



# Cartography M.Sc.

## Master thesis

### Developing Gaze-based Map Interactions

A User Study of Eye-controlled Navigation

Brandon Serrao



2020

# **Developing Gaze-based Map Interactions**

A User Study of Eye-controlled Navigation

submitted for the academic degree of Master of Science (M.Sc.)  
conducted at the Department of Aerospace and Geodesy  
Technical University of Munich

Author: Brandon, Serrao  
Study course: Cartography M.Sc.  
Supervisor: Dr.-Ing. Christian Murphy (TUM)  
Reviewer: Dr. Paolo Raposo (UT)

Chair of the Thesis  
Assessment Board: Prof. Dr. Liqiu Meng

Date of submission: 22.10.2020

## **Statement of Authorship**

Herewith I declare that I am the sole author of the submitted Master's thesis entitled:

**“Developing Gaze-based Map Interactions - A User Study of Eye-controlled Navigation”**

I have fully referenced the ideas and work of others, whether published or unpublished. Literal or analogous citations are clearly marked as such.

Munich, date: 22.10.2020

Brandon, Serrao

## Abstract

Digital map interfaces facilitate the interactions by which users connect with the information contained within the map. These interactions directly impact the efficiency and experience of map use. The volume and complexity of geodata continues to increase, and maps are their high-bandwidth medium for communicating the insights derived. So, the speed of the interactions with which a viewer can explore the medium is crucial. Using the same eye-trackers involved in cartographic usability studies, the eyes' near-instantaneous motion and perception can be wielded as a quick and conscious channel for interactivity.

Towards this, this master's thesis aimed to find suitable and efficient ways to facilitate map interaction using direct eye control. The map interactivity investigated was navigation by panning and zooming. This was done by comparing a mouse and keyboard (interface #1) to a pair of gaze-based interfaces; one interface (#2) used gaze location and the keyboard, while the other (#3) used winks, blinks and gaze-aware screen regions. A task-based user study assessed the user-reported usability, experience, and task load, as well as the users' preference and feedback. The gaze-based interfaces received poorer usability scores, in parts due to interface instability and users' varying ability to wink. User experience and reception of gaze-interactions were more positive, with strong preferences for the

interactivity of #2, indicating a strong potential of gaze-based navigation should the technical issues be overcome.

## Acknowledgements

To Juliane for your encouragement, your confidence in my abilities, this life-changing opportunity, and your superb role as coordinator for this program,

To my thesis supervisor Dr. Murphy, for your understanding, helpful feedback and guidance throughout my writing,

To my friends and fellow classmates – especially Nikita and Verena for their grounding presence,

To those who supported me by participating in or recruiting for my user study,

To my mother, family, and friends across an ocean who continue to believe in and support me,

And to the culture of my home, which has prepared me for and kept me throughout these times,

I say thank you.

## Table of Contents

Abstract.....	4
Acknowledgements.....	6
List of Figures .....	10
List of Tables.....	11
Introduction.....	12
Motivation and Problem Statement.....	12
Research Objectives .....	13
Delimitations:.....	13
Research Questions .....	14
Research Background .....	15
Introduction.....	15
A – Human Visual System.....	16
Function and Behaviour in Brief.....	16
Eye Movements .....	19
Summary and Conclusions .....	20
B – Eye-Based Human-Computer Interaction .....	21
Introduction to HCI Research.....	21
Operating Principles of Eye-Tracking Systems .....	22
Eye-Tracking Applications and Categories .....	24
Summary and Conclusions .....	26
C – Cartographic Interaction Considerations .....	28
Expectations of Map Interaction .....	29

Theoretical Bases of Cartographic Interaction.....	32
Summary and Conclusions .....	35
D – Cartographic Eye-Tracking Research.....	37
Usability and User Experience Assessment Tools.....	37
Similar Applications and Case Studies .....	40
Summary and Conclusions .....	44
Methodology.....	46
Case Study.....	48
Hardware and Software Utilized .....	50
Implementation Details .....	51
Eyetracker API Access and Modifications.....	51
Interface Operation and Description.....	53
Experiment Setup .....	57
Task Design .....	59
Questionnaire Structure.....	61
Results .....	64
Task Performance .....	67
Usability/UX Assessment.....	69
Task Load .....	69
System Usability .....	70
User Experience .....	71
Post-Study Feedback.....	73
Rankings.....	73
Issues and Observations .....	75



Discussion.....	78
Research Questions.....	82
Conclusion.....	85
Outlook.....	86
References .....	88

## List of Figures

Figure 1 - Cross-sectional Diagram of the Eye (Miromoto and Mimica, 2015).....	16
Figure 2 - Density Curve showing Distribution of the Retina's Photoreceptors (Chen et al. 2017).....	17
Figure 3 - Ninio's extinction illusion (Ninio and Stevens, 2000) and a Scintillating Grid Illusion.....	18
Figure 4 - A simulated example of gaze path across a scene (Bergstrom & Schall, 2014) ...	19
Figure 5: Components of Video-based Eye- and Gaze-Tracking (Hansen & Qiang Ji, 2010)	23
Figure 6 - Hierarchy of Eye-tracking System Categories (Duchowski, 2017).....	24
Figure 7 : Components of Cartographic Interaction (Roth 2012).....	28
Figure 8 - Stages of Interaction, as adapted by Roth (2013), combining models of interaction, visual thinking and cartographic taxonomy (Norman; MacEachren & Ganter, 1990; Roth, 2013) .....	32
Figure 9 - The scale structure of the UEQ.....	39
Figure 10: Variants of Gaze-Adaptive Map Legends by Göbel et al. (2018).....	41
Figure 11: Gaze-supported Foot Navigation, using foot-rockers and foot-pedals (Klamka et al. 2015).....	43
Figure 12: Conceptual Diagram showing the basis for the interaction sets implemented ...	46
Figure 13: Gazepoint GP3 Eye-tracker (www.gazept.com).....	50
Figure 14: Hardware Setup for User Testing.....	50
Figure 15: Structure of the PyGaze package .....	51
Figure 16: Example of Detail in Whitbread's Map of London at Different Scales.....	54
Figure 17: Experiment Setup, User Testing the Eyes-Only Interface.....	57
Figure 18: Number of Participants in each Age Range .....	64
Figure 19: Averages for User Task Completion Times (in seconds) for each Interface .....	67
Figure 20: Minimum Recorded Trial Completion Times (in seconds).....	68

Figure 21: Average Completion Times (in seconds) - Per Task Comparison to Mouse & Keyboard Interface.....	68
Figure 22: Results of the NASA Task Load Index Questionnaire.....	69
Figure 23: System Usability Scores, showing Acceptable Usability Score Threshold.....	70
Figure 24: Results of the Short User Experience Questionnaire (UEQ-S), showing different aspects of user experience and the overall qualities.....	71
Figure 25: User Ranking of Interface By Preference.....	73
Figure 27: User Ranking of Panning Interactivity.....	74
Figure 27: User Ranking of Zooming Interactions.....	74
Figure 28: Numbers of Participants Reporting Problems Encountered While Using the Interfaces.....	75

## List of Tables

Table 1: Expected Web Map Interactions identified from Interviews with Cartographic Professionals (Tolochko, 2016).....	31
Table 2: Table of Map Navigation Methods for each Developed Interface.....	49
Table 3: Sample of Data for three Users - Task Completion Times for each Interface (all times in seconds) .....	65

## Introduction

### Motivation and Problem Statement

As eye-tracking technology grows in affordability and competency, the use of the human gaze as an input mode for computer interfaces becomes increasingly viable. The measurement and characterization of eye motion is already extensively used in psychological and usability studies (Brychtova & Coltekin, 2016), which are considered ‘passive’ uses of the technology (Duchowski, 2017). On the other hand, applications of gaze-aware human-computer interfaces have included disability aids (Fejtová et al., 2009) and VR/AR control mechanisms (Orlosky et al., 2015). Eye-tracking in these systems may serve as either the primary or an additional input mode.

Despite some persisting challenges such as the “Midas touch”, extensive research exists on gaze-based pointing techniques, gaze-aware interfaces and eye-typing (Lutteroth et al., 2015; Majaranta et al., 2009; Zhai et al., 1999). While the technical ability to perform navigation, selection and input using these eye-tracking based technologies is clearly present, there is a dearth of research done on implementing these for use in map interactions. The majority of cartographic research using eye-tracking is concerned with evaluating user behavior, improving usability and communication (Dong et al., 2018; Duchowski, 2017).

Digital maps as a communication medium have a high bandwidth, and decoding this information may require significant time and effort. A gaze-aware map could dynamically adapt its content for search (Göbel et al., 2018) or accelerate map navigation using gaze-cursors (Porta et al., 2010). However, attempting to use vision for both perception and control may adversely impact the user's ability to use and comprehend the map. The current lack of research on gaze-aware map interaction hinders our ability to develop eye-based map interfaces.

## Research Objectives

This thesis aims to find suitable and efficient ways to facilitate user-map interactions with direct eye control. Doing so would entail the following objectives:

- a) Identifying the fundamental user interactions with the digital map surface and its associated UI elements,
- b) Determining the possibly beneficial operations in the field of eye-based human-computer interaction,
- c) And finally, assembling a subset of eye-controlled interactions from (b) that allow the interactions of (a) and evaluating their performance

## Delimitations:

- This research is focused on the interactions of a single user of a desktop

computer interacting with digital cartographic representations.

- This research does not seek to optimize the eye-control methods implemented. Any performance evaluations and comparisons will be in the context of usability and suitability for map-tasks.
- The developed interactions are intended to be explicit control methods and/or gaze-aware interactions, not passive eye-tracking and recording for user modelling, offline analysis and other similar support systems.

## Research Questions

- 1) What pairing of fundamental digital map interaction and eye-controlled interaction could produce a usable eye-based map interface?
  - a) What are the necessary or expected interactions for digital maps?
  - b) Which existing eye-based interaction methods facilitate an analogous model of interactivity?
- 2) Can an implementation of gaze-based control and/or awareness provide beneficial map interactions?
  - a) Which map use tasks can be performed/assisted using the selected gaze-based controls/awareness?
  - b) What is the user experience of operating the developed interfaces?
  - c) Which adopted eye-controlled interface(s) perform efficiently?
  - d) What challenges are there in the adoption and usability of such eye-based map interfaces?

## Research Background

### Introduction

The motion of the human eye and the workings of vision have been studied throughout the last century, outlined by (Duchowski, 2017), as researchers seek to understand the nature of visual attention. This has relied on eye-tracking methods, using measurements of a viewer's eyes to derive the direction of their gaze. Reconstructing the path the gaze takes through a scene has allowed researchers to infer the trail of user attention, to assess what guides or distracts, and so develop better user experiences. (Duchowski, 2017)

Eye-tracking technologies currently can record eye motion in real time and are implemented in systems whose interface may either adapt to or be controlled by the human visual system. Eye-controlled interfaces have seen real-world implementation in mobility aids and other assistive technologies for the physically impaired (*EyeMine*, n.d.; Plesnick et al., 2014; Vickers et al., 2010). This mode of explicit control has been explored through the development of various interfaces, in some cases attempting to emulate existing input devices such as computer mice and keyboards (Majaranta et al., 2009; Porta et al., 2010).

This explicit eye-based control has seen limited study in the field of cartographic research. Cartographic literature utilizing eye-tracking for usability assessment and understanding map interaction are plentiful by

comparison. This is noticeable in the overview by (Kiefer et al., 2017) of recent eye-tracking literature in spatial research, which cites significantly fewer works on gaze-based interactions.

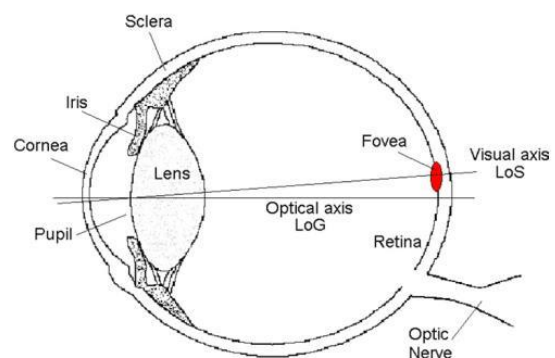
## A – Human Visual System

### Function and Behaviour in Brief

The conventional camera and the human eye share similar mechanisms and anatomy, with Figure 1 showing the arrangement of the eye's components. Both take in environmental light and focus it onto a photosensitive region at the rear, resulting in some form of image. Whether this is physically recorded on film or constructed in the mind, image quality is affected by how well the focusing and sensing of the light is performed.

Light entering the eye is refracted by the lens in order to be focused onto the retina. The lens can adjust its thickness in order to control focusing distance, resulting in sharper vision when focused onto the retina's surface. Photosensitive cells of the retina, called “rods” and “cones”, are sensitive to dim and brighter light respectively.

These are interconnected with other cells such as ganglions, which in concert perform more nuanced sensory functions. The stimulation and transmitted neural signals to the



*Figure 1 - Cross-sectional Diagram of the Eye (Miromoto and Mimica, 2015)*



brain form the basis of human sight. (Duchowski, 2017)

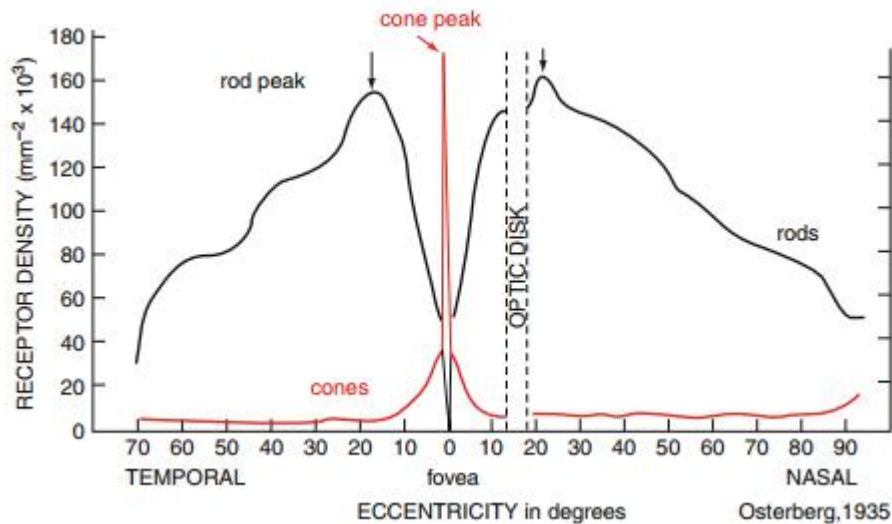
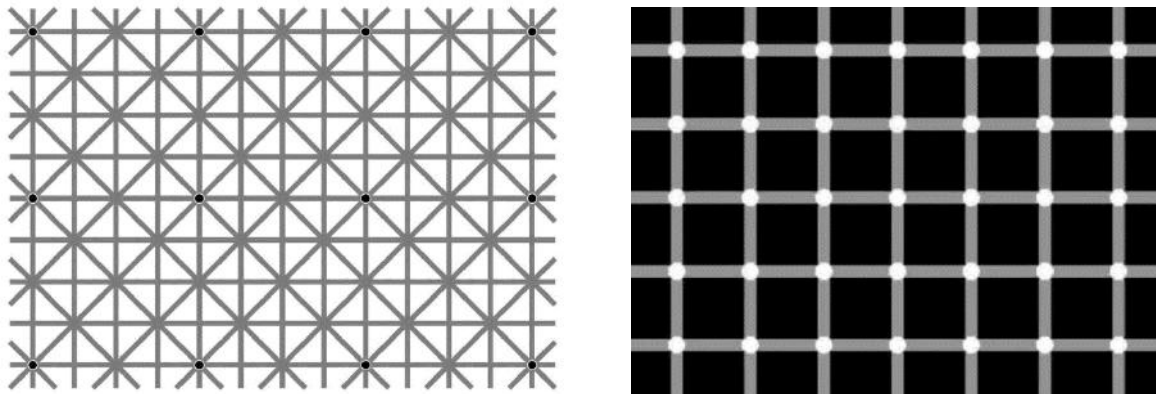


Figure 2 - Density Curve showing Distribution of the Retina's Photoreceptors (Chen et al. 2017)

As seen in Figure 2, the density and population of these receptors vary across the retina, so visual perception of a scene depends on where the light falls upon the retina. The central retinal region, called the macula, is most responsible for colour and detailed vision. Within this region is the fovea, which contains the peak density of cones and no rods. Cones contribute most towards the spatial resolvability of human vision, so this region provides the most detailed vision, diminishing further into the peripheral retinal area. (Chen et al., 2016)

The resolving power of vision can be quantified by the “visual acuity” metric, which is often evaluated based on the ability to distinguish high contrast features. Several factors impact acuity including spacing, light

intensity, target feature contrast, and motion. Retinal location of the light is another factor, with visual acuity diminishing further away from the visual axis which is centred on the fovea (highlighted in Figure 1). It degrades exponentially after 5° from this axis, and beyond 30° the visual field has limited resolving power. However, this peripheral/parafoveal vision is more useful at providing motion perception, being more sensitive to it than foveal vision. (Duchowski, 2017)



*Figure 3 - Ninio's extinction illusion (Ninio and Stevens, 2000) and a Scintillating Grid Illusion*

One consequence of these differences are optical illusions. Two such examples are provided in Figure 3, where the presence of spots at the intersections of lines appear to vary depending on where the viewer is looking. While the exact mechanisms that produce these illusions are still debated, these grid illusions are a visceral example of differences between foveal and peripheral visual perception. (Thomson and Macpherson, 2018)

## Eye Movements

A typical response of the human visual system could involve the eye perceiving an object of interest entering the peripheral field of view, followed by eye movement to bring the object into foveal vision for inspection. This repositioning of the eye is achieved through several simpler types of eye motion; saccades, smooth pursuits and fixations being key for locating attention. There are smaller tremors and micro-movements made by the eye, even during fixations which are uncontrolled and not perceived by the viewer. This functions to combat the saturation effect which leads to fading visual perception. The macro-movements together can be taken as the path of the viewer's attention through a scene. (Duchowski, 2017)

The different types of eye movement have purposes and characteristics. Saccades are the quick, large-scale motions that reposition the fovea to a new region of the field of view. They are quick, only lasting between 10ms and 100ms, with little to no alterations seemingly possible during execution. Fixations are the small-scale motions that keep the retinal focus on a stationary area of interest. Pursuit motion

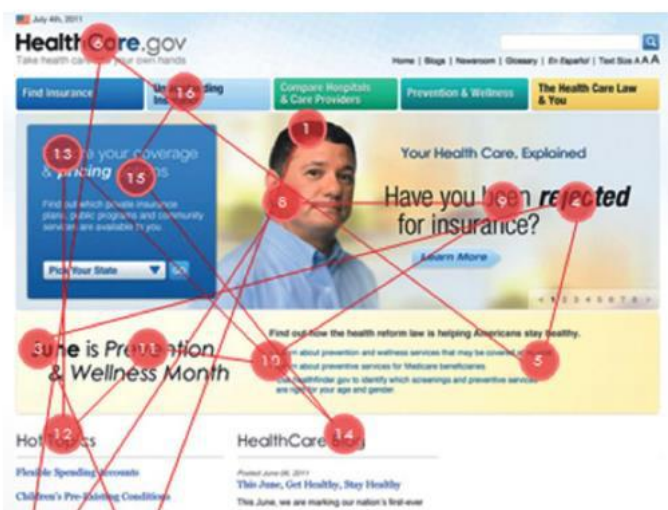


Figure 4 - A simulated example of gaze path across a scene (Bergstrom & Schall, 2014)

involves smoothly maintaining foveal vision on moving objects. Notably it is possible to deliberately refrain from moving the eye, with the viewer instead paying conscious attention more to the parafoveal vision. This choice of behaviour aside, the motion of a viewer's gaze across a scene can indicate the path of user attention. Figure 4 is an example of a scanpath – a sequence of fixation locations and implied saccades between them, as the viewer's gaze scanned the scene. (Duchowski, 2017)

### Summary and Conclusions

The human eye functions and behaves similarly to a camera, with limited perception of detail outside of the focused area in a scene. However, this peripheral region of vision is more sensitive to moving stimuli, to which it can move fairly quickly to focus on. So in developing eye-based interactions, one should consider possible distractions to the eye during the operation of the interface. This could be features like localized animation of the displayed content. Conversely, peripheral motion could be used to inform the viewer of changes in content or lead them to designed interaction opportunities. The apparent inability to change course during saccadic motion and the change blindness during this means this is not a suitable time for changing map content or offering interaction. Pursuit motion offers a direct method of leading the viewer through content via smooth animations.

## B – Eye-Based Human-Computer Interaction

### Introduction to HCI Research

Computers have gone from specialized tools for the trained individual to general purpose tools for the layman. This ubiquity birthed a need for insights into how people interfaced with computing devices. Human-computer interaction (HCI) is a young field of research, with the first conferences and books happening in the last 40 years. Its roots prior to this lay in a broad range of topics concerning psychology, management, and software engineering. The major contributions of its research are the development and empirical testing of artifacts; artifacts being the toolkits, prototypes and interfaces developed. (Lazar et al., 2017)

In the book *Human-Computer Interaction*, Dix (2004) speaks of the a users' input and output channels; the senses of the body (vision, hearing, etc.) being information receptors, and “effectors” such as the voice, limbs, and eye position communicate. Computer keyboards and mice have been the standard input peripherals for mundane desktop computing, hence the hands are the expected effectors. HCI research has continued to explore the use and integration of the other effectors; voice and bodily gesture recognition are commonplace features in modern mobile and entertainment devices. (Dix, 2004)

HCI research has typically used smaller datasets due to needing to gather their own data and participants. Larger studies exist thanks to Big Data techniques for designing, administering, analysing large-scale

experiments. Long-term studies are in short supply as well, likely due to the limited potential relevancy thanks to computing technologies' rapid advancement. The number and variety of participants in HCI studies are generally lower than in some of the fields HCI research draws from. HCI research is concerned with studying the interaction, and not necessarily about characterizing all persons. The user, however, is key part of the interaction thus this research is often user-centric. When research participants are selected, more concern is on their representativeness of the intended user-base, than the sampling method such as in social sciences. (Lazar et al., 2017)

### Operating Principles of Eye-Tracking Systems

Duchowski (2017) describes several different types of eye-tracking and measurement methods. Techniques using cameras were Photo-, and Video-OculoGraphy (PoG and VoG) which both derive the relative orientation of the eye from observations of a corneal reflection, while Electro-OculoGraphy measures electrical signals of the skin around the eye. Another type measures the eye motion via a contact lens device. These only provide relative orientation of the eye, so obtaining the location of gaze in a scene requires further measurements, from using an additional head-tracker for example. Hansen & Qiang Ji (2010) review and summarize eye-detection and tracking techniques at the time. Some of the eye-detection schemes covered

involve using models of eye-shape, locating distinctive eye-features (eg. eye corners) or eye-behaviours (eg. blinking).

Active eye-tracking methods seem the most common in recent literature; “active” referring to a non-ambient light used to illuminate the eyes. Figure 5 shows the general process of applications based on these eye-tracking methods. Eye features are measured, with key features including the pupil and the light source reflection. In table-mounted systems, the light source is in a fixed position and aimed at the eyes while they are observed by camera(s). Image processing is done, usually in real-time, to detect and measure the reflection and pupil location. A viewer’s gaze location on a flat display surface can be determined from these after initial calibration. The multiple feature measurements allow differentiating eye- and head-movement with this method. The light source is usually infra-red, which is invisible to the human eye. So the method is less physically and visibly intrusive. It is also relatively cheaper and easier to deploy thanks to using inexpensive cameras and image

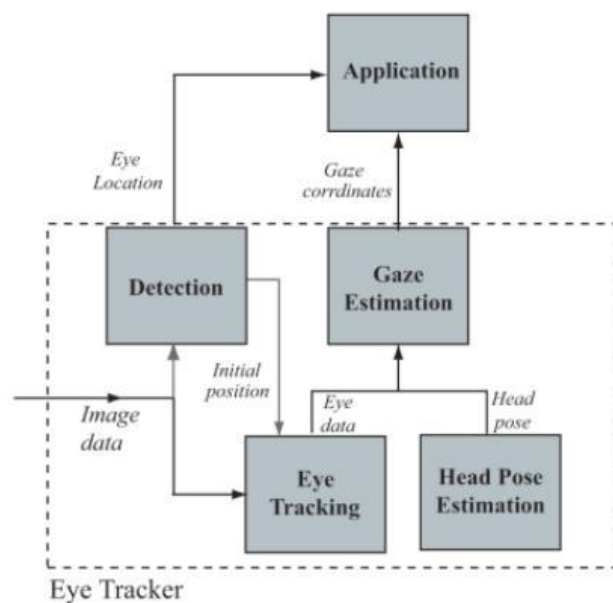


Figure 5: Components of Video-based Eye- and Gaze-Tracking (Hansen & Qiang Ji, 2010)

processing hardware. (Hansen & Qiang Ji, 2010; Lazar et al., 2017)

Pupil tracking plays a major role in the how well these systems perform, and this relies on how well the pupil is illuminated. Several factors impact the brightness sensed by the tracker. Increasing distance between the eye and eye-tracker result in weaker reflections. Head rotation can occlude the pupil, as well as changing the angle, affecting brightness and shape. Eyeglasses continue to present a challenge, causing multiple and dimmer reflections. External or ambient light sources can cause interference, and so have currently limited their use in outdoor applications. (Hansen & Qiang Ji, 2010)

### Eye-Tracking Applications and Categories

A useful categorization of eye-tracking is provided in Figure 6, separating applications into “diagnostic” and “interactive”. Diagnostic applications provide data on viewer attention from which further insights are derived. Interactive applications use the eye-tracker as an input device to empower the eye as an “effector” in a human-computer interaction, as noted by Dix (2004). This use of the eye in a conscious way to influence or entirely conduct an interaction is the interest of this thesis.

Selection-type techniques use deliberate eye-behaviour and gaze-location to

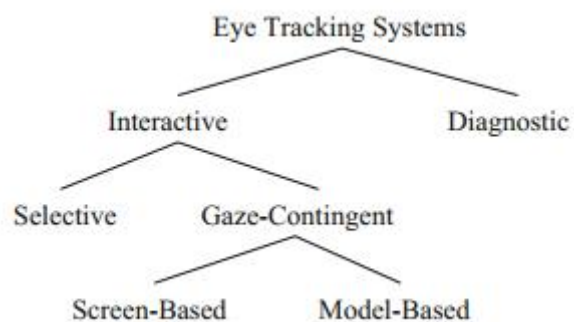


Figure 6 - Hierarchy of Eye-tracking System Categories (Duchowski, 2017)



select and trigger interaction objects within the interface. Eye-pointers or cursors are a common application in the literature, where the user gaze moves selection around the interface. These often use virtual keyboards or other selection areas that have pre-defined functions. Triggering the function or confirming the selection has been achieved through multiple methods. Dwell time is a common method, where the desired function is triggered by maintaining gaze fixation for a predetermined extent of time. A type of switch could also provide the needed confirmation trigger, depending on the extent of the user's motor control. Physical switches can work, or other detectable motion or gesture can be used as triggers. This includes the blinking of the eyes, glances in specific directions, or combinations of eye gestures. (Duchowski, 2017; Majaranta & Rähkä, 2002)

Gaze-contingent displays by comparison use gaze more passively, to modify or adapt displayed content and interactivity based on where the viewer is looking. A common motivation and usage is to conserve computational power in rendering visual scenes by degrading the level of detail in peripheral regions. Figure 6 distinguishes between Screen- and Model-based, where this manipulation is carried out on the displayed image or the graphical model itself. (Duchowski, 2017)

Interaction methods that do not rely on the hands are relevant to individuals with impaired motor functions, such as quadriplegics or those undergoing rehabilitation (van Middendorp et al., 2015; Yasumura et al., 2006). Accordingly, HCI research and projects have used eye-tracking to

enable such individuals. For example, the software EyeMine was created for physically impaired persons to play the game *Minecraft* by using an eye-tracker (*EyeMine*, n.d.). It uses the user's gaze to carry out selections. Another video game example is an eye-control enabling modification for the videogame Portal 2, where it served the similar function of gaze-controlled aiming from a first-person perspective (Bergstrom & Schall, 2014). Commercial products exist using eye-tracking for wheelchair control for the motor-impaired (*Eyedrivo*, n.d.; *EyeGaze*, n.d.), and eye-typing and text-to-speech (*EyeGaze*, n.d.). Other applications cater to all users in visual tasks. For reading, the *eyeBook* used eye-tracking for interpretation of reader's gaze location to provide an enhanced reading experience, by triggering sounds and showing imagery based on what part of the text was being read (Biedert et al., 2010).

### Summary and Conclusions

When presented with desktop computers, the human hands are the expected method of affecting communication with the computer, though alternative modes such as bodily gestures and speech-based communication have been developed. In this area of research, smaller numbers of research participants are common as their suitability is more valued than the sampling method used. From these, the intended audience should be well considered when developing and testing eye-based map interactions.

The function of selection-type eye-tracking techniques aligns more closely with the intention to provide the user with deliberate interactivity. This is apparent in its common applications which emulate existing interaction devices. Gaze-contingency implies a predictability or predetermined nature of user behaviour, which lends itself to subtler, adaptive, and less explicit interaction opportunities. So selection-type techniques may be more applicable to facilitating expected cartographic interaction.

Out of the eye-tracking technologies mentioned, active table-mounted systems appear the least intrusive. The barrier to entry can be low – a single calibration, and allow for some motion of the user’s head. For a non-impaired user of desktop digital maps, the minimal hardware setup may aid adoption and encourage use compared to systems employing electrodes or head-mounts. The issue of light interference in outdoor usage is less applicable to the indoor settings of desktop map use, where light sources are definite and can be manipulated. However, the issues caused by eyeglasses may impact eye-tracking performance so their usage during general eye-tracker studies should be noted.

## C – Cartographic Interaction Considerations

Looking at the growth of capabilities of digital geovisualization, MacEachren & Monmonier (1992) commented on the shift of (geo)graphic visualizations. This shift being from their role as tools communicating insight and as starting points for scientific inquiry, to tools that could allow use of the eyes for “extracting patterns from chaos”. Digital geovisualization could allow visual thinking and interaction with maps at the speed of the interface and of the analyst creating the graphics. MacEachren & Monmonier (1992) also note the increasing creation rate, quantity and complexity of geodata; so the interface to explore this data and its efficiency would be important, as mediator of cartographic interaction between human and map (Roth, 2013a). Figure 7 indicates the components of cartographic interaction, with Roth (2012) showing interface-centred research as one segment of a science of cartographic interaction.

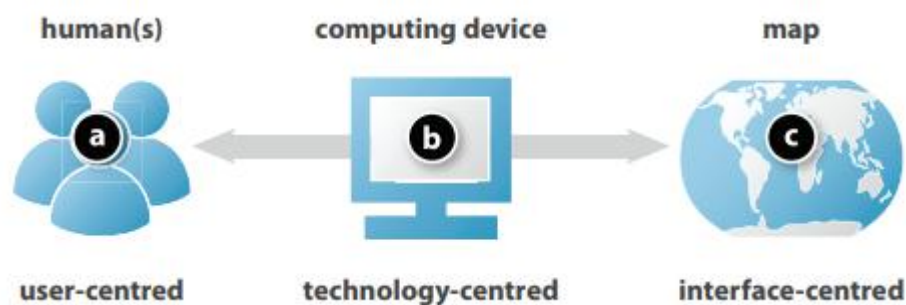


Figure 7 : Components of Cartographic Interaction (Roth 2012)

In *'Interacting with Geospatial Technologies'* (2010), Haklay provides guidelines to aid geospatial developers in creating spatial data interfaces, in

light of the increasing popularity and adoption of digital interactive map applications by non-experts. Noting the usability difficulty of many GIS applications till that time, he connected the need for HCI considerations and for specific spatial interface guidelines to be applied. User background, experience, physical and mental capabilities, prevailing interface conventions and expectations are all considerations in the interaction design.

### Expectations of Map Interaction

Haklay (2010) describes some guidelines for GIS interface design based on the WIMP (*Window, Icon, Menu, Pointer*) scheme of GUIs. Outlined below, these included:

- Consistency, throughout the interface's design and with existing conventions,
- Provide feedback to user actions, which is obvious and immediate,
- Preventing errors, by giving UI elements sufficient size and lucid symbology,
- An easily learnable conceptual model of interaction,
- Allowing for easy reversal of actions,
- Reducing short-term memory load,
- Aiming for universal usability, across ranges of user experience and physical capabilities.

With respect to which conventions are expected in interactive cartographic interfaces, Tolochko (2016) describes those in an expert study

focused more on map presentation and exploration. The interviews examined the reported workflows of cartographic professionals for creating interactive maps, the design conventions and ways of assessing success. Of the insights gleaned was a reductive approach to interface design; to aim for providing a minimal toolset that still allowed the user to fulfil their goal. The successful interface aimed for would have a balance between utility and usability.

For the purpose of map exploration of digital maps, several interactions and implementation recommendations were produced and are summarized on the following page in Table 1. It was noted that the intended purpose of the map should inform how much interactivity is provided. A map for allowing and encouraging exploration should allow greater interaction, than a map for presentation which should be more constrained and focused.

Table 1: Expected Web Map Interactions identified from Interviews with Cartographic Professionals (Tolochko, 2016)

<b><u>Interaction Identified and its Purpose</u></b>	<b><u>Implementation Note</u></b>
<p><b>Panning and Zooming</b></p> <p><i>Reveal and explore more information contained at different scales and beyond current map extents</i></p>	<ul style="list-style-type: none"> <li>▪ Constrain quantity of zoom levels based on map purpose</li> <li>▪ Provide a recentering or zoom-to-extents button</li> <li>▪ Follow conventional implementations for similar interfaces</li> <li>▪ Strong expectation exists to do this</li> </ul>
<p><b>Retrieve</b></p> <p><i>Allowing deeper examination and exploration of the data visualized</i></p>	<ul style="list-style-type: none"> <li>▪ Use popups for showing small amounts of info; use panels for linked visuals and larger amounts of info</li> <li>▪ Users have a strong expectation to click around, examine, and be provided with some feedback</li> </ul>
<p><b>Search &amp; Filter</b></p> <p><i>Finding specifics faster than manual navigation and search</i></p>	<ul style="list-style-type: none"> <li>▪ Not necessary for guided experiences</li> <li>▪ Geolocate current device location is of increasing importance on desktop</li> </ul>
<p><b>Overlay</b></p> <p><i>Comparison of input or available data layers</i></p>	<ul style="list-style-type: none"> <li>▪ Put a limiter on number of simultaneously overlaid layers</li> <li>▪ Ensure the legibility of the underlying data when overlaid</li> </ul>

## Theoretical Bases of Cartographic Interaction

Roth (2013) reviewed research on cartographic interaction examining its characteristics and factors in how, when and why to provide it. Also presented are different models and described types of cartographic interaction, to inform thought on how to perform it. It compares and combines the visual thinking model of Maceachren & Ganter (1990), which describes an iterative pattern-matching process of the mind, and Norman's stages of interaction. Shown in Figure 8, this combination shows the shifting of the user's sensory focus through the stages: from cognition, to perception, to motor functions. With this comparison, one can distinguish between the interaction and the interface. Similarly noted was Davies' (1998)

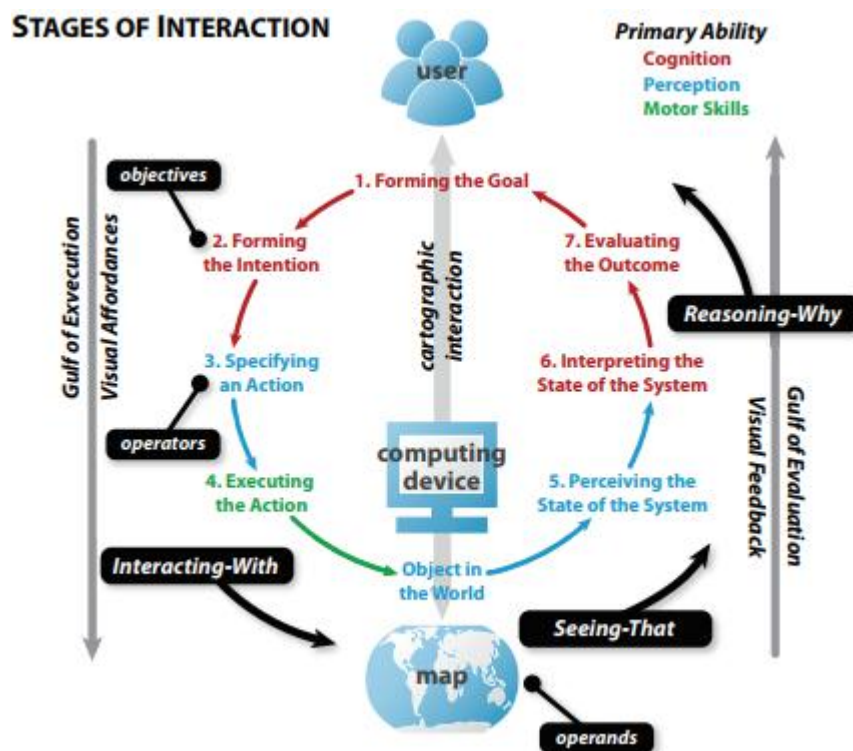


Figure 8 - Stages of Interaction, as adapted by Roth (2013), combining models of interaction, visual thinking and cartographic taxonomy (Norman; MacEachren & Ganter, 1990; Roth, 2013)



differentiation between “work” interactions and “enabling” interactions with a GIS system, which (Roth, 2013b) also noted as high-level operators in cartographic interaction. “Work” operators constitute the actions that actually perform the user goal, while “enabling” that are those facilitating but non-productive activities. Making these distinctions are said to be of value in experiment design for cartographic interfaces.

Regarding the implementation questions including ‘*Why?*’ and ‘*How much?*’ cartographic interaction, too much freedom or too many options within the interface can hinder performing user tasks. There is an amount of interface complexity and flexibility which is beneficial, and is also dependent on characteristics of the user. More experienced users can be afforded more functionality. An interface presenting with more functions, more complexity, more paths to achieve the same goal can add additional load to new users. This aligns with the minimalist design maxim noted by Tolochko (2016). Which functionality to be provided should also cater to the user context and goal in interacting with the map, such as exploration or sense-making. The studies show that always adding cartographic interaction does not always add value. However, that interactivity can play a part in motivating user interaction – innovative cartographic interactivity can grab user attention and encourage use. Interactivity that gives users a feeling of control or offer new potential behaviours or features can motivate users, and this motivation can fuel users to learn to use the interface. Motivation is an important factor in

the success of an interface, as it can inspire users to come to grips with more complex designs. (Roth, 2013a)

Concerning “*How?*” to provide cartographic interactions for direct manipulation of maps, Roth (2013) notes the need for visual symbols and feedback from the interface. These ‘visual affordances’ help inform the user of the changes affected in the map or data, and serve as indications of the style and possibilities of available interactions. Some affordances are the use of visual metaphors, contextual hints, and dynamic behaviours like animations that provide visual feedback. In interactions that directly manipulate the map – where the user can “probe, drag and adjust”, the interface can suffer from a lack of visual affordances (Roth, 2013a).

Such interactions are often used to implement Pan & Zoom operations, used for browsing and exploring maps. Harrower & Sheesley (2005) mention these implementations types, including Scroll Bars, Grab and Drag, Zoom & Re-center, and Zoom Box. Different implementations are more suited for different types of navigation, but some also provide valuable affordances. For example, scroll bars can provide a sense of scale and relative position in the map space.

User expectation of the style and function of map interaction were and continue to be informed by conventions of popular mapping services. So this can hinder some implementations’ effectiveness and adoption; continuous, unconstrained, pointer-focused Pan & Zoom is a modern expectation so incremental zooming and keyboard-based navigation methods can meet

resistance and discomfort from users. The interface is designed prioritizing that it is used, less so its efficiency, and Roth (2013) mentions that this could potentially result in maps designed to meet convention and not for purpose. (Harrower & Sheesley, 2005; Roth, 2013a)

### Summary and Conclusions

With the growth in quantity and complexity of geodata, the efficiency of the interface facilitating its exploration becomes more critical. Alongside speed, the adherence to interface conventions and expectations are also important for the usability of a system. Among the guidelines for graphical interface design, core recommendations include providing feedback and limiting frustration caused by usage errors and lacking reversibility. So the interaction model and its learning process should not be cognitively heavy. So in developing gaze-based interactions, simple and/or analogous models of interactions should be favoured, alongside an aim for providing a minimal toolset for successful interaction. Which interactions are facilitated, and to what degree, should be informed by the intended purpose of the map.

From the expected interactions identified by Tolochko (2016), Panning and Zooming is noted to be an expected core interaction. The interaction methods expected strongly correspond to those of similar interfaces, and a need for reversibility in the form of re-centering/zooming-to-extents. Notably some typical implementations of Pan & Zoom operations lack visual affordances, such as the de facto standard of pointer-focused

navigation. General purpose eye-pointer applications have been developed, so the implementation gaze-based panning and zooming should be possible in a similar fashion.

## D – Cartographic Eye-Tracking Research

### Usability and User Experience Assessment Tools

An interface mediates communication between map and user, so a well-designed interface would offer good usability and be a satisfying interactive experience. Usability of a system can be understood as its “fitness for purpose” which considers the intended audience or users, the tasks to be performed and the characteristics of the physical, social and organizational environment it will be used in (Brooke, 1996).

When testing usability, user performance is measured in terms of accuracy and efficiency metrics while performing some experimental tasks. Some typically used metrics are task completion time and error rate, which indicate efficiency and accuracy respectively. Skill retention rate and learning time are also considered metrics of human performance. International standards on usability state the metrics to be effectiveness, efficiency and satisfaction (ISO, 2019) and (Lazar et al., 2017) state frequently measured variables in research are:

- Efficiency,
- Accuracy,
- Subjective Satisfaction,
- Ease of Learning and Skill Retention Rate,
- Physical and Cognitive Demand.

They also note that efficiency and accuracy are highly related and should be both be measured. This task-centric model of human performance

is an old standard, so eases communicating results in research; it forms the foundation of evaluation and is still relevant today (Lazar et al., 2017).

The System Usability Scale (SUS) is a usability testing tool with a 10-item long questionnaire, which can describe the ease of use of any system. It is noted for its ease and speed, has been frequently used in research and shown to meet the need for a tool to quickly measure subjective usability from users (Bangor et al., 2008). Acceptable usability scores from the system center around a value of about 68 and very good usability above a value of 80 (Lewis & Sauro, 2018). An interface with good usability will allow high performance in the tasks they are designed to facilitate. (Haklay, 2010)

Work performance does not necessarily reflect the human side of interactivity, which contains the important aspect of how it feels to work with the system. Interfaces may require much time to learn, have designs that frustrate and stress users or impact how the message is received. These aspects can be measured subjectively using questionnaires and surveys. Pre-test questionnaires, after task completion, and final post-test questionnaires are opportunities used often. The NASA Task Load Index (NASA TLX) questionnaire is one of the most widely used to assess the cognitive load that is experienced by users in interacting with an interface. A greater load on users means the interface will demand more effort and focus from users, leading to increased fatigue and is less desirable. (Haklay, 2010; Lazar et al., 2017)

Beyond the demand felt using the interface, what the user gets out of the experience must be considered. This is important for their satisfaction and ultimately the success and adoption of the interface. The hedonic qualities of the experience are not measured by the previously mentioned approaches, and rely on qualitative questionnaires to measure aspects like novelty and stimulation. The User Experience Questionnaire (UEQ) is widely used and offers a standardized structure for measuring these user experience qualities. A shorter version, the UEQ-S, is also available for use alongside other assessment tools (Schrepp et al., 2017b). Seen in Figure 9, the structured scales are related and convey aspects like: the interface's clarity and learnability (Perspicuity) and its feeling of performance (Efficiency and Dependability), how engaging it feels (Stimulation and Novelty) and its overall impression (Attractiveness). The produced values are expected to compare to previous iterations of the same product, but a developed benchmark does exist for evaluation of new products. (Schrepp et al., 2017a)

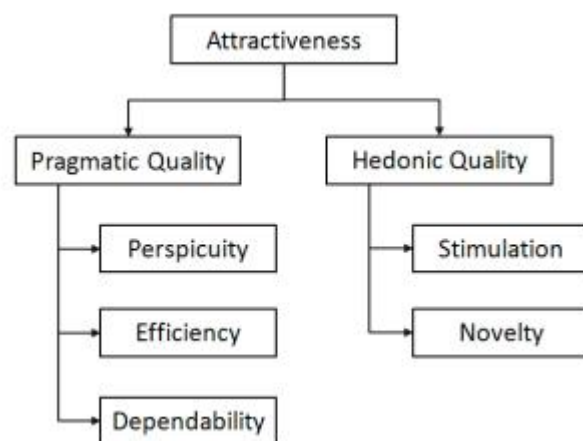


Figure 9 - The scale structure of the UEQ

## Similar Applications and Case Studies

The two main types of cartographic interface devices have been pointers and keying devices, typically computer mice and keyboards. Beyond this, there are others which use or incorporate alternative interfaces devices like microphones, body- and eye trackers (Roth, 2013a). In cartography, eye-trackers have been used as diagnostic systems for research while literature specifically examining eye-based cartographic interaction appears less numerous.

A diagnostic usage of eye-tracking for cartographic research is by (Brychtova & Coltekin, 2016), where they measured the effect of colour distance and font on map reading of static digital maps. An eye-tracker was used to for determining scanpaths and allowing for Area of Interest (AOI) analysis. In a within-subject experiment, they used very simplified static maps of US states and participants were given a search task: to click on the correct state with their mouse within a time limit. The order of tasks were randomized to mitigate the learning effect and participants were asked about prior knowledge and visual impairments. Findings showed that better label-background contrast allowed for faster visual searching, and that users spent an insignificant amount of time dwelling on labels (~130ms) compared to pointing and clicking on the found target (~1300ms).

Eye-tracking applications can be divided between explicit and implicit interactivity, according to the correlation between the user's intention and the system's behaviour. This difference and examples of both explicit and implicit



gaze-based cartographic interactions are noted in the overview by (Kiefer et al., 2017) on eye-tracking in spatial research. One implicit interaction involved creating markers at previously gazed-at map coordinates, to function as orientation cues. Another cited example adjusted displayed information based on the location of the user's gaze, the concept referred to as 'Geofoveation'. Machine learning is used on the user's gaze in another case, to infer the spatial task being performed. These are considered implicit interactions as they are not directly intended consequences of the actions performed by the user. The more explicit interaction noted is (Giannopoulos et al., 2015) which gives confirmation feedback when the pedestrian user looks at the correct street during instructed pedestrian navigation. (Kiefer et al., 2017)



Figure 7: Fixed adaptive legend FA: The legend highlights symbols that have been looked at.

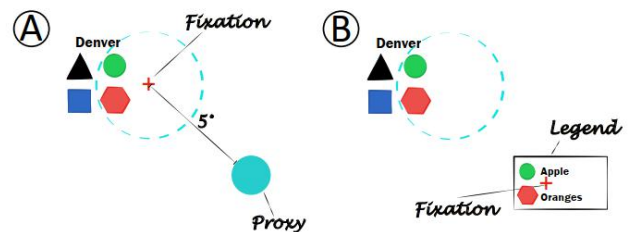


Figure 8: Dynamic adaptive legend (DA): (A) The legend proxy is located at bottom right, at 5° from fixation. (B) User looks at the proxy which immediately unfolds the legend showing the reduced content.

Figure 10: Variants of Gaze-Adaptive Map Legends by Göbel et al. (2018)

Other cartographic eye-tracking applications exists, such as the gaze-aware, adaptive legends (Göbel et al., 2018). The legend would change its contents and/or location based on the previously viewed map features, with

further dynamic behaviour regarding its visibility and placement (shown in Figure 8). User testing of their two variants of adaptive legend showed a decrease in time spent searching through legend contents. User acceptance of the fixed-position legend was higher, while frustration and increased load was felt for the other. It was stated to be likely due to inconsistencies in triggering the interaction to expand the legend with gaze, and unpredictability in behaviour of the dynamically positioned legend.

The works of Stellmach, Dachsel and others (Stellmach et al., 2011; Stellmach & Dachsel, 2012a, 2012b) cover several implementations of gaze-supported navigation behaviours, which use some external device as an additional mode of input. In their research on the potential of gaze-supported interaction, they tested several gaze-supported panning and zooming mode; using a handheld smartphone's touchscreen and tilt-sensing functionality, they tested several combinations of tilt/touch/gaze-based navigation (Stellmach & Dachsel, 2012a). Their term used for zooming to/from the current gaze location was 'gaze-pivot zooming'. Across the different interface combinations, gaze-panning involved the gaze location on the displayed map moving to the screen centre when triggered; zooming was activated and modified by either touching the smartphone's screen or tilting it. From their results, gaze-panning and 'gaze-pivot zooming' received positive feedback and ratings by participants, even compared to the baseline mouse navigation with scrolling zoom. Meanwhile, the touch-based zooming was also positively received though some haptic feedback mechanism was desired – that this

would save the user from having to look away from the main screen. Finally, the tilt-based zooming was considered complicated and tiresome and so not well received.

Their criticism of gaze-only interfaces is the additional operation time from the dwell-time activation method, cited as a hindrance to accessing the full potential of gaze-interaction (Stellmach et al., 2011). This is alongside the cited inaccuracies when using eye-measurement and gaze, such as jittering. In (Stellmach & Dachsel, 2012b) their approach to compensate for these are: enlarged interface elements, and the combining of gaze and manual input. So in addition to eliminating dwell-time, multimodal interfaces also circumvent the unintentional triggering of gaze-interaction, referred to as the ‘*Midas touch*’ effect (Stellmach et al., 2011). In the study by Göbel et al. (2018) they deliberately keep explicit interaction in their interface to a minimum due to this challenge.



*Figure 11: Gaze-supported Foot Navigation, using foot-rockers and foot-pedals (Klamka et al. 2015)*

Further supporting the performance of multimodal gaze-navigation, (Klamka et al., 2015) utilized foot pedals with different degrees of motion as

the additional interaction mode. In this study too, the reception by participants was quite positive and the performance of gaze-supported foot navigation was comparatively well compared to conventional mouse-based navigation. Both the interfaces of Klamka and Stellmach et al. provide continuous and variable zooming and panning rates which the user has some form of control over through this additional interaction mode. The touch- and tilt-based interaction of Stellmach et al. had the rate of change tied to the distance from some reference touch location or orientation; Klamka's pedals, foot-joystick and foot-rocker have degrees as well. In both studies, this level of control is remarked on and appreciated by users.

### Summary and Conclusions

The success of an interactive eye-tracking system would be based on a good usability and user experience. Usability is frequently assessed based on measures of users' task performance and the load experienced by them to achieve their goals. A good user experience is the other key aspect, and is assessed by subjective questions about the satisfaction and engagement felt during use. So to judge the success of a developed gaze-interface, the usability and user experience should be measured.

Non-conventional, multimodal interactions incorporating eye-tracking have been developed and tested, where the interactions are explicitly facilitating the user changing the displayed content. For the navigation of spatial data, gaze-supported Pan & Zoom interaction has been shown to work

competently and successfully when another mode of user input is allowed. This multimodal gaze-interaction approach is preferred by the research, due to gaze-only interaction's instabilities and fundamental problems like the 'Midas touch'. From this, it can be expected that any gaze-only interface will perform less satisfyingly and/or efficiently than multimodal gaze-supported interfaces when following similar methods.

## Methodology

From the expected map interactions noted in the previous chapter, map navigation by panning and zooming was chosen to be the implemented interactivity for this study. There is the need for constantly improving data exploration methods to meet the ever-increasing quantities and complexity of geodata. With the near-instantaneous targeting and selection potential of eye-

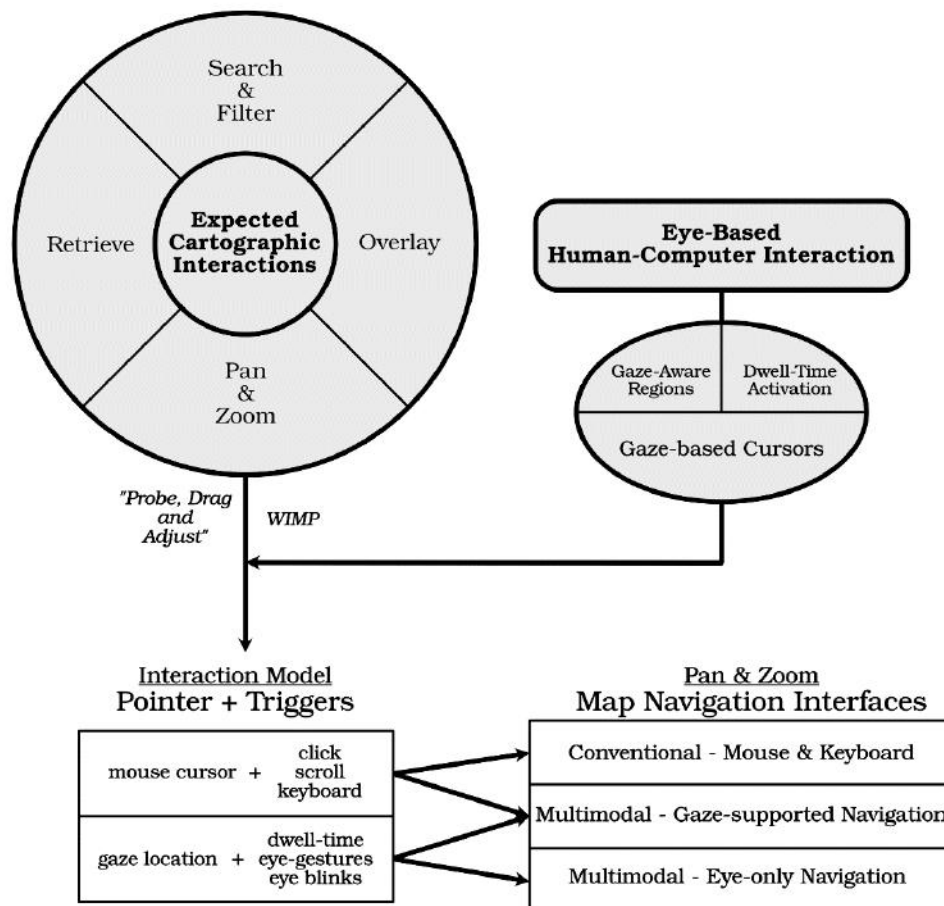


Figure 12: Conceptual Diagram showing the basis for the interaction sets implemented

based interaction, tying cursor-based navigation to the eye offers the prospect of minimizing time spent on the 'enabling' interactions involved. Cursor-based Pan & Zoom interaction is a well-established expectation for digital map interfaces. The existence of eye-controlled cursors means that the underlying model of interaction to perform Pan & Zoom is possible. Visible in Figure 12, the analogies are drawn between mouse- and gaze-based 'pointer' and 'trigger' interaction; the location of a user's gaze can be used with either a bodily or hardware triggering action to perform the same interactivity. Hardware triggers could be buttons/keys/switches, with the mouse and keyboard being those expected for a desktop scenario. Triggers available to gaze-supported interactions can be physical switches, even keyboards, or explicit gestures like blinking.

The conceptual diagram shown in Figure 12 shows the selected subsets of cartographic interaction and eye-based interaction techniques which are tied together to create the interfaces developed in this thesis. The basis of the pairing is the underlying interactivity model of pointer-based interaction, where a pointing method is paired with a selection trigger to allow for the user to interact with the displayed content. This interactivity is well established in the WIMP interaction style ("windows, icons, menus, pointer"), which meets the recommendation for a comprehensible and familiar model of interactivity when creating successful user interfaces. The "probe, drag and adjust" expected interaction, noted by Roth, is cursor-based and so functions on this underlying interaction model. Described in the literature are the equivalent

uses of users' gaze in place of a hardware-controlled cursor, and gaze behaviours as simple switch mechanisms. So from this, the interfaces can arise based on the degree and type of control which is desired to be available to the user: Conventional device-based, Gaze-guided, or fully hands-free and eye-controlled interfaces.

### Case Study

Towards assessing the performance of gaze-based map interactions, a comparison of different developed interfaces was done based on the approach by (Göbel et al., 2018). In their study, how potentially beneficial the tested functionalities were was assessed by carrying out a task-based user experiment. This follows a similar approach, basing the potential benefit of the developed interactions on self-reported performance and user experience, along with in-situ observation. Combining the conventional and gaze-based components of the interaction model, three interfaces for facilitating map navigation were created for testing:



<b>Implementation</b>	<b>Panning</b>	<b>Zooming</b>	<b>Re-centering</b>
<b>#1 - Mouse &amp; Keyboard</b>	Left Click to pan to location;	Scrolling zoom;	The <i>Spacebar</i> key to re-center move view to map center (no zoom level change)
	The Arrow Keys & WASD Keys for fixed increment panning	The <i>Q/E</i> keys to zoom out/in	
<b>#2 - Gaze-Supported Navigation</b>	The <i>W</i> key panned map to gaze-location	The <i>Q/E</i> keys to zoom out/in to gaze-location ( <i>pivot zoom</i> )	The <i>Spacebar</i> key (same as above)
<b>3# - Eyes-Only Navigation</b>	Gaze-aware map border areas ( <i>dwell-time activated</i> )	Keeping an individual left/right eye closed ( <i>winking</i> ) to zoom out/in of gaze location	Long Blink (both eyes)

Table 2: Table of Map Navigation Methods for each Developed Interface

## Hardware and Software Utilized

For this study, the eye-tracker used was the Gazepoint GP3 eye-tracker and the accompanying *Gazepoint Control software*. This choice is mainly due to it being readily available, courtesy the hosting university. As described in the

literature, when using table-mounted active eye-trackers, interference from strong light sources should be considered and mitigated as possible. Interference and reflections from worn eyeglasses should also be noted, and adjustments to seating/display/eye-tracker position done if possible.



Figure 13: Gazepoint GP3 Eye-tracker  
([www.gazept.com](http://www.gazept.com))



Figure 14: Hardware Setup for User Testing

A standard pair of external mouse and keyboard were connected to the host laptop and used for exclusively for conducting the experiment. An

external monitor (24inch, 1920px x 1200px, 16:10 aspect ratio) was used as the primary display for the entire duration of user testing.

## Implementation Details

### Eyetracker API Access and Modifications

The implementation of the interactive interfaces was done in the Python 2.7 programming language. Interfacing with the Gazepoint GP3 eye-tracker was done using PyGame, and the PyGaze package (Dalmaijer et al., 2014) which is a wrapper of code around the OpenGaze API. A simple diagram of the package's logical structure is shown in

Figure 16, showing the wrapper layers between the eye-tracker's data itself and the Python environment (in green).

PyGame is a package for writing video games using Python, and was used to implement graphical changes and interface interactivity; PyGaze also uses PyGame for its basic interactivity. (*PyGaze*

| *Documentation*, n.d.)

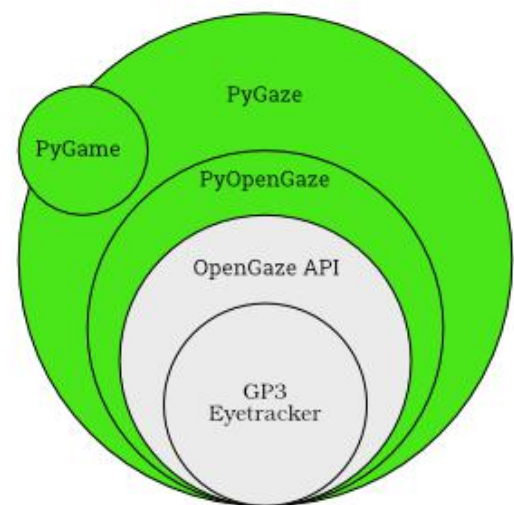


Figure 15: Structure of the PyGaze package

Modifications and additions were made to some of the underlying code of the PyGaze package; additional functions were added to allow exposure and storage of eye open/closed state for use. The changes to PyGaze were in *opengaze.py* :

- *Left\_winking* , *right\_winking* variables - added to the *OpenGazeTracker* class
- *get\_winking\_states* – a new function returning *left\_winking* and *right\_winking* values
- the existing function named *sample* - modified so that, in addition to normally updating the current sample of gaze location, it would also update the *left\_winking* and *right\_winking* variables stored in the eye-tracker Python object.

The main purpose of the modification is to make available the extra information that describes the state of the eyes. When the modified sample function is called, it updates additional variables inside of the eye-tracker Python object. These variables are a pair of Boolean values that correspond to whether or not the readings are considered valid, which are provided by the *Open Gaze API (Open Gaze API Version 2.0, 2013)*. Invalid readings of left and right eye gaze locations occur, for example, when the eye tracker cannot find the eye in view.

When the eye-tracker loses a lock on an eye, it is visible when using the *Gazepoint Control* viewer as a red screen, and this corresponded to measurements from the device during interface development. The interface implementation logic attempts to determine winks and blinks according to sustained, invalid signals over a certain duration. This was an attempt to

compensate for noise and other occasions where single invalid measurements would be recorded.

One common, systematic occurrence during interface development would be when the left eye winked, or left then re-entered view of the eye-tracker. Upon re-acquisition of the left eye, invalid state readings of *both* eyes are sent, despite having not closed or lost sight of the right eye. In some instances, the right eye would be placed as the left eye and may remain switched even when the left eye has been re-acquired. During testing, it appeared as though this was not a sustained signal – only a single data record from the device would wrongly indicate this state. It was surmised that this may behaviour may have been caused by the eye-tracking software operating on an assumption of the ‘primary’ eye being the left eye, and in some situations reassessing only eye in view (the right) as the ‘primary’. This apparently inherent behaviour in the eye-tracking software necessitated that the interfaces use duration-based logic for interpreting blinks and winks by the user.

## Interface Operation and Description

### *The Map*

The three developed interfaces are accessed by running a Python script using a single .bat file, to simplify usage. The map used for testing is taken as a web map tile directory of raster image tiles – these are parsed and stored as

PyGame surfaces for each zoom level available. The scale magnification between zoom levels is 2x.

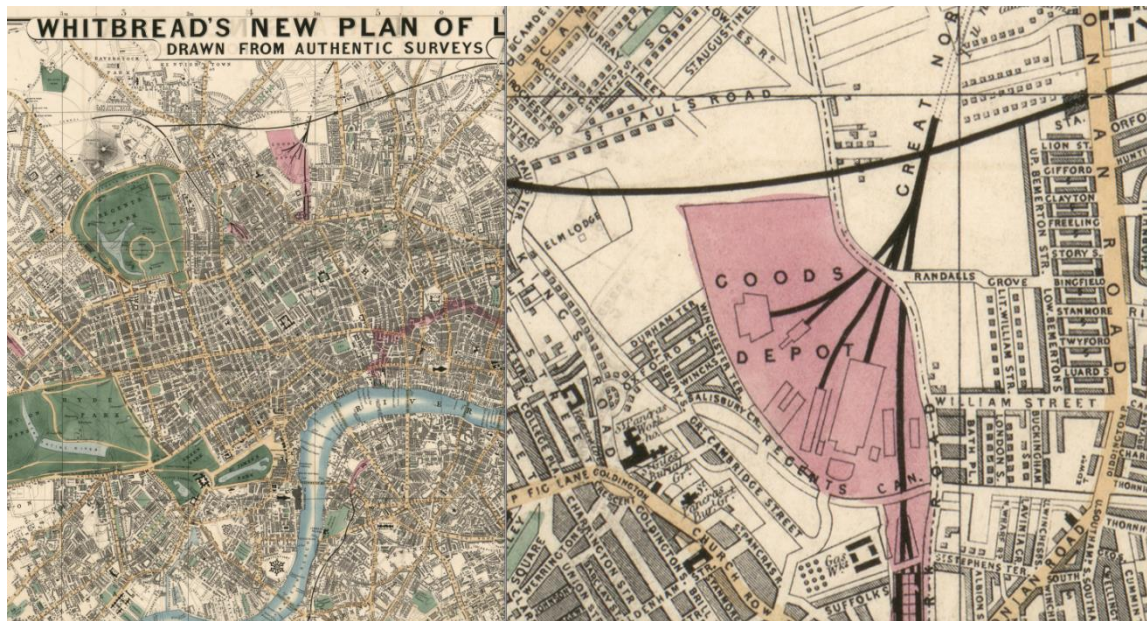


Figure 16: Example of Detail in Whitbread's Map of London at Different Scales

The map chosen for use in the user test is “*Whitbread's New Plan of London: Drawn from Authentic Surveys*”. It was selected as it is provided free for research purposes and being high resolution, dense with detail while still being clear and legible. This is enough to make it a large search space for potential user tasks that would encourage inspection and navigation at different scales, which was desired to encourage use of the developed interaction. Recalling that the intention of this study was to assess the developed interactions qualitatively based on usability and user experience, the impact of familiarity with the real location and the absolute task completion times were expected to be less important for analysis.

Panning functionality occurs in increments of one fifth of the corresponding map dimension, and occurs with a smooth panning animation controlled by a parameter set after trial and error. Zooming is performed without transition animation – the map is redrawn while keeping the zoom focus on some specific screen position, depending on the interface in use.

The following are descriptions of the interface functionality, as additions to the information of Table 2:

#### *Interface #1 – Mouse & Keyboard*

As described in Table 2, this interface is intended to be a baseline and to approximate to a conventional digital map interface. Mouse and keyboard panning was implemented in conventional ways. A major caveat was the lack of drag-panning with the mouse, which was not implemented due to lack of programming skill/time constraints.

#### *Interface #2 – Gaze-Supported Navigation*

This interface used the gaze location as an invisible pointer to direct panning and zooming. No visible gaze cursor is provided as displaying an erratically updating cursor over the map content was considered distracting, and not a necessary visual affordance. A semi-accurate location was thought to be sufficient once the desired location was still within the display area and near the zoom focus. Gaze positions beyond the bounds of the display screen were still treated as valid, allowing for potential zoom/pan based on off-screen

positions. All triggered interactivity - zooming, pan to gaze location, re-centering - were triggered via the specified keyboard keys, and were functionally identical to the mouse-click panning and keyboard-zooming of Interface #1.

### *Interface #3 – Eyes-Only Navigation*

Zooming was guided as with Interface #2, but triggered by the closing of a single eye. This performed the equivalent of a mouse-scrolling zooming for as long as the eye remained closed. The threshold duration set for the system to consider a closed eye signal to be a valid trigger was around 0.5 seconds. Re-centering was triggered using an intentional, voluntarily longer blink with both eyes simultaneously, with a threshold increased to around 0.75 seconds. These durations were set as somewhat longer than an involuntary blink. This was based on trial and error testing during development, arriving at what felt like the best responsiveness to deliberate blinks, aiming to mitigate unintended interaction while trying to preserve responsiveness.

Panning functionality was performed using the black border region the display screen, visible in Figure 17. Both orthogonal and diagonal panning was possible. These regions trigger panning in their corresponding direction when fixated on by the user, based on PyGaze's internal fixation detection threshold. These fixation-sensitive regions extended an equal area beyond the



border of the display screen, to allow for a larger effective trigger area, aiming for better responsiveness by mitigating the impact of jitter.

## Experiment Setup



*Figure 17: Experiment Setup, User Testing the Eyes-Only Interface*

The hardware and peripheral setup was as seen in Figure 17, with the monitor being in a fixed location. The user input devices, eye-tracker, and seating distance were all adjustable to the user's comfort. Before using interfaces #2 and #3, users were asked to get into a comfortable position

where they could access the needed devices. That resting position was where they felt comfortable and the eye-tracker had full view of their eyes. If the participant was wearing eyeglasses at a later time of day, attempts were made to adjust the eye-tracker and screen to angles and levels that would minimize the reflection of overhead lighting into the eye-tracker. This was not always successful - no single arrangement was ideal, as different participants' eyeglasses had different inclinations so changing their reflecting angles. In tests where the reflection issue was nearly solved, monitor height was a bigger influence than eye-tracker adjustment.

Additional experiment procedures were necessary to comply with the university regulations on hygiene and conduct due to COVID-19. All participants and the researcher conducting the experiment were required to keep their masks, as well as maintaining a 1.5m distance. Due to this, all observations of the participants face and screen content were done from the perspective in Figure 17. All spoken instructions were also issued as clearly as possible through the mask.

User tests consisted of three parts:

1. Briefing and Background Questionnaire
2. Interface Testing (in order 1, 2, 3)
  - a) User Tasks (randomized Trial Order)
  - b) Usability/UX Questionnaires
3. Post-Study Questionnaire and Debriefing

## Task Design

As the map interactivity being assessed was navigation (Pan & Zoom), the tasks created needed to encourage or guide the user towards using those interactions. To do this, three separate task sets (referred to as ‘trials’) were created, each consisting of three tasks in order. The tasks in each trial, and so the required navigation interactions, were intended to correspond to the following unordered model:

- *“Find a feature apparent from an overview of map, but requires close inspection to successfully complete”*
  - encouraging the user to use zooming interactions to inspect items of interest
- *“Find small-scale information, at a location far from the last task’s end, in the neighbourhood of a larger-scale feature”*
  - deliberately forcing reorientation of the user, requiring combinations of pan and zoom interactions
  - localized search requiring both user visual attention and control near simultaneously or in quick succession
- *“Follow an extended/linear feature to a described target/destination”*
  - exercising smooth or continuous control of panning function over larger amounts of geographic space

These model tasks would comprise each trial, and so allow for a minimum amount of use of the developed interactions. The trials used in this study are as follows:

(1) Trial #1

- (a) “Find the *stamp* on the *border* of the map, and tell me the date inside.”
- (b) “Find the *park* containing the ‘*Serpentine River*’.”
- (c) “There is a *railway* to the north of *Victoria Park* – what *red area* is it connected to?”

(2) Trial #2

- (a) “Please find the ‘*South Eastern Railway*’.”
- (b) “That Railway ends at a *red* terminal area – what is the closest *bridge*?”
- (c) “Please name the hills next to *Regents Park*.”

(3) Trial #3

- (a) “What *church* is within the *red outlined area* of the city?”
- (b) “Please find the ‘*Abbey*’ between the *Thames River* and *St.James Park*.”
- (c) “Please find the tunnel that crosses the Thames River.”

The trial order was randomly chosen for each participant, while the interface order was always #1~#3. This was meant to counter-balance differences in difficulty and its effect on interface feedback. As the number of participants obtainable was unknown, for every participant one of the six permutations of trial order was chosen until each was used.

## Questionnaire Structure

### *Background Questionnaire*

These gathered demographic information (age range and gender), vision quality and the use of glasses, and information on their cartographic background. For example, the questions “*How comfortable are you when using digital maps?*” and “*Have you created and/or designed digital maps for other persons to use?*” aimed to capture the expected level of competence and familiarity of users. High confidence and experience in map creation would signal the cartographic literacy of the user, which might colour their feedback on system novelty, engagement and performance.

User comfort, familiarity and knowledge of digital map creation was also asked in order to define the user-base being represented by the participants. It was anticipated that most participants would be cartographically literate, being connected to the Chair of Cartography at the host university. This would have been due to the social constraints during the time this study was being conducted. The maximum duration of the user test was targeted at 45

minutes, to make participation more palatable. This also motivated the use of standardized, multiple choice, shortened questionnaires where possible.

### *Usability & User Experience (U/UX) Questionnaires*

Described in the background research (see Section D) are these three questionnaires that were administered after each interface had been tested by completing all the tasks successfully. The System Usability Survey (SUS) and NASA Task Load Index (TLX) assess the usability aspects of the interface, while the shortened User Experience Questionnaire (UEQ-S) does so for user experience. All of these use numbered scales, with the SUS and UEQ-S being 5-point and the TLX being 10-point. Participants filled these questionnaires freely in the order they chose. The following are the concerns of the each of the questionnaires and some of their questions:

- The TLX - the demand experienced by the user – *“How hard did you have to work to accomplish your level of performance?”*, *“How mentally demanding was the task?”*
- The SUS – users’ thoughts on aspects of the interfaces’ usability – *“I thought it was easy to use”*, *“I needed to learn a lot of things before I could get going with this interface”*
- The UEQ-S – quick descriptors of the users’ experience with the interface – *“Obstructive / behindernnd”* vs. *“Supportive / unterstützend*

For the UEQ-S, both the English and German adjectives were provided as the questionnaire has been translated for use in both. Providing both was an attempt to minimize differences in interpretation and so gather more accurate responses.

### *Post-Study Questionnaire*

This questionnaire asked about the user's ranked preference of interface used, panning method, and zooming method. These rankings were meant to corroborate the analysis of the usability and user experience questionnaires. It also enquired about problems experienced during operation in order to capture any overarching issues with specific interactions implemented - "*Did you have any difficulties operating any of the interfaces? If 'Yes', please choose or state what they were below:*". All sections of this final questionnaire allowed for open feedback as well.

## Results

The user study was carried out during the period of the 28<sup>th</sup> of September to the 12<sup>th</sup> October 2020 at the Technical University of Munich. The computer display and user setup was arranged perpendicular to and away from the window; window blinds were used to minimize sunlight entering the room that would interfere with the eye-tracker. Some user tests occurred after sunset as well, in which case overhead fluorescent lights were used.

In total, 16 persons (8 male, 8 female) took part in the experiment. Due to an error in administering the first user test, user feedback for only Interface #3 (Eyes-Only Navigation) was collected. The age range of participants was predominantly of early to mid-aged adults, visible in Figure 8. Out of

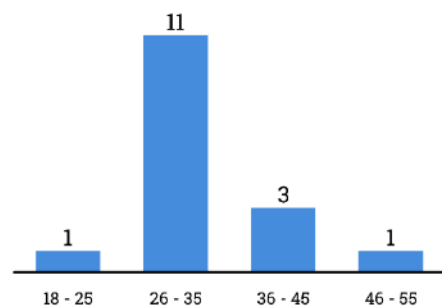


Figure 18: Number of Participants in each Age Range

the 11 participants that indicated they use corrective eyewear when using computer screens, 8 of them wore eyeglasses during the user test. This meant half of all participants wore eyeglasses during testing.

The cartographic background of participants was as strong as expected; 15 users indicated using digital maps regularly or more, and 14 users indicated experience making digital maps for use.



Users tested each interface in a fixed order: 1) Mouse & Keyboard, 2) Gaze-Supported Navigation, 3) Eye-Only Navigation. Interface #1 was considered a baseline from which to compare the other interfaces. Both interfaces #2 and #3 ran an eye-tracker calibration before starting. After the start of each interface and before issuing the first task, the user was given a spoken explanation of how the interface functionality would operate and allowed to test the interactions for a maximum of one minute.

Tasks were spoken aloud and given individually, and no time limit was stated. Which trial (the predetermined set of tasks) was given for each interface test was decided randomly, examples of which are visible below in the second and third columns of Table 3.

*Table 3: Sample of Data for three Users - Task Completion Times for each Interface (all times in seconds)*

User #	Interface #	Trial #	Task 1	Task 2	Task 3
2	1	3	3	72	74
	2	2	98	6	85
	3	1	58	21	160
3	1	1	13	13	61
	2	2	9	25	29
	3	3	8	27	74
4	1	1	8	14	13
	2	3	85	68	81
	3	2	310	87	104

Timing started at the end of announcing the task, and stopped once the user had correctly found the target; in the case of tasks that did not require a verbal answer, the experiment supervisor confirmed that the target was visible on screen. If the user required clarification or repeating the task, this was done during the recorded time. Along with the open-ended time limit, the reported times should not be interpreted as precise measurements of performance. Considering instances where announcement of a task needed repeating or clarification for the user, this could reasonably add up to 5 seconds of uncertainty to the task completion time. So, these measured times are instead intended to potentially corroborate other findings.

## Task Performance

All participants were able to successfully complete all tasks given to them. As Mouse & Keyboard map interaction is conventional, it was expected that task completion times using it would be the shortest overall, as well as the least variable. So interface #1 was considered the baseline for comparison.

<b>(Baseline) - Interface #1 - Mouse and Keyboard</b>											
Trial #	Avg. Trial Time	Count	CV	Difference to Baseline	Task 1	CV	Task 2	CV	Task 3	CV	
1	52	7	37%	0	12	120%	22	35%	18	110%	
2	32	4	27%	0	14	41%	12	35%	6	28%	
3	121	4	31%	0	10	76%	38	80%	72	39%	
<b>Interface #2 - Gaze-Supported Navigation</b>											
Trial #	Avg. Trial time	Count	CV	Difference to Baseline	Task 1	CV	Task 2	CV	Task 3	CV	
1	130	3	69%	78	37	118%	40	113%	53	55%	
2	83	6	66%	51	32	106%	20	42%	31	91%	
3	102	6	78%	-19	25	124%	27	75%	52	93%	
<b>Interface #3 - Eyes-Only Navigation</b>											
Trial #	Avg. Trial time	Count	CV	Difference to Baseline	Task 1	CV	Task 2	CV	Task 3	CV	
1	211	5	29%	159	40	57%	28	67%	143	50%	
2	259	5	73%	227	105	113%	96	94%	57	86%	
3	160	5	46%	40	26	57%	51	40%	83	61%	

Figure 19: Averages for User Task Completion Times (in seconds) for each Interface

The recorded task completion times show this to have held true overall, as shown in Figure 19. Interface #1 completed the trials the fastest (Column 2) with one exception - outperformed by interface #2 in task #3 by 19 seconds on average. Considering high values of the coefficient of variation (CV) across that trial by interface, it would indicate that there were users who completed

that trial very quickly. Looking at the minimum tasks times (see Figure 20) it shows a task completion time 46 seconds faster than the conventional interface.

Completion time for individual tasks and entire trials were noticeably higher for the non-conventional interfaces, especially interface #3. Figure 19 also shows the difference in the average time taken to complete entire trials as 'Difference to Baseline'. Based on this comparison to the Mouse & Keyboard interface, the Eyes-Only Navigation method appears to have taken longest on average to complete tasks using. However, the variability and limited number of measurements (Column 3) discourages drawing conclusions from this data alone, as noted earlier.

The per-task time differences (see Figure 21) also shows the trend of

Minimum Times		Trial		
		1	2	3
Interface	#1	34	20	69
	#2	52	38	23
	#3	125	35	73

Figure 21: Minimum Recorded Trial Completion Times (in seconds)

Average Time Difference to Baseline				
Interface	Trial #	Task 1	Task 2	Task 3
#2 - Gaze-Supported Navigation	1	25	18	35
	2	18	8	25
	3	15	-11	-20
#3 - Eyes-Only Navigation	1	28	6	125
	2	91	84	52
	3	15	13	11

Figure 21: Average Completion Times (in seconds) - Per Task Comparison to Mouse & Keyboard Interface

interface #2 having better completion times than interface #3. It may also suggests faster task completion using interface #2 in some cases, but the variability still discourages drawing strong conclusions.

By looking at the fastest trial completion times (see Figure 20) it shows interface #2 performing significantly faster

than the baseline in one single trial. With only weak trends visible from inspecting these times, the minimum trial time suggests the possibility that interface #2 could perform comparatively quickly to the conventional interface.

## Usability/UX Assessment

### Task Load

The NASA TLX rates the interfaces based on the users' self-reported experience of demand felt when using the interface to perform the tasks. The results are shown in Figure 22, along three different types of demand (*Mental, Physical, and Temporal*) and *Performance, Effort, and Frustration* reported by the user. The most conventional interface - #1 Mouse & Keyboard, was expected to be the most familiar and comfortable to users and so experience the least demand and highest competence. The results aligned well with this expectation, as it was rated lowest across all dimensions of demand, effort and frustration, and the highest self-reported performance. Gaze-supported

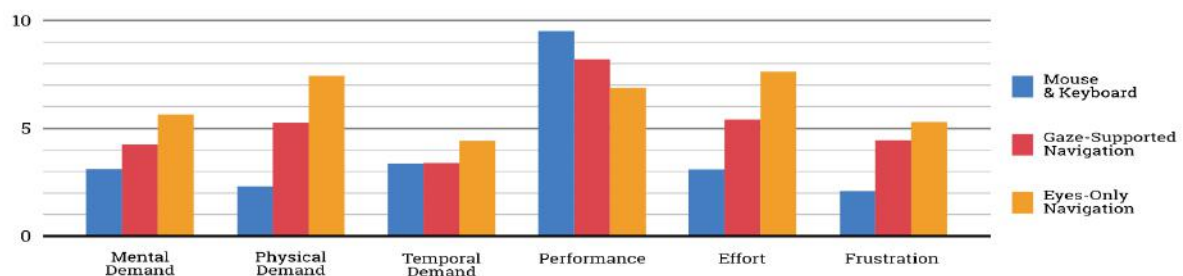


Figure 22: Results of the NASA Task Load Index Questionnaire

and Eye-only navigation both performed worse, with Eye-only navigation being rated as more demanding across all dimensions. Eye-only navigation also rated highly in physical demand and effort, 7.4 and 7.6 respectively, while the others were approximately rated 5 or lower.

### System Usability

As the conventional interface, Mouse and Keyboard was expected to rate at least acceptably on the System Usability Scale. An acceptable usability score on the SUS scale is around 68, shown as the dashed line on Figure 23. The Mouse and Keyboard achieve a SUS score of 81 from user feedback, which comfortably rates it as having good usability. By contrast, the SUS scores attained by Gaze-Supported and Eyes-Only navigation were 65 and 54. This rates Gaze-supported navigation as having marginally acceptable usability, while Eye-Only navigation has significantly poor usability.

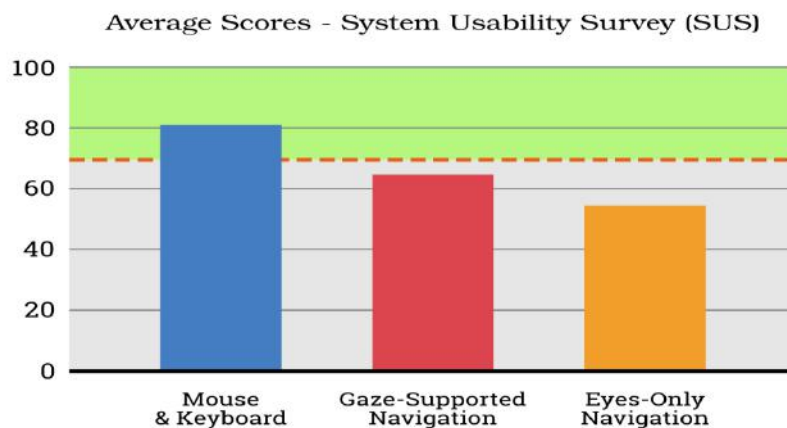


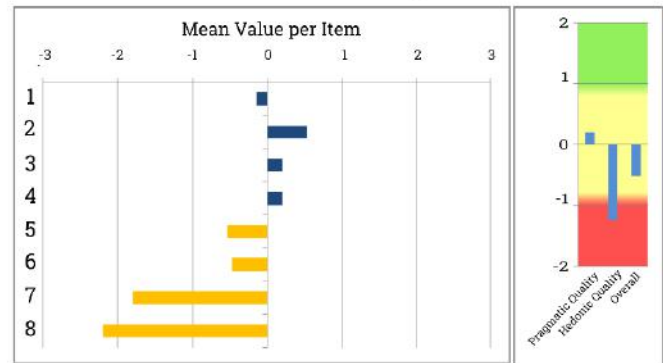
Figure 23: System Usability Scores, showing Acceptable Usability Score Threshold

## User Experience

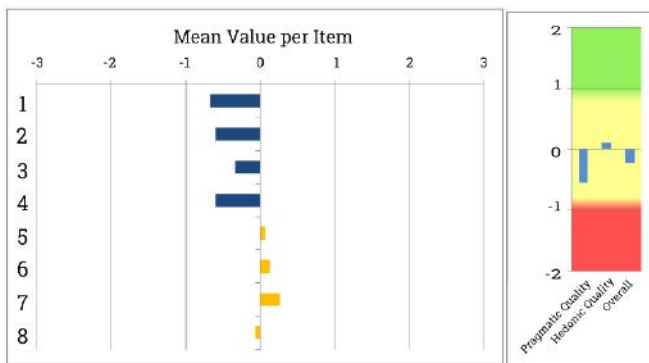
User Experience Questionnaire (UEQ-S)  
Items and their Related Scales

Scale	Item #	Negative	Positive
Pragmatic Quality	1	obstructive	supportive
	2	complicated	easy
	3	inefficient	efficient
	4	confusing	clear
Hedonic Quality	5	boring	exciting
	6	not interesting	interesting
	7	conventional	inventive
	8	usual	leading edge

Interface #1  
Mouse and Keyboard



Interface #2  
Gaze-Supported Navigation



Interface #3  
Eyes-Only Navigation

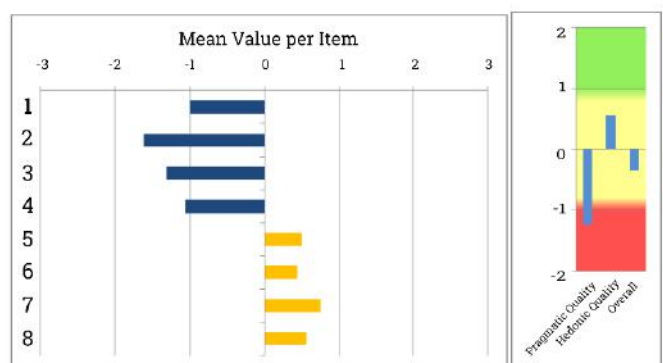


Figure 24: Results of the Short User Experience Questionnaire (UEQ-S), showing different aspects of user experience and the overall qualities

This shorter form of the UEQ rates the interfaces on their functional and pleasing qualities, termed *Pragmatic* and *Hedonic* Qualities. In Figure 24, the descriptive adjectives and which quality they are ascribed to is shown, as well as the final calculated values for *Pragmatic* and *Hedonic* Qualities and an overall rating of the user experience. Noted though is that extreme ratings, being values beyond +/- 2 on the scales, are uncommon as users avoid

extreme scale ends. So values beyond +/- 1 should be thought of as considerable (green/red regions), while below that considered neutral (yellow region). As the baseline interface and the most conventional, it is unsurprising that interface #1 has very high values in negative hedonic descriptors of 'conventional' and 'usual' - this is definitively what the interface is. The user experience was rated as expected: overall neutral with significantly poor hedonic quality. The overall experience of the Gaze-Supported Navigation is neutral. Eye-Only navigation attained some mildly positive hedonic quality while the goal-oriented experience of using the interface was significantly poor.



## Post-Study Feedback

### Rankings

These figures show how users ranked the interfaces and the interactivity that was available, in order of descending preference.

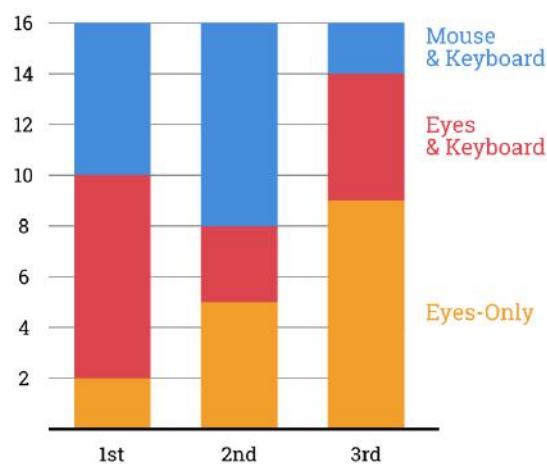


Figure 25: User Ranking of Interface By Preference

Figure 25 shows the responses to a question asking users to rank the interfaces used in order of their preference, from most preferred (1<sup>st</sup>) to least preferred (3<sup>rd</sup>). Many users had a strong inclination toward interface #2 – as many as preferred the conventional interface. A significant amount of users placed the same highly favoured interface #2 as last choice as well. This left the conventional method as a strong backup choice in 2<sup>nd</sup> place, while the Eye-Only interface #3 was comprised around 50% of users' last resort.

The individual interactions for Panning and Zooming were ranked as well. Figure 26 shows the preferences of zooming interactions, with

conventional mouse-based scrolling zoom dominating the first preference. The ranking of blink-based zooming follows the same trend as the Eye-Only interface it is part of, being relegated strongly towards last choice. The “Zoom Keys” method is about equally represented throughout, which could indicate a general indifference to the method or perhaps a misunderstanding of what was meant by the option.

Figure 27 shows the panning interactions’ ranking, where the Gaze-Supported Panning (Gaze Location + Key Press) has a strong first choice preference. This strongly corresponds to the interface preferences in Figure 25, though for panning the preference appears to have a more solid first place position. The keyboard-based directional panning (Directional Keys) appears to trend similarly to its interface in Figure 25, suggesting it was a preferred fall-back method. Preference about the mouse-based panning to clicked location was evenly spread out. The gaze-aware screen borders were a strong last choice, following the trend of interface #3.

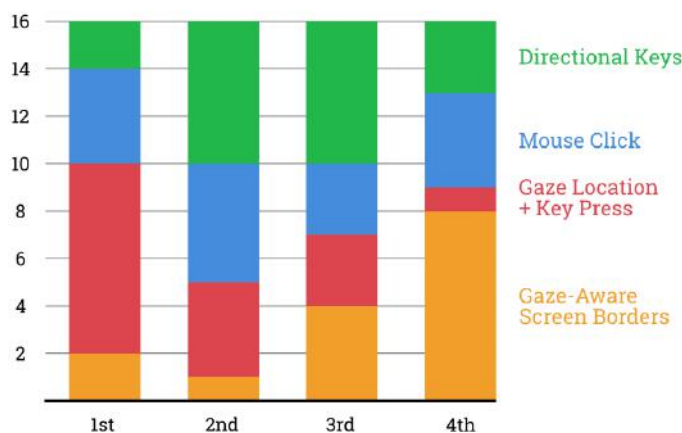


Figure 27: User Ranking of Panning Interactivity

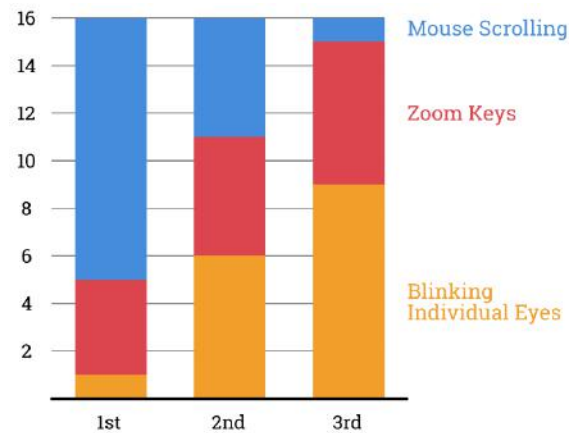
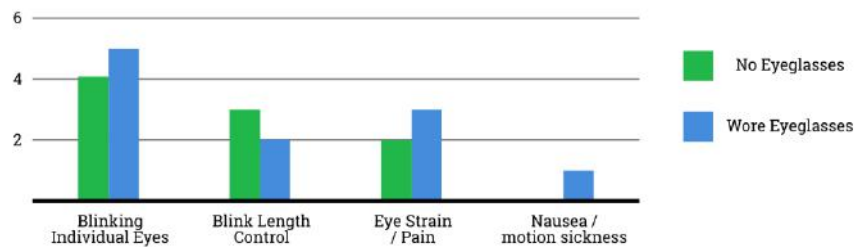


Figure 27: User Ranking of Zooming Interactions

## Issues and Observations



*Figure 28: Numbers of Participants Reporting Problems Encountered While Using the Interfaces*

Out of the 16 participants, 12 reported experiencing at least one of the anticipated problems placed at the end of the post-study questionnaire. Looking at Figure 28, none of these specific problems appear distinctly associated with the wearing of eyeglasses during the user test.

However, 3 of the 8 participants wearing eyeglasses had poor calibration that could not be remedied by adjusting seating position, computer screen or eye-tracker orientation. Then, the poor calibration was accepted and the user allowed to use the interface. In one case, it was noted that the reflection of the overhead fluorescent lights off of the eyeglasses would coincide with the corneal reflection seen by the eye-tracker. A solution in some cases was altering the computer screen's height.

The most common problem across both types of participants was the inability or difficulty in 'winking'/blinking individual eyes. Some persons were could only close their left or right eye independently, hindering their ability to use the zooming interaction of interface #3. Some participants commented

that they would confuse which eye corresponded to which zooming function – in or out. One attempted solution independently arrived at by participants was to block the desired eye with their hand – this worked with varying stability and reliability.

An instability or erratic behaviour of interface #3 appeared to be a switching or failure of the wink-based zooming interaction, for reasons that were not readily apparent. In some cases, left and right eye zooming behaviour would switch during use, and may not revert back to the original intended behaviour. In development of interface #3, the eye-switching behaviour was noted, and so could have been the main cause of this.

Users were observed to move their head in the desired direction of panning while using interface #3. This appeared to occur due to frustration at failing panning using the gaze-aware border regions. It rarely occurred during use of interface #2 - usually in tests where there were responsiveness issues.

Some users lean in to computer screen during use, and head motion could lead to the eye-tracker losing view of the user's eyes. Within the developed interfaces, there was no overlay or indication of when the user was within view or in a good position. This led to frustration and extra time spent in tasks, attempting to regain a correct position. So an overlay or indicator allowing the user to reposition themselves in view should be considered necessary for interfaces aiming to provide both eye-control and freedom of movement.

So, contributors to unreliable and unpredictable behaviours are:

- poor calibration,
- loss of eye-tracker lock on eyes due to the user moving out of eye-tracker view,
- reflected overhead lighting swamping the corneal reflection.

Despite whatever challenges faced when using the developed interfaces, all users expressed that the user study and experience of explicitly controlling a map using their eyes was either interesting or exciting. Specifically mentioned was the satisfying responsiveness and speed of interface #2, while criticizing the unreliability of interface #3 and new physical demand to wink.

## Discussion

Of the three interfaces tested, the conventional interface (interface #1) was expected to be superior in task performance and in user-reported usability. This expectation was based on how ubiquitous and well-established it is. The results of the user tests indicates generally poorer performance by the developed interfaces compared to the conventional interface of mouse and keyboard. Task completion times and their variation were worse the more explicit eye-control the interface relied on.

The level of demand to operate the interface #1 was also the lowest, as expected, shown by it having the best (lowest) scores in the NASA Task Load Index results. This corresponds to its highest rated aspect in user experience being 'easy', while being very 'usual' and definitively 'conventional'. As for its preference by users, it was a strong 1<sup>st</sup> and 2<sup>nd</sup> choice out of the three interfaces; users' strong preference for it was expected, but the degree to which it was replaced by interface #2 speaks to the positive reception by users. For zooming, it was the clear 1<sup>st</sup> choice for most users, perhaps due to obstructive natures of the other zoom triggers - interface #2 using multiple keystrokes versus a fast visceral scrolling zoom, and interface #3 being unreliable, unstable and frustrating due to the troublesome winking requirement. This would correspond to those interfaces having appropriate preferences, with interface #2 having an even spread while the unreliable

interface #3 heavily left as last resort. However, mouse-based panning did not have an established strong preference, instead being spread somewhat evenly throughout. This is likely not entirely due to the strength and quality of panning with interface #2, but instead that the implementation of mouse-based panning (move to clicked location) is not the actually expected method. As noted by Roth (2013a) in commenting on the expected ability to ‘probe, *drag*, and adjust’, mouse-based panning using dragging is the stronger expectation. Some participants did immediately attempt to perform this interaction when using interface #1. Should this dragging interactivity be available, it would be reasonable to expect a similarly strong preference to it as with the scrolling zoom. Further credence to this comes from the strong preference for conventional panning methods as a 2<sup>nd</sup> choice, around 70% of participants.

However, the performance and reception of the developed interfaces were significantly positive both in feedback and measurements. While the gaze-based interfaces did not score well in the usability assessment, interface #2 scored a value of 65 on the SUS scale, near to value of 68 considered an acceptable level of usability. Several users demonstrated faster or equivalent times using that interface, compared to the baseline set with interface #1 (see Figure 19). While users reported a somewhat greater demand to operate it, they also rated it as performing almost as well as the conventional interface. Along with the favourable comments from users about the feel of operating it, it performed better than the conventional interface in the user experience

assessment dimensions by being less boring and more novel by comparison. The user rankings show encouragingly large amounts of users (50%) preferred this interface as their first choice; its panning method was also ranked a first choice by half of users. These are strong indicators that for map navigation, the gaze-supported navigation method may be a competent and satisfactory alternative to current mouse-based navigation.

The performance of the Eyes-Only interface was more lacklustre – there was a lower preference for its interactivity, slower task completion speeds and greater overall load experienced by users. Important points to note for this interface are the technical problems and challenges that likely negatively impacted its reception: the instability in the behaviour of the interface's response, and the inability of users to wink. Users not being able to enact the trigger for zooming functionality was naturally frustrating and would have fed in to the longer task times, higher demand, significantly low goal-oriented qualities of the user experience, and its lower ranking in preferred interface. Physical demand and effort are noticeably higher in the Task Load Index, corresponding to these difficulties in operation as two-thirds of users reported difficulty blinking/winking their eyes. When facing trouble attempting to trigger the panning interactions via the border regions, users would often exhibit the intuitive behaviour of stretching or tilting their head out in the desired panning direction. Despite these challenges, this Eye-Only interface received higher scores in the stimulating and engaging aspects of the user experience assessment. This corresponds with the user comments on the



novelty and interesting experience of controlling the interface with their eyes. Notably, though its preference by users was low, 7 users still rated it above last preference (2 as 1<sup>st</sup>, 5 as 2<sup>nd</sup>). Overall, while the interface was considered more novel and had a worse usability and experience, the preferences of users indicate that there was supportive interest in such interaction methods. A more stable and user-compatible version of the interface may receive better reception if more universally executable eye-actions are used, such as some specific eye-gestures.

Between panning and zooming, it appeared that eye-based panning methods showed more promise by users showing greater acceptance of interface #2. This is likely skewed by the unsatisfactory method of mouse-based panning provided, and the troubles around triggering the zooming functions of interface #3. Perhaps with a more reliable performance, by using a more universally capable eye-gesture, an eye-only zooming interaction would be better received.

Key factors resulting in poor usability were the:

- interactivity implementation (eye-tracker behaviour/quality; interface programming)
- user characteristics (eyeglasses; viewing behaviour)

Despite the poor usability and user experience, all participants completed all tasks given to them using the interactivity implemented. This demonstrates that gaze-supported and/or explicitly eye-controlled map

navigation is possible, and reinforces that it can be comparable in speed given ideal circumstances. This would agree with the work on gaze-supported navigation by (Stellmach & Dachsel, 2012a). Eyes-only map navigation is of greater immediate value to the motor-impaired and they are a user-base whose mobility constraints minimize the negative impact from user motion.

A currently relevant general purpose application is also now in hygiene-critical public settings, such as interactive mall maps, allowing for contactless and hands-free information services. The time cost of user calibration can now be considered worth it in exchange for the security of being contactless.

### Research Questions

This thesis aimed to find ways to facilitate directly eye-controlled user-map interactions, by assembling eye-controlled interactions into a map interface and evaluating its performance. The answers arrived at to the research questions involved in fulfilling this goal are described below:

- 1) What pairing of fundamental digital map interaction and eye-based interaction could produce a usable eye-based map interface?

The literature reviewed showed a set of expected cartographic interaction expected and from this, the Pan & Zoom interactivity was chosen as the scope of thesis. This choice was based on its typical implementations using an interaction model which has analogies in eye-based techniques, specifically gaze-based pointers. As shown in the Methodology section, the operation of

pointer-based map navigation is conceptually separable into a “pointer” and “trigger” operations, which could be matched with equivalent eye-based control methods. So it followed that these eye-controlled operations alone, or mixed with the conventional hardware interfaces (Mouse & Keyboard), could provide the interaction necessary for an eye-controlled or gaze-assisted map navigation interface.

2) Can an implementation of gaze-based control and/or awareness provide beneficial map interactions?

From the results of the user study, specifically the user task completion times, it has been shown that both of the developed interfaces facilitated map navigation. This was true for all users, including those who experienced poor or missing functionality during their tests. From the successful completion by all participants, it can be stated that gaze-based navigation interfaces can provide the necessary interactivity for this map use.

With respect to the interactions being beneficial, the individual performances by users as well as the usability assessment show that there is potential for an experience on par with the conventional Mouse and Keyboard interface. While the conventional interface performed more consistently and quickly in task completion time, as expected, in some cases users performed tasks similarly or better with one of the gaze-based interfaces. Alongside its several faster user performances, interface #2 was assessed to have near acceptable usability, to be only slightly more demanding, while more

stimulating to users. Users demonstrated a very strong preference for it, marginally more than the conventional interface. Users also found the fully eye-controlled interface #3 significantly more stimulating as well and still showed some preference to it despite the operational difficulties. These takeaways are strong indicators that gaze-based interactivity has a strong potential for benefits to operational performance as well as user engagement.

Lastly, this user study highlighted a couple challenges to successful gaze-based navigation interfaces. The presence and interplay of users' glasses and the usage scenario, specifically lighting and orientation of the eye-tracker, were significant impediments to successful interface operation. Troubleshooting was not always possible, and these factors may not be easily controllable or predictable in wider use cases. The other challenge was the assumption about user capabilities, specifically being able to wink both eyes individually. While this was specific to this user study in interface #3, the general consideration is that the physical method of affecting the interaction should be one your user base is capable of.

Acknowledging these considerations, it is shown that an implementation of gaze-based interaction can indeed provide the map navigation required, and shows possible benefits to performance and engagement.

## Conclusion

To assess the possibility of beneficial eye-controlled cartographic interaction, two interfaces were developed to perform Pan and Zoom interactivity on a static map. One interface used gaze location to guide interaction triggered using a keyboard (Gaze-Supported navigation), while the other used eye blinks and gaze-aware map regions to be fully eye-controlled (Eyes-Only navigation). These interfaces were compared to a conventional interface, which used a mouse and keyboard combination.

A task-based user study was conducted, with users testing all three interfaces, and providing feedback to assess each interface's usability and user experience using standardized questionnaires (SUS, NASA TLX, UEQ-S). The conventional interface performed as expected, while the developed gaze-based interfaces were assessed more poorly by these standardized questionnaires. Assessment of user feedback showed a strong preference for the gaze-supported navigation interface, and a disfavour towards the Eye-Only interface. For gaze-supported navigation the contributing factors for this preference likely were: the lack of dragging to pan with mouse, and positive aspects of the user experience when using the gaze-supported interaction. Regarding the poor evaluation of the Eyes-Only interface: based on observations and user assessments, this was understood to be due to both the instability of the gaze-based interfaces and the users' difficulty in performing the triggering actions (winking). Despite the challenges in using

the interfaces, users were able to complete all interaction tasks successfully, indicating that explicit eye-controlled map navigation is possible; user preferences and achieved performance strongly indicate potential benefits of gaze-involved navigation, in achievable operational speed and user stimulation.

## Outlook

Considering the promising performance results and reception by users, further studies evaluating gaze-based interaction for navigation and geodata exploration are likely to be beneficial. More skilful and comprehensive implementations of gaze interactivity matched with larger, more controlled experiments could lead to feasible alternatives to mouse-cursor navigation. Of the expected map interactivities, the *'Retrieve'* interactions are similarly tied to the pointer-based interaction model; research into geodata comprehension and analysis could be done following a methodology similar to this thesis. Competent gaze-navigation could find adoption by professionals, offering improved speed in cartographic workflows or geodata exploration.

Contactless interfaces are relevant in hygiene-sensitive settings, such as medical procedures or increasingly so in public spaces; interactive map services in malls and the like are arguably less safe due to the reality of widespread infectious diseases. Competent gaze-enabled map interfaces in these situations could offer increased accessibility to the physically impaired or differently abled persons, who may be unable to perform larger scale

physical gestures. Health concerns aside, an alternative suite of efficient eye-based interactivity can make the mouse-reliant GIS and cartographic workflows accessible to the physically impaired.

## References

- Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, 24(6), 574–594.  
<https://doi.org/10.1080/10447310802205776>
- Bergstrom, J. R., & Schall, A. (2014). *Eye Tracking in User Experience Design*. Morgan Kaufmann.  
<http://proquest.tech.safaribooksonline.de/9780124081383>
- Biedert, R., Buscher, G., & Dengel, A. (2010). The eyeBook – Using Eye Tracking to Enhance the Reading Experience. *Informatik-Spektrum*, 33(3), 272–281. <https://doi.org/10.1007/s00287-009-0381-2>
- Brooke, J. (1996). SUS: a “quick and dirty” usability. *Usability Evaluation in Industry*, 189.
- Brychtova, A., & Coltekin, A. (2016). An Empirical User Study for Measuring the Influence of Colour Distance and Font Size in Map Reading Using Eye Tracking. *The Cartographic Journal*, 53(3), 202–212. <https://doi.org/10.1179/1743277414Y.0000000103>
- Chen, J., Cranton, W., & Fihn, M. (Eds.). (2016). *Handbook of Visual Display Technology*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-14346-0>
- Dalmajer, E. S., Mathôt, S., & Van der Stigchel, S. (2014). PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods*, 46(4), 913–921. <https://doi.org/10.3758/s13428-013-0422-2>
- Davies, C. (1998). Analysing “work” in complex system tasks: An exploratory study with GIS. *Behaviour & Information Technology*, 17(4), 218–230.  
<https://doi.org/10.1080/014492998119427>
- Dix, A. (Ed.). (2004). *Human-computer interaction* (3rd ed). Pearson/Prentice-Hall.



- Duchowski, A. T. (2017). *Eye Tracking Methodology: Theory and Practice* (3rd ed.). Springer International Publishing.
- Eyedrivotomic*. (n.d.). Eyedrivotomic. Retrieved October 8, 2020, from <https://eyedrivotomic.org/>
- EyeGaze*. (n.d.). Eyegaze Inc. Retrieved October 9, 2020, from <https://eyegaze.com/users/spinal-cord-injuries/>
- EyeMine*. (n.d.). Retrieved April 16, 2020, from <https://www.specialeffect.org.uk/eyemine>
- Giannopoulos, I., Kiefer, P., & Raubal, M. (2015). GazeNav: Gaze-Based Pedestrian Navigation. *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 337–346. <https://doi.org/10.1145/2785830.2785873>
- Göbel, F., Kiefer, P., Giannopoulos, I., Duchowski, A. T., & Raubal, M. (2018). Improving map reading with gaze-adaptive legends. *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications - ETRA '18*, 1–9. <https://doi.org/10.1145/3204493.3204544>
- Haklay, M. M. (Ed.). (2010). Interacting with Geospatial Technologies. In *Interacting with Geospatial Technologies* (pp. i–xv). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470689813.fmatter>
- Hansen, D. W., & Qiang Ji. (2010). In the Eye of the Beholder: A Survey of Models for Eyes and Gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3), 478–500. <https://doi.org/10.1109/TPAMI.2009.30>
- Harrower, M., & Sheesley, B. (2005). Designing Better Map Interfaces: A Framework for Panning and Zooming. *Transactions in GIS*, 9(2), 77–89. <https://doi.org/10.1111/j.1467-9671.2005.00207.x>

ISO. (2019). *ISO 9241-210:2019*. International Standards Organization.

<https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/07/75/77520.html>

Kiefer, P., Giannopoulos, I., Raubal, M., & Duchowski, A. (2017). Eye tracking for spatial research: Cognition, computation, challenges. *Spatial Cognition & Computation*, 17(1–2), 1–19.

<https://doi.org/10.1080/13875868.2016.1254634>

Klamka, K., Siegel, A., Vogt, S., Göbel, F., Stellmach, S., & Dachsel, R. (2015). *Look & Pedal: Hands-free Navigation in Zoomable Information Spaces through Gaze-supported Foot Input*.

<https://doi.org/10.1145/2818346.2820751>

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research Methods in Human-Computer Interaction*. Elsevier Science & Technology.

<http://ebookcentral.proquest.com/lib/munchentech/detail.action?docID=4851896>

Lewis, J. R., & Sauro, J. (2018). *Item Benchmarks for the System*. 13(3), 10.

Maceachren, A. M., & Ganter, J. H. (1990). A pattern identification approach to cartographic visualization. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 27(2), 64–81. <https://doi.org/10.3138/M226-1337-2387-3007>

MacEachren, A. M., & Monmonier, M. (1992). Introduction. *Cartography and Geographic Information Systems*, 19(4), 197–200. <https://doi.org/10.1559/152304092783721303>

Majaranta, P., Ahola, U.-K., & Špakov, O. (2009). Fast gaze typing with an adjustable dwell time. *Proceedings of the 27th International Conference on Human Factors in Computing Systems - CHI 09*, 357. <https://doi.org/10.1145/1518701.1518758>

- Majaranta, P., & R ih a, K.-J. (2002). Twenty Years of Eye Typing: Systems and Design Issues. *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*, 15–22.  
<https://doi.org/10.1145/507072.507076>
- Open Gaze API Version 2.0*. (2013). [https://www.gazept.com/dl/Gazepoint\\_API\\_v2.0.pdf](https://www.gazept.com/dl/Gazepoint_API_v2.0.pdf)
- Plesnick, S., Repice, D., & Loughnane, P. (2014). Eye-controlled wheelchair. *2014 IEEE Canada International Humanitarian Technology Conference - (IHTC)*, 1–4.  
<https://doi.org/10.1109/IHTC.2014.7147553>
- Porta, M., Ravarelli, A., & Spagnoli, G. (2010). *CeCursor*, a contextual eye cursor for general pointing in windows environments. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10*, 331. <https://doi.org/10.1145/1743666.1743741>
- PyGaze | documentation*. (n.d.). Retrieved October 13, 2020, from  
<http://www.pygaze.org/documentation/>
- Roth, R. E. (2013a). Interactive maps: What we know and what we need to know. *Journal of Spatial Information Science*, 2013(6), 59–115. <https://doi.org/10.5311/JOSIS.2013.6.105>
- Roth, R. E. (2013b). An Empirically-Derived Taxonomy of Interaction Primitives for Interactive Cartography and Geovisualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12), 2356–2365. <https://doi.org/10.1109/TVCG.2013.130>
- Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017a). Construction of a Benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(4), 40. <https://doi.org/10.9781/ijimai.2017.445>
- Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017b). Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(6), 103. <https://doi.org/10.9781/ijimai.2017.09.001>

- Stellmach, S., & Dachsel, R. (2012a). *Investigating gaze-supported multimodal pan and zoom*.  
<https://doi.org/10.1145/2168556.2168636>
- Stellmach, S., & Dachsel, R. (2012b). Look & touch: Gaze-supported target acquisition. *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems - CHI '12*, 2981.  
<https://doi.org/10.1145/2207676.2208709>
- Stellmach, S., Stober, S., Nürnberger, A., & Dachsel, R. (2011). Designing gaze-supported multimodal interactions for the exploration of large image collections. *Proceedings of the 1st Conference on Novel Gaze-Controlled Applications*, 1–8. <https://doi.org/10.1145/1983302.1983303>
- Tolochko, R. (2016). *Contemporary Professional Practices in Interactive Web Map Design*.
- van Middendorp, J. J., Watkins, F., Park, C., & Landymore, H. (2015). Eye-tracking computer systems for inpatients with tetraplegia: Findings from a feasibility study. *Spinal Cord*, 53(3), 221–225.  
<https://doi.org/10.1038/sc.2014.219>
- Vickers, S., Istance, H., & Smalley, M. (2010). EyeGuitar: Making rhythm based music video games accessible using only eye movements. *Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology - ACE '10*, 36.  
<https://doi.org/10.1145/1971630.1971641>
- Whitbread, J., Jeffreys, W. & Fillmore, M. (1853) Whitbread's new plan of London: drawn from authentic surveys. [London: Published by J. Whitbread, mapseller] [Map] Retrieved from the Library of Congress, <https://www.loc.gov/item/2015591076/>.
- Yasumura, M., Itoh, K., Ichikawa, A., & Okamoto, A. (2006). People with Motor and Mobility Impairment: Human Computer Interaction, Rehabilitation. In K. Miesenberger, J. Klaus, W. L. Zagler, & A. I. Karshmer (Eds.), *Computers Helping People with Special Needs* (pp. 902–904). Springer. [https://doi.org/10.1007/11788713\\_131](https://doi.org/10.1007/11788713_131)

