University of Zurich, Department of Geography
Geographic Information Visualization and Analysis (GIVA)

Assessing route memorizing and visual search performance of focus and context maps

GEO 511 - Master's Thesis

Author	Lorenz Bosshardt 08-714-933
Advisor	Dr. Arzu Çöltekin University of Zurich
Co-Advisor	Dr. Kai-Florian Richter University of Zurich
Faculty Member	Prof. Dr. Sara Irina Fabrikant University of Zurich

Submission	June 2014

Abstract

Planning a way from A to B was never as easy as nowadays. Online route planners relieve wayfinding by providing map based route information. However, map reading and information interpreting still remain challenging. A conventionally used method to highlight a way is often just a simple line overlying a base map. Yet, it is questionable if this guiding technique appropriately leads the map reader to the important information. Therefore, the present thesis proposes a route highlighting approach based on the idea of facilitating information retrieval by providing only full information in a focus zone along the way and diminishing it in less relevant map areas (context zone). To understand how such a focus and context highlighting method works and whether it facilitates map reading, a human subject study was conducted. 24 participants took part in an experiment, involving the recoding of users' gaze movements. The study comprised a route memorizing task where subjects were requested to retrace the learned route on another base map and a visual search task concerning detection of-off route information. The within-subject design implied two routes from different map excerpts. Each map was depicted with the focus and context highlighting as well as with the common line highlighting method. The result showed no performance difference between the two route highlighting methods regarding both tasks. On the contrary, eye-tracking data revealed a partly different gaze behavior, indicating that the proposed focus and context method draw slightly more attention on route adjacent objects. Based on these results, the study provides suggestions for future assessments of focus and context maps in order to improve this kind of wayfinding assistance.

Acknowledgements

First of all, I would like to thank my advisor Dr. Arzu Çöltekin and co-advisor Dr. Kai-Florian Richter who gave me precious assistance and supported me throughout the entire work process.

Special thanks go to Annina Brügger and Laureen Bauer who took the time to revise the thesis.

Further, I am deeply grateful to all participants who voluntarily took part in the experiment.

Last but not least, many thanks to my friends at university for having had challenging table tennis matches and good discussions during recreative breaks.

Contents

List of Figures

List of Tables

1 Introduction

1.1. Motivation

Online map services have taken their places among the most popular cartographic products in recent years. They are mostly freely accessible and provide a vast range of functionalities. Furthermore, the digital technology has enabled new ways of usage that offers users information on the go for an increasing number of tasks. Regarding interactive map tools, one of the great benefits of online maps is the navigation function that allows users to find a route between any given points. In addition to written or verbal instructions, almost all automatic route planners also provide graphical information in form of a map. Therefore, route planning is not anymore a challenge for users themselves, since route information is getting more and more accurate and can even be adapted to individual preferences. Nevertheless, even when the cost of preparing the queried path is enormously reduced, map reading and applying spatial knowledge to real world environments still remains challenging.

Using an online route planner to find a way from A to B, the computed route is very often given as a visually highlighted path. Seemingly, the most common highlighting method is simply a line, overlaying the suggested route, thus, distinguishing it from the rest of the map's path network. Figure 1.1 is an example of such this common route highlighting method. Throughout this thesis, this line highlighting method is abbreviated as "LH". Freksa (1999) considers these maps as a kind of customized wayfinding support where the base map provides necessary context information. Besides visual wayfinding support, a route can also be found having verbal or written path descriptions. In contrast to maps, these kinds of navigational support often only refer to specifically chosen instructions that must be taken along the route itself, whereas maps provide more than only sequential knowledge. Nevertheless, for showing a route on a map, many aspects of the map may not be necessary to emphasize the path. Hence, a reference map might be too general for simple wayfinding tasks and therefore not economically optimized.

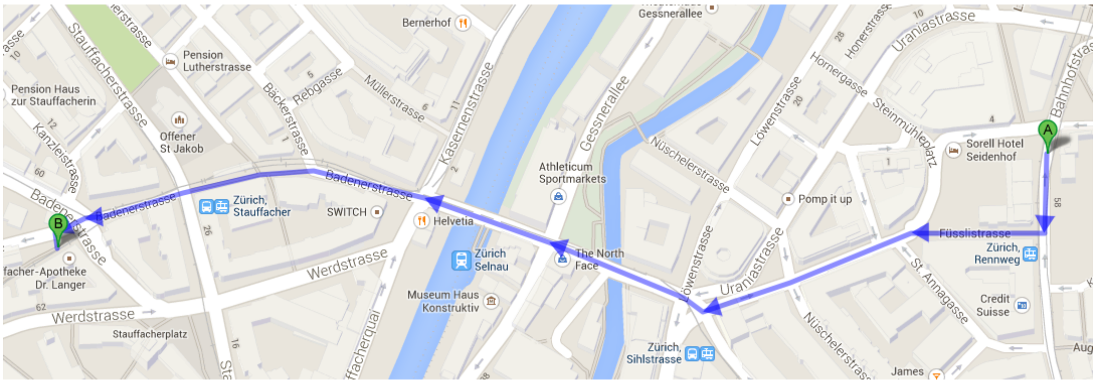


Figure 1.1.: Common route highlighting method as provided by Google (2010)

An approach that combines the advantage of a map depiction with the advantage of the selectiveness in verbal route description are *sketch maps*. These maps depict those aspects, which would be named in a verbal direction giving, neglecting all the others. Hence, they can be regarded as visually enhanced verbal route descriptions, where important spatial features are emphasized. This kind of heavily compiled maps can only be used to go from A to B on one particular route, thus, they are made just for one particular task. Figure 1.2 shows an example of such a sketch map.

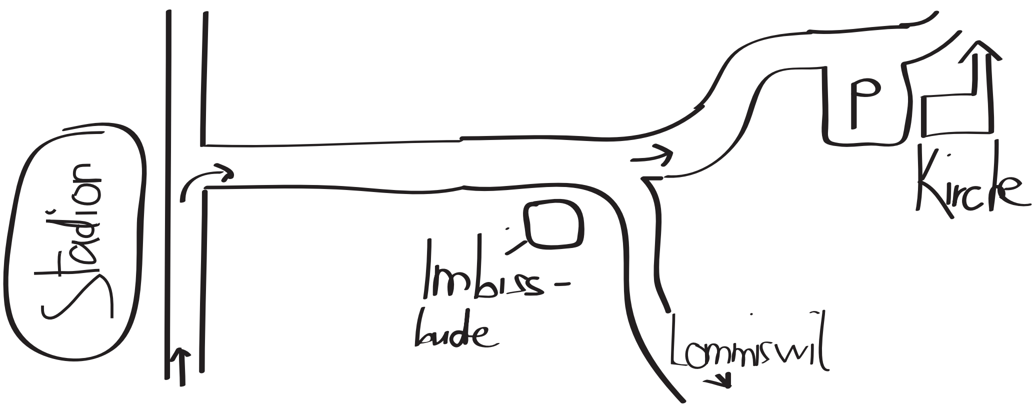


Figure 1.2.: Example of a sketch map

A second approach would be so called *schematic maps*. Theses maps provide assistance for a given class of wayfinding tasks. Transportation network maps such as the iconic London Underground Map[1] (Figure 1.3), are good examples of these kind of maps, where domain specific aspects are depicted in a more or less maintained spatial relationship, whereas non-domain aspects are ignored. In contrast to sketch maps, schematic maps typically maintain more spatial relations and are more complete in regard to entities that are depicted. Nevertheless, no crisp boundary between these two maplike representations can be drawn.

Figure 1.3.: London tube map: example of a schematic map

A third concept that also deals with task specific maps is called *focus maps* (Zipf and Richter, 2002). Like sketch and schematic maps, the goal of these maps is to give map based assistance in wayfinding tasks. As the name says, these maps put the focus on specific areas or selected map features or both. Besides information in the focus areas, context information is also provided. Therefore, information can be graduated according to its importance, creating maps with different levels of focus zones. Doing so reduces the overall amount of information on the display and funnels the reader's attention to the relevant information at the same time. Different focus areas can be emphasized using various highlighting methods, intensely depicting features in the focus zone(s) and diminishing them in less relevant areas.

[1]London Underground Map: http://www.tfl.gov.uk/maps/track/tube

To come back to the representations of any common online route planner (Figure 1.1), the most important region for navigational map reading is primarily the area around the actual route itself. The map underlying the highlighted route serves as a basic orientation aid, providing reference points and clues about the setting and context of the region the route is embedded in. Bearing in mind the concept of focus maps, it can be questioned if the common LH method is the most appropriate route guiding aid. Depending on how the base map is designed, an overload of information mainly in areas further away from the route could occur. The base map might be too general thus supplying too many pieces of information for any specific task. Since human working memory has a limited capacity, excess information or its disorganization can cause difficulties in map reading, leading to a decrease in recognition and performance (Kemps, 1999). On the other hand, a lack of information makes it hard to characterize a route at all.

To find out more about how the LH methods work, a simplified highlighting method, based on the concept of focus map by Zipf and Richter (2002), is proposed for this study. Similarly to the common LH method, the proposed method puts the actual path in the foreground. However, the proposed approach distinguishes between focus and context regions. Due to the division between focus and context, the proposed highlighting method is abbreviated as FC method. Figure 1.4 shows an example of a FC map which is based on Goolge Maps (2014).

Figure 1.4.: Example of a focus and context map based on Google Maps (2010)

In such a map, the context zone is everything in the display apart from a buffer area around the actual route. In the context area, color saturation and lightness are reduced, resulting in a black and white map, which is kept in darkish tones of grey. The buffer zone that describes the route and its proximate area is therefore denoted as focus area which is depicted in full color. In short, the proposed approach utilizes with the visual variable "color" (Bertin, 1983); dividing the map into two zones, while only providing full color information in a buffer area around the route.

1.2. Research Questions

The proposed highlighting method (FC) is applied for task specific maps, where the goal of a user's map consultion is given in advance. In contrast to the abovementioned line highlighting method, the so called focus and context approach tries to reduce visual map complexity by dimming out color in the non-relevant areas. Nevertheless, having theoretically argued for the proposed method, the question arises how it is actually used in map reading task. In order to compare the difference of retained information between LH and FC method, *map memorizing* performance can be analyzed. Besides wayfinding and memorizing tasks, map reading always involve some kind of *visual search*. Having the LH and FC method, it is also interesting to see whether there is a difference in search performance in the context area between these two highlighting methods. The following research questions comprise these thoughts:

- How does the FC method differ from the LH method in terms of route memorizing performance (mapping the memorized route on another kind of map)?
- How does the FC method compared to the LH method influence the use of the context (off-route) information in a memorizing task?
- How does the FC method contrasted to the LH method influence the use of the context (off-route) information in a visual search task for off-route objects?

In order to tackle these questions, a human subject study is proposed involving the recording of users' eye movements.

1.3. Thesis Structure

The following Chapter addresses the research background, considering causes of map complexity and how visual information is perceived and processed. Additionally, the concept of focus and context as applied in geovisualization will be presented followed by references to previous research on the topic of map memorizing tasks.
Chapter 3 is about the methods used to tackle the research questions and will give insight into the experiment design. Chapter 4 presents the general results from the experiment, whereas Chapter 5 addresses the result gained from the eye-tracking data. Chapter 6 discusses the results and refers to the research questions as well as to the research framework. The conclusion is presented in Chapter 7. Important material used for the experiment can be looked-up in the Appendix.

2 | Research Framework

Maps have been a popular product for centuries providing an overview of our environment in a visually pleasant and understandable way. Working out a two-dimensional representation of the earth using symbols to highlight relations between objects, entities and themes can be seen as the core work of a cartographer. Developing visualizations for different scales is the key issue of the map generalization process. The cartographer's job is to qualitatively decide what objects have to be depicted in which way with the goal to reduce visual complexity while simultaneously providing the relevant information. Beside geometrical considerations, this process is also steered by many semantically driven decisions. However, this issue still presents an ongoing challenge to geovisualization, especially with regard to digital map products where zooming functions require continuous adjustability of scale.

Creating a map that meets the needs of users by giving relevant information in a preferably easily accessible way does not depend on map design alone. A factor that also influences the reading process is the user itself. Having users with different expertise in map reading and interpretation as well as varying spatial abilities is an important determinant in the cartographical communication process. Furthermore, the purpose of the map consult is also vital with regard to how information is encoded. Whether it is a visual search task or a navigational task changes the ease of reading a given map.

However, when assessing a map's usefulness only certain criteria can be considered. Since maps are rarely produced for specific target groups and are used in a number of tasks, it is hard to take into account these two factors. Therefore, the amount of information and its organization is especially important. These issues are comprised in the term *map complexity*.

2.1. Map Complexity

Assessing and measuring a map's complexity has always been a relevant issue in cognitive map design research (Montello, 2002). Finding the factors that determine complexity can help improve map design and provide rules for the generalization process. However, in order to measure complexity it first has to be defined more accurately. Several studies (Brophy, 1980; MacEachren, 1982) primarily distinguished between visual and intellectual complexity. Visual or graphic complexity describes the syntactic aspect of a map. It is mainly a matter of how we perceive and process visual information as a result of how the cartographer makes use of visual variables (Bertin, 1983) to modulate the graphic content in a display. On the other hand, intellectual complexity describes the subsequent stage of information processing, based on a user's background knowledge, its individual interpretation and conclusion as a result of previous information. Hence, it refers to the semantic aspects of a map. Intellectual complexity is difficult to analyze because it varies from user to user and can hardly be controlled in an experiment. Possibly, this is why previous research seems to have been focused on measuring visual rather than intellectual complexity. Nevertheless, the distinction between these two kinds of complexity is not plain sailing (e.g. Tufte 1989, Robinson 1952).

Quantifying graphical information is an ongoing topic in visualization science where several studies exist that tried to find a meaningful measurement for visual map complexity. For instance, counting objects or object types (Oliva et al., 2004; Schnur et al., 2010), nodes and lines (MacEachren, 1982) or measuring the distribution of map objects are examples of these efforts. A computational approach is proposed by Fairbairn (2006), in which he suggested data compression rate as a promising index for determining complexity. The advantage of such quantification lies in its effectiveness. Harrie and Stigmar (2010) tested several of the abovementioned methods in terms of map legibility and compared this to subjective ratings (perceived complexity). As it turned out, the metrics 'number of objects', 'number of points' as well as 'object line length' corresponded better to subjective judgement than 'object area'. Two further measurements were proposed by Rosenholtz et al. (2007) aiming for quantifying the visual clutter in a display. 'Feature congestion' is based on the analogy that the more congested a display is, the more difficult it would be to add a new guiding attribute. The determining factors which describe the actual features of the clutter were color, orientation and luminance. The other clutter measurement, 'Suband Entropy', is based on the efficiency with which an image can be encoded while maintaining perceptual quality (Rosenholtz et al., 2007).

Besides the above mentioned clutter measurement methods, visual complexity can also be assessed by looking at how people perceive an image when searching for a specific target. Excess or disorganization of information in visual stimuli impairs object recognition and decreases visual search performance. Hence, for simple search experiments, set size (number of 'items') in relation to search reaction time can be used as a measure for search difficulty. As further discussed in chapter 2.3, the items' features and their arrangement interact with the set size. As stated by Duncan and Humphreys (1989), this interaction effect can be denoted as signal and noise problem. Their *attentional engagement theory* describes search efficiency as function of target-distractor (TD) differences and distractor-distractor (DD) differences. The former relation increases search efficiency whereas the latter one is a contributor to lower search efficiency. However, the load of information that can be processed in a map reading task always depends on the capacity of working memory. Especially relevant for this study are the visual processing abilities of the short term memory.

2.2. Working Memory

The term 'working memory' (WM) is often used synonymously with the notion short term memory. As the name says, an important characteristic is its temporary memory storage. Furthermore, the system is responsible for information encoding and manipulation (Baddeley and Hitch, 1974). Its counterpart is the long term memory (LTM), where learning and reasoning processes take place for solving more complex tasks. Before any visual information is stored in LTM, it first needs to be held in WM. Therefore, the larger the capacity of the short term memory, the faster data can be learned and stored in the LTM.

Baddeley and Hitch (1974) proposed a three-component model of the working memory. The *central executive* is the main part of the system that is responsible for the control of cognitive processes comprising three different functions. First, the selectiveness and inhibition function helps to focus on the task, sorting relevant pieces of information from distracting ones. Second, the shifting function assists switching between retrieval strategies and/or different tasks. The third function helps to bind the information in an ordered way and updates information by replacing old one by new one (Baddeley and Hitch, 1974). Moreover, the *central executive* is relevant for the coordination of the two subsystems (also named "slave systems"), *the phonological loop* and the *visuo-spatial sketchpad*. The phonological loop's functions are to store and rehearse auditory verbal information. Visual information can also be entered in this part of the WM when silently articulating the read information, which often happens when reading a text. The visuo-spatial sketchpad holds visual and spatial information. Besides its short term storing function, it is responsible for the construction and manipulation of images. Regarding spatial knowledge, construction of mental maps is a result of the process going on in the visuo-spatial sketchpad.

For the purpose of this thesis, the visuo-spatial sketchpad is of particular interest, since visual search tasks as well as route memorizing tasks are both processed in this part of the working memory. The taxonomy by Baddeley and Hitch (1974) is sometimes named differently with the visuo spatial-sketchpad being referred to visual short-time memory (VSTM) or visual working memory (VWM).

2.2.1. The Capacity of the Visual Working Memory

Along with the memory's structure the capacity of the visual short time memory has always been a key issue in cognitive psychology. To examine memory span, 'change detections tasks' are one of the commonly used methods (Luck and Vogel, 1997). Here, an observer is exposed to two arrays of items with a short time interval between. The second time the array appears, it has to be indicated whether one item was different compared to the first array or the two arrays were identical. Adding more items increases the difficulty of the task continuously until a certain point a limit can be observed.

An early work addressing the question about the short term memory's capacity was accomplished by Miller (1956). In his work about the magical number seven' he noticed a memory span of about seven simple elements (called chunks). Since then, much work was done trying to uncover the secrets about a possible threshold value regarding VSTM's capacity. Some researchers suggested that capacity is set by a strict number of items regardless of their complexity (Awh et al., 2007; Vogel et al., 2001; Irwin, 1992). Others reported that with increasing item complexity, the number of items that can be memorized decreases (Alvarez and Cavanagh, 2004; Eng et al., 2005; Olson and Jiang, 2002). Complexity of an item is being defined by the number of features (e.g. color, pattern, orientation) it is composed of.

Besides 'single feature' versus 'multi feature objects', Kemps (1999) emphasized the importance of structural complexity in visual memorizing tasks. Structural complexity is a matter of how the different objects are aligned in a display. He concluded, that next to the quantity of objects it is also their arrangement that has an influence on memory recall ability. A controlled experiment showed that memory performance of objects positioned in a matrix was higher than objects positioned in random order. His findings might refer to the 'Gestalt principles' (Banerjee, 1994) which are closely linked to perceived complexity. Here, 'pattern goodness' is a result of symmetrical patterns determined by Gestalt factors such as similarity, symmetry, continuation, simplicity, closure, proximity and other forms of regularity. As a rule of thumb, it can be said that the more regularity an object shows, the less complex it is perceived (Attneave, 1955). Similarly, the more ordered a group of objects appears, the easier it is to remember objects of that group.

Several concepts to ease map reading and alleviate VSTM capacity by reducing complexity are presented in chapter 2.4. However, before doing so, the following chapter discusses how information enters the visual short term memory and which factors guide this process.

2.3. Perceptual Processing and Visual Attention

According to Ware (2004) perceptual processing can roughly be structured into three main stages, describing a simplified model of how we perceive visual information.

Stage one describes low-level information extracting in which features from every part of the visual field are processed simultaneously. This parallel processing is fairly rapid and it is largely independent of our attentiveness. Therefore, this stage is also described as bottom-up processing, where only low-level properties of the visual scene are processed. Stage two is influenced by available information from stage 1 but also steered by the reader's attention. Driven by particular search queries, the visual scene is divided into regions and patterns having the same characteristics such as similar colors or same texture. Hence, this stage of pattern-finding is a combination of bottom-up and top-down feature processing. Stage three describes the highest level of perception where objects are held in visual working memory as a result of active attention. Objects are constructed from patterns perceived in stage two and deliver answers to visual queries. This sequential goal-directed stage is denoted as high-level form of processing where only a few objects can be held at a time. Figure 2.1 shows this three stage model as described by Ware (2004).

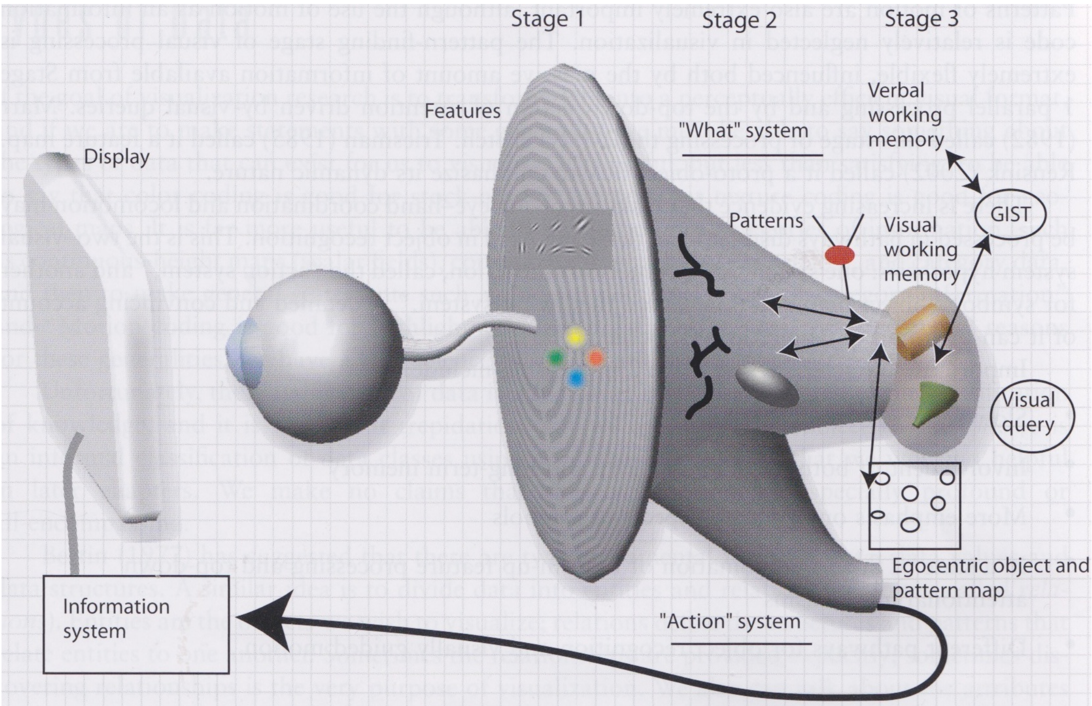


Figure 2.1.: Three stage model of visual information processing by Ware (2004)

In earlier research, visual information processing was mostly classified in a two stage architecture. A central work suggesting this division was Treisman's (1980) *feature integration theory.* In a first stage, preattentive processing helps to extract simple features while at a second stage attention comes into play for extracting more complex properties (Treisman and Gelade, 1980). The preattentive processing stage is characterized by the fact that set size does not influence reaction time regarding visual search for a specific feature target. Treisman (1980) argued, that these unique feature targets can be detect automatically without focusing attention. Hence, a specific feature such as a color value would "pop out" from the display. Whenever a reader knows something about a target's characteristics, search is also based on top-down information. In this case, a so called 'feature map' (Treisman, 1985) helps the reader to rapidly spot unique features. Compared to Ware (2012) this process is similar to the second stage, where bottom-up and top-down information are available. The second processing stage described in feature integration theory is characterized by the absence of a unique feature target, thus no pop out effect takes place. To find a specific target, a reader has to scan and focus attention to single locations on the map. For example, searching for a green circle among blue circles and green squares requires to look for a conjunctive target defined by shape and color (Lloyd, 1997). In this case, no unique feature map helps to rapidly detect the target, which means that it is not spotted during the preattentive processing stage.

Treisman's (1980) model of visual processing has been discussed and evolved over the years (Wolfe et al., 1989; Wolfe, 1994; Di Lollo et al., 2001) while most subsequent models kept the two-stage architecture. However, as Wolfe (1998) stated, a simple definition of guiding attributes based on reaction time is not as obvious as originally supposed. Wolfe and Horowitz (2004) claimed that some attributes do guide attention more than others but do not need necessarily be properties of the preattetive stage of perceptual processing. They provided a list (see Table 2.1) of attributes organized according to the likelihood that they are sources of guidance of attention.

Besides size, orientation and motion color is listed as an undoubted guiding attribute. Therefore the likelihood that color guides attention in a visual search task is very high while it might or might not provoke a pop out effect among the visual display. In the latter case, the visual search is less dependent on color, meaning that a total absence hardly influence search performance. However, color values are not only used as guiding features, they are even more important as information carrier. Since colors in maps are often used to codify nominal data, an absence of them reduces information that is normally important to distinguish different map entities. Color reduction in maps can therefore be regarded as one kind of map generalization where information is reduced

Undoubted attributes	**Probalbe attributes**	**Possible attributes**	**Doubtful cases**	**Probable non-attributes**
Color	Luminance	Lighting direction	Novelity	Intersection
Motion	Vernier offset	Glosiness	Letter identity	Optic flow
Size	Stereoscopic depth	Expansion	Alphanumeric category	Three dimensional volumes
Orientation	Pictoral depth cues	Number		Color changes
	Shape	Aspect ratio		Faces
	Line termination			Your name
	Closure			Semantic category
	Topological status			
	Curvature			

Table 2.1.: Attention guiding attribute adapted from Wolfe and Horowitz (2004)

in order to simplify the overall map depiction.

2.4. Focus + Context

With the rise of digital technology, the possibility to improve cartographic products by providing information fitted to the user's purpose opened up. Studying one specific map area on different scales traditionally required the use of several physical maps. So whenever a tourist wanted to orientate himself in a city he was using a large scaled city map. However, as soon as he wanted to know something about the surrounding of the city he had to pick up a smaller scaled map. Combining several scales on one sheet was an expensive effort. One of the earliest suggestions to tackle this issue was the concept of the 'magnifying glass' by Snyder (1987). The approach magnifies the area of interest (focus area) while maintaining visibility of the contextual regions. In recent years, the idea of having multiple scales in one map became fairly popular. Guerra and Boutoura (2001) as well as Fairbairn and Taylor (1995) used this method to assist map reading in urban environments.

Similar to the magnifying glass, many studies proposed the fisheye view as a solution to the abovementioned problem (Harrie et al., 2002; Sarkar and Brown, 1992; Sarkar et al., 1993). While, they appear to be easy to use (Gutwin and Skopik, 2003; Gutwin and Fedak, 2004) they are highly expensive to generate. Slightly modified versions such as the focus+context+glue approach (Yamamoto et al., 2009) tried to deal with these problems with a glue zone where additional map generalization takes place. Besides generating polyfocal visualizations with distortion techniques (spatial method) focus and context areas can also be created using other methods. Kosara et al. (2002) proposed a method called *Semantic Depth of Field*, which is based on the idea of guiding the reader to the import information. Context information is blurry depicted whereas the non- blurred areas point to the relevant information. In contrast to the formerly mentioned spatial methods the latter one is an example of a cue based visualization method.

Figure 2.2 shows two example of the abovementioned methods as applied in geovisualization. The left image is an example of a focus and context map using blur to generate context areas, while the right image shows an approach that is based on geometrical distortion techniques which integrates several scales on one display.

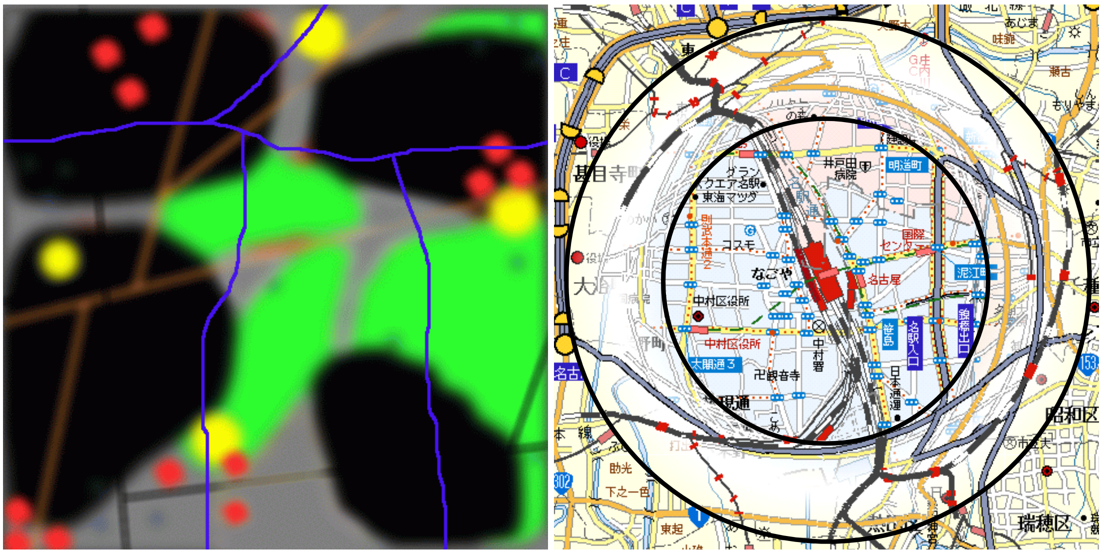


Figure 2.2.: Focus and Context applied in maps - examples from Kosara et al. (2002) (left) and Yamamoto et al. (2009) (right)

On the whole, focus and context visualizations makes it possible to show more, more detailed and more targeted information at the same time. Further, it assists users in locating the relevant information in a wider context. While the concept of focus information is fairly clear, that of context information is somehow harder to grasp. One of the early works proposing a definition for context information in computing applications was Schilit et al. (1994). According to them, three important aspects of context exist: where you are, who you are with and what resources are nearby. However, it is still difficult to grasp this idea in terms of focus information in geovisualization. A more comprehensive definition of context is given by Dey (2001, pg. 5).

> *"Context is any information that can be used to characterise the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves."*

In terms of map reading uses, the application is the map itself whereas entities are visual variables on the map. Regarding map interactions with the real environment, the entities are wider defined including the user and as well as physical objects. Whether an entity is considered as relevant or not depends on what the user intends to do. Dey (2001, pg. 5) further defines the term 'context awareness', where exactly this issue is emphasized:

> *"A system is context-aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task."*

In short, focus and context applications always have to be regarded in respect to the task they are going to be used for. Even if this sounds rather trivial it is important to keep in mind when assessing the usability of such reading support methods. Hence, in order to adequately examine focus & context visualization, map reading tasks have to be identified. Board (1978) strongly emphasized the importance of clarifying the purpose of the map study before any experimental evaluation about map effectiveness and efficiency is to take place. Therefore, he proposed a checklist of map reading task. Table 2.2 shows these tasks grouped by three basic map functions.

Navigational tasks are all about orientating oneself on the basis of the map. Focus and context visualization as presented above is mainly concerned with these kinds of tasks. Measurement tasks deal with comparisons between different map objects. These kinds of tasks are essential regarding the reading of thematic maps. However, all kinds of measurement tasks are somehow combined with search tasks. The latter are even more

Navigation	Measurement	Visualization
Search	Search	Search
Identify and locate own position on map	Identify	Identify
Orientate on map	Count	Describe
Search for optimum route on map	Compare	Compare/ Recognize
Search for landmarks en route	Estimate	Delimit
Recognize landmarks en route	Interpolate	Verfify
Search for destination	Measure	Generalize
Identifiy destination		Prefer / Like
Verify		

Table 2.2.: Map reading tasks adapted from Board (1978)

crucial regarding the last category of reading tasks. Visualization emphasized tasks mainly involve map scanning and searching for visual patterns among map entities. The purpose thereby is to get a mental representation of the depicted map environment.

Creating task specific maps requires the selection of information to meet the map's purpose. In doing so, different levels of knowledge are distinguished. Compared to a general reference map, some of the presented knowledge in a focus and context map can 1) be omitted, 2) remains the same or 3) be distorted (Barkowsky and Freksa, 1997). Selection of required knowledge, resulting in 'aspects', is therefore a crucial step in the construction process of such a map. Since space on a map is limited, these aspects compete witch each other for specific map locations. Therefore, map aspects should be ranked in a depictional precedence with regard to the map reading tasks (Barkowsky and Freksa, 1997). Thereby, it is important that the used aspect ranking matches with the user's assumption about the depictional precedence. If this is not the case, map reading could be impended and lead to misinterpretations (Berendt et al., 1997). Schematic maps such as transportation network maps (see Figure 1.3) are good examples of task specific maps where only a few aspects are depicted. To achieve cognitive adequacy (Klippel et al., 2005) many aspects are heavily distorted or completely omitted. On the other hand, domain specific aspects such as network intersections are clearly emphasized. For instance using these maps for measuring task (Board, 1978) is not very helpful and will lead to wrongly interpreted information due to wrongly made assumption about the depictional precedence.

As we have seen so far, constructing focus and context visualization involves a selection process which is mainly based on the kind of task the map is proposed for. The selected features determine which information is relevant for the task by highlighting specific instances of these features to guide the users' reading process (Richter et al., 2008). So while, it cannot be said which features are important but in general, it can be distinguished between maps that focus on specific objects and maps that put the focus on specific areas (Richter et al., 2008).

A good example of a focus and context map implementation, which is both strongly task-orientated and considers objects as well as specific areas, is given by Neis and Zipf (2007). The authors' proposed method assists map reading for wayfinding tasks using mobile maps. The concept is based on the idea of a focus map by Zipf and Richter (2002) as discussed in chapter 1.1 and the relevance of landmarks recognition in wayfinding tasks (e.g. Golledge 1999). Areas of higher interest (focus zones) are primarily denoted as buffer zones around the path. Context regions are depicted in lighter colors and the degree of generalization is higher than the one in the focus areas. Furthermore, a second buffer around the path detects local landmarks which can be visualized more prominently. Figure 2.3 shows an example of a preliminary version of Neis's and Zipf's (2007) implementation. The left image presents a conventional navigational map whereas the right is the result of the above described focus approach. Besides the obvious division between focus and context area, there are also two local landmarks (stop sign and church) emphasized nearby the route.

Figure 2.3.: Conventional navigational map compared to focus and context approach by Neis and Zipf (2007)

2.5. Map Learning Tasks

Map reading and information retrieving is steered by many factors and involves several process stages as discussed in the previous chapters. Similar to other learning tasks, map learning produces a mental representation of the visual stimuli. In terms of map learning, this mental representation is also referred to as cognitive map. However, as Thorndyke and Stasz (1980) stated, map learning differs in two ways from purely spatial or visual learning tasks (e.g. memorizing pictures, faces, etc.). First, learning a map is often more complex than learning from other visual stimuli. Besides names, shape and various features of objects, a map requires to spatially locate different entities absolutely as well as relatively. For instance, learning a highlighted path from A to B is not just about memorizing the shape of the route. References to route type, street names, intersections as well as adjacent buildings and object have to be made. In short, conceptual as well as spatial information has to be acquired by applying various decoding and acquiring strategies. Second, learning from a map is characterized by exposing the subject to all information simultaneously. In other learning tasks, information is often presented sequentially. A map learner has to decide what information is important, in which order it has to be processed and how much time to spend on single objects. Due to the many decisions that have to be made when learning a map, Thorndyke and Stasz (1980) lists map-learning tasks in one line with natural learning situations.

2.5.1. Route Knowledge versus Survey Knowledge

Spatial knowledge has been a field of interest for cognitive and environmental psychology for a long time. A basic and recurring distinction in many studies concerning the acquisition of spatial knowledge is the difference between configurational and route knowledge (e.g. Siegel and White 1975; Stern and Leiser 1988). Route knowledge is typically characterized by the ability to refer to various locations along a route. These locations are sequentially memorized without the knowledge of general interrelationships. In contrast, configurational knowledge, also called survey knowledge, entails a more holistic view, comparable to knowledge gained from a bird's eye view (Hirtle and Hudson, 1991). Hence, it also implies Euclidian relationships. Golledge et al. (1985) suggested that any acquisition of spatial knowledge begins with the gathering of information about single landmarks, whereupon route knowledge is acquired, and as a final step, all pieces of information are put together, building up survey knowledge.
However, other research suggested that configurational knowledge is not learned after route knowledge but rather in conjunction with the latter in some cases (e.g. if map users learn several routes through the same area (Moar and Carleton, 1982)). Hirtle and Hudson (1991) found that route knowledge can be acquired both through route presentation and map presentation. Survey knowledge, however, is more accurately gained from map than from route presentation.

Many studies experimentally tested route knowledge by exposing subjects to a sequence of images, for example using a slide presentation (e.g. Hirtle and Hudson 1991). Doing so, participants primarily build up their knowledge based on visual inputs. However, route knowledge can also be considered as directional information given in written or verbal form. Such knowledge is normally used when being asked for wayfinding assistance by people who are navigating in unfamiliar areas. Besides directional information, it comprises decision points referring to landmarks along the path (e.g. turn left at the supermarket, then the third street to the right). Meilinger and Knauff (2008) experimentally compared directional route knowledge with survey knowledge in a memorizing task where subject had to retrace the learned route in real environment. Either they were learning the route with a map or merely with directional advises. As it turned out, no difference was observable between the two methods concerning navigational performance in real environment. The null effect might be the result of a similar memorizing strategy. Meilinger and Knauff (2008) argue that map learners might have translated the route into verbal directions to themeselves, eventually having memorized fairly similar information as the route learners. This verbal re-coding theory is backed up by similar experimental investigations (Schlender et al., 2000; Pazzaglia and De Beni, 2001). The findings by Meilinger and Knauff (2008) go along with the ones by Hirtle and Hudson (1991), who stated that route knowledge can also be acquired by studying a map. Garden et al. (2002) also reported verbal translation of visually perceived information. The authors conclude that learning a route on a map involves different aspects of the working memory. Besides the use of visuo-spatial components other storing components might be used. Additionally, they found that the use of different components of the working memory also depended on the subjects' spatial ability. Subject with higher spatial abilities tend to rely heavily on the visuo-spatial working memory in contrast to subjects with lower spatial abilities.

2.5.2. Spatial Abilities & Route Memorizing Tasks

Research about spatial abilities is plentiful and has been conducted for dozens of years. Many studies have focused on gender difference involving psychometric tests. A general observation in most psychometric literature revealed that males on average perform better than females in some of the most frequent spatial tests (Montello et al., 1999). One of these tests was originally proposed by Vandenberg and Kuse (1978). Their fairly well known mental rotation test revealed sex differences among different age categories. Following research backed up the higher performance of males regarding Vandenberg's test (e.g. James and Kimura 1997). Silverman and Eals (1992) proposed another psychometric test, which in contrast to the abovementioned observations in gender difference revealed rather surprising outcomes. The test required subjects to memorize an array of 27 objects, which was then replaced by another array, comprising the same objects but with 14 of them having switched positions. Subjects were asked to indicate which objects changed places and which remained at the same location. Surprisingly, the authors reported that females outscored males on the object location memory test (Silverman and Eals, 1992). Interested by this outcome, McBurney et al. (1997) examined a test based on the commercial Memory game, where the authors found the same pattern among genders. Similarily, James and Kimura (1997) repeated the object location memory test and confirmed women's superiority in remembering objects at specific locations.

Besides gender differences in spatial ability, much research has addressed gender differences in map learning tasks. Various studies have suggested that males and females differ in their learning strategy in route learning tasks. Miller and Santoni (1986), who found that women reported more landmarks when asked to give written path description, gave early support for this suggestion. Similarly, McGuiness and Sparks (1983) found that undergraduate females draw more landmarks on a previously learned campus map compared to their male colleagues. Further, evidence for gender difference in orientational strategies has come from studies of verbal way description. Ward et al. (1986) reported that males used more metric distance and cardinal direction terms than females when giving way descriptions of a previously learned route. Additionally, Gwinn et al. (2002) found that men, too, heavily rely on landmarks but employ them differently. Along with general observations, males use landmark for direction and distance judgments whereas females tend to memorize them as labels along a route.
Generally speaking, it is supposed that females rather use a landmark-based strategy whereas males use a geometricly-based strategy in route learning tasks (Choi and Silverman, 1996; Galea and Kimura, 1993; Miller and Santoni, 1986; Dabbs Jr et al., 1998). In other words, women acquire route knowledge whereas men are more likely to gain survey-based knowledge from a map. To conclude, many studies support gender differences in map learning regarding map reading strategies. Nevertheless, a consistent superiority of one gender can not be stated (Montello et al., 1999).

Map reading and route learning tasks are well-studied subjects in cognitive psychology whereas only few studies exits from the point of view of geovisualization. As shown, much research is focused on analyzing differences between women and men, whereas task-orientated research (e.g. route learning / memorizing) putting the actual map design in the center of interest remains rare. Many maps used in the abovementioned studies are considerably simplified and held in a rather schematic design. Hence, they do not really consider a maps' variety in terms of more comprehensive information visualization displays. As discussed at the very beginning of this chapter, a map is a complex visual medium, which should be considered when assessing map reading performances.

3 | Methods

3.1. Participants

On the whole, 26 participants were recruited to take part in the experiment. The participation was voluntarily and not being compensated. Out of 26 attendees, only 24 recordings were suitable for further analysis ($N = 24$, 10 females, 14 males). Three recordings were not usable because of a very low eye-tracking sample, meaning that any eye movement analysis would not be possible. 18 participants were in the age category between 25-31 years, whereas 4 where in the category between 18-24 years and two participants were between 50-60 years old. 14 subjects wore optical aids such as glasses or contacts whereas no one stated to have any color deficiency.

Concerning the location of the map stimuli (see chapter 3.3.2), participants were asked if they were familiar with the city of Sheffield (UK) and Birmingham (UK). One participant indicated having lived for a short period of time in Sheffield, whereas she asserted not having recognized the map excerpt afterwards. In general, participants stated to be familiar with Google Maps (GM)[1] but rather not familiar with Open Street Maps (OSM)[2] (see Figure 3.1). 9 subjects have never used OSM and 10 claimed to use it rarely. Furthermore, almost half of the participants indicated to use a route planner on a monthly basis whereas no one has never used such a tool (see figure 3.2). Since online route planners provide information in different ways, one question addressed the helpfulness of such communication tools when looking for a way to be described. Figure 3.3 clearly shows that visual way descriptions (maps) are highly preferred over written or verbal advises, which are both similarly rated.

[1]Google Maps: https://www.google.ch/maps
[2]Open Street Maps: http://www.openstreetmap.org

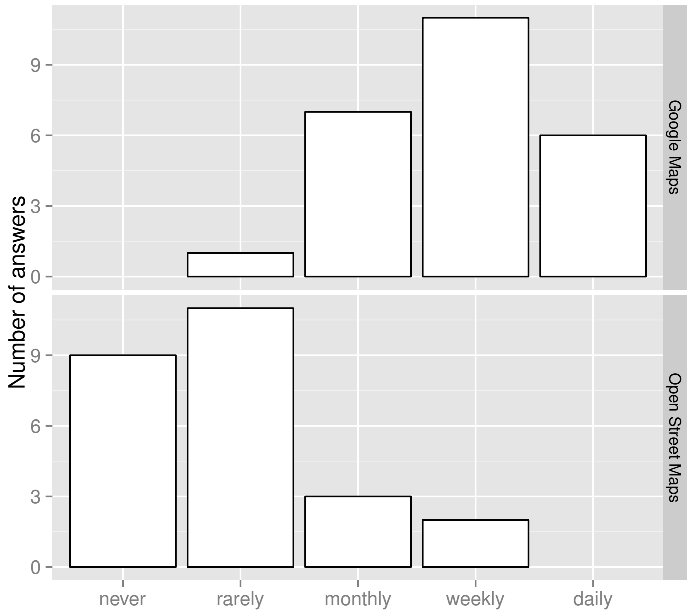


Figure 3.1.: Frequency of use of Google Maps versus Open Street Maps

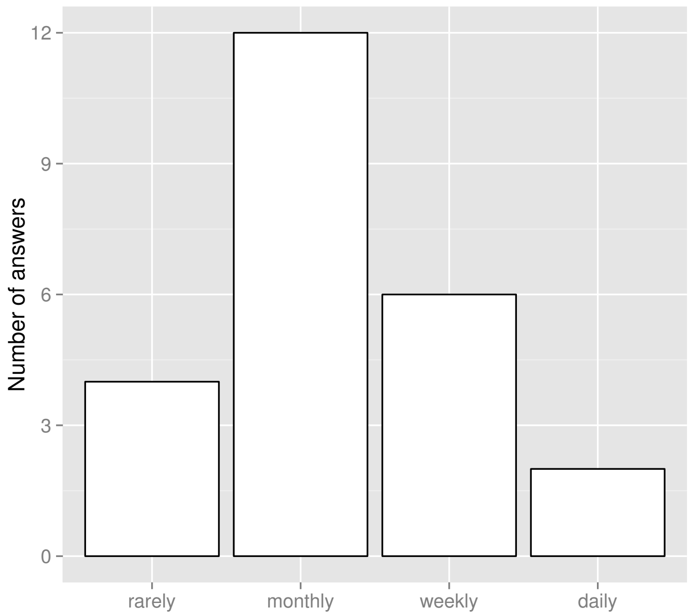


Figure 3.2.: Frequency of use of online route planners

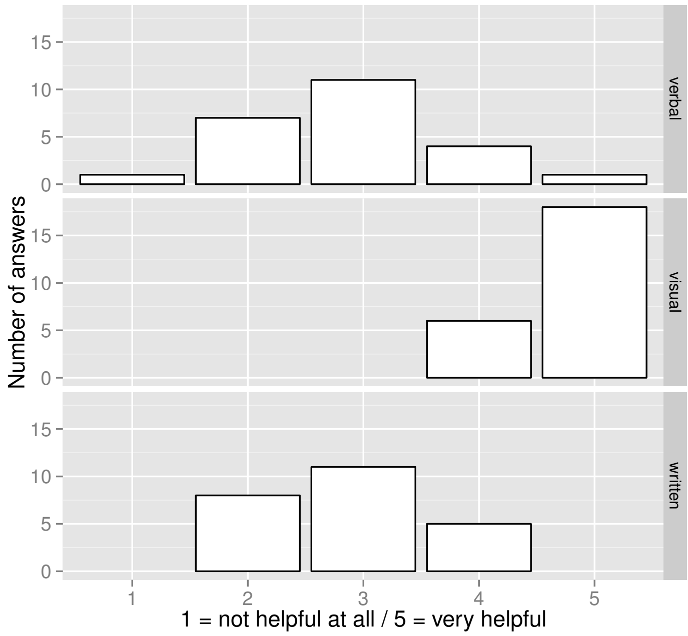


Figure 3.3.: Rating of helpfulness of different route description methods in comparison

3.2. Apparatus

The experiment was conducted using Tobii's [3] eye-tracker apparatus TX300. The built in infrared sensor has a sampling rate of 300 Hz with the variability $<0.3\%$ to recognize fixations. The sensor comes along with a screen unit with a display measuring 23inch and a screen resolution of 1920x1080 pixels, having an aspect ratio of 16:9. The screen unit has a built-in webcam which was used to record each participant's experimental session. The used software package to run the experiment and to handle the data was Tobii studio 3.2.1. Besides a designing and recording function to present the stimuli, the application also provides replay and visualization as well as statistical tools to analyse the collected data.

The eye-tracking device (see Figure 3.4) is situated in the eye-movement lab at the Department of Geography of the University of Zurich. The environment such as the room illumination, screen and chair position as well as the background illumination of the computer screen were held identically throughout the entire study. This ensures the same influence of external factors for all participants.

[3] Tobii Technology: http://www.tobii.com

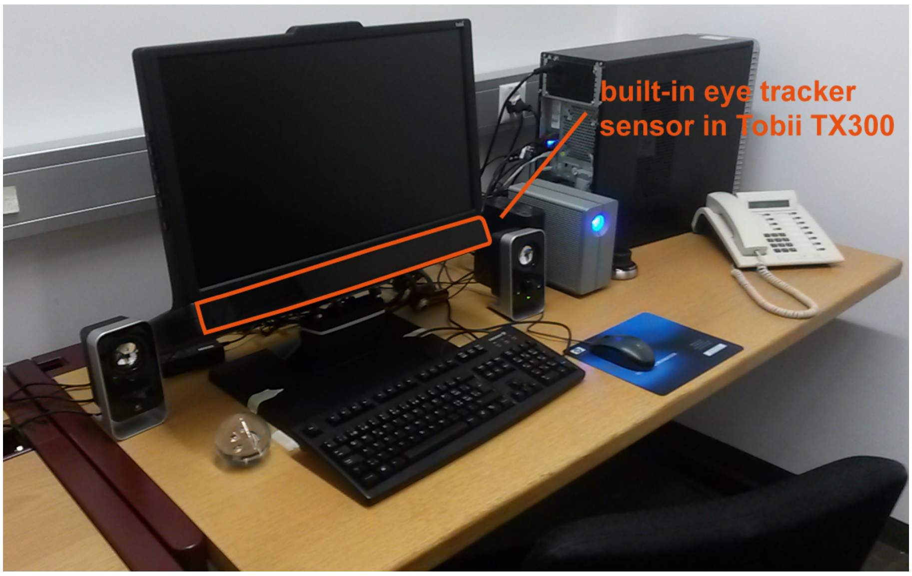


Figure 3.4.: Eye-tracker apparatus in the eye-movement lab at the University of Zurich

3.3. Materials

3.3.1. Object Location Memory Test

Spatial abilities such as the memorization of objects' location is a typical skill becoming important in map reading tasks. As discussed earlier, many psychometric test exits assessing various spatial abilities. For the reason of this thesis, the object location memory test as originally proposed by Silverman and Eals (1992) was embedded in the experiment with the goal of providing a reference value for the assessed measurements (see Section 3.8). The object location memory test was an online version of the original test. The test is running on a website[4] which is freely accessible and where it is possible to set up and run an own experiment for scientific purposes. The results can be downloaded from the server in form of an excel sheet.

The experiment starts with a warm-up run, where participants have one minute time to study a display depicting an array of ten objects. After that, they see the same objects but some of them have changed position. In order to denote whether an object has changed position, participants can click on the respective object to mark its changed location. The practice run repeats until the participant recognizes all changed objects. Following to the warm-up run starts the actual test. This time, participants are asked

[4]Object location memory test: http://opl.apa.org/Experiments/About/AboutObjectLocationMemory.aspx

to memorize an array of 27 objects whereat the time limit is set to one minute again. No matter whether all changes are recognized the first time, every participant has to run through all five identical trials. The left image in Figure 3.5 shows the array of objects asked to be memorized whereas on the right image the same objects with partly changed location is depicted.

The individual scores are reported as numbers of a) hits (correctly identified object that has moved), b) correct rejections (correctly identified as a non-moved object), c) false alarms (incorrectly indicated as an object that has moved) and d) miss (incorrectly indicated that object has not moved). Out of the 27 objects, 14 objects moved and 13 remained at the same location.

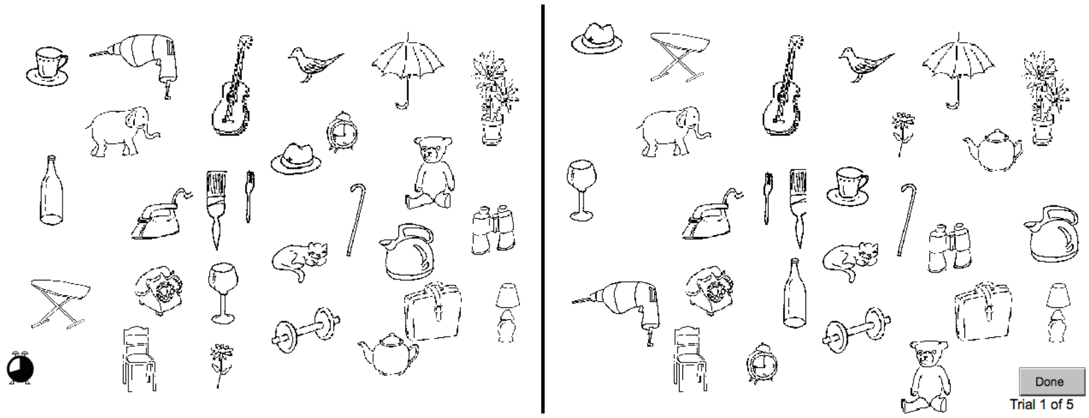


Figure 3.5.: Object location memory test as originally proposed by Silverman and Eals (1992)

In order to analyse the data, the authors suggest a discrimination index (DI) based on Banks (1970), where percentage of hits (HitRate) and the percentage of false alarms (FalseAlarmRate) are contrasted to each other. The following formula describes the DI:

$$DI = \frac{1}{2} + \frac{(HitRate - FalseAlarmRate) * (1 + HitRate - FalseAlarmRate)}{4 * HitRate(1 - FalseAlarmRate)}$$

An index value of 1 indicates a *HitRate* of 100% and a *FalseAlarmRate* of 0%. The DI is calculated for every trial. In order to determine the overall performance for each participant, the mean value across all five trials was calculated.

3.3.2. Map stimuli

The map stimuli used for this experiment are based on Google Maps (Google, 2010). Due to the prominence of GM across online map services and because of its comprehensive route planning tool, the proposed highlighting methods are based on an existing map design, that is already frequently used for navigational map reading tasks. The maps were presented in the default roadmap design as originally provided by the producer. Two map excerpt from the city centers of Birmingham (UK) and Sheffield (UK) were chosen as the basic stimuli.
For the subsequent design manipulations, a screenshot from both map locations were taken with a size of 1900x1068 pixels. Any changes made on the map screenshots were realized with Adobe Illustrator & Adobe Photoshop (CS 4, 14.0.0[5]). For both map excerpt, a short route between two points of interest (in both cases a hotel and a restaurant/bar) were chosen. The two routes can be compared based on the number of segments each route comprises. A segment is defined as a part of the route, that lies between two turnings (crossings). Regarding navigational tasks, no directional decisions have to be made when passing a segment. The map excerpt of Sheffield comprises 11 segments, whereas the one of Birmingham comprises 9 segments. Turnings where adjacent segments lead outside the map excerpt were not considered as turnings.

Each map excerpt was presented in two ways, showing the route either with the line highlighting method or with the focus & context method. Overall, four map stimuli were made. For the LH method, a simple line path (with a stroke of 11 points and a transparency of 40%) was overlaid the route, indicating the start and the end point with a pin icon having an "A" and a "B" on it (see Appendix Figures A.1 & A.3). The same routes with the focus & context highlighting were designed with a buffer along the way, having a width of 40 points and a fuzzy transition between the focus and the context zone (see Appendix Figures A.2 & A.4). The width of the focus zone was determined to be as large as possible but simultaneously the route should be highlighted as unambiguous as possible. As mentioned earlier, the context area of the map was decolored and the illumination of all shades of grey were reduced by 40 percent. The adjusting values were the result of a qualitative comparison with the

[5]Adobe: www.adobe.com

goal of indicating the path as clearly as possible and making the context information as readable as possible. For the drawing tasks (see 3.4.1) two map excerpts from Open Street Map (OpenStreetMap-contributors, 2012) depicting the same regions were chosen (see Appendix Figures A.2 & A.4). As the GM screenshots, the OSM images had the same resolution and were displayed in the original design provided by the producers.

3.4. Experimental Design

3.4.1. Tasks

The main part of the experiment involving the stimuli described in the previous chapter included two different tasks. in order to tackle the proposed research questions (see chapter 1.2). Both tasks were solved with each stimulus. The stimuli as well as the respective instructions were presented as an html page in a web browser. The pages were linked together with a "forward button" on each page, which could be used to proceed to the next page whenever instructions were read or tasks were completed. All information that was necessary to understand what the tasks were about was given in written form.

Task 1

The instruction in task 1 described a fictional scenario, telling participants that they are in the UK for a visit. Initial point was always a hotel from which they had to find a nearby restaurant. Since roaming fees are enormously high in foreign countries, they would look up the way to the desired location using a computer in the hotel lobby. For each map, the first assignment was to memorize the route from point A (hotel) to point B (restaurant). Study time was limited to a maximum of three minutes whereas there was no time indication on the display. By doing this, subjects should not feel time-pressured. The upper limit was determined by several dry runs where test participants did not reach this time limitation. However, if this three minutes was exceeded the experimenter would have told the participants to stop and proceed by clicking the forward button. Instructions on the next display told participants that they had found a city map on the way to the destination whereas they were requested to find the memorized route on this map. The actual task was to draw the route on the city map (clipping from Open Street Maps) with the mouse as fast and accurate as possible. To confirm task completion, they had to proceed by clicking on the forward button.

Task 2

Task 2 was presented always coming straight after task 1. Instructions tied in with the scenario described for task 1. They were told to be back in the hotel, consulting the same map as seen before. This time, they checked the map to find a certain place (Point of interest) in order to go there and take a picture of the instructed POI. In both maps, the target was located off the route with the intention of measuring the readability of the context information. The actual task instruction was to find the point on the map and click on the target in order to confirm task completion. By clicking on the forward button, the same map as the one they had just seen in the previous task appeared.

Figure 3.6 shows the targets for both map locations. Since the locations are off the route, these targets appear non colored in FC maps. As the images show, the target for the Birmingham map was the "Birmingham Opera" indicated by a single POI. For the Sheffield map, the "Tudor Square" had to be found. In contrast to the Birmingham Opera, the Tudor Square is represented by an entire area. Both places were labeled with the same font size. In order to curtail the area where it had to be searched for the target, participants were given initial assistance. For the Birmingham map participants were instructed that the target lies southwards from pin B, whereas for the Sheffield map they were told to search northwards from pin A.

Figure 3.7 shows the sequence of task 1 and task 2. The boxes stand for the single displays. Apart from the time limitation regarding map studying, runtime was controlled by the participants. Whenever instructions were read or task was completed they could click on the forward button to get on the following display.

3.4.2. Within-Subject Design

For this experiment, a within-subject design was chosen. In a within-subject design, a participant (subject) is exposed to all independent variables, which means that every subject in this experiment saw every stimuli (Martin, 2008). In contrast to a between-

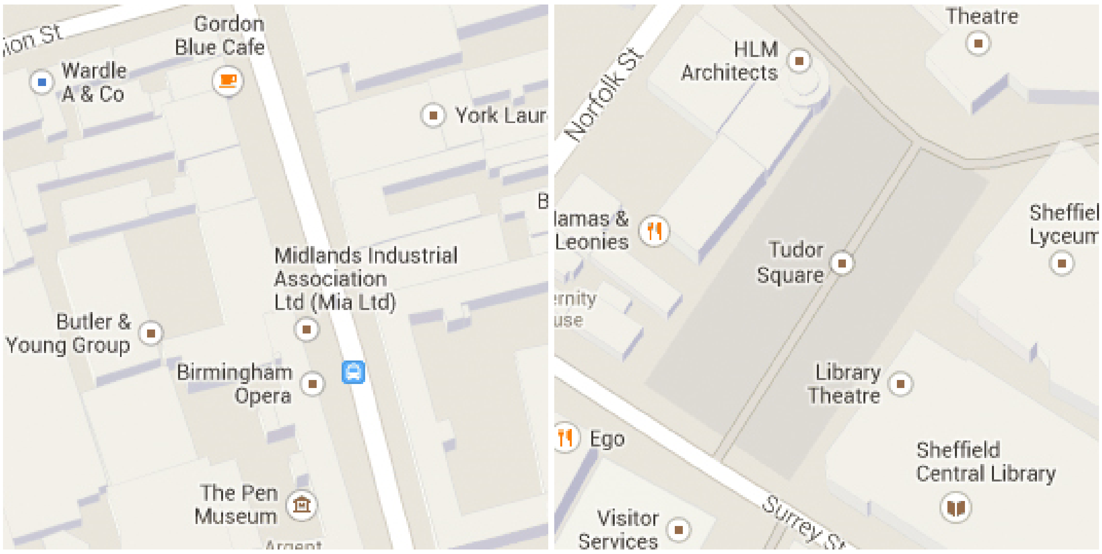


Figure 3.6.: Search targets in task 2 - Birmingham Opera (left) & Tudor Square in Sheffield (right) on map excepts from Google (2010)

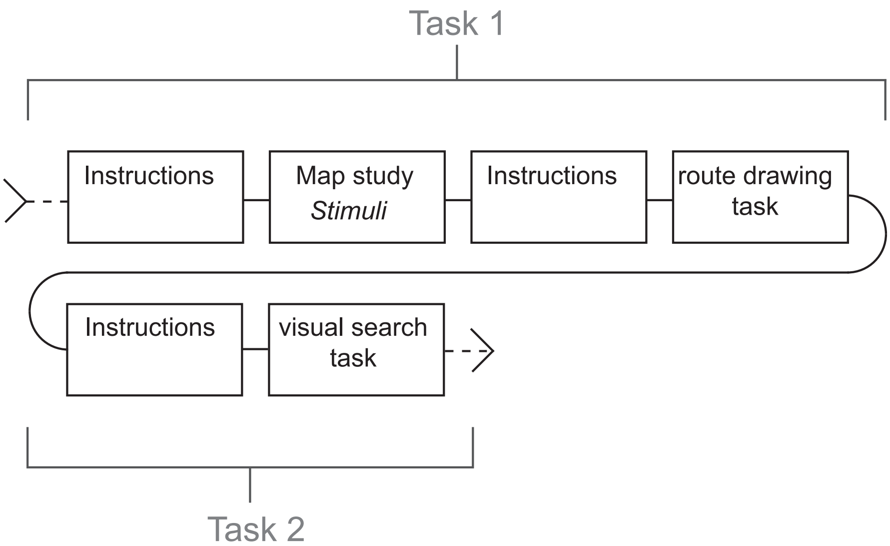


Figure 3.7.: Order of task 1 and task 2

subject design, where the independent variable is at least manipulated between two participants, the within-subject design holds a very practical advantage: fewer participants are needed.

Additionally, bearing in mind the differences between men and women in spatial abilities (e.g. James and Kimura, 1997; McBurney et al., 1997), the chosen within-design is a good way to minimize the gender effects as well as the individual differences between participants. Otherwise, using a between-subject design requires balancing the number of females and males throughout all groups. Nevertheless, there are some drawbacks emerging from the nature of a within-subject design. Since every stimulus is exposed to every participant, there must be several trials following each other. That involves the risk that participants get somehow influenced by the previously shown stimulus, therefore the results can get skewed. These kinds of side effects are called order effects, referring to the order the independent variables are presented (Martin, 2008). For the experiment proposed in this thesis, order effects are a big issue. A map with FC highlighting followed by a LH highlighted map (or vice versa) might have influence on the gaze pattern and the encoding strategy the participant applies. The even bigger problem comes with the nature of the memorizing task. Since both map locations were shown twice, each participant was solving the tasks once on a previously seen map. Obviously, a route that has been seen twice, albeit depicted with different highlighting method, will heavily influence the outcomes for the second seen map. The order effect describing this massive impact is called learning effect. To gain control over learning effects and potential other order effects, trials and tasks in a within-subject design have to be counterbalanced.

3.4.3. Counterbalancing

The concept of counterbalancing is based on the assumption that some confounding order effects might influence participants' behaviour and therefore the dependent variables. Since several trials are necessary to expose every stimulus to a participant, the order of the stimuli gets crucial. Because nobody knows exactly the size and the appearance of order effects, the concept of counterbalancing describes methods to vary the order of the stimuli between participants. In the case of this experiment, a complete counterbalancing scheme was applied.

Having two routes from two maps (different locations) and each of them depicted with two highlighting methods, there are four visual displays (stimuli) overall. In order to get a complete counterbalancing scheme, these four visual displays have to be rotated in a way that each stimulus gets once in a unique position. The total number of possible arrangements is simply four factorial (4!), resulting in 24 different combinations. Naming the four stimuli A, B, C and D, Table 3.1 shows the scheme of complete counterbalancing which was applied for this experiment.

ABCD	ABDC	ACBD	ACDB	ADBC	ADCB
BACD	BADC	BCAD	BCDA	BDAC	BDCA
CABD	CADB	CBAD	CBDA	CDAB	CDBA
DABC	DACB	DBAC	DBCA	DCAB	DCBA

Table 3.1.: Complete counterbalancing scheme

When applying a counterbalancing scheme like the one in Table 3.1, it still has to be assumed that the influence having A followed by B is the reverse of the vice versa order. This assumption is called symmetrical transfer (Poulton and Freeman, 1966) and it is crucial to approve this assumption in order to justify the chosen within-subject design. For this thesis, the symmetrical transfer between all four stimuli is assumed.

3.4.4. Factorial Design

For this experiment several independent variables exist. On the one hand, there are the two tasks which will be analyzed separately. On the other hand, there are the two highlighting methods, each presented with two different routes (map locations). Therefore, FC and LH method can be considered as the two levels of the same variable. However, the two routes must be regarded as separate variables which will be treated accordingly. In short, a 2 x (2 x 2) matrix results where only comparisons between two means can be made. Figure 3.8 shows this matrix which has in total eight fields also named as the experiment's factors.

Participants' behavior due to different independent variables is captured as dependent variables which describes the actual experimental effect. According to the research questions (see 1.2), several dependent variables exist. Task performances (measured in efficiency and effectiveness) as well as the eye-tracking metrics (see 3.5.3) are the measured output of the experiment. The time used to complete task 1 and 2 is a measurement for the efficiency whereas effectiveness in task 1 is expressed by the accuracy

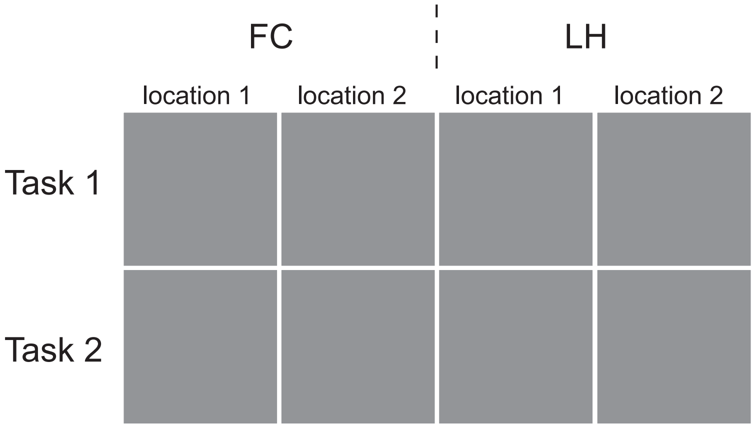


Figure 3.8.: Factorial design

given by the amount of segments that are rightly recognized as part of the route. Accuracy in task 2 was also analyzed by looking at the precision of the target notification. More details about measuring accuracy can be looked up in chapter 3.5.2.

3.4.5. Procedure

The user study was conducted in the eye-movement laboratory of the Department of Geography at the University of Zurich. Before every test session, software and hardware were checked in order to ensure a trouble-free proceeding. Once the participants arrived, they were welcomed and requested to sit down at a separate table in the room. After a short explanation about the content and procedure of the experiment, participants were asked to read the consent form (see Appendix A). This form comprised statements about the content and purpose of the study as well as it informed about the use and publication of participant's anonymized data. To confirm the read information, they were asked to sign the consent form but were also informed about the possibility to withdraw their consent at any time after the experiment took place. Subsequently, participants needed to fill out a questionnaire (see Appendix Figure A.7), asking about general demographics data (sex, age) and visual impairments as well as some question about the field of study. Hereafter, they changed place, switching to the eye-tracker working station where they were requested to get in a comfortable and stable position on the chair. Following this, a calibration was done in order to determine whether the eye-tracking sensor was able to detect the participants' eyes for the entire screen size. Before the actual eye-tracking experiment started, participants were told that the webcam will record video as well as audio of the entire session. They were also told that the following tasks do not assess their individual performance but rather the usability of the tested element. Furthermore, they were encouraged to ask if they had any questions during the study but requested not to speak loudly by themselves during the memorizing parts of the experiment. The reason therefore is to minimize the possibility that participants re-code the visual stimuli in verbal instructions.

Ready for the testing, the online version of the object location memory test ran in a webbrowser as described in Section 3.3.1. Since the online test was self-describing, participants did not need any further explanations from the experimenter. Once they finished the object location memory test, the experimenter interrupted the session and closed the browser window. This automatically started the full screen mode of another browser window containing the start display of the main experiment (see chapter 3.4.1). As the previous test was, the main part of the experiment was self-explanatory which did not require any further interruptions by the experimenter.
The end of this part of the experiment was indicated by a display, telling that the experiment is now finished. Subsequently participants were requested to switch to the desk they were at the beginning to fill out another questionnaire about preference of the seen highlighting methods as well as about possible memorizing strategies they have applied (see Appendix Figure A.10). At the very end, participants were given a piece of chocolate as a thank-you gift. The entire experiment session took about 40 minutes in average. Figure 3.9 shows the abovementioned elements of the procedure.

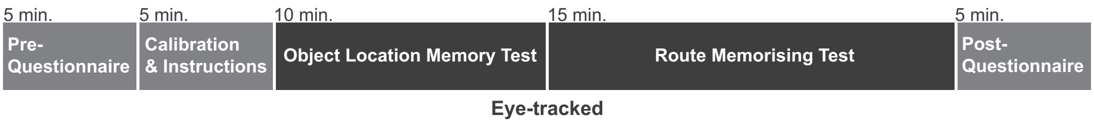


Figure 3.9.: Procedure with rough time specifications

3.5. Statistics

3.5.1. General Statistics

Data analysis, statistical computing and graphics were conducted using the open source software R[6]. Performance for task 1 and 2 was measured by time and accuracy. Hence, ratio as well as ordinal data were used for the analysis. According to the factorial design, statistical tests compared two related conditions with two levels each. All analysis in this study test for differences between these related conditions. Whether a parametric test or a non-parametric test was chosen is based on data's level of measurement and whether data is normally distributed or not. In order to test the assumption of normality, the Shapiro-Wilk test was performed. Whenever the assumptions of non-parametric data were hurt, the Wilcoxon signed-rank test was conducted. In such a case, the median (*Mdn*) is given for the compared data sets. The result of this test is reported by the p-value (p) as well as the effect size (r). In case of normally distributed data, the dependent t-test was conducted. The output of this test is given by the mean (M) and standard error (*SE*) for both groups, the value of t (t) as well as the effect size (r). To compare the scores of accuracy, a non-parametric test for ordinal data was used. In these cases, the sign-test was performed.

Results are considered statistically significant at a p-value ± 0.05. Statistical trends were considered when p-values lie between p=0.05 and p=0.1.
All error bar charts depicted in the following sections represent the mean with the 95% confidence interval. These intervals were adjusted for the within-subject design according to Loftus and Masson (1994).

3.5.2. Defining Accuracy

Regarding task 1, performance was measured by task completion time as well as by accuracy. To define accuracy, path complexity expressed through the number of route segments (as described in chapter 3.3.2) was used. For each participant and stimuli a hit rate was calculated. The hit rate indicates how many path segments were correctly recognized out of the total number of segments the respective route comprised. Besides

[6] R project for statistical computing: http://www.r-project.org/

the hit rate, the recognition of the start and end point was also considered as an important indicator of how accurate a route is drawn. On balance, a point rationing scheme was applied, accounting both mentioned factors. Table 3.2 shows this scheme, specifying the constraints for the respective score. 5 indicates the maximum score, whereas 1 stands for the lowest possible score.

Point	Constraints
5	start and end point AND hit rate = 100%
4	start and end point AND hit rate ≥ 60%
3	start and end point AND 60% ≥ hit rate ≥ 40% OR start or end point AND hit rate ≥ 60%
2	start or end point AND 60% ≥ hit rate ≥ 40% OR neither start nor end point AND hit rate ≥ 60%
1	everything else

Table 3.2.: Point rationing scheme

Task completion regarding the visual search task was indicated by a mouse click on the respective target. However, it must also be determined whether participants had even clicked on the right target and if so, how accurate they had been. Therefore, a bounding box defining the target area was placed around the actual target. Since the target on the Sheffield map was an entire square, size and location of the bounding box were defined by the square's dimension. For the target on the Birmingham map, the bounding box comprised the POI as well as the label. Clicks within the respective target area were contemplated as accurate.

3.5.3. Eye-tracking Metrics

Eye-tracking data help to understand subjects' perception of the visual stimuli whereat assumptions about the information retrieval can be made. To do so, quantitative analysis were done using a handful of metrics that seem fairly relevant in the context of this experiment. The following section lists these metrics which will be brought up again when presenting the gaze data analysis in Chapter 5.

Fixation duration (FD) as well as visit duration (VD) are time measures, describing how long a participant looked at a predefined area of interest (AOI). FD describes the length of fixations in seconds within an AOI whereas VD is the total time in seconds for every time a participant had looked within an AOI, starting with a fixation inside the defined AOI but ending with one outside the AOI. Another kind of time metric is the time to first fixation (TFF) of a particular AOI. Additionally, when participants were asked to click somewhere inside an AOI, time to first fixation minus time to first mouse click can be calculated. This measurement will be analyzed regarding the visual search task. Besides time metrics, fixation counts (FC) and visit counts (VC) are measured. FD is the total number of fixation within an AOI whereas VC is a measurement adding up the number of visits for a particular AOI.

All metrics described above are ratio scale data meaning that statistical data analysis and reporting will be done according to Section 3.5.1.

4 General Results

This chapter provides the general results from the experiment as well as outcomes from the questionnaire users were answering subsequently to the main experiment. According to the experiment design, the chapter reports results from task 1 and task 2 separately. Moreover, a separate section refers to the order effects which were particularly prominent due to the study design. Findings from the eye-tracking analysis can be found in Chapter 5. In order to easily distinguish between the independent variables, the two different routes are named by their location (Sheffield and Birmingham). These locations are sometimes abbreviated by "S" and "B". Accordingly, the abbreviations "FC" and "LH" refer to the respective highlighting method.

For all results, the two main effects were reported firstly. This concerns differences between the highlighting methods (LH - FC) regardless of the map locations and differences between map locations (Birmingham - Sheffield) regardless of the highlighting methods. The main effects are the results of averaging the dependent variables across the respective level of the independent variable. Any noteworthy interaction effects are reported subsequently to the main effects.

4.1. Task 1

In task 1 map reading time (studying time) was measured. The upper limit of three minutes per stimuli was not reached by any of the participants. Regarding the main effects, all data sets are normally distributed. When grouping data by highlighting method, a dependent t-test revealed no statistical significant difference between studying time using FC maps (M = 46.63, SE = 4.88) and studying time using LH maps (M = 48.14, SE = 4.77), $t(23)$ = .25, p = .816 , r = .05. Figure 4.1 shows the box plot as well as the error bar chart for this analysis. Looking at the second main effect, also no statistical significant difference between studying time using the Birmingham maps ($M = 44.64$, $SE = 4.25$) and studying time using the Sheffield maps ($M = 50.11$, $SE = 3.86$) can be reported, $t(23) = 1.61$, $p = .121$, $r = .31$. Figure 4.2 shows the respective plots for the second main effect. Any pairwise comparison (interaction effects) show no deviation from the main effects and are therefore not further described.

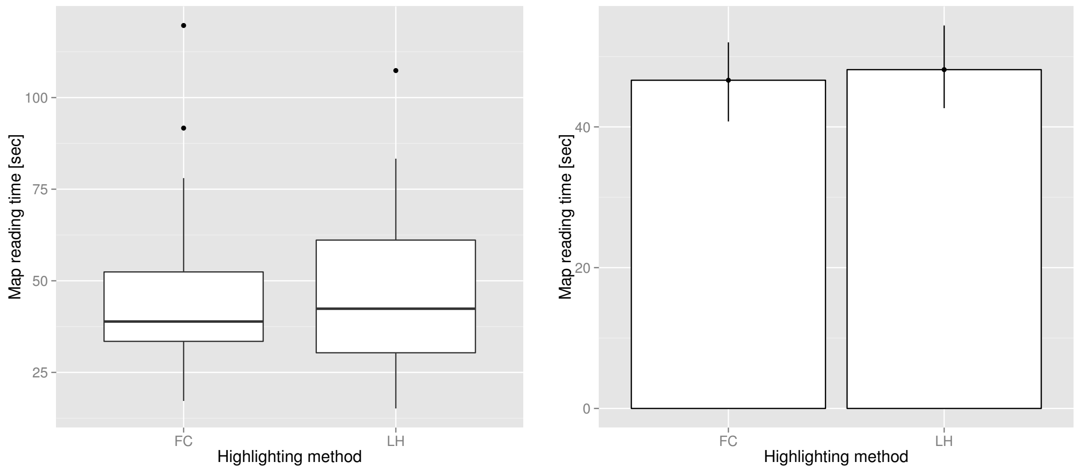


Figure 4.1.: Map reading time grouped by highlighting method (left: box plot; right: error bar chart)

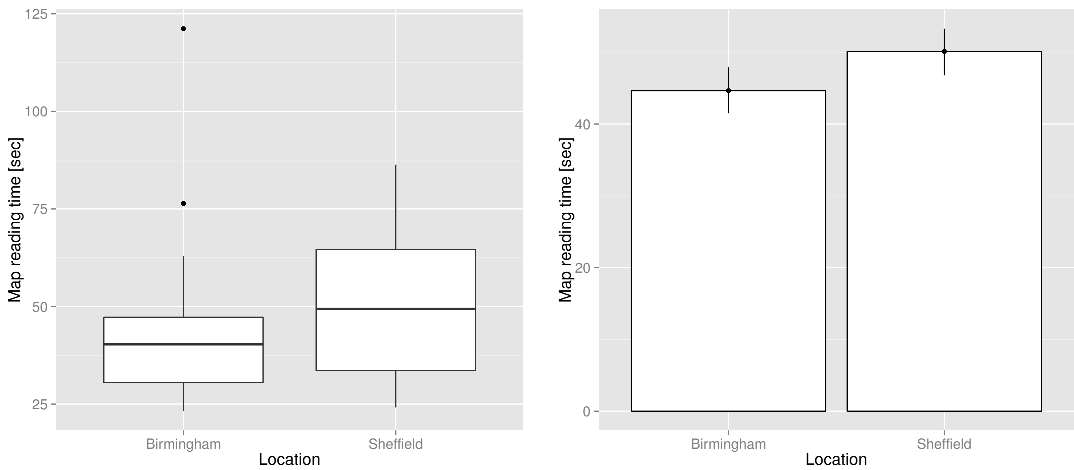


Figure 4.2.: Map reading time grouped by location (left: box plot; right: error bar chart)

Performance in task 1 is a measurement of accuracy and time needed to draw the route on another map. First, drawing time analysis will be presented followed by the accuracy analysis. All data sets concerning drawing time are normally distributed and therefore tested for differences using the dependent t-test. In contrast, the accuracy

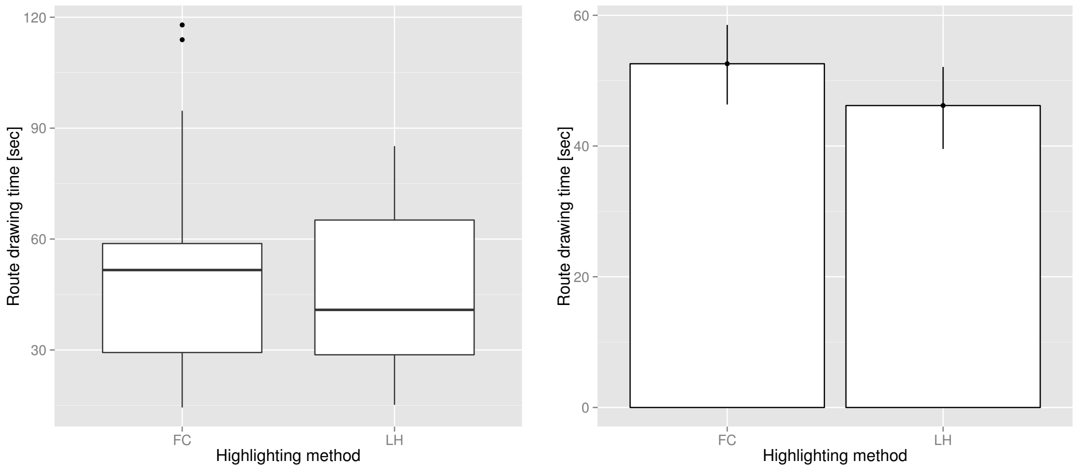


Figure 4.3.: Route drawing time grouped by highlighting method (left: box plot; right: error bar chart)

measurement is on an ordinal measurement level and therefore analyzed with the non-parametric sign-test.

Drawing the routes presented on the FC maps (M = 52.59, SE = 5.56) did not needed significantly more or less time than doing so for the routes presented on the LH maps (M = 46.19, SE = 4.56), $t(23)$ = -1.0, p = .327 , r = .21. Hence, no difference can be reported comparing drawing times grouped by highlighting method.

On the other hand, a difference in drawing time can be found when looking at the data grouped by map location. The routes presented on the Birmingham maps (M = 38.37, SE = 4.64) took participants significantly less time to draw than the ones on the Sheffield maps (M = 60.41, SE = 5.32), $t(23)$ = 3.62, p = .002 , r = .61. Figure 4.3 and 4.4 show box plots and error bar charts regarding both main effects. The error bar chart in the latter figure clearly shows the difference of route drawing time between the two map locations.

Pairwise comparisons between locations show the same situation as for the respective main effect. Thus, comparing any highlighting method on the Sheffield map with any other on the Birmingham map results in a significant difference. However, interaction effects regarding comparisons between highlighting methods at each location separately does not show any significant difference. Figure 4.5 shows the box blot and error bar chart for all four stimuli separately where the abovementioned interaction effects become evident.

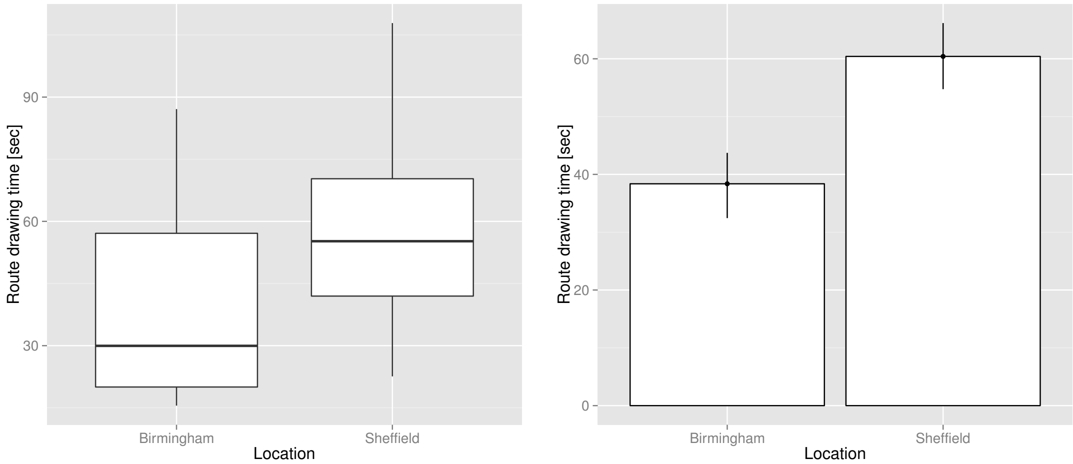


Figure 4.4.: Route drawing time grouped by location (left: box plot; right: error bar chart)

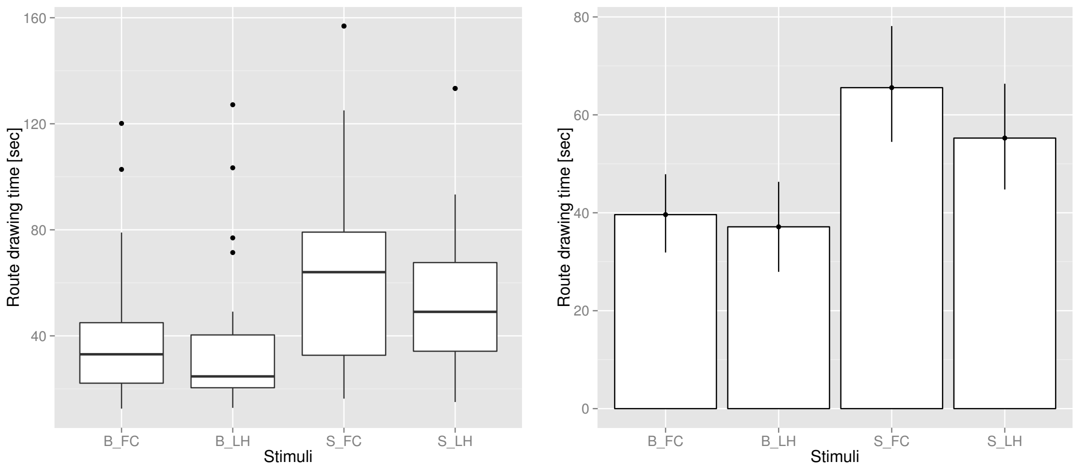


Figure 4.5.: Route drawing time presented for all four stimuli separately (left: box plot; right: error bar chart)

Looking at the main effects regarding drawing accuracy, no statistical significance can be observed considering the differences between FC maps ($Mdn = 4$) and LH maps ($Mdn = 4$), $p = .211$. Figure 4.6 depicts the box plot of both datasets. Similar to the findings from the drawing time, a highly significant difference between the Birmingham maps ($Mdn = 5$) and the Sheffield maps ($Mdn = 3$) can be reported, $p = .001$ (see Figure 4.7).

When looking at the interaction effects between the highlighting methods, no significant difference is observable. Neither is there a statistical trend observable comparing FC and LH methods for the Birmingham map, nor is there any for the Sheffield one. Though, pairwise comparisons between locations are highly significant in all possible cases. Figure 4.8 presents the box plots showing drawing accuracy for all four stimuli separately.

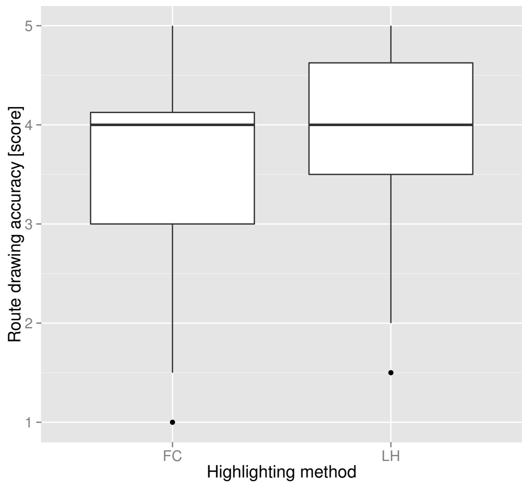


Figure 4.6.: Route drawing accuracy grouped by highlighting method

Summing up, no statistical significances regarding any effect could be found for map studying time. On the contrary, significant differences were found for map drawing time as well as for drawing accuracy when data were grouped by location. Both results indicate, that the route on the Birmingham map needed less time to read and was easier to retrace than the route on the Sheffield map. Differences between the highlighting methods can not be reported for all kind of comparisons regardless type of measurement.

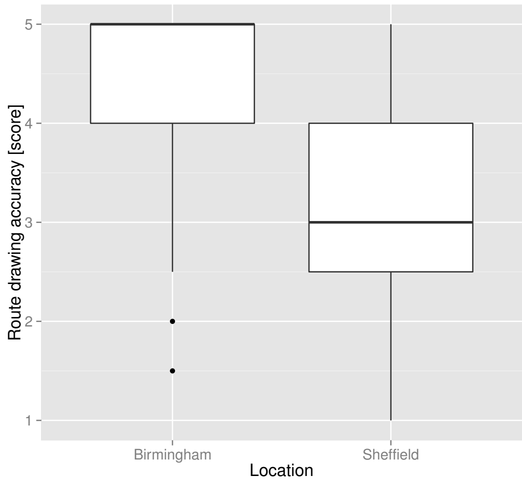


Figure 4.7.: Route drawing accuracy grouped by map location

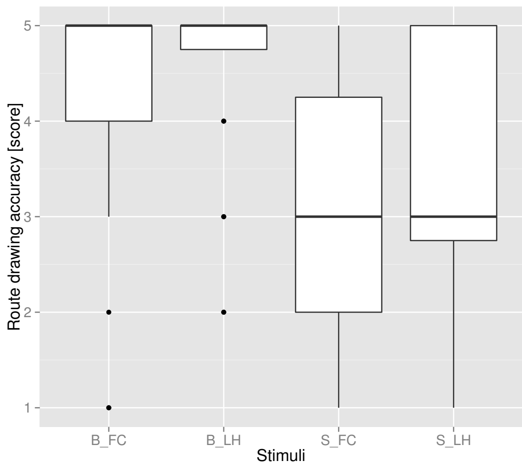


Figure 4.8.: Route drawing accuracy presented for all four stimuli separately

4.2. Task 2

Task 2 required to find an off-route location on the maps. Data sets describing the main effects are all normally distributed. Regarding search times between highlighting methods regardless of the map location reveals that search times using FC maps (M = 21.61, SE = 3.74) do not differ from search times using LH maps (M = 21.41, SE = 3.51), $t(23)$ = -0.03, p = .973 , r = .01.

Regarding data grouped by map location reveals that search time for the target on the Birmingham maps (M = 17.56, SE = 2.24) were significantly lower than the search time for the targets on the Sheffield maps (M = 38.37, SE = 3.36), $t(23)$ = 2.03, p = .049 , r = .39. Figure 4.9 presents data grouped by highlighting method whereas Figure 4.10 depicts data aggregated by map location.

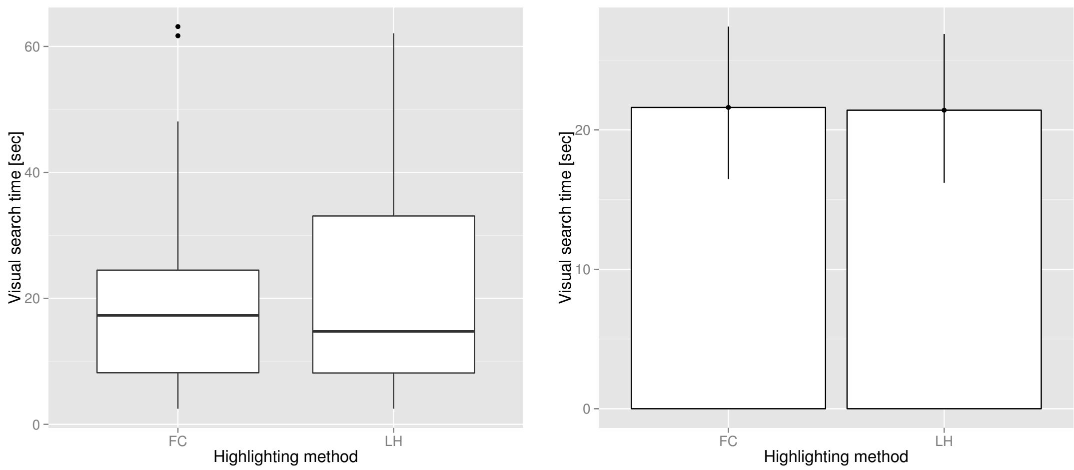


Figure 4.9.: Task 2 completion time grouped by highlighting method (left: box plot; right: error bar chart)

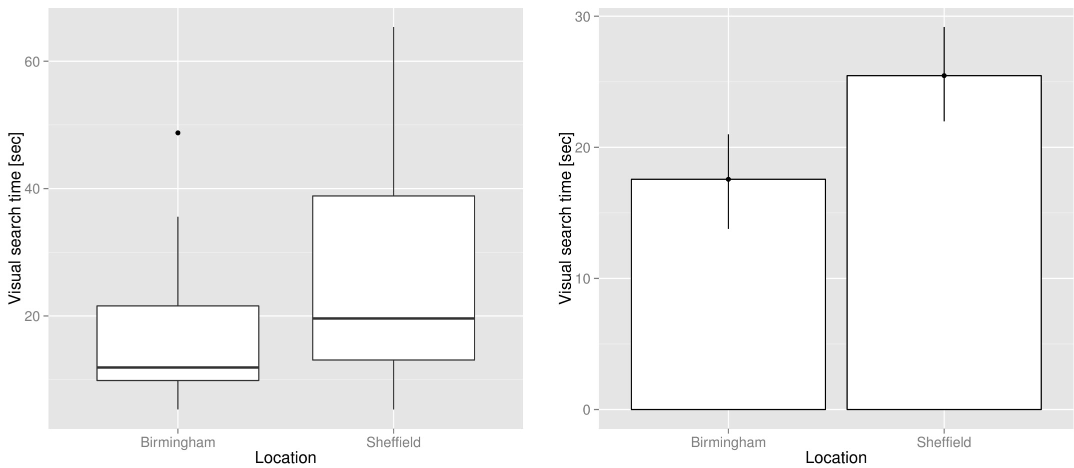


Figure 4.10.: Task 2 completion time grouped by map location (left: box plot; right: error bar chart)

When looking at the results across all stimuli, the box plot in Figure 4.11 reveals a rather large sampling variance across all stimuli. Furthermore, all outliers depicted in the box-plot describe high values, indicating a positive skewed distribution across all stimuli. Hence, confidence intervals are fairly large in all single data-sets pointing out

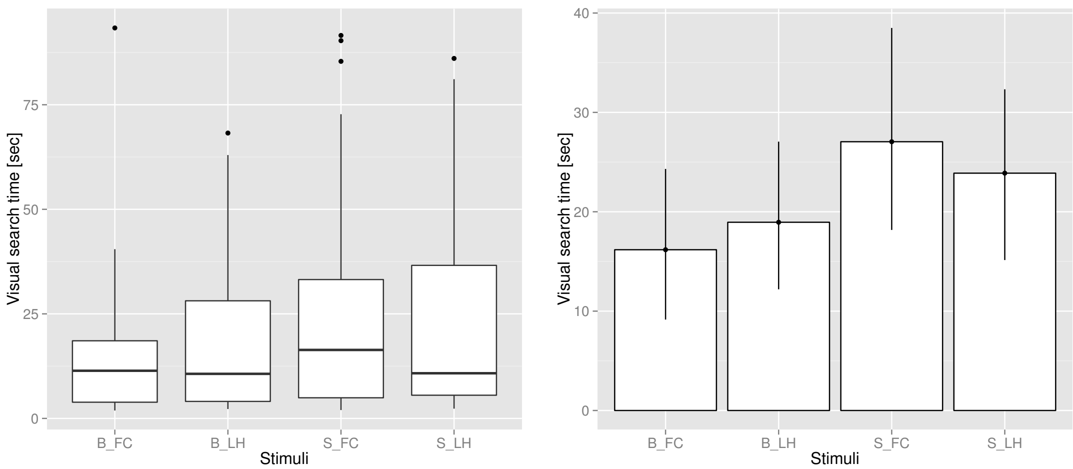


Figure 4.11.: Task 2 completion time presented for all stimuli separately (left: box plot; right: error bar chart)

no significant difference when looking for interaction effects between the single stimuli. This applies also for any differences between map locations.

The abovementioned distributional effect might be a result of the task order. Since the same target was searched twice on the same map, it is questionable whether the second time the target was detected was an actual search task at all. Regardless of the highlighting method, participants probably just remembered where the target was. Hence the second measure could merely indicate the speed in mouse handling. Therefore, a trimmed analysis considering only first time searches is conducted.

Regarding data grouped by highlighting method, no significant difference in search time between FC maps (M = 36.87, SE = 6.32) and LH maps (M = 35.96, SE = 4.84) is found, $t(11)$ = -0.17, p = .872, r = .05. Grouping the data by map location reveals no significant difference between search time on the Birmingham maps (M = 29.38, SE = 4.74) and search time on the Sheffield maps (M = 43.45, SE = 7.62), $t(11)$ = 1.75, p = .107, r = .47. Figure 4.12 and 4.13 depict the results for the respective trimmed analysis. Yet, it is important to state that in consequences of aggregating data this way halves the total number of observation (N = 12).

Regarding pairwise comparisons, no significant differences can be found when looking for any interaction effects. However, comparing only the FC maps reveals a statistical trend, indicating that search time for the target on the Birmingham map (Mdn = 19.64) was lower than for the target on the respective Sheffield map (Mdn = 32.17),

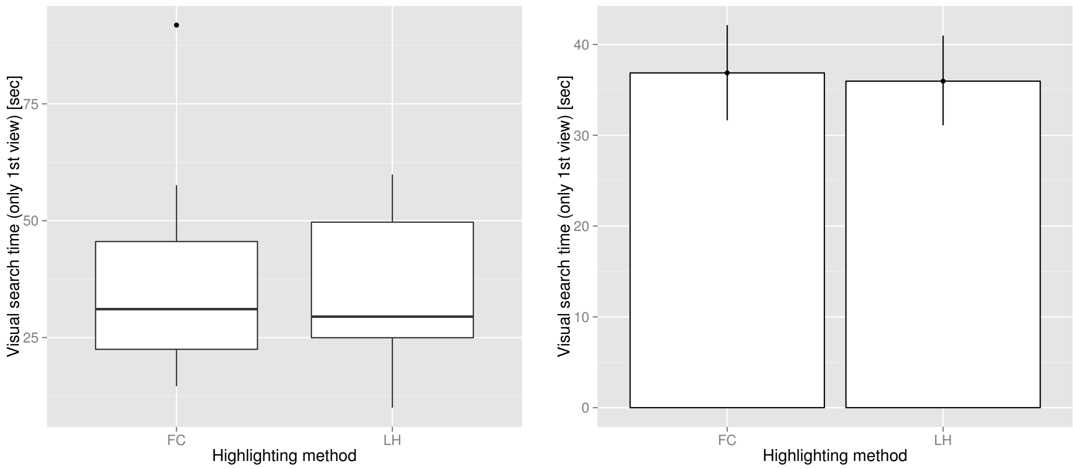


Figure 4.12.: Task 2 completion time (only 1st time searches) grouped by highlighting method (left: box plot; right: error bar chart)

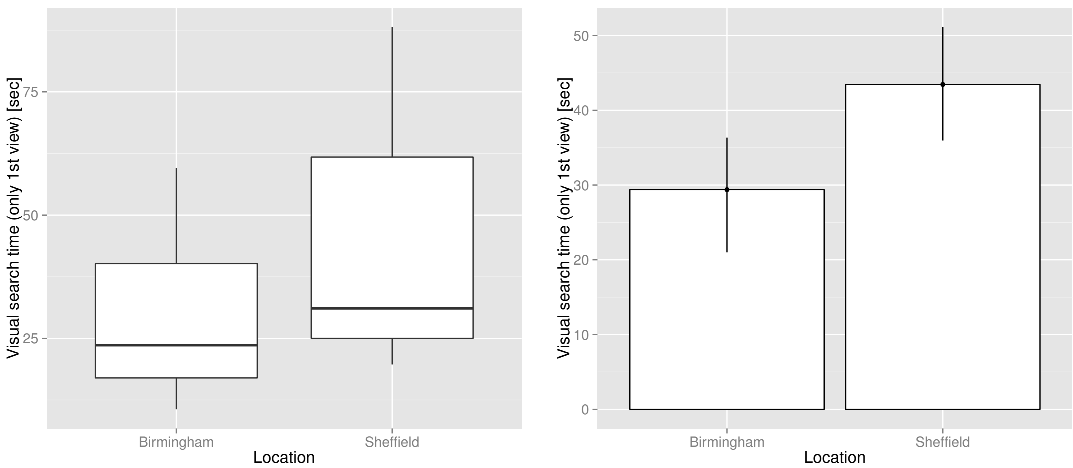


Figure 4.13.: Task 2 completion time (only 1st time searches) grouped by map location (left: box plot; right: error bar chart)

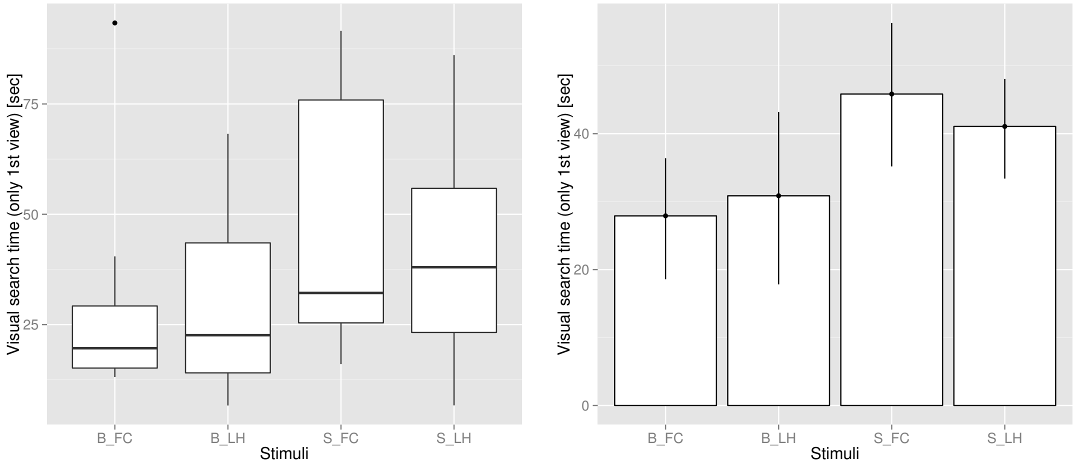


Figure 4.14.: Task 2 completion time (only 1st time searches) presented for all stimuli separately (left: box plot; right: error bar chart)

p = .093 r = -.34. Figure 4.14 shows the box plot and error bar chart concerning the trimmed analysis for all stimuli separately. It is evident that less outliers regarding results for all stimuli are present.

In short, task 2 revealed no difference in visual search time between the two highlighting methods. A difference was only found between the map locations, showing that less time was needed to find the target on the Birmingham maps than on the Sheffield ones. However, this result was only found for the main effect. Pairwise comparisons did not show any significant differences. A further analysis where only first time searches were considered showed a statistical trend between FC methods among locations. However, the latter analysis comprised only 12 data entries per stimuli.

4.2.1. Accuracy of the visual search task

As explained in Chapter 3.5.2, accuracy was measured using a bounding box around the respective search target. Regarding the target on the Birmingham map, 40 out of 48 clicks were made within the defined area. However, in eight cases participants did not click on the target at all. For the Sheffield target, 42 clicks were counted within the target area whereas in 6 cases no clicks at all were made. In all cases where participants did not click on the target, they fixated it, meaning that they simply forgot to confirm the end of the search process. Therefore, data for task 2 could have been extracted by using gaze movements. In brief, whenever participants clicked on the target, they clicked within the defined target area. In all other cases, accuracy can not be evaluated conclusively.

4.3. Performance by DI

A further analysis concerns task performances with regard to performance of the object location memory test. Since there is no correlation between the two performance measurements with participants DI values (see 3.3.1) regarding the entire datasets, a group analysis is proposed. For this purpose participants' DI were sorted and graduated in four groups. The upper quartile and the lower quartile define the two groups which will be compared (for each group, $N = 6$).

Time data for task 1 as well as for task 2 are normally distributed, meaning that a dependent t-test is used to examine differences between the two groups.

Regarding drawing time, testing for a difference reveals that the low DI group ($M = 57.09$, $SE = 12.51$) does not differ significantly to the high DI group ($M = 44.02$, $SE = 7.74$), $t(5) = .67$, $p = .53$, $r = .27$. Similarly, for the drawing accuracy a sign-test reveals no significant difference between the low DI group ($Mdn = 3.88$) and the high DI group ($Mdn = 4.25$), $p = .225$.

For task 2, the same general output can be observed. A dependent t-test shows no significant difference in search times between the low DI group ($M = 20.88$, $SE = 5.19$) and the high DI group ($M = 15.87$, $SE = 3.57$), $t(5) = .62$, $p = .561$, $r = .27$.

Figure 4.15 and 4.16 show the box-plots for both tasks. However, these results must be regarded cautiously when bearing in mind the counterbalanced design applied for this study. It is not just that the sample size of the groups is rather low ($N = 6$), but also the order of the stimuli across both groups is not balanced anymore.

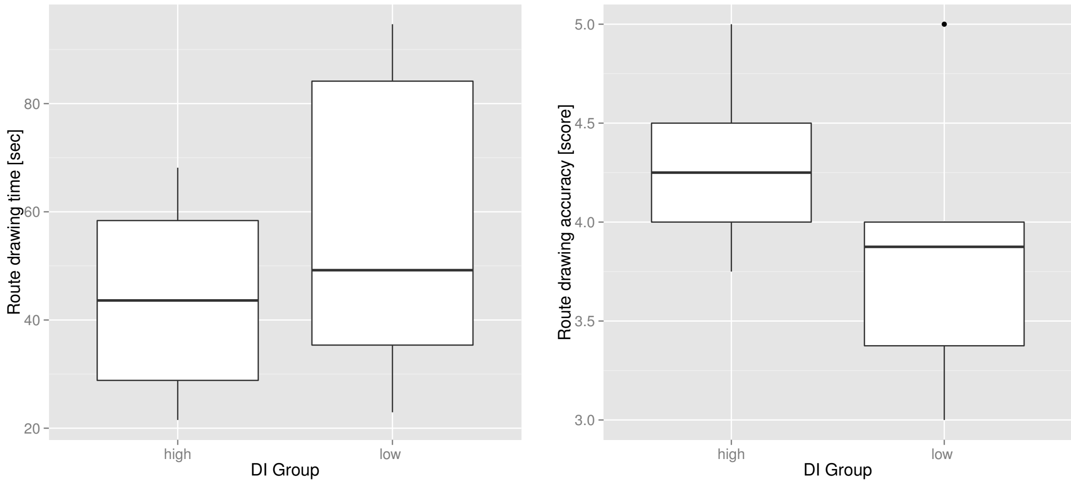


Figure 4.15.: Box-plots comparing high and low DI group for task 1

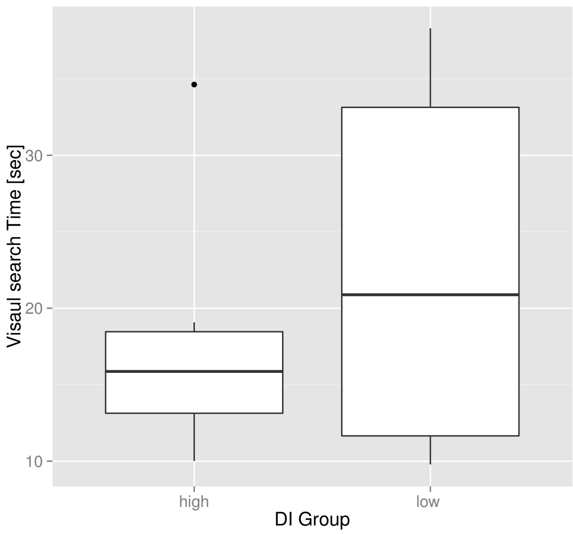


Figure 4.16.: Box-plots comparing high and low DI group for task 2

4.4. Order effects

As seen in almost all previously presented results, distribution of data is rather large. This comes as no surprise, then order effects heavily affected the outcomes. Due to the counterbalancing scheme, these effects should have been equalized throughout the tasks. Nevertheless, it is worth to have a closer look at how strongly order effects influenced the results. Since both routes were presented twice, learning effects regarding the route locations were at strongest. To measure these effects, differences between the first exposure of a route contrasted to the second time the route was exposed were analyzed.

Data regarding learning effect about studying time (task 1) are normally distributed for the Birmingham maps. A dependent t-test reveals a highly significant difference between the first exposure (M = 59.81, SE = 5.36) and the second exposure (M = 49.08, SE = 7.77), $t(23) = 2.61$, $p = .016$, $r = .48$. On the other hand, data regarding learning effect concerning studying time for the Sheffield maps are not normally distributed, meaning that a Wilcoxon signs-rank test was performed. Yet, the output is quite similar. The difference between first time exposure (Mdn = 57.59) and second time exposure (Mdn = 34.69) are also highly significant for the Sheffield maps, p = .008 r = -.54. The effect size r for both map locations is fairly high, meaning that participants spent significant less time for memorizing the route the second time they were exposed to it. Figure 4.17 shows the error bar charts for both locations.

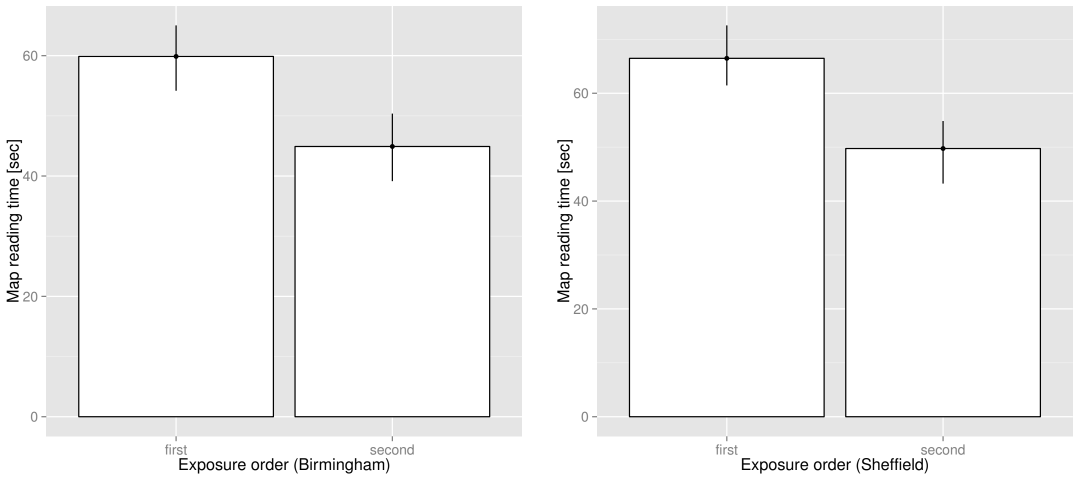


Figure 4.17.: Error bar charts showing learning effects of map studying time for both map locations

The sampling distribution of the differences between first and second time the route was drawn is not normally distributed for both locations, meaning that the non-parametric Wilcoxon signed-rank test was applied. For the Sheffield map, the first time the route was drawn (Mdn = 74.96) differs significantly from the second time it was done (Mdn = 43.82), p = .003 r = -.59. Interestingly, no such difference between the first time (Mdn

=33.05) and the second time (*Mdn* = 33.09) the route was drawn on the Birmingham map is observable, p = .229 r = -.25. Figure 4.18 shows error bar charts regarding learning effects for both locations. Comparing the plots, the difference in learning effect between the two maps gets visually evident.

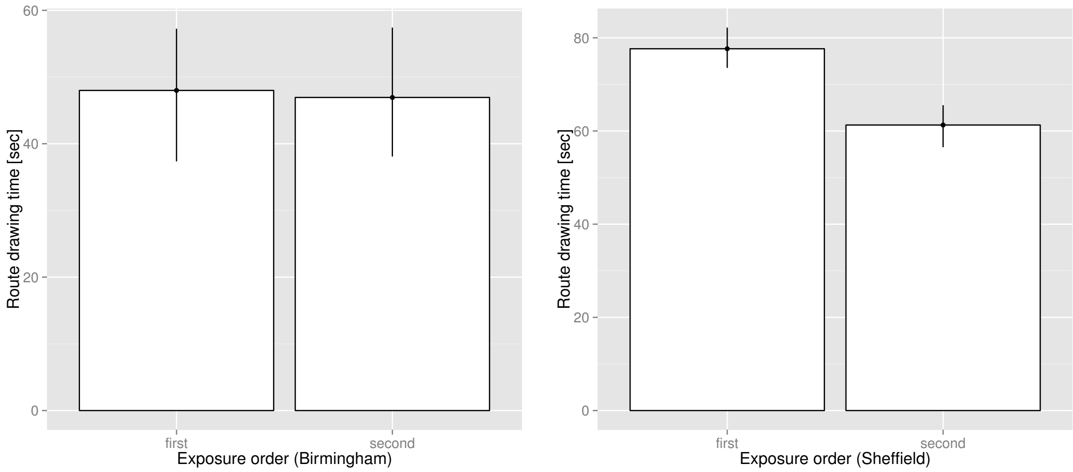


Figure 4.18.: Error bar charts showing learning effects of drawing time for both map locations

Learning effects concerning drawing accuracy were not found for both route locations regardless of highlighting methods. Neither was there a difference between first and second time the Birmingham routes were drawn (*Mdn* = 5 & *Mdn* = 5, p = 1.23) nor a considerable learning effect regarding the Sheffield routes (*Mdn* = 3 & *Mdn* = 3, p = 1). The box plot in Figure 4.19 shows these findings, indicating that no significant performance improvements were made when the route was drawn the second time. Regarding the Birmingham route, these findings are hardly surprising since the median score for drawing accuracy were fairly high even for the first time the route was drawn.

In task two, one might think that learning effects are rather remarkable for both map locations. As stated earlier, the second time the target was searched on the same map can hardly be considered as visual search at all. Nevertheless, learning effects are only evident for one location. Differences between the two time measurements are normally distributed for the Birmingham map but not for the Sheffield map. The first search on the former one (M = 30.81, *SE* = 5.07) does not differ significantly to the second search (M = 18.55, *SE* = 5.32), $t(23)$ = 1.65, p = .111, r = .33. However, learning effects regarding the visual search on the Sheffield map is noteworthy. The first time the target was searched (*Mdn* = 33.47) differs significantly to the second time it was

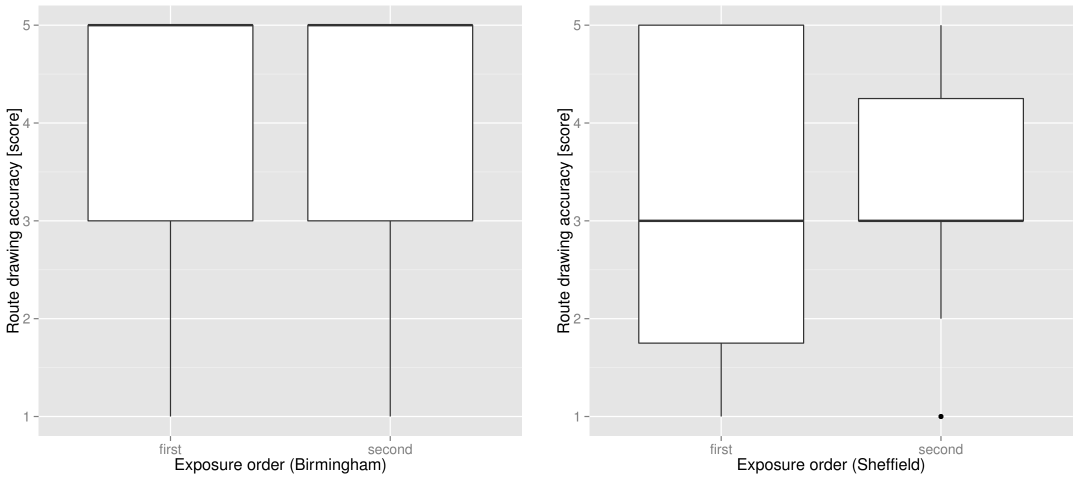


Figure 4.19.: Box plots showing learning effects of drawing accuracy grouped for both map locations

searched (Mdn = 6.61), p = .001 r = -.68. Figure 4.20 depicts the error bar charts for both locations, showing that confidence intervals for the Birmingham location are considerably larger.

All order effects presented above are main effects grouped by map location. Examining order effects grouped by highlighting method would not be as meaningful, since the dominant order effect is clearly a result of seeing the same map excerpt twice, regardless of the highlighting method.

Besides analyzing the actual learning effects, further examinations about the differences between these effects were done. Having learning effects grouped by map location, it is interesting to see whether they differ between the Birmingham and Sheffield map location throughout the tasks. Yet, no difference regarding studying time can be found between the Birmingham map (M = 30.44, SE = 4.79) and the Sheffield map (M = 29.77, SE = 7.18), $t(23)$ = -.11, p = .912, r = .02. Since only one map location showed learning effect regarding drawing time and no effect at all were observable for drawing accuracy, it is not substantive to check these data sets for learning effect differences. The same statement applies for learning effects regarding speed of search in task 2.

Differences between learning effects aggregated on map locations can further be analyzed in order to find out whether the order of the highlighting method influences the size of the learning effect. Yet, no significant result can be reported for this analysis

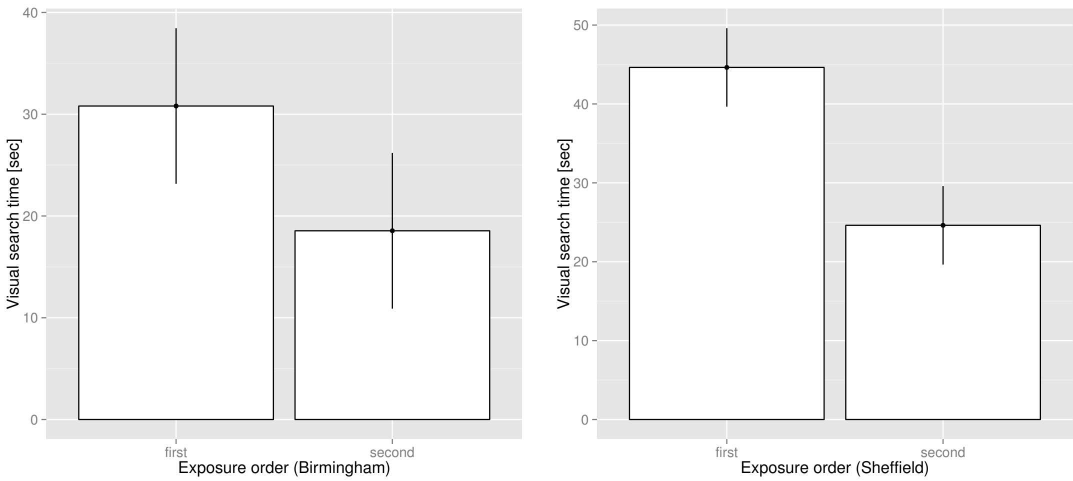


Figure 4.20.: Error bar charts showing learning effects of task 2 for both map locations

regarding measurements in both tasks. Hence, if the route is first presented first as FC map or as LH map does not make a difference.

Regarding the different stimuli orders, it is further questionable if seeing the same map locations straight after each other results in a higher learning effect than in the case of seeing the two map locations not in a row. But again, no significant differences or trends were observable for this analysis. Nevertheless, it has to be remarked, that for the latter analysis number of data is halved ($N = 12$).

4.5. Results from the post-questionnaire

The post-questionnaire comprised two rating scale questions which were answered subsequently to the main experiment. The first question addressed the highlighting methods, where subjects were asked to rate each highlighting method in terms of usefulness regarding route memorizing support. Figure 4.21 reports this outcome. A Wilcoxon signed-rank test reveals a statistical trend indicating that subjects rated the LH method ($Mdn = 4$) slightly higher than the FC method ($Mdn = 3$), $p = .074$ $r =$ -.36. However, the effect size r for this trend is rather small. For the avoidance of doubt, no one rated any of the methods as "not helpful at all".

The second question concerned the usefulness of map features in terms of route memorizing support. Therefore, participants were asked to rate five different features: street

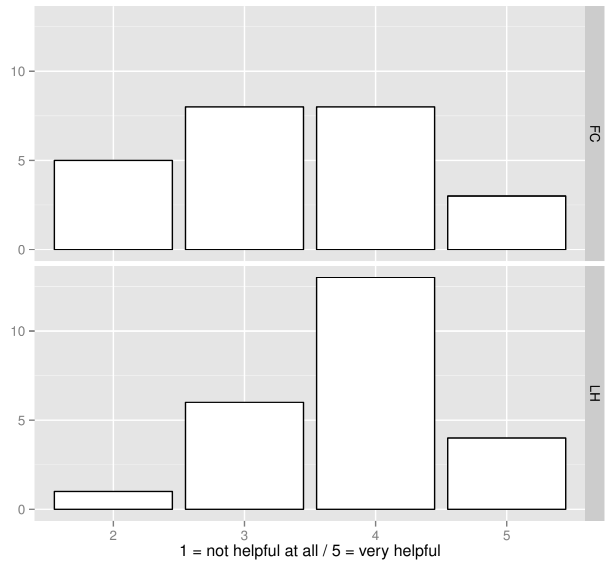


Figure 4.21.: Rating of highlighting methods

names, shape of route, prominent objects, shape of prominent object and labeled POIs. The meaning of prominent objects were further explained by giving examples like main street, parks and water bodies. Figure 4.22 reveals the outcome of this question. Most participants considered street names (*Mdn* = 5) as well as the shape of the route (*Mdn* = 4) as the two most useful features when memorizing a route. Moreover, nobody denoted these features as "not helpful as all". Regarding usefulness of prominent objects (*Mdn* = 4) and labeled POIs (*Mdn* = 3), participants denoted the former one as slightly more useful than the latter one. As a least helpful support for route memorizing, participants indicated the shape of prominent buildings (*Mdn* = 3). Figure 4.23 shows the box plots for all features. In addition to the given features, participants could also report any other map elements that helped them to memorize the route. One person stated having used different way marking methods whereas for another one street angles were helpful to memorize the route.

The results from the second question apply primarily for GM's map design. Prominent objects as well as the shape of prominent buildings are two map features that are heavily depending on how they are depicted and where the map excerpt comes from. Therefore, rating results for these two map features have to be regarded cautiously in terms of external validity.

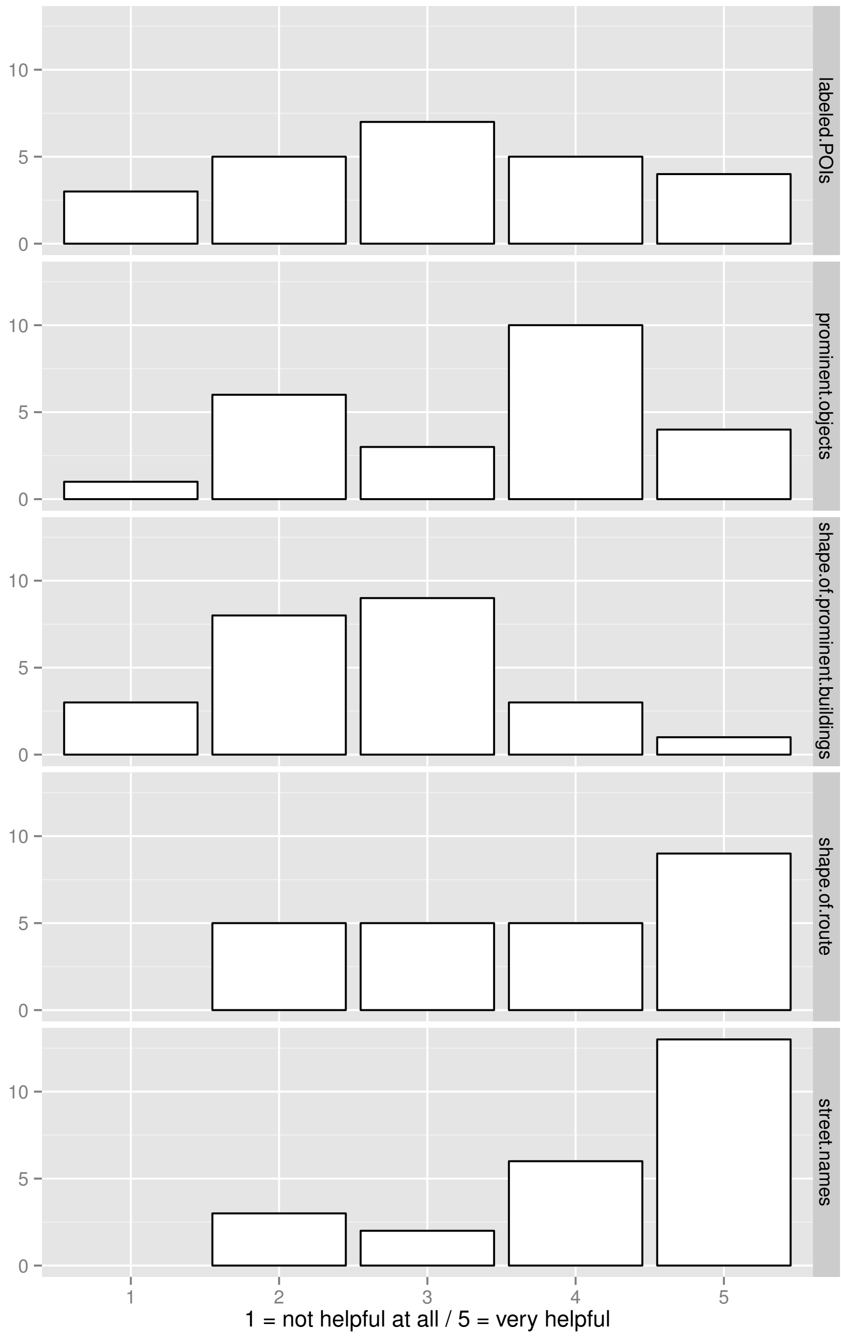


Figure 4.22.: Rating of different map features in term of memorizing support

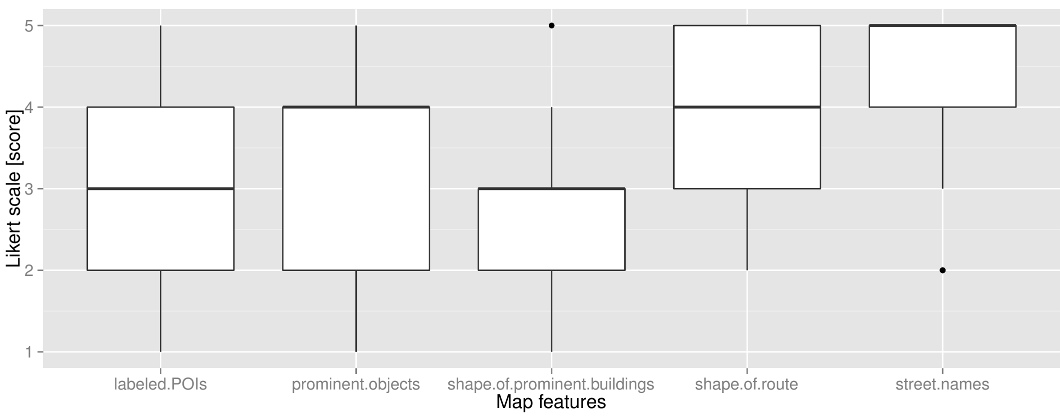


Figure 4.23.: Box plots showing ratings of usefulness of different map features

5 | Eye-tracking Results

This chapter discusses the findings from the eye-tracking data. First, an analysis regarding task 1 will be presented, followed by an examination bringing deeper insight into the gaze behavior concerning the visual search task. For all analysis, AOI's for the two map excerpts were defined separately. Therefore, the main effect concerning difference between map locations is not reported for task 1 due to differently shaped and sized areas.

5.1. Task 1

Task 1 required participants to memorize a route without having given any further purposes to do so. In order to examine if highlighting methods influence gaze behavior, eye tracking data are analyzed for the route memorizing part of task 1. Therefore, an AOI which is identical with the focus zone is defined for each stimulus. Figure 5.1 and 5.2 show this zone for both map locations drawn in the LH maps. Accordingly, the the same AOIs are defined for the respective FC maps. In the following, this area of interest will be abbreviated with the letter "F".

The main effect analyzed for the F zone concerns the differences between FC and LH method regarding each map location separately. The metrics used to examine this effect were relative fixation duration, relative visit duration as well as relative fixation counts and relative visit counts. The relativeness refers always to the total studying time (ST) a participant took to memorize the route.

A dependent t-test shows a statistical trend when comparing relative FD between FC maps (M = .52, SE = .03) and LH maps (M = .49, SE = .03), $t(23)$ = -2.03, p = .054, r = -.39. This indicates that participants spent proportionally more time looking on the focus zone with the FC map than with the LH map. Figure 5.3 shows the box-plot and the error bar chart which are relevant for this analysis.

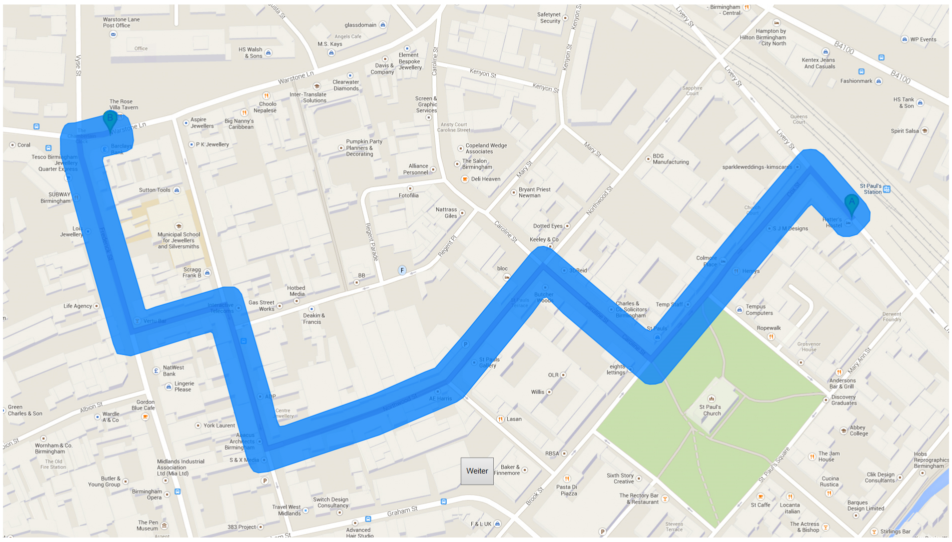


Figure 5.1.: F - AOI surrounding the Birmingham route

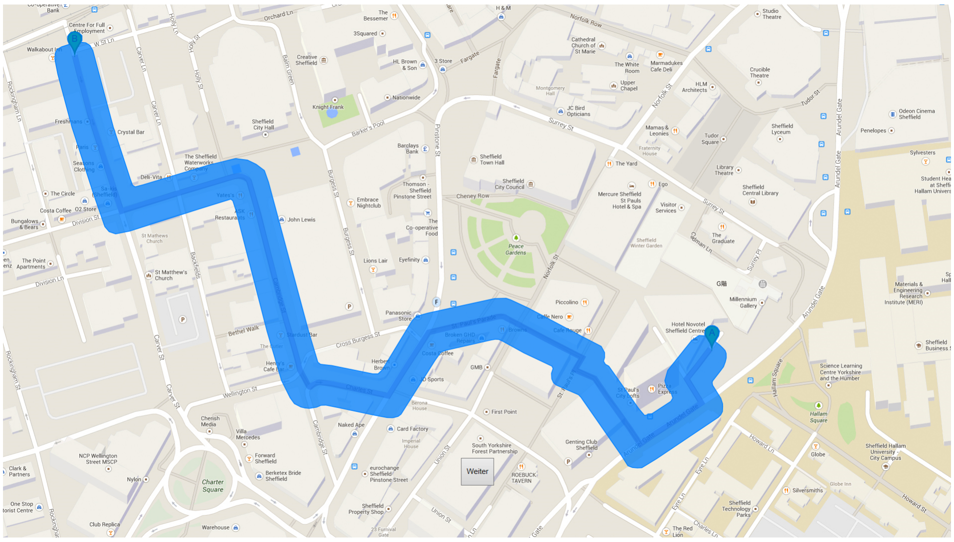


Figure 5.2.: F - AOI surrounding the Sheffield route

When looking at the same data but presented for every stimuli separately, the error bar charts in Figure 5.4 reveals a bigger difference between highlighting methods on the Birmingham map than on the Sheffield one. Nonetheless, pairwise comparisons for the Birmingham map uncover no significant difference in relative FD between the FC map (M = .54, SE = .03) and the LH map (M = .49, SE = .03), $t(23)$ = -1.69, p

= .103, r = -.33. A similar outcome is found concerning the Sheffield location where relative fixation duration regarding the F-AOI does not significantly differ between the FC map (M = .51, SE = .03) and the LH map (M = .49, SE = .03), $t(23)$ = -.76, p = .456, r = -.16.

In brief, a difference regarding the main effect for relative FD can be observed in form of a statistical trend between FC maps and LH maps indicating a proportionally longer fixation duration when looking at FC maps. However, this trend can not be stated when looking at the interaction effects for each map location separately.

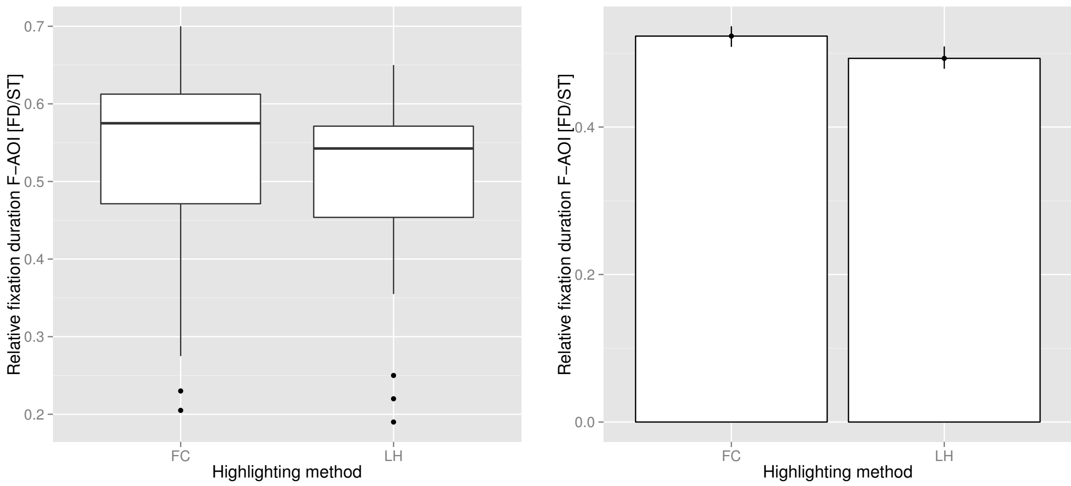


Figure 5.3.: Relative fixation duration grouped by highlighting method (left: box plot; right: error bar chart)

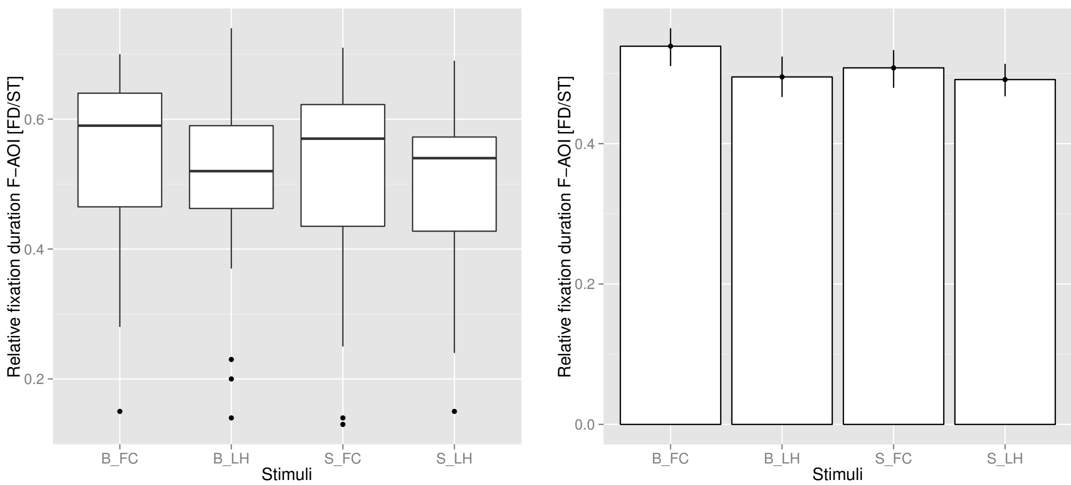


Figure 5.4.: Relative fixation duration presented for all stimuli separately (left: box plot; right: error bar chart)

Besides fixation duration, visit duration was also used for this analysis. As with the former one, only VD proportional to the total studying time is examined. In contrast to FD, VD adds up the time spent within an AOI including the time to the next fixation outside the AOI. Therefore, the more often a participant switched between the respective AOI and its surrounding, the wider the gab between these measurements gets. The difference between the two data sets is normally distributed, meaning that a dependent t-test is applied for this analysis. It points out that on the FC maps no longer relative visit duration for the F-zone (M = .60, SE = .0.3) is observable compared to the LH maps (M = .57, SE = .02), $t(23)$ = -1.67, p = .109, r = -.33. Figure 5.5 shows the box plot and error bar chart concerning this analysis.

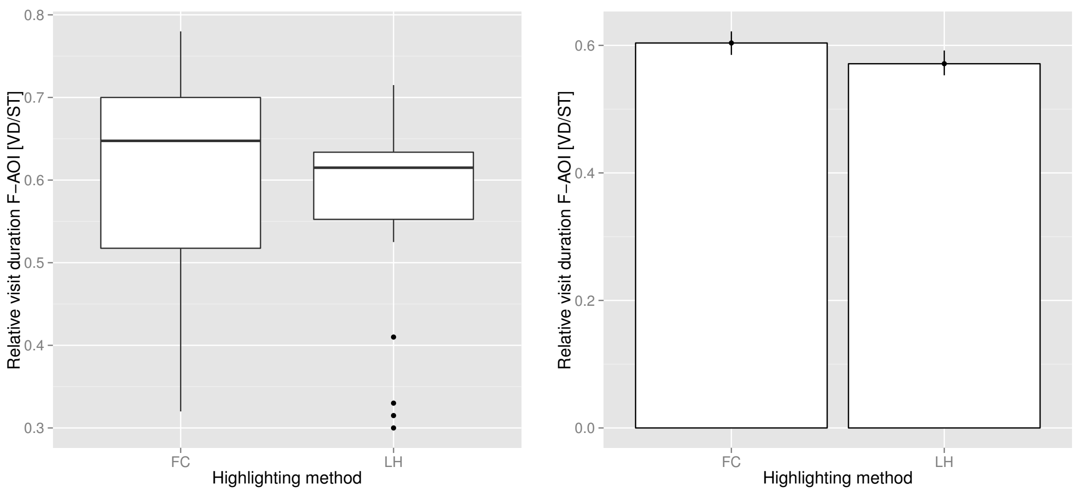


Figure 5.5.: Relative visit duration grouped by highlighting method (left: box plot; right: error bar chart)

Even when no difference is found regarding the main effect for relative visit duration an interesting result can be reported for one of the interaction effects. Considering only the F-zone on the Birmingham map, participants' relative visit duration was significantly longer on the FC map (Mdn = 0.66) than on the LH map (Mdn = 0.6), p = .023, r = .46. In addition, it has to be stated that data for the former comparison are not normally distributed whereas the one for the following comparison are normally distributed. A dependent t-test revealed no significant difference between FC method (M = .58, SE = .03) and LH method (M = .57, SE = .02), $t(23)$ = -0.64, p = -.528, r = .132. Figure 5.6 depicts these outcomes.

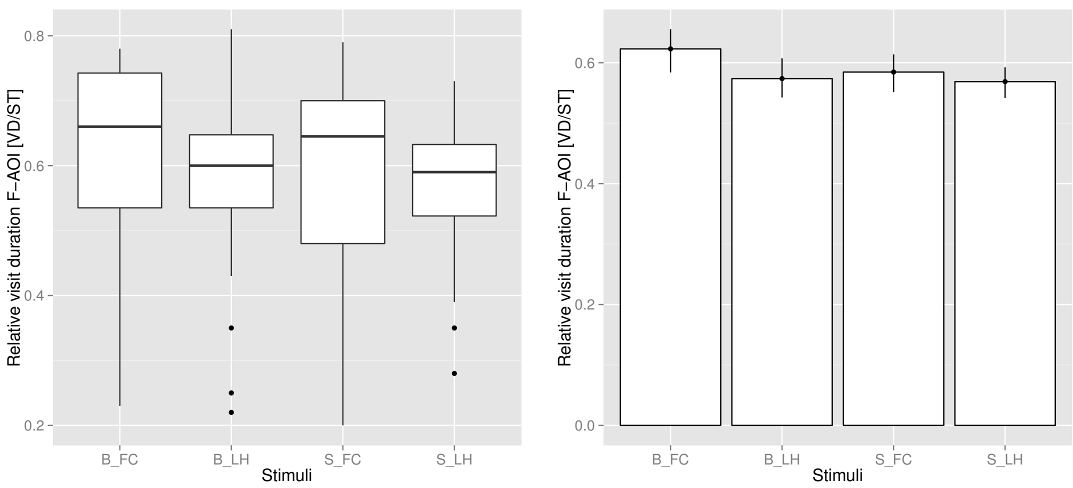


Figure 5.6.: Relative visit duration presented for all stimuli separately (left: box plot; right: error bar chart)

Complementary to the time metrics used in the analysis described above, relative fixation counts as well as visit counts are further examined. Data sets regarding the following analysis are all normally distributed. Comparing fixation counts between FC maps (M = .52, SE = .02) and LH maps (M = .52, SE = .02) revealed no significance difference, $t(23)$ = -0.53, p = .602, r = -.11. Relative fixation counts regarded for each location separately show the same general outcome; no significant result can be reported, neither for a difference on the Birmingham map nor one on the Sheffield map.

In contrast to fixation counts, visit counts give insight in how many times participants switch between the defined AOI and its surrounding areas. However, regarding the main effect, relative visit counts do not differ significantly between FC maps (M = .52, SE = .02) and LH maps (M = .55, SE = .02), $t(23)$ = -1.46, p = .157, r = -.28. Same general result is found when looking at the interaction effects for pairwise comparisons between highlighting methods at both locations. Figure 5.7 shows the box plot as well as the error bar chart depicting data for all stimuli separately.

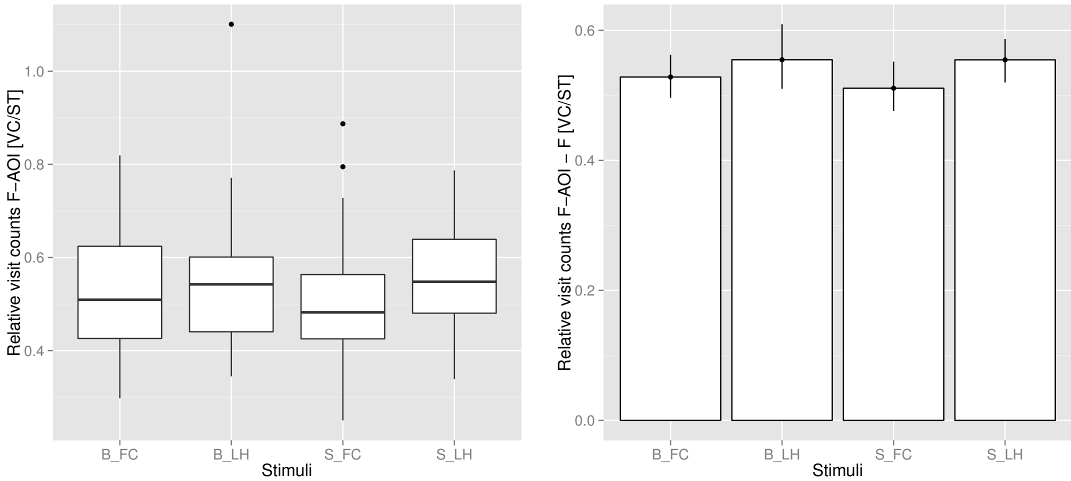


Figure 5.7.: Relative visit counts presented for all stimuli separately (left: box plot; right: error bar chart)

5.2. Task 2

Gaze behavior regarding the visual search task is certainly heavily depending on the position of the target in respect of the indicated search starting point. Nevertheless, to make the two locations comparable an equally sized area of interest around the search targets was defined. Figure 5.8 and 5.9 show the area for both maps. In the following, this AOI will be abbreviated with the letter "T".

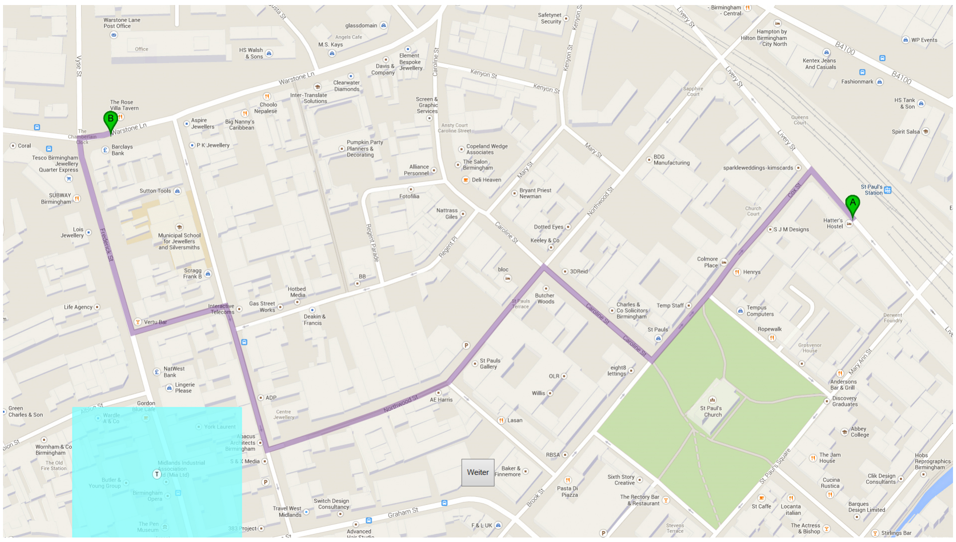


Figure 5.8.: T - AOI placed around the search target (Birmingham map)

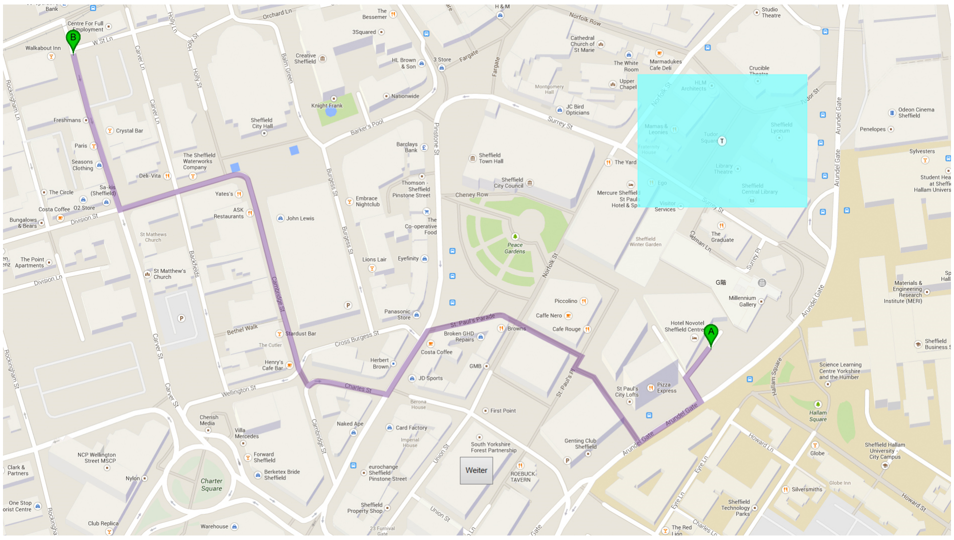


Figure 5.9.: T - AOI placed around the search target (Sheffield map)

The eye-tracking metric used for this analysis was the time to first fixation (TFF) compared with task completion time (TCT). The difference between these two measurements reveals the time a participant needed to successfully locate the target after having had the first fixation within the T-zone.

Data for the main effect regarding the difference between highlighting methods are normally distributed. A dependent t-test brings out no significant result between FC maps (M = 15.63, SE = 3.13) and LH maps (M = 13.84, SE = 2.39), $t(23)$ = -40, p = .691, r = .08. On the other hand, data regarding the difference between the two map locations are not normally distributed. Hence, a Wilcoxon signed-rank test was used indicating that time difference for the T-zone on the Birmingham map (Mdn = 9.26) is slightly lower than the one on the Sheffield map (Mdn = 13.66) which results in a statistical trend, p = .078 r = -.36. This means, that after the first fixation within the T-zone, participants needed less time to detect the actual target on the Birmingham map than on Sheffield map. Figure 5.10 shows the main effect regarding the difference between highlighting methods whereas Figure 5.11 depicts the one for differences between the two maps. When looking at the interaction effects no significant difference can be reported. The described difference (statistical trend) between the two map locations cannot be found for any simple effect, neither when comparing the same highlighting method between locations nor for any simple cross comparison concerning location and highlighting method.

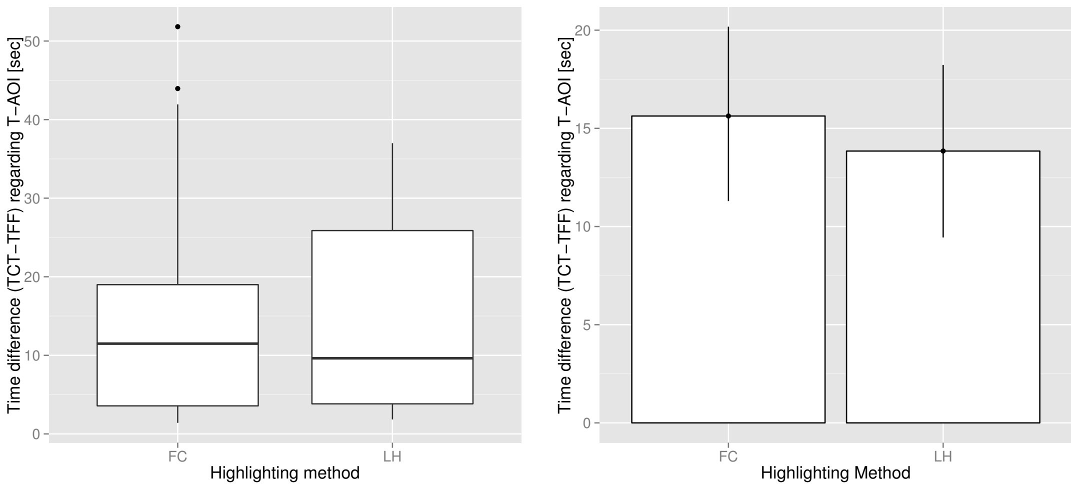


Figure 5.10.: TCT-TFF grouped by highlighting method (left: box plot; right: error bar chart)

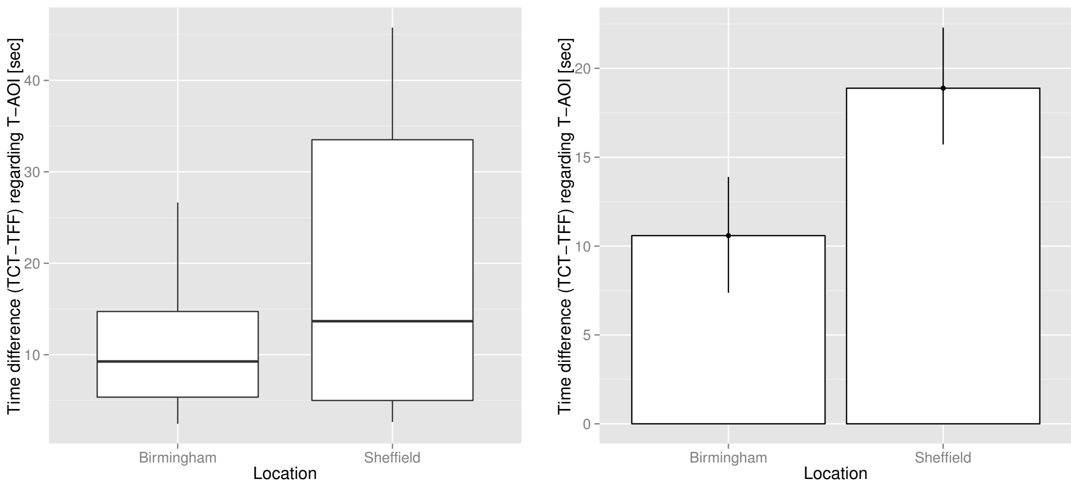


Figure 5.11.: TCT-TFF grouped by locations (left: box plot; right: error bar chart)

6 | Discussion

The present study was designed to investigate the differences between a common route highlighting method and a focus and context map highlighting approach. To do so, a human subject study was conducted involving a route memorizing and retracing task as well as a visual search task. This chapter sums up and discusses the results in view of the research questions and provides interpretations for the respective findings. Furthermore, the outcomes are put in context with previous research on this topic. Besides the main results, Subsection 6.2 discusses exploratory observations made while analyzing the data. In closing, study limitations are addressed.

6.1. Research Questions

> *How does the FC method differ from the LH method in terms of route memorizing performance (mapping the memorized route on another kind of map)?*

Efficiency

Results regarding the efficiency in route drawing showed no significant difference between the two highlighting methods. This outcome was observable for the aggregated data as well as for pairwise comparisons where methods were analyzed for each location separately. Accordingly, the tested highlighting methods might not have a difference influence on the retracing speed.

Effectiveness

Effectiveness was measured by accuracy of the retraced route on another kind of map. Similar to the outcomes of the efficiency analysis, no significant difference in accuracy between the two highlighting method was revealed. Hence, it can be said, that the tested highlighting methods might not have a different influence on the accuracy of a retraced route.

Both performance measurements seem not to be differently influenced by the map highlighting methods. In order to retrace the route on another map as accurate and fast as possible, several intermediate steps are necessary. First, the map excerpt must be recognized by references to memorized map objects, whereby first parts of the route can be detected. From there, the route can be recreated using any memorized map entities that characterized the route. From this point of view it gets apparent that besides the actual map depiction (e.g. different highlighting methods) individual map reading and memorizing strategies most likely influence retracing performance. It might be the case, that strategies are rather important regardless of the highlighting method. Previous research on the topic of route memorizing abilities mainly distinguished between a landmark based strategy, where preferably route knowledge is acquired, and a geometrically based strategy strategy where rather survey knowledge is extracted (Siegel and White, 1975). Since many studies found clear evidence for differences in strategy usage regarding route memorizing tasks (Dabbs Jr et al., 1998) it can be assumed that similar strategies had been applied in the present experiment. However, applying different strategies might not lead to difference in route memorizing performance as Galea and Kimura (1993) stated.

A second hypothesis is that the chosen encoding strategies are depending on the highlighting method but do not influence memorizing performance. As Hirtle and Hudson (1991) found out, route knowledge can also be acquired by studying a map. Since the FC method is only providing full information for objects adjacent to the route, it could have been possible that landmarks in this zone are preferentially selected as memorizing support. Along with this assumption, it might be thinkable that landmarks lying just next to the route are used to memorize rather directional information. Consequently, FC highlighting might encourage subjects to extract route knowledge. Another impact on memorizing strategies might have verbal re-coding of visually perceived map information (Garden et al., 2002). Presumably, participants having silently translated the route into verbal directions could have also acquired rather directional information. Anyway, the more aspects of the working memory are involved in the memorizing process, the more information can be stored. The reason therefore lies in the fact that subjects who actively applied visual and verbal strategies acquired information referring to more than one storage component (Baddeley and Hitch, 1974). However, no evidence is present confirming the use of verbal re-coding strategies and even if so, it would be still questionable what kind of knowledge is preferably acquired.

Besides testing memorizing performance, a question revealed participants' satisfaction for each highlighting method. A statistical trend indicated a slightly higher satisfaction rate for the line highlighting method. This result could be explained by the degree of familiarity towards the LH method. According to the pre-questionnaire, frequency of use of Google Maps is fairly high among almost all participants meaning that the line highlighting method is better known than the proposed FC highlighting approach. To put it briefly, participants may have rated according to their familiarity, implying that the more common and used a method is, the higher the satisfaction rate gets

> *How does the FC method compared to the LH method influence the use of the context (off-route) information in a memorizing task?*

In order to answer this research question, context information was defined as all information lying outside the focus region, which is given by the buffer around the route. The use of information was further analyzed by examining several eye-tracking metrics.

Results for relative fixation duration in the respective zone showed a statistical trend, indicating that subjects looked proportionally longer at the focus zone when memorizing the route depicted with the FC highlighting than with the LH method. However, no such effect was observable when comparing the methods for each location separately. The results concerning the main effect for the proportional visit durations showed no difference between the highlighting methods. Hence, it can be generally said that the tested highlighting methods might not have a different influence on relative visit durations. Nonetheless, results regarding the map locations separately revealed significant longer visit durations for the focus zone on the Birmingham map when studying the route using FC highlighting compared to the LH method. No such difference regarding the visit durations could be observed for the route on the Sheffield map.

Furthermore, fixation counts as well as visit counts were analyzed for the defined areas. Both measurements indicated no difference between the highlighting methods. This result was found for the main effect as well as for the interaction effects regarding locations separately. Accordingly, the tested highlighting methods may not differently influence fixation counts as well as visit counts when trying to memorize a route on such a map.

Interestingly, a difference between highlighting methods regarding fixation duration was found but no such result can be reported concerning fixation counts. Having a longer fixation duration but similar fixation counts denotes, that participants spent more time in average gazing at a particular point than switching on and off the route. This is the case when participants studied the route with FC highlighting. It can be assumed, that the highlighting methods influence gaze behavior differently and therefore diverse encoding strategies are applied. The FC method rather funnels the reader's attention towards the focus area where more time is spent and single fixations take longer in average. These findings go along with the abovementioned hypothesis, yet, nothing can be said about the kind of information that is retrieved. However, the effect size for the this observation is rather small, thus, any interpretations have to be regarded cautiously.

Another interesting finding concerns visit duration. Visit duration is defined as the time length a participant spends within an AOI summed up until the first fixation outside this AOI. Results show that visit durations were significantly longer on the focus zone when regarding the FC highlighting method than with the LH method on the Birmingham map. One can now assume that either more switching between focus zone and context zone took place or participants' gaze remained longer within the AOI (e.g. fixation duration). But neither visit counts nor relative fixation duration outcomes show similar characteristics between the two methods. However, the further away the first fixation point outside the AOI appears, the longer the visit duration gets without changing visit counts or fixation duration. Therefore, not only visit counts and fixation duration but also the distance from the last fixation point within to the first fixation point outside the AOI, influence visit duration. This might be an explanation for the abovementioned findings.

> *How does the FC method contrasted to the LH method influence the use of the context (off-route) information in a visual search task for off-route objects?*

As the result showed, speed of visual search for off-route targets did not differ between the FC highlighting and the LH method. The same general outcome was observed for comparisons regarding each location separately. Furthermore, when only first time searches were considered, general result did not change. Hence, it can be stated that whether the target was searched on the ordinarily colored map or on the black and white variant did not affect search performance. Consequently, the color of the targets objects, as depicted in Google Maps, does not guide the reader towards these POIs. Hence, the given targets do not implicitly hold information that supports detecting process. For instance, looking for a water body or a park holds implicit color information whereat top-down knowledge could be used for faster detection (Treisman and Gelade, 1980). However, regarding GM's map design, some POIs such as cafes, restaurants or bus stops are colored more prominently which would probably lead to faster detection.

6.2. Exploratory Observations

Besides the main effect regarding differences between highlighting methods, data were always grouped by map location too. Concerning memorizing performance, huge significant differences were found. First off, retracing speed for the route on the Birmingham map excerpts was much lower than for the route on the Sheffield ones. Second, drawing accuracy was significantly higher for the route on the Birmingham maps. In short, memorizing and retracing the Sheffield route seems to be more difficult than the Birmingham route. In the present study, complexity of a path is primarily defined as the number of segments a route comprises. Even though the Sheffield route consists of more segments, the observed effect might be the result of additional factors. For instance, conjoining segments regarding the Birmingham route are rather orthogonally aligned which is probably easier to remember than having much wider angles between segments. Furthermore, the start point on the Birmingham map is probably more salient and better characterized on Google Maps as well as on OSM than for the Sheffield route. In general, the fewer characteristics are necessary to describe a route, the easier it might be to memorize and retrace it. Albeit the clear performance difference between the two routes, participants did not need significantly more time to memorize the one on the Sheffield map. Hence, subjects perceived route complexity might not have been different between the two routes.

Differences in performance were not only found for task 1, also in task 2 a significant difference between search times for the two targets could be reported. In this case, the target on the Birmingham map was faster detected than the one on the Sheffield map. However, a second analysis considering only the first time searches indicated no such difference. Nevertheless, it is arguable if this comparison is even meaningful due to a slightly different task setting. To better compare the two targets to each other, an equally sized bounding box around the target was used to determine the difference in time between first fixation within the bounding box and task completion time. As it turned out, the target on the Birmingham map was still slightly faster found after first fixation in the defined area. This outcome was surprising with regard to the fact that on the Sheffield map, the target itself (square) defines a bigger part of the bounding box. Furthermore, the label of the target on the Sheffield map stands rather alone meaning that less distracters and noise are present. Hence, the outcome is hardly explainable by common visual search theories. Perhaps, the finding could be explained as an issue of language comprehension meaning that participants were not aware of searching for an actual square.

A result that confirms the abovementioned findings is given by the analysis of order effects. Since no such effect could be reported regarding drawing speed as well as accuracy for the route on the Birmingham map, it might be the case that this route was probably too easy to memorize. On the contrary, strong learning effects were found for the Sheffield route.
Another observation regarding order effects concerned speed of visual search on the Birmingham map. In contrast to the target on the Sheffield map, search time for the Birmingham map did not significantly differ between first and second time the target was searched. This means, that despite of having found the target once, the second search probably still required to scan the region once again. Hence, the target on the Birmingham map does not provoke a pop-out effect (Treisman and Gelade, 1980) even when pre-knowledge about the target is present.

6.3. Limitations and Outlook

According to the findings made in the present study, focus and context highlighting does not differ to the common line highlighting method with regard to memorizing and visual search performance. However, the tested FC method is only one example of a focus and context implementation. Other approaches like for example blurring out context regions (Kosara et al., 2002) might lead to different results. Moreover, decoloring of map areas also reduces map information. However, the kind of information that is omitted depends on the design of the base map. Since the standard design of Google Maps does not use color in an extensive way (e.g. compared to OSM), the amount of information in a black and white version is not drastically reduced. It is therefore arguable, if focus and context highlighting based on color is a meaningful way to reduce visual map complexity for such a map design. Hence, applying the suggested FC approach to other maps might influence memorizing performance differently.

As the outcomes further revealed, decoloration of context area resulted in partly different gaze behaviors. Nonetheless, the question remains open whether the highlighting method affected participants' encoding strategy. Whenever it did so, it would be interesting to know if subjects adapted their reading strategy according to the respective kind of route depiction. Another factor that may have an influence on the reading strategy is the path's complexity. In order to determine how difficult the depiction of a route is, the individually perceived level of path complexity might be an alternative measurement, which could have been integrated in the present study.

Route memorizing performance was determined by assessing retracing speed and accuracy. However, doing so requires more than just memorizing abilities. It also involves reading and interpreting the map on which the route has to be drawn. This might be one reason why subjects' performances did not correspond with the outcomes of the object location memory test. As originally designed to examine sex differences (Silverman and Eals, 1992), the test does not reflect the same abilities as the ones needed in the present experiment. Memorizing map objects is probably more demanding than memorizing location of simple objects (Thorndyke and Stasz, 1980). Presumably, just having asked participants which objects they remembered could have correlated better with the pre-test outcomes.

Future research addressing the assessment of focus and context maps could include the analysis of users' reading strategies in order to determine whether highlighting methods have any influence or not. In addition, besides testing maps putting the focus on specific areas, it might be interesting to examine the effect of having put the focus on rather single map objects or a testing a combination of object and area based focus as suggested by Neis and Zipf (2007). As stated in the *Research Framework*, the actual map design should be put in the centre of interest and therefore, task-orientated research (e.g. route learning/memorizing) is important. The suggested approach of focus and context maps was performed as such, but can obviously be developed and further extended.

7 | Conclusion

The present study proposed a simplified focus and context approach where maps are divided into a context and focus region describing the way between two points. The basic idea was to provide navigational assistance while reducing visual complexity by decolorating less relevant areas. In contrast to the common line highlighted method, the suggested approach should help to more easily funnel the reader's attention towards the actual path. The study evaluated the proposed approach by contrasting it to the line highlighting method. In particular interest was whether there is a different influence on route memorizing performance regarding the two highlighting methods. Therefore, after memorizing a path participants were requested to retrace the way on another map. Performance was measured in drawing accuracy and task completion time. In addition, the methods were tested for visual search performance for finding off-route objects.

As the results revealed, no significant difference could be found regarding memorizing performance between the highlighting methods. Neither did they differ concerning retracing accuracy, nor for task completion time. These results were consistent for two distinctive routes from different locations. Findings regarding visual search performance for off-route objects did as well not differ between the tested stimuli. Hence, the use of the visual variable "color" in Google Maps might not be affecting search speed regarding simple POI detection.
Bearing in mind the results from the satisfaction rating and contrasting it two the abovementioned outcomes concerning participants performances, one might recommend the line highlighting method on a preferential basis. However, the tested focus and context highlighting is only a simplified approach which also needs to be examined for navigational tasks in real environment. As the eye-tracking data revealed, a partly different gaze behavior between the two highlighting methods was the case. Hence, it

might be thinkable that the highlighting methods guided subjects attention. Although street names as well as the shape of the route were generally rated as most helpful features when memorizing a path, it might still be possible that different features were used for different highlighting methods.

All things considered, the idea of guiding the map reader to task relevant information by distinguishing map content between focus and context elements has great potential. The present study can be regarded as a preliminary assessment of a focus and context highlighting approach where information is graduated on the basis of its location by means of color. Testing the influences of highlighting methods on different map reading strategies as well as assessing focus and context maps having put the focus on map objects remain a matter of future research.

Bibliography

Alvarez, G. A. and Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, 15(2):106–111.

Attneave, F. (1955). Symmetry, information, and memory for patterns. *The American journal of psychology*, 68(2):209–222.

Awh, E., Barton, B., and Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological science*, 18(7):622–628.

Baddeley, A. D. and Hitch, G. J. (1974). Working memory. *The psychology of learning and motivation*, 8:47–89.

Banerjee, J. (1994). Gestalt theory of perception. *Encyclopaedic Dictionary of Psychological.*

Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin*, 74(2):81.

Barkowsky, T. and Freksa, C. (1997). Cognitive requirements on making and interpreting maps. In *Spatial Information Theory A Theoretical Basis for GIS*, pages 347–361. Springer.

Berendt, B., Rauh, R., and Barkowsky, T. (1997). Spatial thinking with geographic maps: An empirical study. In Czap, H., Ohly, P., and Pribbenow, S., editors, *Herausforderungen an die Wissensorganisation: Visualisierung, multimediale Dokumente, Internetstrukturen*, pages 63–74. ERGON-Verlag, Würzburg.

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

Board, C. (1978). Map reading tasks appropriate in experimental studies in cartographic communication. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 15(1):1–12.

Brophy, D. M. (1980). Some reflections on the complexity of maps. *Technical papers of ACSM 40th Annual Meeting*, pages 343–352.

Choi, J. and Silverman, I. (1996). Sexual dimorphism in spatial behaviors: Applications to route learning. *Evolution and Cognition*, 2(2):165–171.

Dabbs Jr, J. M., Chang, E.-L., Strong, R. A., and Milun, R. (1998). Spatial ability, navigation strategy, and geographic knowledge among men and women. *Evolution and Human Behavior*, 19(2):89–98.

Dey, A. K. (2001). Understanding and using context. *Personal and ubiquitous computing*, 5(1):4–7.

Di Lollo, V., Kawahara, J.-i., Zuvic, S. M., and Visser, T. A. (2001). The preattentive emperor has no clothes: a dynamic redressing. *Journal of Experimental Psychology: General*, 130(3):479.

Duncan, J. and Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological review*, 96(3):433.

Eng, H. Y., Chen, D., and Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, 12(6):1127–1133.

Fairbairn, D. (2006). Measuring map complexity. *The Cartographic Journal*, 43(3):224–238.

Fairbairn, D. and Taylor, G. (1995). Developing a variable-scale map projection for urban areas. *Computers & Geosciences*, 21(9):1053–1064.

Freksa, C. (1999). Spatial aspects of task-specific wayfinding maps - a representation-theoretic perspective. In *Visual and spatial reasoning in design*, pages 15–32. Key Centre of Design Computing and Cognition, University of Sydne.

Galea, L. A. and Kimura, D. (1993). Sex differences in route-learning. *Personality and individual differences*, 14(1):53–65.

Garden, S., Cornoldi, C., and Logie, R. H. (2002). Visuo-spatial working memory in navigation. *Applied Cognitive Psychology*, 16(1):35–50.

Golledge, R. G. (1999). Human wayfinding and cognitive maps. In *Wayfinding behavior: Cognitive mapping and other spatial processes*, pages 5–45. The Johnes Hopkins University Press, Baltimore, Maryland.

Golledge, R. G., Smith, T. R., Pellegrino, J. W., Doherty, S., and Marshall, S. P. (1985). A conceptual model and empirical analysis of children's acquisition of spatial knowledge. *Journal of Environmental Psychology*, 5(2):125–152.

Google (2010). Google maps. https://www.google.com/maps.

Guerra, F. and Boutoura, C. (2001). An electronic lens on digital tourist city-maps. In *Proceedings of the 20th International Cartographic Conference*, pages 1151–1157.

Gutwin, C. and Fedak, C. (2004). A comparison of fisheye lenses for interactive layout tasks. In *Proceedings of Graphics interface 2004*, pages 213–220. Canadian Human-Computer Communications Society.

Gutwin, C. and Skopik, A. (2003). Fisheyes are good for large steering tasks. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 201–208.

Gwinn, H. M., Fernando, S., James, S., and Wilson, J. F. (2002). Do landmarks help or hinder women in route learning? *Perceptual and motor skills*, 95(3):713–718.

Harrie, L., Sarjakoski, L. T., and Lehto, L. (2002). A variable-scale map for small-display cartography. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(4):237–242.

Harrie, L. and Stigmar, H. (2010). An evaluation of measures for quantifying map information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(3):266–274.

Hirtle, S. C. and Hudson, J. (1991). Acquisition of spatial knowledge for routes. *Journal of Environmental Psychology*, 11(4):335–345.

Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(2):307.

James, T. W. and Kimura, D. (1997). Sex differences in remembering the locations of objects in an array: Location-shifts versus location-exchanges. *Evolution and Human Behavior*, 18(3):155–163.

Kemps, E. (1999). Effects of complexity on visuo-spatial working memory. *European Journal of Cognitive Psychology*, 11(3):335–356.

Klippel, A., Tappe, H., Kulik, L., and Lee, P. U. (2005). Wayfinding choremes—a language for modeling conceptual route knowledge. *Journal of Visual Languages & Computing*, 16(4):311–329.

Kosara, R., Miksch, S., and Hauser, H. (2002). Focus+context taken literally. *Computer Graphics and Applications, IEEE*, 22(1):22–29.

Lloyd, R. (1997). Visual search processes used in map reading. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 34(1):11–32.

Loftus, G. R. and Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic bulletin & review*, 1(4):476–490.

Luck, S. J. and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657):279–281.

MacEachren, A. M. (1982). Map complexity: Comparison and measurement. *The American Cartographer*, 9(1):31–46.

Martin, D. W. (2008). *Doing psychology experiments*. Thomson/Wadsworth, Belmont, California.

McBurney, D. H., Gaulin, S. J., Devineni, T., and Adams, C. (1997). Superior spatial memory of women: Stronger evidence for the gathering hypothesis. *Evolution and Human Behavior*, 18(3):165–174.

McGuiness, D. and Sparks, J. (1983). Cognitive style and cognitive maps: Sex differences in representations of a familiar terrain. *Journal of Mental Imagery*, 7(2):91–100.

Meilinger, T. and Knauff, M. (2008). Ask for directions or use a map: A field experiment on spatial orientation and wayfinding in an urban environment. *Journal of Spatial Science*, 53(2):13–23.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81.

Miller, L. K. and Santoni, V. (1986). Sex differences in spatial abilities: Strategic and experiential correlates. *Acta Psychologica*, 62(3):225–235.

Moar, I. and Carleton, L. (1982). Memory for routes. *The Quarterly Journal of Experimental Psychology*, 34(3):381–394.

Montello, D. R. (2002). Cognitive map-design research in the twentieth century: Theoretical and empirical approaches. *Cartography and Geographic Information Science*, 29(3):283–304.

Montello, D. R., Lovelace, K. L., Golledge, R. G., and Self, C. M. (1999). Sex-related differences and similarities in geographic and environmental spatial abilities. *Annals of the Association of American Geographers*, 89(3):515–534.

Neis, P. and Zipf, A. (2007). Realizing focus maps with landmarks using openls services. In Mok, E. and Gartner, G., editors, *Proceedings of the 4th International Symposium on Location Based Services and TeleCartography*, HongKong Polytechnic University. Department of Land Surveying and Geo-Informatics.

Oliva, A., Mack, M. L., Shrestha, M., and Peeper, A. (2004). Identifying the perceptual dimensions of visual complexity of scenes. In Forbus, K., Gentner, D., and Regier, T., editors, *Proceedings of the 26th Annual Cognitive Science Society*, pages 1041–1046. Austin, Texas.

Olson, I. R. and Jiang, Y. (2002). Is visual short-term memory object based? rejection of the "strong-object" hypothesis. *Perception & Psychophysics*, 64(7):1055–1067.

OpenStreetMap-contributors (2012). Open street maps. http://www.osm.org.

Pazzaglia, F. and De Beni, R. (2001). Strategies of processing spatial information in survey and landmark-centred individuals. *European Journal of Cognitive Psychology*, 13(4):493–508.

Poulton, E. and Freeman, P. (1966). Unwanted asymmetrical transfer effects with balanced experimental designs. *Psychological Bulletin*, 66(1):1.

Richter, K.-F., Peters, D., Kuhnmünch, G., and Schmid, F. (2008). What do focus maps focus on? In *Spatial Cognition VI. Learning, Reasoning, and Talking about Space*, volume 5248, pages 154–170. Springer.

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

Sarkar, M. and Brown, M. H. (1992). Graphical fisheye views of graphs. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 83–91.

Sarkar, M., Snibbe, S. S., Tversky, O. J., and Reiss, S. P. (1993). Stretching the rubber sheet: A metaphor for viewing large layouts on small screens. In *Proceedings of the 6th annual ACM symposium on User interface software and technology*, pages 81–91.

Schilit, B., Adams, N., and Want, R. (1994). Context-aware computing applications. In *Mobile Computing Systems and Applications, 1994. WMCSA 1994. First Workshop on*, pages 85–90. IEEE.

Schlender, D., Peters, O. H., and Wienhöfer, M. (2000). The effects of maps and textual information on navigation in a desktop virtual environment. *Spatial Cognition and Computation*, 2(4):421–433.

Schnur, S., Bektaş, K., Salahi, M., and Çöltekin, A. (2010). A comparison of measured and perceived visual complexity for dynamic web maps. In *Proceedings of 6th International Conference on Geographic Information Science 2010*.

Siegel, A. W. and White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in child development and behavior*, 10:9–55.

Silverman, I. and Eals, M. (1992). Sex differences in spatial abilities: Evolutionary theory and data. In Barkow, J. H., Cosmides, L., and Tooby, J., editors, *The adapted mind: Evolutionary psychology and the generation of culture*, pages 533–549. Oxford University Press.

Snyder, J. P. (1987). "magnifying-glass" azimuthal map projections. *The American Cartographer*, 14(1):61–68.

Stern, E. and Leiser, D. (1988). Levels of spatial knowledge and urban travel modeling. *Geographical Analysis*, 20(2):140–155.

Thorndyke, P. W. and Stasz, C. (1980). Individual differences in procedures for knowledge acquisition from maps. *Cognitive Psychology*, 12(1):137–175.

Treisman, A. (1985). Preattentive processing in vision. *Computer vision, graphics, and image processing*, 31(2):156–177.

Treisman, A. M. and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136.

Vandenberg, S. G. and Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and motor skills*, 47(2):599–604.

Vogel, E. K., Woodman, G. F., and Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 27(1):92.

Ward, S. L., Newcombe, N., and Overton, W. F. (1986). Turn left at the church, or three miles north a study of direction giving and sex differences. *Environment and Behavior*, 18(2):192–213.

Ware, C. (2004). *Information visualization: perception for design*. Morgan Kaufmann.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic bulletin & review*, 1(2):202–238.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, 9(1):33–39.

Wolfe, J. M., Cave, K. R., and Franzel, S. L. (1989). Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human perception and performance*, 15(3):419.

Wolfe, J. M. and Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5(6):495–501.

Yamamoto, D., Ozeki, S., and Takahashi, N. (2009). Focus+glue+context: an improved fisheye approach for web map services. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 101–110.

Zipf, A. and Richter, K.-F. (2002). Using focus maps to ease map reading — developing smart applications for mobile devices. *Künstliche Intelligenz, Special Issue Spatial Cognition*, 4(02):35–37.

A | Appendix

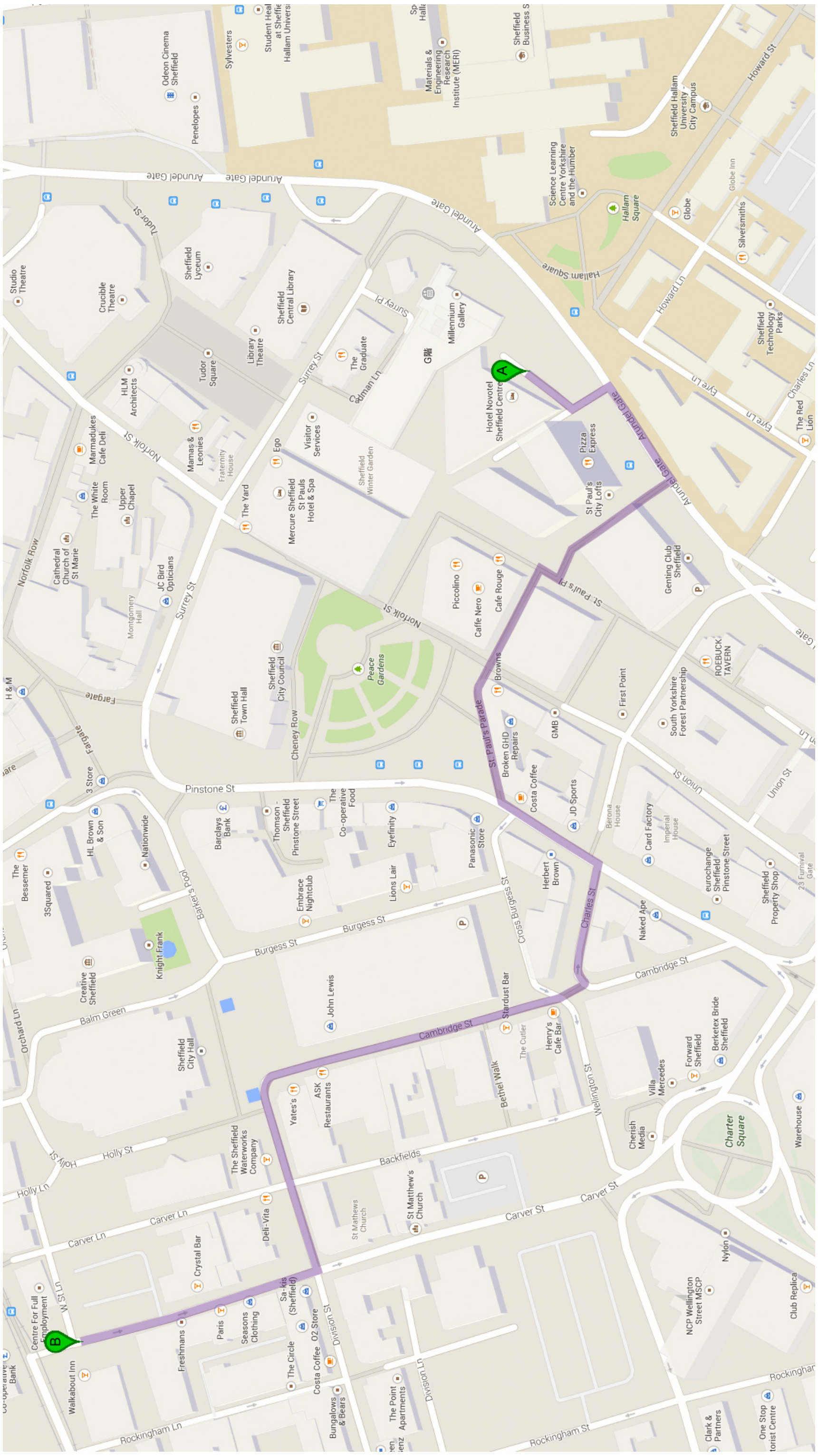


Figure A.1.: GM excerpt of Sheffield - LH highlighting

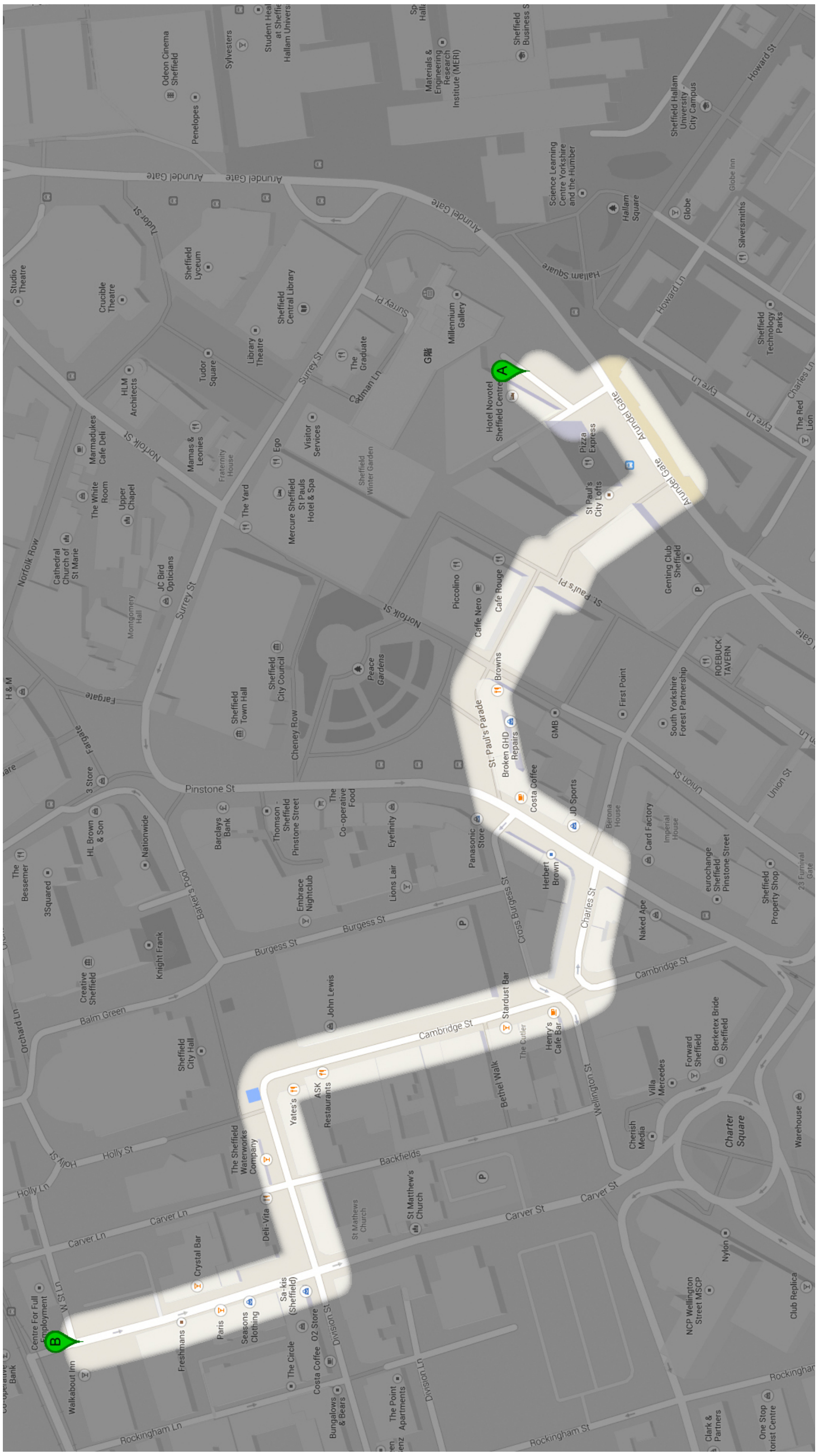


Figure A.2.: GM excerpt of Sheffield - FC highlighting

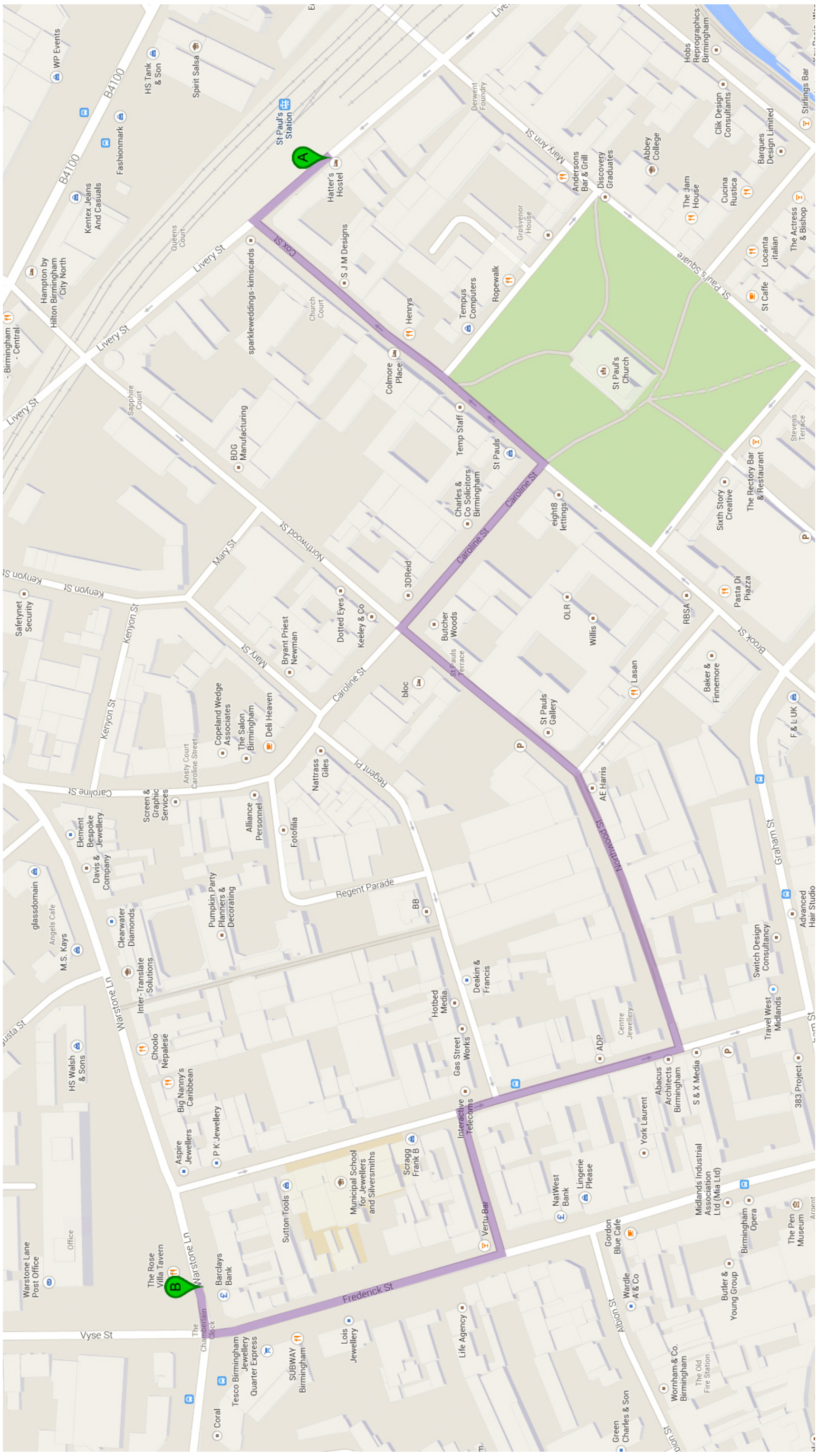


Figure A.3.: GM excerpt of Birmingham - LH highlighting

Figure A.4.: GM excerpt of Sheffield - FC highlighting

Figure A.5.: OSM excerpt of Sheffield

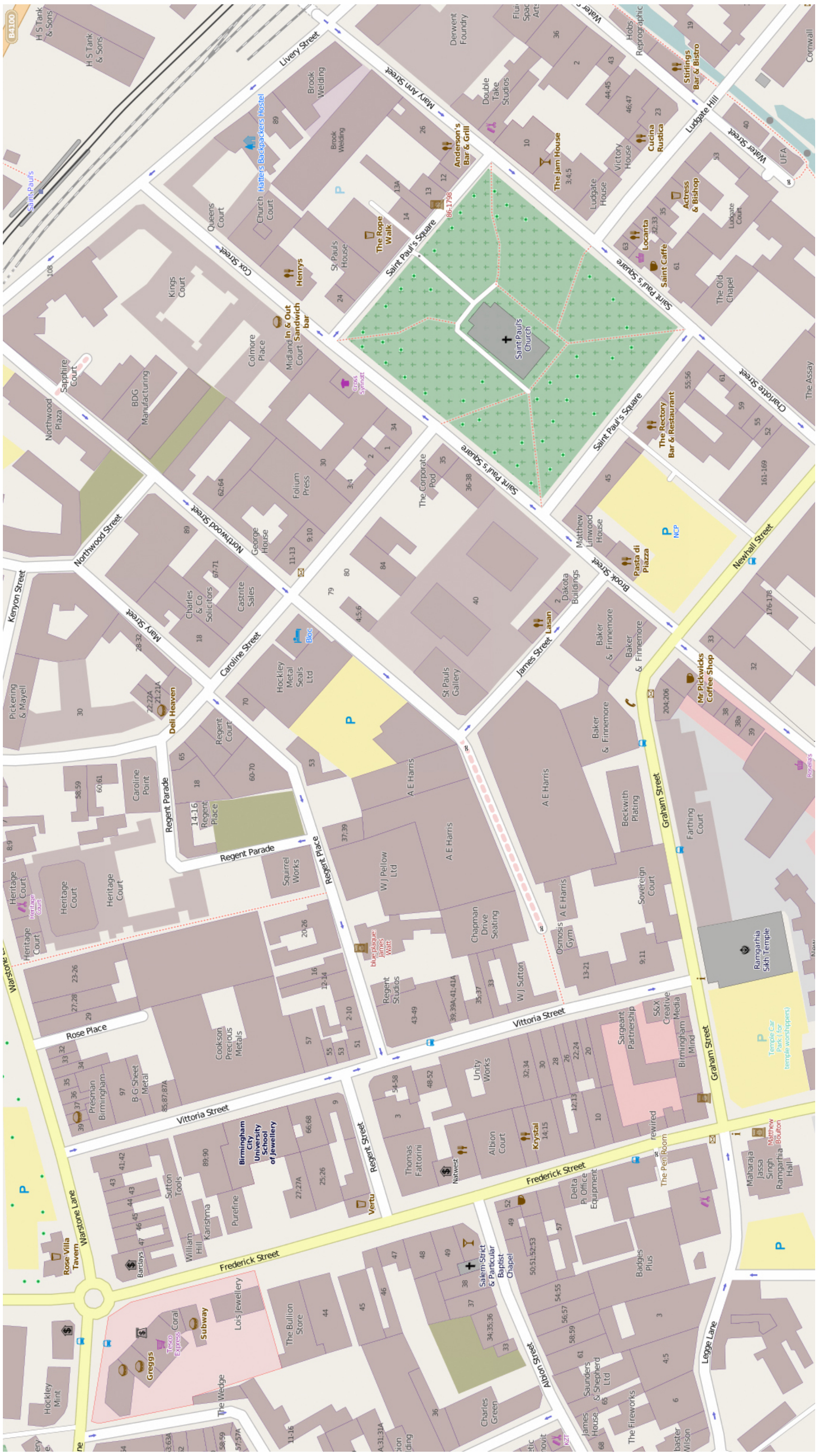


Figure A.6.: OSM excerpt of Birmingham

Universität Zürich – **Teilnehmerinformationen und Einwilligungsformular** **Evaluation von Weg-Markierungs-Methoden für Online-Karten**
25.02. – 14.03. 2014
Teilnehmernummer:

Zweck der Studie

Sie sind eingeladen, an einer Studie über die Evaluation von Routenmarkierungsmethoden für online Karten teilzunehmen. Ziel ist es, mehr Informationen über den Einfluss solcher Methoden herauszufinden. Die Studie ist Teil meiner Masterarbeit am Geographischen Institut der Universität Zürich.

Beschreibung der Studie und mögliche Risiken

Falls Sie sich entscheiden an dieser Studie teilzunehmen, werden sie im Anschluss gebeten, einen kurzen Fragebogen zu allgemeindemographischen Angaben, auszufüllen.
Im Folgenden werden Sie verschiedene Aufgaben am Computer lösen, währenddessen Ihre Interaktion mit dem Computer mittels einer Kamera, eines Mikrofons und eines Eye-Trackers aufgezeichnet werden. Der Eye-Tracker ermöglicht es, Ihre Augenbewegungen ohne jeglichen Körperkontakt aufzuzeichnen. Dazu wird nicht sichtbares Licht im nahen Infrarotbereich verwendet. Diese Methode ist für die Augen unschädlich und während dem Test für Sie nicht merkbar.
Die gesamte Studie wird ungefähr 40 Minuten dauern.

Datendiskretion und Publikation

Jegliche Information, die während der Studie in Verbindung mit Ihnen gebracht werden kann, wird vertraulich behandelt und nur mit Ihrer ausdrücklichen Erlaubnis an Dritte weitergegeben. Mit der Unterschrift unter dieses Dokument erlauben Sie mir, die Ergebnisse des Versuchs mehrmals zu publizieren. Dabei werden keinerlei Informationen veröffentlicht, die es ermöglichen, Sie zu identifizieren.

Entschädigung

Die Teilnahme an dieser Studie wird nicht vergütet. Mögliche anfallende Kosten für die Teilnehmenden an dieser Studie werden nicht zurückerstattet.

Ergebnisse der Studie

Falls Sie über die Ergebnisse der Studie informiert werden möchten, können Sie im Nachhinein Ihre Anschrift hinterlassen. Die gewünschte Kopie wird Ihnen dann zugestellt.

Einwilligung an der Teilnahme

Ihre Entscheidung, an der Studie teilzunehmen oder nicht, wird etwaige zukünftige Beziehungen mit der Universität Zürich nicht beeinträchtigen. Entscheiden Sie sich dafür, an der Studie teilzunehmen, steht es Ihnen jederzeit frei, die Teilnahme ohne Begründung abzubrechen. Sollten Sie Fragen haben, zögern Sie bitte nicht, mir diese zu stellen. Sollten zu einem späteren Zeitpunkt Fragen aufkommen, werde ich (Lorenz Bosshardt, 079'366’96’76, lorenzbosshartd@gmail.com) diese gerne beantworten. Sie erhalten eine Kopie dieses Dokuments.

Seite 1 von 2

Universität Zürich – **Teilnehmerinformationen und Einwilligungsformular** **Evaluation von Weg-Markierungs-Methoden für Online-Karten**
25.02. – 14.03. 2014
Teilnehmernummer:

Mit Ihrer Unterschrift bescheinigen Sie, oben stehende Informationen gelesen und verstanden zu haben und willigen ein, unter den dort beschriebenen Bedingungen am Experiment teilzunehmen.

....
Unterschrift des Teilnehmers

....
Unterschrift des Experimentleiters

....
Vor- und Nachname in Blockschrift

....
Vor- und Nachname in Blockschrift

....
Ort und Datum

Universität Zürich – **Teilnehmerinformationen und Einwilligungsformular** **Evaluation von Weg-Markierungs-Methoden für Online-Karten**
25.02. – 14.03. 2014
Teilnehmernummer:

WIEDERRUF DER EINWILLIGUNG

Hiermit möchte ich meine Einwilligung, an der oben beschriebenen Studie teilzunehmen, widerrufen.

....
Unterschrift des Teilnehmers

....
Ort und Datum

....
Vor- und Nachname in Blockschrift

Mit dem Widerruf der Einwilligung beeinträchtigen Sie in keiner Weise Ihre Beziehungen mit der Universität Zürich. Der Widerruf kann jederzeit und ohne Angabe von Gründen beantragt werden. Den Widerruf der Einwilligung bitte an *Lorenz Bosshardt, Friesstrasse 42, 8050 Zürich* senden.

***1. Teilnehmer-Nr.: (einzufüllen vom Experimentleiter)**

***2. Alter**

- ○ 18 -24
- ○ 25 - 31
- ○ 32 - 38
- ○ 39 - 45
- ○ 45 - 51
- ○ 51+

***3. Geschlecht**

- ○ weiblich
- ○ männlich

***4. Wurde dir jemals von einem Arzt gesagt, dass dein Farbsehvermögen beeinträchtig ist?**

- ○ ja
- ○ nein

***5. Trägst du irgendwelche Sehhilfen (z.B. Kontaktlinsen oder Brille) für dieses Experiment?**

- ○ ja
- ○ nein

Figure A.7.: Pre-Questionnaire pg. 1

***6. Wie oft nutzst du diese Kartendienste?**

	nie	selten	monatlich	wöchentlich	täglich
Google Maps	○	○	○	○	○
Open Street Maps	○	○	○	○	○

***7. Wie oft brauchst du eine Routenplaner-Funktion (wie z.B. dargestellt im untenstehenden Bild) bei Online Kartendiensten?**

- ○ nie
- ○ selten
- ○ monatlich
- ○ wöchentlich
- ○ täglich

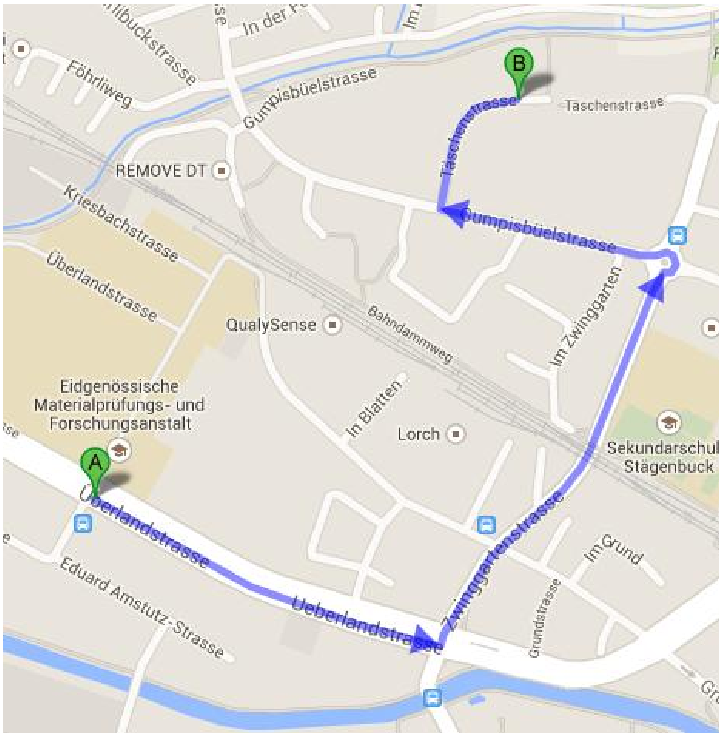


Figure A.8.: Pre-Questionnaire pg. 2

***8. Angenommen, du suchts einen unbekannten Weg in Gehdistanz, wie würdest du die Art der Navigationshilfe betreffend deiner Präferenzen einordnen?**

	1 (nicht hilfreich)	2	3	4	5 (sehr hilfreich)
mündliche Wegerklärung	○	○	○	○	○
schriftliche Wegerklärung	○	○	○	○	○
visuelle Erklärung (Karte)	○	○	○	○	○

***9. Wie gut kennst du die Stadt "Sheffield" in Grossbritannien?**

○ bestens bekannt (Ich wohne in Sheffield oder habe dort eine Zeit lang gewohnt)

○ bekannt (Ich war als Tourist in Sheffield)

○ "virtuell bekannt" (Ich studierte aus beruflichen oder persönlichen Gründen Karten der Stadt)

○ nicht sehr bekannt (Ich würde die Stadt auf einer Karte wiedererkennen)

○ überhaupt nicht bekannt (Ich habe keine Vorstellungen, wie diese Stadt aussieht)

***10. Wie gut kennst du die Stadt "Birmingham" in Grossbritannien?**

○ bestens bekannt (Ich wohne in Birmingham oder habe dort eine Zeit lang gewohnt)

○ bekannt (Ich war als Tourist in Birmingham)

○ "virtuell bekannt" (Ich studierte aus beruflichen oder persönlichen Gründen Karten der Stadt)

○ nicht sehr bekannt (Ich würde die Stadt auf einer Karte wiedererkennen)

○ überhaupt nicht bekannt (Ich habe keine Vorstellungen, wie diese Stadt aussieht)

***11. Wie viele Stunden hast du von Gestern auf Heute geschlafen?**

○ < 3 Stunden

○ 3-5 Stunden

○ 5-7 Stunden

○ 7-9 Stunden

○ > 9 Stunden

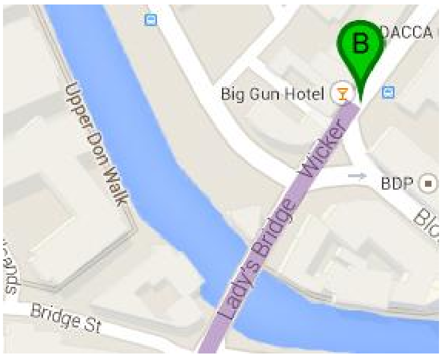
Figure A.9.: Pre-Questionnaire pg. 3

***1. Teilnehmer-Nr.: (einzufüllen vom Experimentleiter)**

Du hast zwei verschiedene Arten der Routenmarkierung gesehen. Die zwei untenstehenden Bilder zeigen dir nochmals die Unterschiede.

Methode A

Methode B



2. Wie hilfreich waren diese Methoden für dich, um dir die Route zu merken?

	1 (nicht hilfreich)	2	3	4	5 (sehr hilfreich)
Methode A	○	○	○	○	○
Methode B	○	○	○	○	○

Figure A.10.: Post-Questionnaire pg. 1

3. Wie hilfreich waren für die dich die folgenden Kartenelemente / Objekte, um dir die Route einzuprägen?

	1 (nicht hilfreich)	2	3	4	5 (sehr hilfreich)
markierte Punkte (z.B. Café, Restaurant, Läden, Hotels etc.)	○	○	○	○	○
Strassennamen	○	○	○	○	○
Form der markierten Route	○	○	○	○	○
Form von markanten Gebäuden	○	○	○	○	○
andere Objekte (Hauptstrassen / Bahnlinien / Pärke / Flüsse / Seen)	○	○	○	○	○

4. Andere Objekte / Kartenelemnte die dir halfen, den markierten Weg zu merken?

Figure A.11.: Post-Questionnaire pg. 2

Personal declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Zurich, 27 June 2014

Lorenz Bosshardt