All external sources are explicitly acknowledged in the thesis.

I acknowledge that AI tools were employed solely for grammar correction, syntax adjustments, and sentence refinement during the completion of this thesis. No reasoning was performed, or scientific content generated by AI.

Simeon Lüthi

Zürich, April 25th, 2025