Exploring the Cognitive Processes of Map Users Employing Eye Tracking and EEG

Merve KESKIN

ISBN:
Printed in Belgium by University Press, Ghent

Cover design of the book and the chapters:

Faculty of Sciences
Department of Geography
Academic year 2020-2021

Thesis submitted in accordance with the requirements for the degree of Doctor of Science: Geomatics and Surveying (Joint PhD of Ghent University & Istanbul Tecnhical University).

Proefschrift aangeboden tot het behalen van de graad van Doctor in de Wetenschappen: Geomatica en Landmeetkunde (Joint PhD van de Ghent University & Istanbul Tecnhical University).

Merve KESKIN

Supervisors

Prof. dr. Philippe DE MAEYER
Ghent University, Faculty of Sciences, Department of Geography

Prof. dr. Ahmet Ozgur DOGRU
Istanbul Technical University, Department of Geomatics Engineering

Dr. Kristien OOMS
Ghent University, Faculty of Sciences, Department of Geography

Members of the examination committee

Prof. dr. Nesibe Necla ULUGTEKIN
Istanbul Technical University, Department of Geomatics Engineering

Prof. dr. Arzu COLTEKIN
Institute for Interactive Technologies FHNW, HCI/Extended Reality

Prof. dr. Haosheng HUANG
Ghent University, Department of Geography

Prof. dr. Ali Melih BASARANER
Yildiz Technical University, Department of Geomatics Engineering

Dr. Caner GUNEY
Istanbul Technical University, Department of Geomatics Engineering

Dr. Klaas BOMBEKE
Ghent University, Media, Innovation and Communication Technologies, Department of Communication Sciences,

Chair
Prof. dr. Veerle VAN EETVELDE
Ghent University, Faculty of Sciences, Department of Geography

Acknowledgement
This research was supported by TUBITAK (The Scientific and Technological Research Council of Turkey) [grant number: 2214a-1059B141600039].

Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA: Analysis of Variance
AoI: Area of Interest
BSS: Blind Source Seperation
EEG: Electroencephalogram
EFRP: Eye-fixation-related-potentials
EOG: Electrooculography
ERD: Event-related desynchronization
ERP: Event-related potentials
ERS: Event-related synchronization
ET: Eye tracking
FAA: Frontal Alpha Assymetry
fMRI: Functional Magnetic Resonance Imaging
GIS: Geographic Information System
GL: Good learners
HCI: Human Computer Interaction
ICA: Independent Component Analysis
ISO: International Organization for Standardization
LTM: Long-term memory
M: Mean
MED: Median
POR: Point of Regard
PSD: Power Spectral Density
RPL: Relatively poor learners
SD: Standard Deviation
TCP/IP: Transmission Control Protocol/Internet Protocol
TTL: Transistor-transistor logic
UCD: User-Centered Design
UI: User Interface
UX: User Experience
WM: Working memory

Word of gratitude

It feels extremely surreal that I have come to the end of my PhD journey. These past years have been a rollercoaster ride between two universities, two countries, and later between two homes. It has definitely been a challenging experience both emotionally and professionally but I have always tried to look on the bright side on this duality and make the most of it. I have learnt a lot during this period not only about my PhD but also about life and myself in general. There has been times that I did not believe myself and lost my self-esteem, I made mistakes, I criticized myself harshly, I cried, but, in the end I have found a way to gain my confidence back, right the wrongs, put a smile back on my face and keep going no matter what happens. Of course, I owe a great deal of this to the wonderful people I have met along the way. They always helped me get up when I fall, cherish my success and happiness, encouraged me, inspired me and supported me. This section is dedicated to express my sincere gratitude to the people who helped paving me the way to accomplish my PhD. I feel so lucky to have all of them and without their presence and unconditional support, I would not have come this far.

One person I would like to thank in particular is my supervisor at ITU, prof. Ahmet Ozgur Dogru, who has not abandoned me for more than ten years, since I was an undergraduate student. He believed in me with my PhD subject when no one did, he created many opportunities for me, introduced me to the right people, gave me a constant moral and professional support, put me in his priority list when there was one hell of work to do in his basket. No words can describe how thankful I am for the things he has done for me. He is more than a supervisor to me; he is my role model and the biggest reason of where I am right now. At this point, I would also like to express my appreciation to my supervisor at UGent, prof. Philippe De Maeyer; I consider it a privilege to be his PhD student, to have had a chance to learn from him and to be a part of the CartoGIS research unit during the last four years I have spent at UGent. Philippe supported me in many ways since the day he offered me a research visit at UGent and once I came here, he initiated the joint-PhD program, which changed all my plans for the better. I would like to thank him for the opportunities he provided me in both academic and professional life.

Dr. Kristien Ooms who was co-supervising me at UGent deserves special thanks for her endless support and for being my rock throughout the way. She made sure that everything was all set long before I even came here, followed my progress very closely even after she changed her job and always brought the smartest solutions to the table especially at the times I felt helpless. Her perspective of seeing things and her hard work will always continue to inspire me.

Special gratitude goes to Prof. Dr. Nesibe Necla Ulugtekin who is the unofficial supervisor of my entire academic journey and my mentor in life. With her expertise and wisdom yet her enthusiasm, her belief in young people and her support for them, her critical thinking, her righteousness and fairness, her independent and dissentient spirit, she is everything I look up to and admire.

I would like to thank Prof. Dr. Hendrick Slabbinck from the marketing department of Ghent University for his tremendous support and help in the neuro-lab; Dr. Andrey Nikolaev from the department of experimental psychology at Catholic University of Leuven (KU Leuven) for sharing his knowledge and expertise on eye tracking and EEG while designing the second user experiment. I would also like to thank Dr. Klaas Bombeke from the department of Communication Sciences of Ghent University for sharing his expertise in EEG and assisting me with EEG data analysis and the interpretation of the results.

I am also incredibly grateful for being a part of our international cartographic family, in other words, International Cartographic Association (ICA). I, hereby, would like to thank dr. Corne van Elzakker for his time and interest for my research, the former president of ICA, prof. Georg Gartner for all the constructive discussions we have had, and prof. Sara Fabrikant for inviting me for a reseach visit in her lab at the Department of Geography at University of Zurich.

This PhD would not have been possible without the moral support of my friends here in Belgium and in Turkey. I would like to thank Burak, my former partner in life who transitioned to be my best friend and family through years, for his unconditional love, his continuous belief in who I am/ what I do, and his wisdom and calmness at my panic moments. He patiently picked me up every time I fell and I am so lucky to have him in my life. I also would like to thank Asena, my housemate, my partner in crime, and my sister I met late, for the much-needed distraction she provided, all the fantastic times we had together, and just for making our house a home, for making Ghent Istanbul. It is time to let these golden birds fly away.

I have to thank several other amazing people for being there for me for years; my fabulous soul mate Atakan, my burlesque muse Ece, my oldest and dearest friend Ezgi, my Mandalorian-Atlanta-Sherlock buddy Ziad, for laughing at my jokes, for flattering me, for making me feel like a queen and all but simply for showing some love.

Finally yet importantly, my most special gratitude goes to my mother, who is literally my heroine, for her unconditional love and support, her incredible care and attention while raising me, and all the sacrifices she made for me. Your positivity, your cool attitude, your great smile, your ability of making impossible possible, your ambition and your strength inspire me to be better person like you. You bright up my life from miles away every day and show me that there is nothing we cannot do together. We indeed make a great team and this PhD is our common achievement. Therefore, mom, I dedicate this thesis to you and I look forward to deal with the future uncertainties in life with you.

Ghent, 07/08/2020

"Things don't always go as planned, Mr. Angier. That's the beauty of science."
Tesla (David Bowie), Prestige

Chapter 1:
General Introduction

"Map design can be thought of as mind design; the way a map is designed will influence the views of the world it stimulates or inhibits."

Daniel R. Montello, 2002

This chapter is partially modified from:

Keskin, M., Dogru, A. O., Guney, C., Basaraner, A. M., & Ulugtekin, N. N. (2016, June). Contribution of neuroscience related technologies to cartography. In Proceedings of 6th International Conference on Cartography and GIS, 394-404.

1.1. Background Information

1.1.1 Spatial Cognition and Cognitive Cartography

Cognition is defined as "the intelligent process and products of the human mind includes such mental activities as perception, thought, reasoning, problem solving and mental imagery". Mental image is an important topic in cognition studies. "The mental image is an internal representation similar to sensory experience but arising from memory" (Peterson, 1994, p.35). On the other hand, spatial cognition can be defined as restructuring the space on a mental level and its assimilated reflection. Similar to perception, spatial cognition is related to both physical environment and the abilities of individuals regarding socio-cultural, economic and political characteristics of their daily life. As Boulding (1956) and Lynch (1960) introduced spatial cognition can be explained by two terms: (i) spatial image (other terms used instead are mental map, cognitive map) and (ii) cognitive mapping. In this context, spatial image refers to a cognitive representation of the nature and attributes of the spatial environment, whereas cognitive mapping is "…a process composed of a series of psychological transformations by which an individual acquires, codes, stores, recalls and decodes information about the relative locations and attributes of phenomena in their everyday spatial environment" (Downs & Stea, 1973, p.7). Thus, human spatial behavior is dependent on the individual's cognitive map of the spatial environment. Because cognitive mapping is considered as the basic component of human adaptation and a cognitive map is a must for everyday environmental behavior. These environmental behaviors occur when seeking answers for 'where certain valued things are' and 'how to get to where are from where we are' (Downs & Stea, 1973, p.313).

Due to the above definitions, cognitive mapping process is characterized by how individuals make use of the information existing in different environments and how it affects spatial behavior (coping with complex data, interpretation, etc.). Hence, the perception of the environment plays a significant role for improvement of spatial images. It is important that individuals perceive and comprehend a foreknown concept with its spatial relationships; that is how cognitive maps are established (Güç, et al., 2012). In this context, a cognitive map is not necessarily a map, it is actually an analogy of process, not product which a cartographic map. In fact, a cartographic map itself has a huge impact on our spatial images and our concept of cognitive maps (Boulding, 1956; Downs & Stea, 1973). Montello (2002) also stated that map design is about human cognition design and according to him, this can be termed as "intuitive map psychology". In spite of the situation before, the intuition of maps become a formal part of cartographic education in 20th century. By this way, cartographers proposed that intuitions about map cognition can be developed by borrowing theories and methods of psychology in a more systematic way. This approach created a new research area called cognitive cartography. Applying cognitive theories and methods to understand maps, and map applications to understand cognition are subjects of cognitive cartography (Montello, 2002; Keskin, et al., 2016).

Abstraction of physical reality shows differences in geometry and cognition. In geometry, the spatial world can be described in terms of points, lines, and areas. On the contrary, in cognition, basic entities usually are not points; they may be theentire physical objects like books or chairs. Figure 1.1 demonstrates how to conceptualize a physical object in geometry and in cognition. The cognitive apparatus is flexible as to which level in a huge lattice of part-whole relations to select as 'basic level'; it can also focus either on the relation between an object and a configuration of objects, or on the relation between an object and its parts (Freska, 2013). For this reason, to increase the understandability of a cartographic visualization, we need to link the geometric fundamentals to the cognitive processes. Within map design, the relationships between geometry and cognition should be taken into an account as the premise criteria. The understandability of a cartographic representation is highly correlated with the effectiveness of its design. For instance, the visual variables that will be used in the design can be decided based on the strategy of Shneiderman's (1996) visual information mantra; overview first, zoom&filter, details on demand.

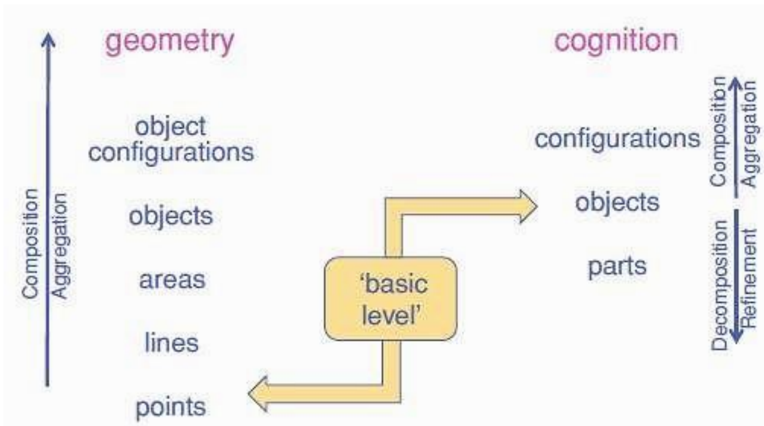


Figure 1.1. Left: In geometry, aggregation of structures from atomic point entities. Right: In cognition, no matter geometrical complexity or meaningfulness, basic entities can be decomposed into more elementary entities or aggregated into more complex configurations; without invoking elementary constituents (Freska, 2013).

1.1.2. Principles of Cartographic Design for increasing Spatial Cognition

As Jones (2010) mentioned, in 1996, the Society of British Cartographers published a list of five principles which should guide good map design;

- Concept before Compilation
- Hierarchy with Harmony
- Simplicity from Sacrifice
- Maximum Information at Minimum Cost
- Engage the Emotion to Engage the Understanding

For all those principles guiding good map design, the user of the map is a very important factor throughout the design procedure. A map should invoke our spatial attention and evoke things in our spatial memory with its design. Regarding to that, understandability of a cartographic representation is highly correlated with the usefulness of its design. These fundamentals of map-use lead us to user-centered design (UCD). The basic philosophy of UCD is to put the user at the center of the development process by focusing the needs of the real people who will use their product and what do they want to achieve with it (Haklay & Nivala, 2010). There are several approaches and guidelines for UCD, however, the principles of Norman (1990) & Gould and Lewis (1985) are the most widely adopted ones (Table 1.1).

Table 1.1. UCD principles

Norman's principles	Principles of Gould and Lewis
1. Use both knowledge in the world and knowledge in the head. **2.** Simplify the structure of tasks. **3.** Make things visible. **4.** Get the mapping right. **5.** Exploit the power of constraints, both natural and artificial. **6.** Design for error. **7.** When all else fails, standardize.	**1.** Early focus on users and tasks *Observe users and engage them to the design process.* **2.** Empirical measurement *Measure different elements of the design – from the recording of users' reaction and performance with a prototype to the evaluation of the final version.* **3.** Iterative design *Go back and redesign if problems are identified in a user testing. Keep in mind 'design, test, measure, and redesign' cycle.*

Gould and Lewis's (1985) framework was translated in ISO 13407 into a set of instructions that help achieving user needs by utilizing a UCD approach throughout the whole life cycle of a system (Haklay & Nivala, 2010). The three-step design is an iterative process (Figure 1.2):

1. Step: Find out the user requirements.

Study potential users and the context. Decide on which usability criteria are to be emphasized in the study: effectiveness, efficiency, satisfaction, memorability, and/or minimal errors.

2. Step: Provide design solutions.

The first design prototypes and preliminary mock-ups.

3. Step: Analyze whether the defined user requirements have been met.

Use usability engineering methods.

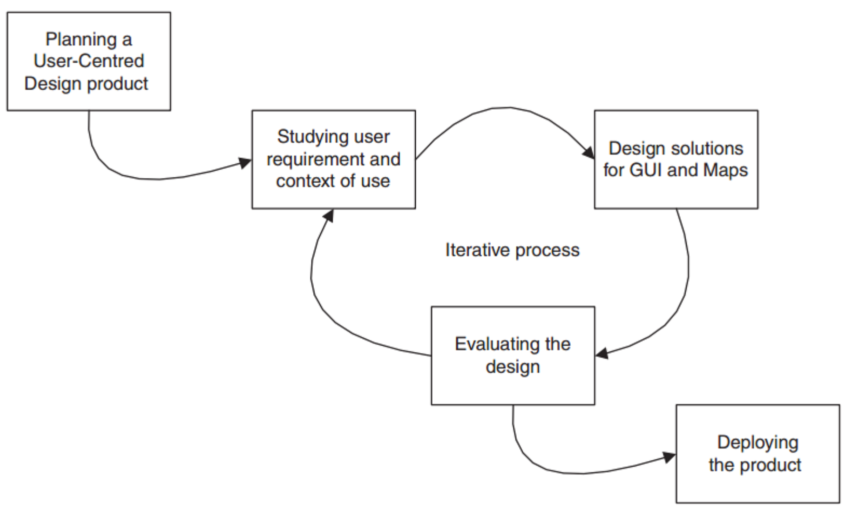


Figure 1.2. UCD cycle for geospatial technologies (Haklay & Nivala, 2010).

If the results indicate that the user requirements have not been achieved, the iterative process goes back to redefining the user requirements. If the requirements are met, the product is deployed. However, evaluations for existing product should continue to make sure the product still satisfies user needs (Haklay & Nivala, 2010).

1.1.3. User Centered Design (UCD) and Usability Research Relationship

While user centered design (UCD) can be described as an approach to the design and development process focusing on gaining a deep understanding of who will use the product (URL 1), usability is a measure of user experience interacting with the design. According to International Organization for Standardization (ISO) definition, usability is the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. In this context, usability standards (ISO 9241-11) are categorized as primarily concerned with:

- the use of the product (effectiveness, efficiency and satisfaction in a particular context of use)
- the user interface and interaction
- the process used to develop the product
- the capability of an organization to apply user centered design (URL 2)

As a component of the product use, effectiveness refers whether users complete tasks, achieve goals with the product. Another issue is efficiency, which is mostly measured with time and about how much effort users require to do a specific task. Lastly, satisfaction can be described as users' conception on the products ease of use. There are also some fundamental factors affecting the use of product such as users, their goals and context of use. Users can be highly trained, experienced or novice. The use purposes of the product may vary based on where and how the product is being used (URL 3). Related to ISO 9241-11, ISO 13407 defines UCD activities for the entire life cycle of interactive computer-based systems including human factors, ergonomics knowledge and techniques as presented in Table 1.2 (URL 2). This standard is the basis for many UCD methodologies.

Above stated standards can also be applied for usability research of cartographic products. In this context, first, user tasks should be prepared and classified based on specific criteria. Throughout the task design procedure, user group, goal, context of use and the tasks expected to be completed by users should be taken into consideration. User groups may diversify in terms of education level, expertise, age, gender, culture, and sensory disabilities, and these factors directly affect spatial perception of individuals as well as the perception of cartographic products (Slocum et al, 2001).

After determining the user group and user goals, the context of use, which is also one of the usability measurement types defined by ISO, should be specified. The context of use is related to visualization environment, visualization method (i.e. whether it is 2D, 3D, static, dynamic or interactive) and the advantages and/or disadvantages of visualization method to choose. For instance, MacEachren et al (1999) introduced four criteria (i.e. immersion, interactivity, information density, intelligence of display objects) for usability evaluation of geospatial virtual environments (GeoVEs). Immersion refers to the sensation of being enveloped by the environment, while interactivity allows manipulating the characteristics of environment components. Information density handles the level of detail in the geovisualization environment, and the last criteria, which is intelligence of display objects, can be described as assisting components that help users to interpret the representation on the screen.

Table 1.2. Usability Metrics defined by ISO 13407:1999, ISO 9241-11:199

Usability measurement type	Definition
Effectiveness	The accuracy and completeness with which users achieve specified goals
Efficiency	The resources expended in a relation to the accuracy and completeness with which users achieve goals
Satisfaction	Freedom from discomfort, and positive attitude to the use of the product
Context of use	Characteristics of the users, tasks and the organizational and physical environments
Goal	Intended outcome
Task	Activities required to achieve a goal

To design user-tasks, it is good to follow a task taxonomy. In literature, there are many task taxonomies developed for cartographic representations (Amar, et al., 2005; Casner, 1991; Kveladze, 2015; Roth & Mattis, 1990; Roth, 2012; Wehrend & Lewis, 1990). For instance, Armstrong & Densham (1995) offers a sequence of steps that normally are followed in formulating and solving the location-selection problem, whereas Knapp (1995) produced a six-step task analysis model. Knapp's task analysis model includes task (what is to be accomplished), goal (why it is to be accomplished), physical actions (how it is to be accomplished), mental actions (thought process while accomplishing it), data (the data set with which it will be accomplished) and visual operators (primitive operators for visual interaction with display) that users would be associated with. In addition, user tasks may incorporate from simple to complex questions based on the visualization and its abilities.

The ability of users on performing user tasks is measured by identifying the methods used to test user behaviors. There are several theoretical and empirical methods to test the usability. Empirical methods allow ideas to be generated and verified by systematic observation and measurements. Systematic observations are standardized, controlled, recorded, repeatable, and publicly verifiable. In that manner, systematic empiricism should be held separate from trial-and-error map design approach developed by cartographers over the years. Similarly, informal experiments held by cartographers to evaluate the effectiveness of a design should be distinguished. As all empirical research on cartography does not include cognitive theories or human subjects, it is not possible to claim that only empirical methods leads the truth. However, it is obvious that combination of empirical research and predictive and explanatory ideas can lead scientific outcomes (CCS, 1995; Montello, 2002). As Slocum et al. (2001) stated that building an effective visualization method is a two-step process which involves theory-driven cognitive research and usability engineering to evaluate existing methods. Theory-driven cognitive research covers studies to understand how users create and make use of mental maps of real world phenomena when they interact with maps. If theories related to these are developed, there will not be much need for user tests of a specific geovisualization technique. On the other hand, usability engineering refers identification of methods to analyze and enhance the usability of software. Usability term here means the ease of use and effectiveness, efficiency and satisfaction, just as mentioned by ISO. "User testing" and "user studies" applied in cartography show significant similarities with the empirical research subjected to usability engineering (Slocum et al., 2001). All in all, the key to developing highly usable cartographic products is employing UCD and employing UCD requires usability research that focuses on user behaviors and needs (URL 1).

1.1.4. User Interface/User Experience (UI/UX) Design for Cartography

User studies conducted in cartography show significant similarities with the empirical research subjected to usability engineering (Slocum et al., 2001). In general, humans use interfaces which corresponds to user interface (UI) design and they experience interactions which are related with user experience (UX) design. The user experience proceeds in stages and each stage should be carefully designed. Interaction is a two-way conversation; it is important to provide affordance and feedback to maintain the dialogue. When a map user interacts with a cartographic design, cartographic interaction occurs. Cartographic interaction is the ability to manipulate the map temporarily to meet user needs (Figure 1.3).

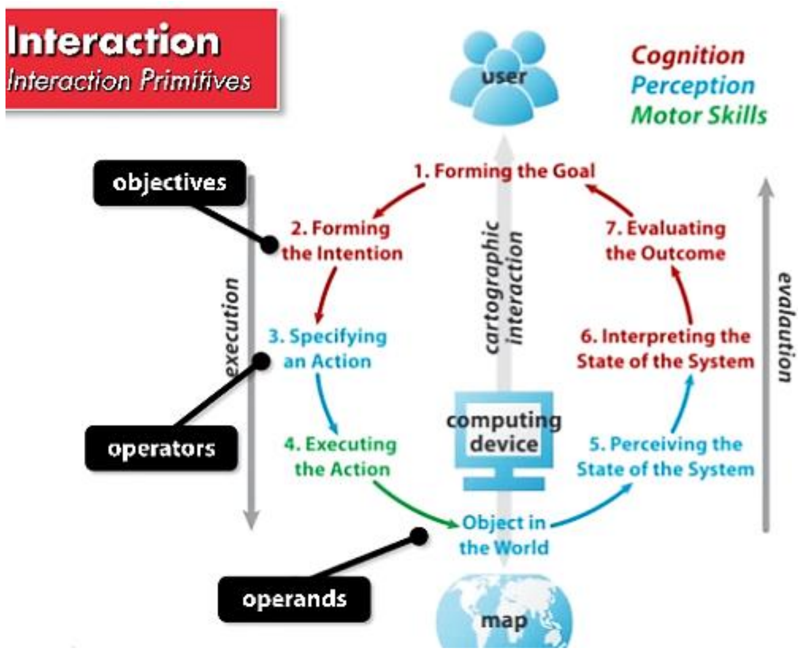


Figure 1.3. Cartographic Interaction Primitives (Roth, 2012)

Roth (2013) listed the six fundamental questions of cartographic interaction as follows:

- **What?** *the definition of cartographic interaction in the context of cartographic research*
- **Why?** *the purpose of cartographic interaction and the value it provides*
- **When?** *the times that cartographic interaction positively supports work*
- **Who?** *the types of users provided cartographic interaction and the way in which differences across users impacts interface designs and interaction strategies*.
- **Where?** *the computing device through which cartographic interaction is provided and the limitations/constraints on cartographic interaction imposed by the device*
- **How?** *the fundamental cartographic interaction primitives and the design of cartographic interfaces that implement them*. 'How?' is the most important question for cartographic interaction because it seeks

for answers for cartographic representation by deciding on which visual variables should be employed for a specific design.

1.1.5. Usability Research Methods

Keeping in mind the empirical and theoretical classification of usability research, practical usability methods can be categorized as qualitative and quantitative (Table 1.3). Qualitative research includes user experience, actions and behaviors including their feelings, opinions and emotions to gain insight into reason or motivation of their reactions. The methods may vary but qualitative research generally seeks for explanation about influences and processes of user experience from discussion, interview or so. Qualitative methods can be listed as; focus group, post-experience interview, think aloud, observation, video recording, audio recording, and screen recording (Van Elzakker, 2004; Kveladze, 2015; Ooms, 2012). These methods are used often for usability of 2D, 3D static, dynamic and interactive visualizations.

Table 1.3. A comparison of quantitative and qualitative methods (URL 4)

Quantitative Methods	Qualitative Methods
Learn about design for perception & cognition	Learn about design for culture & preference
Difficult to translate insights into practice	Insights often constrained by current practices
Produces general insights	Produces specific insights
Insights are superficial	Insights are deep
Identify significant differences	Explain significant differences
May not explain differences	May not identify differences
Harder to design	Easier to design
Easier to analyze	Harder to analyze

Usability evaluation methods such as interview and think-aloud protocols are widely used due to the fact that they require no measurement apparatuses and allow usability experts to measure usability of software systems in a relatively easy way. However, analyzing and evaluating the collected data from these methods are time consuming. Additionally, the results of the analysis may be hard to replicate because the collected data is based on qualitative evaluations. To overcome this limitations, quantitative evaluation methods have been developed (Kimura et. al., 2009). Different from qualitative methods, quantitative research allows measuring, quantifying and counting issues considered in a usability study (Kveladze, 2015). Task analysis, questionnaire, sketch maps, eye tracking and EEG (electroencephalogram) can be counted as frequently used quantitative methods. It is important to decide which method to use in which cases or in which stages of the usability research considering the following four aspects (URL 4):

- Participants: *the people recruited for the study (if empirical)*
- Materials: *the cartographic products that are to be evaluated*
- Procedure: *the process that participants need to complete*
- Analysis: *the way to collect and interpret the data*

1.1.5.1. Monitoring gaze activity in the context of spatial cognition - Eye tracking

Eye tracking is one of the quantitative usability research methods and a frequently used user experience technique for user interfaces and websites (e.g., Djamasbi, et al., 2010; Fleetwood & Byrne, 2006). It allows tracking the movements of the participant's eyes: his point of regard (POR) is registered at a certain sampling rate. From this long list of (x, y) positions, eye movement metrics such as fixations and saccades can be derived. A fixation is a stable POR during a certain time span (at least 80 to 100 ms) and indicates the users' content interpretation at that location. A saccade is a rapid eye movement between two fixations, typically completed in tens of milliseconds. A scan path can be described as a succession of fixations and saccades (Ooms, 2012).

The first eye tracking application in user experience domain was conducted by Fits, et al. (1950), who used motion picture cameras to study the movements of pilots' eyes. Jenks (1973) initiated eye tracking use for cartographic purposes by exploring the scan paths of users looking at a dot map. Although a few more studies followed this first one, no significant eye tracking study on the assessment of the cartographic products have been implemented for almost 10 years because of the following reasons presented by Jacob and Karn (2003):

- technical problems related to capturing the actual eye movements that may cause inaccurate and unreliable results
- complicated and time-consuming data extraction, and
- difficulty in interpretation of extracted data.

However, later advances in eye-trackers and eye tracking software and their decreasing costs concluded that eye tracking technology is useful for interpretation of visual information efficiently while performing a complex visual and cognitive task (Duchowski, 2007; Jacob & Karn, 2003; Ooms, 2012). It is also possible to execute detailed analysis and eye tracking contributed to psychological research on the cognitive processes linked with visual search (Ooms, 2012). For instance, Çöltekin et. al. (2010) studied the users' visual interaction with highly interactive interfaces. The study mainly investigated whether the efficiency of users can be characterized by specific display interaction event sequences, and whether studying user strategies could be incorporated in a way to improve the design of the dynamic displays. Another research intended to combine eye tracking with user logging (mouse and keyboard actions) with cartographic products and referenced screen coordinates to geographic coordinates in order to know which geographic object corresponds to the gaze coordinates at all times. This approach is promising in terms of efficiently studying user behavior with interactive and static stimuli in multiple research fields (Ooms, et al., 2014).

1.1.5.2. Monitoring brain activities in the context of spatial cognition

Developments in medical research have led observing neurons in the brain with a high spatial and temporal resolution. The discovery of place cells also marks an attention to research on spatial cognition. Place cells are located in entorhinal cortex, which is a part of temporal lobe and functionate as the center of memory and navigation. The entorhinal cortex is the main interface between the hippocampus and neocortex (Figure 1.4). Understanding the activity of the place cell in the hippocampus equals to understanding how neurons code complex cognition. Hence, '… *any discussion of the hippocampal neurophysiology of spatial* *cognition needs to start from the fact that spatial location is a primary driver of neural firing patterns in the rodent hippocampus and spatial firing is clearly the best first-order description of rodent hippocampal representations'* (Redish & Ekstrom, 2013, p.30).

There are several researches on spatial representations in the medial temporal lobe. Hippocampal lesions lead various deficiencies on spatial cognition such as impairments in forming spatial relationships and spatial learning abilities. For instance, patients with hippocampal lesions succeeded retrieving a single route to a hidden location, unlike retrieving the location of multiple hidden objects within a spatial environment (Bohbot et al., 1998, 2007). These findings implied that calculations including multiple routes and environments are mostly related to hippocampus.

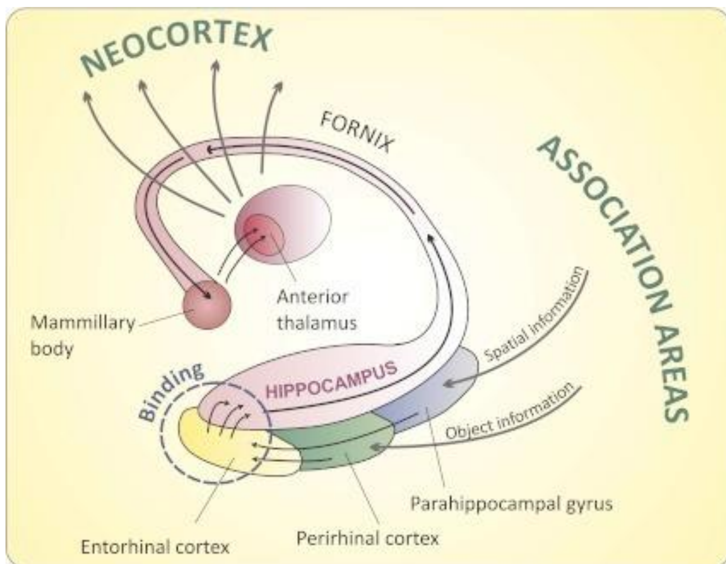


Figure 1.4. Hippocampal system (Kessels & Kopelman, 2012)

Besides hippocampus, the neural basis of human spatial memory depends on parahippocampal cortex and retrosplenical cortex which include visual-spatial scene processing and survey representation (Redish & Ekstrom, 2013). Patients with hippocampal deficiencies still keep on activities involving spatial memory such as locating a recently learned object within a room (Bohbot et al, 1998). In this context, the patient who was widely studied and known by his initials, H.M., played an important role in the development of cognitive neuropsychology. H.M. was a memory disorder patient who had a "bilateral resection of the entire (pyriform–amygdaloid–hippocampal) complex including the hippocampal gyrus extending posteriorly for a length of 8–9 cm from the tips of the temporal lobes in an attempt to cure his epilepsy (Corkin et al, 1997). Some research conducted with H.M. showed that he succeeded in many spatial memory tasks, including knowledge of the layout of his apartment. The posterior parts of his hippocampus, which might be important for spatial processing, were not damaged. However, he was unable to learn, store and retrieve new spatial routes, especially multiple routes. This situation suggested that the parts of his hippocampus that are responsible for spatial processing were damaged (Redish & Ekstrom, 2013). In addition, posterior parahipocampal cortex that receives signals from visual areas, allows allocentric processing of visual-spatial information. For instance, some studies involving the same spatial tasks presented that patients with parahippocampal lesions encountered very serious deficiencies, whereas patients with more profound hippocampal lesions did not show deficits (Bohbot et al, 1998; Ploner, et al., 2000). Another example is that compared with sighted ones, parahippocampal cortex of blind participants who imagine navigating showed less activation (Deutschländer et al., 2009). Due to results of the previous studies, parahippocampal cortex plays a valuable role in visual-spatial processing (Redish & Ekstrom, 2013).

All the above findings were acquired by functional Magnetic Resonance Imaging (fMRI), which allows measuring neural activity indirectly by using blood oxygen–level dependent (BOLD) signal. fMRI facilitates imaging brain activity with a high spatial resolution (~1mm), depending on the oxygen consumption in the activated brain area, and the increment and decrement in metabolism energy. Shortly, it measures the metabolic activity related whole-brain changes. Despite the fact that fMRI is unable to measure the activity of single neurons such as place cell activity, it ensures significant information related to the functions of hippocampus and parahippocampal cortex during navigation. Since it is not an invasive method, it can be employed for healthy participants as well (Redish & Ekstrom, 2013).

Although it has been used rarely in cartography, fMRI (functional magnetic resonance imaging) provided some essential information how user perceive maps and respond map-related tasks. One of the studies implemented fMRI technology showed that the ability to find targets embedded within complex visual environments requires the dynamic programming of visuomotor search behaviors based on fMRI results. The research revealed many significant results. For instance, visuomotor search resulted a greater activation in the posterior parietal cortex and the frontal eye fields in the right hemisphere. Furthermore, the activity in a network of cortical regions have an influence on the search-dependent variance in superior colliculus activity. Saccadic eye movements, covert shifts of attention, and visuomotor search occurred overlapping but not identical zones of activation. Lastly, this study focused on functional anatomy of overt spatial exploration rather than covert shifts of spatial attention (Gitelman et al., 2002) Another research carried out by Lobben et al. (2005) attempted to measure and analyze users' navigational abilities by using fMRI. In this context, individual performances of users and activated brain areas were determined while they perform rotation and sleuthing tasks. The research had some interesting results such as; during sleuthing, activation in the right and left hemispheres were similar. However, rotation task activated the right hemisphere more. In addition, while rotation caused more activation in the lateral occipital gyri which is the center of visual processing, sleuthing activated middle frontal gyrus, postcentral gyrus, and the angular gyrus more than rotation task.

Similar to fMRI, EEG records the electrical activity along the scalp produced by firing of the neurons in the brain with high temporal resolution and relatively low cost. In this method, electrodes placed in specific parts of the brain, which vary depending on which sensory system is being tested, make recordings that are then processed by a computer (Lee et. al., 2009). As a single neuron is not sufficient to draw a measurable potential at the scalp, the aggregation of synchronized neurons is considered during EEG. Since EEG is time-sensitive (i.e. a single millisecond) and a direct measure of the electrical activity in the brain, it resolves the changes depending on the different cognitive processes while performing certain tasks. However, because of the fact that the electrical activity is measured by electrodes placed at few specific regions, it is difficult to locate where in the brain the activity comes from. EEG can be used simultaneously with fMRI so that high-temporal-resolution data can be recorded with high-spatial-resolution (Lee et. al., 2009) (URL 5).

EEG has been used for various medical purposes such as critical monitoring and diagnostic tool in the clinic, sleep or fatigue monitoring (Winslow et al., 2013). Besides medical purposes, it helps addressing the question of how the brain answers to a specific visualization, image or design (Lee et. al., 2009). What generally recorded in EEG is called ERP (event-related potentials) allowing to check an EEG in certain moments when subjects received the event. The EEG signal represents oscillations observed across a wide range of frequencies which are commonly divided into distinct frequency bands (e.g., alpha band: 8–12 Hz, theta band: 4-8 Hz). Spectral analyses of the EEG can be used to compute the band-specific frequency power for given periods of time. Event-related power decreases from a reference to an activation interval are commonly referred to as event-related desynchronization (ERD), while power increases are referred to as event-related synchronization (ERS) (Pfurtscheller & da Silva 1999). These EEG metrics will be explained in detail in Chapter 3, 4 and 5.

1.1.6. Why mixing methods, i.e. integrating eye tracking with brain imaging methods?

Many methods such as eye-tracking, thinking aloud, interview or questionnaire have hitherto been applied in cartographic usability research (e.g. Herbert & Chen, 2015; Kveladze, Kraak & van Elzakker, 2017; Ooms, 2016; Ooms, Dupont & Lapon, 2017). On the one hand, how human brain supports spatial tasks has not sufficiently been researched yet, despite the advances in brain science and spatial cognition. There is especially a lack of research on the sources of individual differences (i.e. expertise, gender, etc.) and the relationship between the organization of spatial thinking and geographic space. On the other hand, there is limited empirical evidence on the user’s cognitive processes involved in map-related tasks, although cartographers hold theoretical knowledge on usability and design issues of maps (e.g. Kimerling, Buckley, Muehrcke & Muehrcke, 2009; MacEachren, 2004; Ooms, et al., 2015; 2016). Many disciplines from sports (e.g. Barfoot, Matthew & Callaway, 2012; Masaki, Hirao, Maruo, Foti & Hajcak, 2018) to marketing (e.g. Ohme, Reykowska, Wiener & Choromanska, 2010; Verhulst, Slabbinck, Vermeir & Larivière, 2018) make use of brain imaging techniques (e.g. EEG, fMRI) to understand the cognitive procedures of individuals. EEG provides direct measures of instantaneous electrical brain activity; it can benefit cartographic user studies, be combined with other quantitative methods, such as eye-tracking (ET) and sketch maps, to gain a better understanding of cognitive abilities and limitations of novice and expert map users. The insights that particularly arose from the differences due to expertise will henceforth contribute to creating effective cartographic products.

Eye tracking studies provide valuable information about use/user issues of cartographic products. Additionally, some other spatial cognition research employed brain imaging techniques to understand user behaviors. Linking the psychological/mental processes underlying the sensations we experience in everyday life to their underlying physiological biochemical processes is a challenging research area. To achieve this goal, we have to understand perception, in other words, how our brain makes sense out of the signals coming from our senses. Not only 'how the perception is processed', but also 'the attention' (how our brain can cope with the huge data and how it selects certain information)' should be understood (Görgen, 2010). There are two types of attention; overt and covert. Overt attention is defined as selectively processing one location over others by moving the eyes to point at that location. Covert attention is defined as paying attention without moving the eyes. While overt attention is externally observed, covert attention process is not visible from outside. To investigate covert attention, one can use psychophysical experiments or apply techniques like EEG or fMRI. For overt attention, one can use eye tracking to find out what actually guides the process and whether physical features of pictures play a role in guiding our attention. In this case, although eye tracking provides information about the gaze location, it does not provide any information about neuronal activity. Likely, EEG or fMRI do not directly provide information about the gaze position (Görgen, 2010). Hence, we need a holistic approach to understanding and interpreting map user behaviors.

There are several different usability studies on cortical processing in response to stimulus presentation. Thus, if an EEG experiment is designed in a way to parse eye tracking data simultaneously, it will be possible to monitor performance of users in visual tasks. By this way, it may lead a significant insight to the map user behaviors. Combined EEG and eye tracking has been used in many research areas. For instance, Alves et al. (2012) introduced a concept to explore customer experiences in service design by considering an augmented customer journey using EEG and eye tracking. Another research intended to enhance cognitive performance in sport by measuring neurocognitive activity and visual focus in real time which can be used to provide immediate feedback to the coach, in 'real world' settings, for optimizing training protocols for the individual athlete or for Eye Movement Desensitization & Reprocessing therapy, focused relaxation, etc. (Bartfut, et. al., 2012).

1.2. Research rationale and synopsis

For a map to become meaningful, it requires a map user, and to improve the understanding of map design, cartographers should be well aware of the principles of human perception and cognition and that design has a great impact on usability (Griffin, 2017). According to Eckert (1908), map logic agrees to the map production rules that has strong influence on cartographic perception. To be able to understand map users' behaviors, it is important to identify the cognitive procedures. A need for research that evaluates the cognitive issues related to map-use, has been identified long time ago, but remains largely unanswered. It is essential that experts in cartography – professional map-makers – understand how the novice users read, interpret, and store the visual information presented to them. With the understanding of map knowledge of users, cartographers can determine how to use the input stemmed from individual differences to enhance the design of maps for specific purpose and user groups, in other words, focus on effective map designs that ideally do not cause a high cognitive load. To do so, EEG and eye tracking can be integrated since both methods suggest statistical and visual analysis, which may reveal significant insights about map user behaviours.

1.2.1. Research objectives & thesis outline

Research Objective 1:

Contribute to the understanding of how different map users process the visual information on digital 2D static maps.

RQ 1: How do expert and novice map users "study and store" the visual information presented on digital 2D static maps?

RQ 2: How do expert and novice map users "recall and use" the visual information previously presented on digital 2D static maps?

Research Objective 2:

Evaluate the potential of brain imaging techniques, integration of EEG with eye tracking for cartographic cognitive/usability research

RQ 3: What is the added-value of EEG in terms of cartographic usability research?

Research Objective 3:

Explore the influence of a subset of visual variables (i.e. location, size, shape, color) in spatial cognition and the use of this input to enhance the design and communication of cartographic products.

RQ 4: How does the participants' attentional behavior vary towards the map elements of interest?

RQ 5: How do we improve the design of maps based on the input collected through user experiments?

Research Objective 1 aims to understand the influence of expertise on the spatial memory abilities and the attentional behavior of expert and novice map users by extending the eye tracking usability research done by Ooms (2012). In this context, while RQ1 focuses on the procedures during the map study phase (i.e. map learning), RQ2 refers to the processes of the retrieval of the previously gathered map-related information (i.e. spatial memory). Research Objective 2 explores the contribution of EEG in cartographic usability research to understand the usability issues of maps and the cognitive issues of map users whereas Research Objective 3 deals with the influence of a subset of visual variables in map-learning and how this input can be integrated into map design.

To cover the above-mentioned research objectives and questions, two user experiments were conducted which will be explained in detail in the following chapters. I will try to provide explanations for the fundamentals of the cartographic user experiment design, specifically with eye tracking and EEG, and answer the research questions through conducted user experiments. In both experiments, gaze and brain activity were recorded simultaneously, however, Experiment 1 had a simple design and an exploratory characteristic, since we would initially assure that the eye tracking and EEG synchronization is of sufficient quality to explore users' cognitive behaviors towards map stimuli. On the contrary, in Experiment 2, a complex and more structured approach was followed as a result of lessons learned from the previous experiment and collaborating with the domain experts, therefore, it was hypothesis driven. Table 1.4 summarizes all the experimental design elements together with the goals and hypothesis of both experiments whereas, Table 1.5 overviews the structure and the methodological aspects of these experiments, which will be explained in detail in the following chapters.

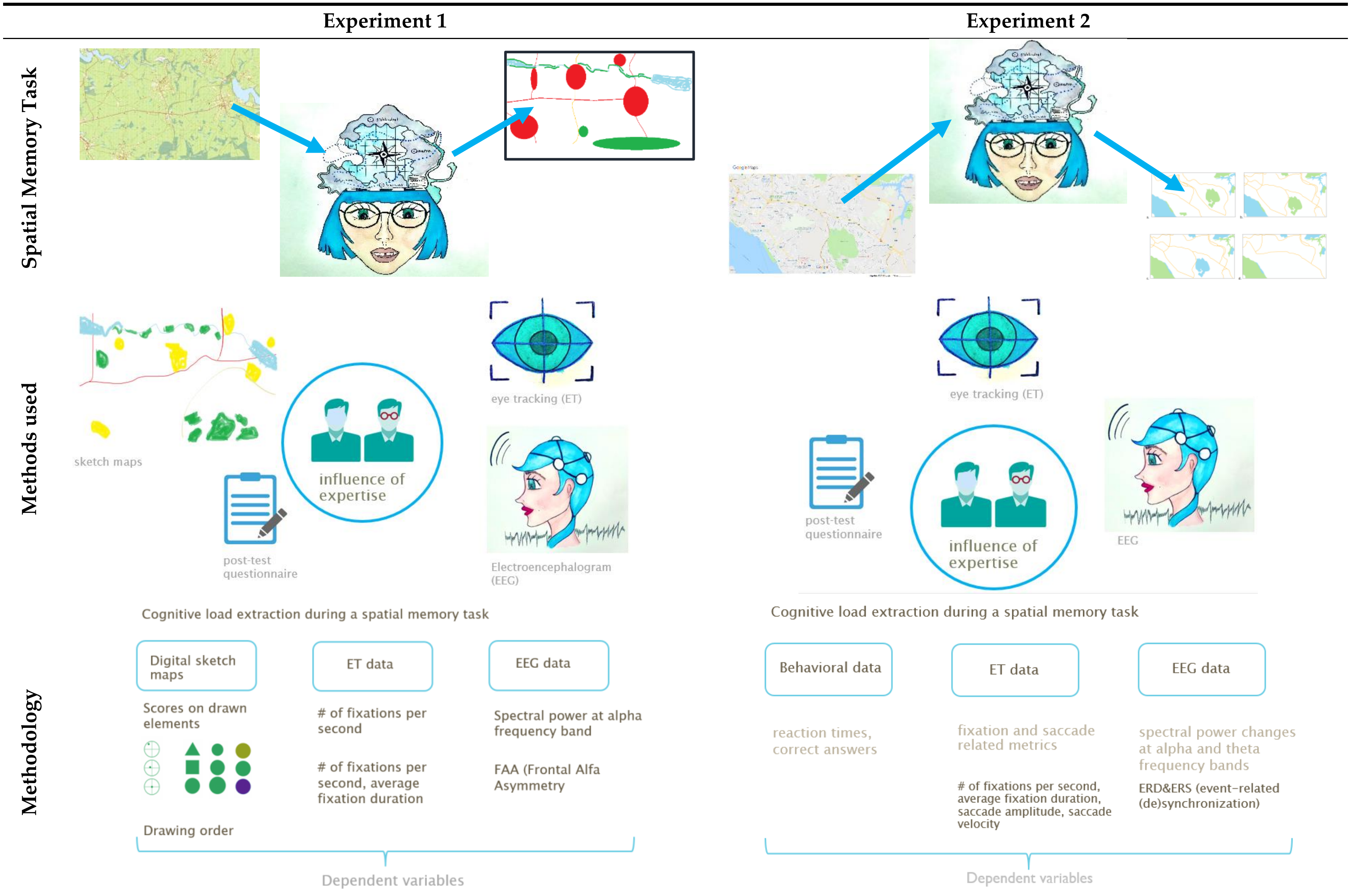
Table 1.4. Summary of the two experimental designs

	Experiment 1	Experiment 2
Research Question	How does cognitive load vary between expert and novice participants while memorizing the main structuring elements of a map stimulus without any time constraints?	How does cognitive load vary between expert and novice participants while memorizing a (part of) map content in a limited study period? How does the complexity/difficulty of the task influence the cognitive load?
Goal	To evaluate the cognitive processes, abilities and/or limitations of map users when they first study a digital 2D static map and retrieve this information later.	To test the effect of task difficulty on behavior, which is the retrieval of the main structuring elements with varying levels.
Hypothesis	We expect that the spatial memory task will cause higher cognitive load in novice participants compared to expert ones.	The tasks involving the retrieval of only linear features will cause lesser cognitive load for both groups compared to the other features. We additionally expect that expert participants would perform better at tasks demanding higher cognitive load.
Participants	56 participants: 24 experts (13 females, 11 males) 30 novices (7 females, 23 males) Age range: 18-35	38 participants: 17 experts (9 females, 8 males) 21 novices (9 females, 12 males) Age range: 25-35
Task procedures	Participants studied one map stimulus for as long as they wanted to memorize all the main structuring elements included in the map they studied. Once they thought they had studied the map long enough, they pressed a certain key and then they had to draw this map from memory by using MS Paint. After drawing the sketch map, participants used a special key to terminate the task.	Randomized block design: Seven blocks representing seven difficulty types. Each block includes 50 trials (i.e., one for each stimulus) focusing on the similarity of: Block 1: The whole map Block 2: Roads and hydrography Block 3: Roads and green areas Block 4: Green areas and hydrography Block 5: Green areas Block 6: Hydrography Block 7: Roads
Independent variables	1 map design type (i.e., digital 2D static topographic map) 1 task difficulty level (i.e., retrieval of the main structuring elements of the whole map stimulus) 2 expertise levels (i.e., experts vs. novices)	1 map design type (i.e., screenshots of Google's road maps) 7 task difficulty levels (i.e., classified as easy, moderate, hard) ~ linear & polygon features within blocks 2 expertise levels (i.e., experts vs. novices)
Dependent variables	Trial durations*, eye movements, EEG (alpha power, FAA), self-reported metrics (i.e., questionnaire)*	Response time of correct answers, eye movements, EEG metrics (ERD-ERS), self-reported metrics (i.e., questionnaire)*

*not mentioned in this chapter, but published in Keskin et al. (2018).

As seen in Figure 1.5, there is no one-to-one correspondence between the research questions and the chapters. Chapter 2, 4, and 5 are the modified versions of the articles published on international peer-reviewed journals. This chapter is dedicated to the literature review related to spatial cognition, cognitive cartography, user-centered cartographic design and usability research methods that are the basis of the thesis rationale. In Chapter 2, Experiment 1 is presented with a focus on eye tracking and sketch maps, therefore leaving the EEG part out. Chapter 3 includes an overview of the technical and methodological issues of a cartographic user experiment design using eye tracking and EEG mostly based on hands-on experience. That chapter also explains how the second user experiment was designed and Chapter 4 concentrates on eye tracking and EEG methods used in both experiments. In this context, two user experiments are compared in terms of their design and the preliminary results of eye tracking and EEG are presented. In Chapter 5, Experiment 2 is explained and full eye tracking and EEG results are provided. Chapter 6 is dedicated to the post-hoc eye tracking analysis based on AoI (area of interest) using the data collected in Experiment 2. Chapter 7 discusses the results of the user experiments as the research questions are revisited and further recommendations are made, whereas Chapter 8 summarizes the results of the previous chapters and presents concluding remarks.

Table 1.5. Summary of two user experiments



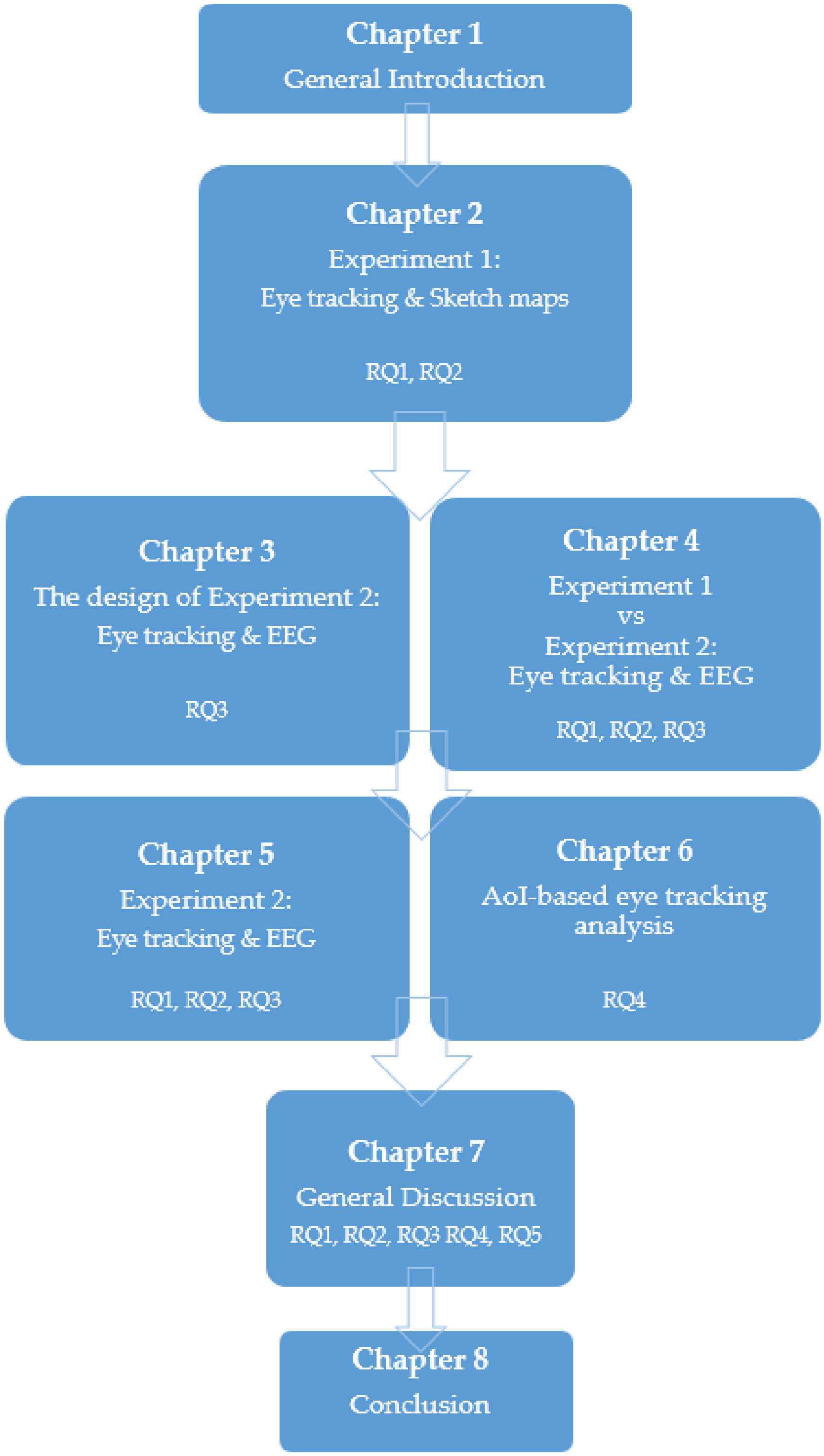


Figure 1.5. Dissertation Outline

References

Alves, R., Lim, V., Niforatos, E., Chen, M., Karapanos, E., & Nunes, N. J. (2012). Augmenting Customer Journey Maps with quantitative empirical data: a case on EEG and eye tracking. arXiv:1209.3155.

Amar, R., Eagan, J., & Stasko, J. (2005). Low-level components of analytic activity in information visualization, in Proceedings of IEEE Symposium on Information Visualization. Baltimore, Maryland, 111-117.

Armstrong, M.P., Densham, P.J. (1995). A conceptual framework for improving human–computer interaction in locational decision making. In Cognitive Aspects of Human–Computer Interaction for Geographic Information Systems M.J. Egenhofer (Eds.), 343–354. Kluwer Academic Publishers, The Netherlands.

Barfoot, K. M., Matthew, M. C., & Callaway, A. J. (2012). Combined EEG and eyetracking in sports skills training and performance analysis. In World Congress of Performance Analysis of Sport IX University of Worcester.

Bohbot, V. D., Kalina, M., Stepankova, K., Spackova, N., Petrides, M., & Nadel, M. (1998). Spatial memory deficits in patients with lesions to the right hippocampus and to the right parahippocampal cortex. Neuropsychologia, 36, 1217–1238.

Bohbot, V. D., Lerch, J., Thorndycraft, B., Iaria, G., & Zijdenbos, A. P. (2007). Gray matter differences correlate with spontaneous strategies in a human virtual navigation task. The Journal of Neuroscience, 27, 10078–10083.

Boulding, Kenneth E., (1956). The Image. Ann Arbor: University of Michigan Press. 150-151.

Casner, S. M. (1991). A Task-Analytic Approach to the Automated Design of Graphic Presentations, 'ACM Transactions on Graphics', 10, 2, 111-151.

Committee on the Conduct of Science (CCS) (1995). National Academy of Sciences, on being a scientist, 2nd ed. Washington, D.C.: National Academy Press.

Corkin, S., Amaral, D. G., González, R. G., Johnson, K. A., & Hyman, B. T. (1997). HM's medial temporal lobe lesion: findings from magnetic resonance imaging. The Journal of Neuroscience, 17(10), 3964-3979.

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

Deutschländer, A., Stephan, T., Hufner, K., Wagner, J., Wiesmann, M., Strupp, M., Brandt, T, & Jahn, K. (2009). Imagined locomotion in the blind: An fMRI study. NeuroImage, 45, 122–128.

Djamasbi, S., Siegel, M., & Tullis, T. (2010). Generation Y, web design, and eye tracking. International Journal of HumanComputer Studies, 68(5), 307-323.

Downs, R. M., & Stea, D. (Eds.). (1973). 'Cognitive Representations' in Image and environment: Cognitive mapping and spatial behavior. Transaction Publishers. 79-86.

Duchowski, A. T. (2007). Eye tracking methodology - Theory and practice. London: Springer. 328(614), 2-3.

Fits, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. Aeronautical Engineering Review, 9(2), 2429.

Fleetwood, M. D., & Byrne, M. D. (2006). Modeling the visual search of displays:mArevised ACT-R model of icon search based on eye tracking data. HumanComputer Interaction, 21(2), 153-197.

Freksa, C. (2013). Spatial computing – how spatial structures replace computational effort. In M. Raubal, D. Mark, & A. Frank (Eds.), Cognitive and linguistic aspects of geographic space (Lecture notes in geoinformation and cartography, 23–42). Berlin/Heidelberg/New York: Springer.

Gitelman, D. R., Parrish, T. B., Friston, K. J., & Mesulam, M. M. (2002). Functional anatomy of visual search: regional segregations within the frontal eye fields and effective connectivity of the superior colliculus. Neuroimage, 15(4), 970-982.

Gould, J. D., & Lewis, C. (1985). Designing for usability: Key principles and what designers think. Communications of the ACM, 28(3), 300-311.

Görgen, K. (2010). Combining Eyetracking and EEG. Publications of the Institute of Cognitive Science, 15.

Güç, B., Gencel, Z., & Karadayı, A. (2012). Mekan-Kullanıcı İlişkilerinin Hastane Örneğinde Sayısal Olarak Modellenmesi (Numerical Modeling of Space-User Relations in the Hospital Sample). SDU International Journal of Technological Science, 4(1).

Haklay, M. M., & Nivala, A. M. (2010). User-Centred Design. Interacting with geospatial technologies, 89-106.

Herbert, G., & Chen, X. (2015). A comparison of usefulness of 2D and 3D representations of urban planning. Cartography and Geographic Information Science, 42(1), 22-32.

Jacob, R., & Karn, K. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In: R. Radach, J. Hyona & H. Deubel (Eds.), The mind's eye: Cognitive and applied aspects of eye movement research (573-605). Amsterdam: Elsevier

Jenks, G. F. (1973). Visual integration in thematic mapping: Fact or fiction? International Yearbook of Cartography, 13, 112-127.

Jones, C. K. E. (2010). Cartographic theory and principles. Interacting with Geospatial Technologies, 37-65.

Keskin, M., Dogru, A. O., Guney, C., Basaraner M. (2016): "Kartografik Gösterimlerin Algılanabilirliğinin Ölçülmesinde Nörobilişsel Yöntemler (Neurocognitive Methods for Cartographic Usability Research)", Harita Teknolojileri dergisi (Electronic Journal of Map Technologies), 8(2), 177-189.

Kimerling, A. J., Buckley, A. R., Muehrcke, P. C., & Muehrcke, J. O. (2009). Map use: reading and analysis. Esri Press

Kimura, M., Uwano, H., Ohira, M., & Matsumoto, K. I. (2009). Toward constructing an electroencephalogram measurement method for usability evaluation. In HumanComputer Interaction. New Trends (95-104). Springer Berlin Heidelberg.

Knapp, L. (1995). A Task Analysis Approach to the Visualization of geographic data., in Cognitive aspects of human-computer interaction for geographic information systems, ed. By Nygers, T. L., Mark, D. M., Laurini, R., and Egenhofer, M. J., 355-371, Kluwer Academic Publishers, The Netherlands.

Kveladze, I. (2015). Space-time Cube Design and Usability. Doctoral dissertation, ITC Dissertation Number 268. ITC, University of Twente, Enschede.

Kveladze, I., Kraak, M. J., & van Elzakker, C. P. (2017, July). The Usability of a GeoVisual Analytics Environment for the Exploration and Analysis of Different Datasets. In International Cartographic Conference (207-221). Springer, Cham.

Lee, H., Lee, J., & Seo, S. (2009). Brain Response to Good and Bad Design. Lecture Notes in Computer Science, 111–120.

Lobben, A. K., Olson, J. M. and Huang, J. (2005). 'Using fMRI in cartographic research', in Proceedings of the 22nd ICC (p. 10), A Corufia, Spain.

Lynch, K. (1960). Image of the city. Cambridge, MA: Technology Press/ Harvard University Press.

MacEachren, A. M., Edsall, R., Haug, D., Baxter, R., Otto, G., Masters, R., Fuhrmann S., & Qian, L. (1999). Exploring the potential of virtual environments for geographic visualization. , in Annual Meeting of the Association of American Geographers, Honolulu, Hawaii, March 23-27, 1999, AAG, 371-380.

MacEachren, A. M. (2004). How maps work: representation, visualization, and design. New York, NY: Guilford Press.

Masaki, H., Hirao, T., Maruo, Y., Foti, D., & Hajcak, G. (2018). Feedback-related electroencephalogram oscillations of athletes with high and low sports anxiety. Frontiers in psychology, 9, 1420.

Montello, D. R. (2002). Cognitive Map-Design Research in the Twentieth Century: Theoretical and Empirical Approaches. Cartography and Geographic Information.

Norman, D.A. (1990). The Design of Everyday Things. New York: Doubleday

Ohme, R., Reykowska, D., Wiener, D., & Choromanska, A. (2010). Application of frontal EEG asymmetry to advertising research. Journal of economic psychology, 31(5), 785-793.

Ooms, K. (2012). Maps, how do users see them?: an in depth investigation of the map users' cognitive processes. Doctoral dissertation, Ghent University, Ghent. Retrieved from https://biblio.ugent.be/publication/3103344

Ooms, K., Coltekin, A., De Maeyer, P., Dupont, L., Fabrikant, S., Incoul, A., & Van der Haegen, L. (2014). Combining user logging with eye tracking for interactive and dynamic applications. Behavior Research Methods, 47(4), 977-993.

Ooms, K., De Maeyer, P., & Fack, V. (2015). Listen to the map user: cognition, memory, and expertise. The Cartographic Journal, 52(1), 3-19.

Ooms, K., De Maeyer, P., Dupont, L., Van Der Veken, N., Van de Weghe, N., & Verplaetse, S. (2016). Education in cartography: What is the status of young people's map-reading skills?. Cartography and Geographic Information Science, 43(2), 134-153.

Ooms, K. (2016, June). Cartographic User Research in the 21st Century: Mixing and Interacting. In T. Bandrov, & M. Konecny, (Eds.), 6th International Conference on Cartography and GIS Proceedings (p. 367). Albena, Bulgaria:.Bulgarian Cartographic Association. Retrieved from http://iccgis2016.cartography-gis.com/6iccgis-2016-proceedings/

Ooms, K., Dupont, L., & Lapon, L. (2017). Mixing methods and triangulating results to study the influence of panning on map users' attentive behaviour. The Cartographic Journal, 54(3), 1-18. doi:10.1080/00087041.2016.1213517

Ploner, C. J., Gaymard, B. M., Rivaud-Pechoux, S., Baulac, M., Clemenceau, S., Samson, S., & Pierrot Deseilligny, C. (2000). Lesions affecting the parahippocampal cortex yield spatial memory deficits in humans. Cerebral Cortex, 10,1211–1216.

Peterson, M. P. (1994). Cognitive Issues in Cartographic Visualization. Modern Cartography Series, 27–43.

Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. Clinical neurophysiology, 110(11), 1842-1857.

Redish, A. D., & Ekstrom, A. (2013). Hippocampus and related areas: What the place cell literature tells us about cognitive maps in rats and humans. İn Waller, D., Nadel, L. (Eds) Handbook of spatial cognition, 15-34. Washington, DC, US: American Psychological Association, x, 309.

Roth, S. F. & Mattis J. (1990). "Data characterization for intelligent graphics presentation," Proc. CHI '90, 193-200.

Roth, R. E. (2012). Cartographic Interaction Primitives: Framework and Synthesis, The Cartographic Journal, 49, 4, 376-395.

Roth, R. E. (2013). Interactive Maps: What we know and what we need to know. The Journal of Spatial Information Science. 6: 59-115.

Shneiderman, B. (1996, September). The eyes have it: A task by data type taxonomy for information visualizations. In Proceedings 1996 IEEE symposium on visual languages. 336-343.

Slocum, T. A., Blok, C., Jiang, B., Koussoulakou, A., Montello, D. R., Fuhrmann, S., & Hedley, N. R. (2001). Cognitive and usability issues in geovisualization. Cartography and Geographic Information Science, 28(1), 61-75.

Van Elzakker, C. P. J. M. (2004). The use of maps in the exploration of geographic data. Institute for Geo-Information Science and Earth Observation, Enschede, The Netherlands

Verhulst, N., Vermeir, I., Slabbinck, H., Mauri, M., Russo, V., & Larivière, B. (2018). A Multi-Method Exploration of Emotions During Service Delivery through Self-Service Technology or by an Employee. In La Londe Conference.

Wehrend, S. & Lewis, C. (1990). A problem-oriented classification of visualization techniques. In Proceedings of the 1st conference on Visualization '90 (VIS '90), Arie Kaufman (Ed.). IEEE Computer Society Press, Los Alamitos, CA, USA, 139-143.

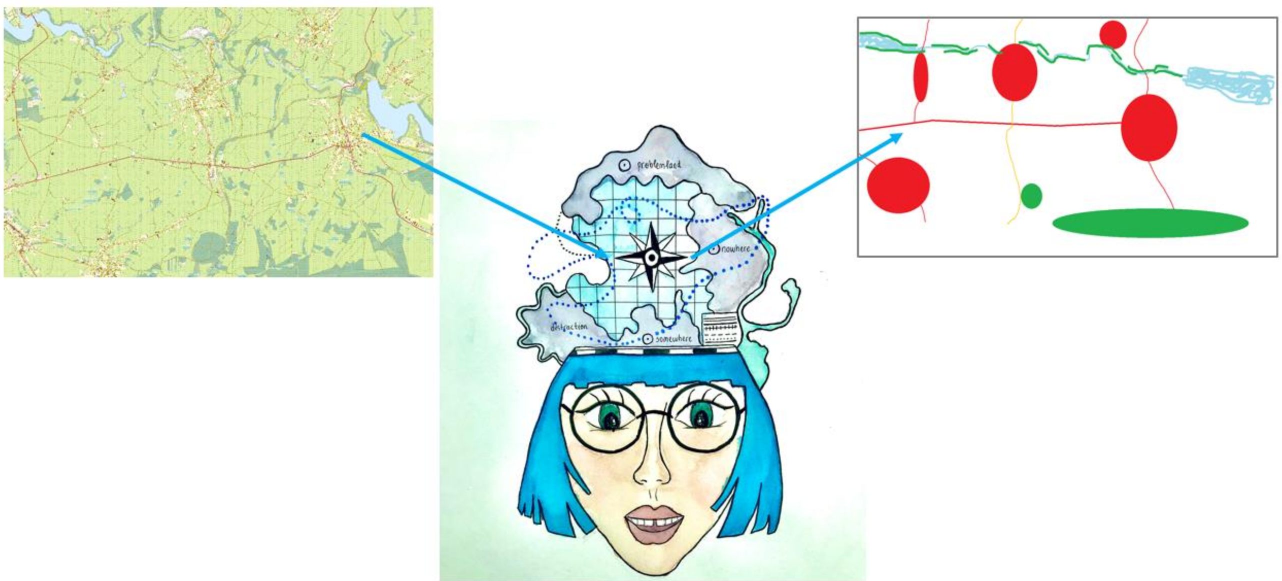
Winslow, B., Carpenter, A., Flint, J., Wang, X., Tomasetti, D., Johnston, M., & Hale, K. (2013). Combining EEG and Eye Tracking: Using Fixation-Locked Potentials in Visual Search. Journal of Eye Movement Research, 6(4), 1-11.

URL 1: http://www.usability.gov/what-and-why/index.html Last accessed on 29.04.2016

URL 2: http://www.usabilitynet.org/tools/r_international.htm Last accessed on 29.04.2016

URL 3: http://www.usabilitynet.org/management/b_what.htm Last accessed on 29.04.2016
URL 4: Roth, R. E. (2015). Mixed Methods for Cartographic & Geographic Research Retrieved from http://www.slideshare.net/reroth/mixed-methods-for-cartographic-geographic-research?qid=ffb84221-b757-445a-a262-6271d5f0d66a&v=&b=&from_search=24 Last accessed on 24.06.2016.
URL 5: https://www.quora.com/What-are-the-main-differences-between-EEG-fMRI-and-SST-Steady-State-Topography Last accessed on 29.04.2016

Chapter 2:
Experiment 1:
Digital sketch maps & eye tracking



"Sketch maps reflect differences
both between worlds and within worlds."

Mark Billinghurst & Suzanne Weghorst, 1995

Abstract. *This chapter explores (a group of) map users' cognitive processes in learning, acquiring and remembering information presented via digital 2D static topographic maps. In this context, we conducted a mixed-methods user experiment employing digital sketch maps and eye tracking. On the one hand, the performance of the participants was assessed based on the order with which the objects were drawn and the influence of a subset of visual variables (i.e. presence & location, size, shape, color). On the other hand, trial durations and eye tracking statistics such as average duration of fixations, and number of fixations per seconds were compared. Moreover, selected AoIs (Area of Interests) were explored to gain a deeper insight on visual behavior of the participants. Depending on the normality of the data, we used either two-way ANOVA or Mann-Whitney U test to inspect the significance of the results. Based on the evaluation of the drawing order, we observed that experts and males drew roads first whereas; novices and females focused more on hydrographic object. According to the assessment of drawn elements, no significant differences emerged between neither experts and novices, nor females and males for the retrieval of spatial information presented on 2D maps with a simple design and content. The differences in trial durations between novices and experts were not statistically significant while both studying and drawing. Similarly, no significant difference occurred between female and male participants for either studying or drawing. Eye tracking metrics also supported these findings. For average duration of fixation, there was found no significant difference between experts and novices, as well as between females and males. Similarly, no significant differences were found for the mean number of fixation.*

Author contributions: *Conceptualization: Merve Keskin, Kristien Ooms; Methodology: Kristien Ooms, Merve Keskin; Software, Validation, Formal Analysis: Merve Keskin; Investigation: Merve Keskin, Kristien Ooms, Ahmet Ozgur Dogru; Resources: Kristien Ooms, Philippe De Maeyer; Data Curation, Merve Keskin; Writing-Original Draft Preparation: Merve Keskin; Writing-Review & Editing: Merve Keskin, Ahmet Ozgur Dogru, Kristien Ooms, Philippe De Maeyer.*

This chapter is modified from:

Keskin, M., Ooms, K., Dogru, A. O., & De Maeyer, P. (2018). Digital sketch maps and eye tracking statistics as instruments to obtain insights into spatial cognition. Journal of Eye Movement Research, 11(3).

2.1. Introduction

Maps convey direct and indirect information about the objects they represent. In addition to information about the location, name, shape, and size of objects, maps provide spatial relationships among these objects. When a person needs to find a geographic phenomenon, select a route, navigate, or estimate a distance, (s)he tends to memorize the relevant direct or indirect information on a map. Together with human (perceptual, cognitive, and visual) abilities, the retrieval of spatial information is strongly correlated with map learning. Map learning is distinguished from other learning concepts because (i) it requires comprehending and memorizing the direct information presented in maps and (ii) all the information to be learned is presented at once. These two characteristics of map learning allow map users flexibility regarding when, how and in which order they execute tasks such as selecting and focusing (Thorndyke & Stasz, 1980). Hence, each user/user group develops different strategies for approaching the spatial information on maps (e.g. Çöltekin, Fabrikant & Lacayo, 2010; Ooms, De Maeyer & Fack, 2014a; Ooms, De Maeyer & Fack, 2015; Schriver, Morrow, Wickens & Talleur, 2008; Voyer, Postma, Brake & Imperato-McGinley, 2007).

This chapter intends to examine map users' cognitive processes of learning, acquiring and remembering information presented via digital 2D static topographic maps. The map users targeted in the research are broadly categorized as novices and experts considering their individual group differences of age, gender, ethnicity and language. The main research question addressed in this chapter is "do novices and experts use different strategies while studying maps and recalling map-related information?". In this context, the experiments are designed based on the principles and strategies defined by Thorndyke & Stasz (1980), Montello, Sullivan & Pick (1994) and Ooms et al. (2015). Various methods (e.g. think-aloud, eye tracking, interview) have been applied to evaluate the recall of map-related information from memory (e.g. Herbert & Chen, 2015; Kveladze, Kraak & van Elzakker, 2017; Ooms, 2016; Ooms, Dupont & Lapon, 2017). Sketch maps are one of these methods, since they concretize the extracted information from a cognitive map (also called a mental image, map image, mental map) through drawing. This concept is further discussed in the literature review in the following section.

User testing methods can be mixed for many reasons such as to enrich the quantitative research in cartography, to better contextualize map design and use/user recommendations, to improve the consistency and detail of results, and to adopt and adapt new approaches to our study design (Ooms, 2016; Ooms, et al., 2017; Popelka, Stachoň, Šašink & Doležalová, 2016; Roth, et al., 2017). In our study, we also use mixed methods of sketch maps, eye tracking (ET) and a post-test questionnaire. Both eye tracking and sketch map methods individually provide a considerable amount of valuable information related to map users. Therefore, the combination of these methods potentially brings advantages to the user study design in terms of methods, materials or user needs and to the evaluation of results, in addition to yielding additional insights about map users' behaviors. In fact, sketch maps and ET can be considered as complementary to one another; for instance, ET metrics can explain an outcome obtained from sketch maps or vice versa. ET is also valuable for the validation of results acquired from one method with those from the other.

2.2. Literature Review

2.2.1 Map learning and cognitive map production

Learning and remembering cartographic information are associated with how the human cognitive system addresses geographic information presented via maps to produce cognitive maps. Especially in the past four decades, cognitive maps have become an intriguing research topic in geography (Downs & Stea, 1977; Portugali, 1996) as well as in neuroscience (O'Keefe & Dostrovsky, 1971; O'Keefe, 1976; O'Keefe & Nadel, 1978) and psychology (Shettleworth, 2010). The discovery of place cells (O'Keefe & Nadel, 1978) and grid cells (Hafting, Fyhn, Molden, Moser & Moser, 2005) stands as evidence that there exists a group of neurons in the brain that are responsible for our cognitive maps and inner navigation. Various studies in cartography have emphasized how we see maps and how we derive meaning from them (Kimerling et al., 2009; MacEachren, 2004; Ooms et al., 2016).

Learning a map involves two interacting cognitive factors: (i) control processes and (ii) the memorial system (Kulhavy & Stock, 1996). The first cognitive factor of map learning refers to matching the map to the prior knowledge existing in the memory and the achievement of the map-learning task. In this respect, prior knowledge can originate from general and specific map knowledge. General map knowledge helps in distinguishing maps from other spatial displays. It enables the encoding of maps and the development of strategies for map learning.

The influence of general map knowledge on map learning depends on the perception of "maplikeness" and the degree of expertise (Dickmann, 2013; Kulhavy, Stock & Kealy, 1993a). Past studies have presented that map learning is more efficient when the stimulus is more maplike (Kulhavy, Schwartz & Shaha, 1983; Kulhavy, Stock, Woodard & Haygood, 1993b) and that experts and novices differ somewhat in terms of their ability to learn and remember information presented via maps (Thorndyke & Stasz, 1980). If an effective spatial behavior requires using vector-like information about distances and directions, this information should be stored as maplike representations (Shettleworth, 2010). O'Keefe and Nadel (1978) proposed that the spatial learning system forms cognitive maps through exploration and a later study claimed that associative learning integrates "all kinds of spatial information spontaneously into a unitary maplike representation" (Shettleworth, 2010, p.288). In addition, Portugali (1996) listed much research showing that without any prior training, children can comprehend aerial photographs at appropriate scale and are able to use them as maps. This outcome proves that maplike behavior is very fundamental in human development and that mapping skills develop much earlier than predicted.

Expertise plays a role equally important to maplikeness in map learning. To recall the locations and configurations of spatial objects from the memory usually requires experience with cartographic products in which topographic and topological information are represented by graphic symbols (Dickmann, Edler, Bestgen & Kuchinke, 2016). Unlike general map knowledge, specific map knowledge stems from the modifications of related information in the long-term memory (LTM) based on the degree of familiarity with particular map representations. These representations are called knowledge-weighted cognitive maps, which are constructed from perceptual stimulus, initial map-learning conditions and the way that the information has been used (Intons-Peterson & McDaniel, 1991).

The second cognitive factor of map learning - the memorial system - addresses the mode of representation and the resources to store and maintain cognitive maps (Kulhavy & Stock, 1996). A map image holds both features represented by visual variables and structural information. The structural information refers to a spatial framework such as geometric and metric relations among features, whereas the visual variables, described by Bertin (1967) are the fundamental units that help distinguishing map symbols and encoding information presented via maps. Visual variables (i.e. position, size, shape, value, color hue, orientation, and texture) play a key role in cartographic design because their use for map symbols has a great impact on visual attention and perception. How these variables are perceived depends on their property (i.e. selective, associative, ordered or quantitative) (for further reading see also Wolfe, 2000).

As Kulhavy & Stock (1996) argued, we should understand whether our cognitive map is just a collection of features and their properties or it encodes structural relationships as well. The answer depends on the similarity of the map and its cognitive map. Clearly, all individuals create their own unique cognitive maps. Cognitive map creation occurs in a fashion similar to Haken's (1977) theory of information and self-organization (synergetics). Synergetics, originating in physics, is a method and a philosophy to explain the formation and the self-organization of individual elements in an open and complex system for the stability and the survival of the whole system (Haken & Portugali, 2016). Let us try to explain cognitive map creation in the human brain, which is also an open and complex system. When the brain receives spatial information through the external world (a physical environment or maps), the cognitive system constructs a cognitive map out of a partial set of features stored in the brain as internal representations. During this procedure, the cognitive system is governed by order parameters, which are the common principles shaped by the interactions among the individual elements of the system (Haken & Portugali, 1996). According to the synergetics theory, atoms form order parameters, and order parameters enslave (govern) atoms (Haken & Portugali, 2016). Therefore, cognitive map construction can only be achieved "when a certain mapping principle, or mapping order parameter, enslaves the various features" through associative memory (Portugali, 1996, p. 14). As a result, the cognitive map is successfully created from the interaction of the internal and external representations of the environment influenced by order parameters (for further reading, see Lakoff, 1987; Edelman, 1992).

Nevertheless, our cognitive system is capacity-limited in terms of encoding new information for storage in LTM and also of retrieving and making use of old information already in memory (Kulhavy & Stock, 1996). As Atkinson & Shiffrin (1968) proposed, memory involves a sequence of three stages; sensory memory, working (short-term) memory, and LTM. Sensory memory holds the information gathered through all our senses for a brief time span and then decays and is lost. A part of the information in the sensory memory is transferred to the working memory (WM). The WM can receive selected inputs from the sensory register, as well as from LTM. WM is active during encoding and storing new information for short time periods or during the retrieval and use of the old information. On the other hand, LTM retains the informative knowledge (memories, things we learn, etc.) permanently, because it has an almost limitless capacity. Once

WM transfers information to LTM, this information can be remembered for longer periods. This transmission is called the learning process and requires rehearsal (Atkinson & Shiffrin, 1968; Kulhavy & Stock, 1996; Ooms et al., 2015) (Figure 2.1). Unlike LTM, WM has a limited capacity in terms of individual items of information called chunks. A chunk is any stimulus that has become familiar, hence recognizable, through experience (Simon, 1989; Cooper, 1998). To be able to draw cognitive maps, the chunks of information obtained from maps must be transferred from LTM to WM.

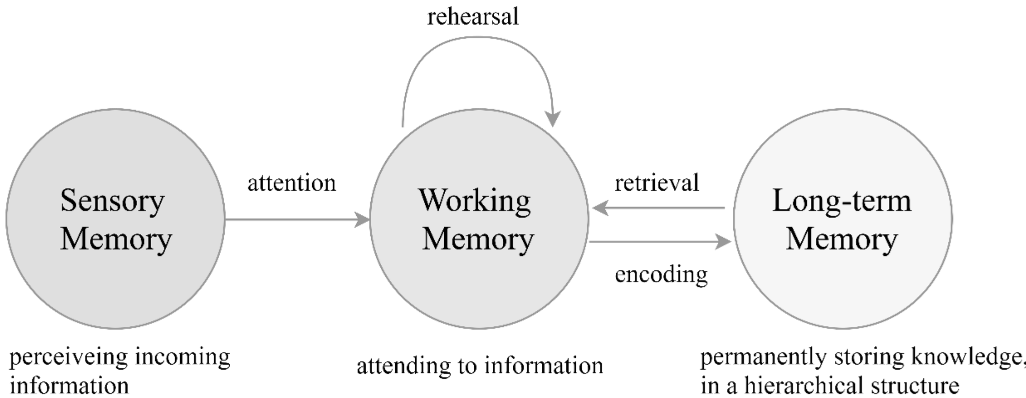


Figure 2.1. The interaction between memory types

Besides WM capacity, object-location memory and landmarks play principal roles for the cognition of the spatial objects, the formation of the cognitive representations and the recall processes of those. According to Tversky (1992), the brain reorganizes the information entirely through (i) hierarchical organization or categorization, (ii) the use of perspective, and (iii) the use of landmarks or cognitive reference points. Once people learn the locations of objects, they can establish the spatial structure of a map to form a mental representation or cognitive map of the environment. Cognitive maps hold information not only about spatial objects, but also the relations and distances between objects, even the absence of spatial objects. The distortions in the spatial object positions and their relations are indicators of hierarchical encoding and perceptual organization (Edler, Bestgen, Kuchinke & Dickmann, 2014). In this context, landmarks and routes are considered as the core units of a spatial representation and are helpful primarily for orientation (Siege & White, 1975; Bestgen, Edler, Kuchinke & Dickmann, 2016).

In cartography, empirical studies focusing on map design and spatial cognition are increasing, however, only a number of them devoted to the exploration of cartographic elements (e.g. visual variables) which play an important role in cognitive map formation (e.g. Stachoň et al., 2013; Edler et al., 2014; Dickmann et al., 2016). Hence, we cannot yet formulate the cognitive map production precisely and the assessment of this procedure is not straightforward. Nevertheless, sketch maps, considering their complexity, can be utilized as one of the sources to examine this process.

2.2.2. Sketch maps

Sketch maps are the reflections of individual cognitive maps. According to Forbus, Usher & Chapman (2004, p.61), sketch maps are defined as "compact spatial representations that express the key spatial features of a situation for the task at hand, abstracting away the mass of details that would otherwise obscure the relevant aspects". Therefore, the interpretation of sketch maps reveals the underlying task-related cognitive process of individuals. A sketch map is also a three-dimensional representation through space, time and sequence because the ordered retrieval of movements within time and space results in our cognitive maps (Huynh & Doherty, 2007). The hierarchical order of nodes and paths drawn on the sketch maps represents the hierarchical order of information (primary-level, secondary-level, and so on) presented on the maps. As Lynch (1960, p. 86) describes, "the sequence in which sketch maps were drawn seemed to indicate that the image develops, or grows, in different ways." The earlier the element is recalled, the more important it is to a person. Lower hierarchical levels correspond to decreasing amounts of spatial information, decreasing frequency of use and greater difficulty of remembering (Golledge & Spector, 1978). Hence, drawing order can yield insights into how these elements are stored in the user's memory. In other words, if an element is drawn earlier, it means that it is more accessible in LTM, thus, retrieved with ease (Ooms et al., 2015).

Sketch maps have been used in several research projects as a data collection method to investigate the cognitive processes of map users (e.g. Bell & Archibald, 2011; Billinghurst & Weghorst, 1995; Forbus et al., 2004; Huynh & Doherty, 2007; Ooms, 2012). Sketch maps are often combined with the think aloud procedure as a complementary data collection method (e.g. Kettunen, Putto, Gyselinck, Krause & Sarjakoski, 2015; Ooms et al., 2015) because thinking aloud gives insights into the user's unfiltered thoughts. Thinking aloud itself, however, has the disadvantage that it also consumes part of the user's memory capacity.

2.2.2.1. Retrieving a sketch map from memory

Spatial memory is controlled by perception-based and memory-based processes (Edler et al., 2014). Sketch maps underlie the map users' cognitive procedures of learning and remembering the information presented via maps. Hence, it is essential to identify the cognitive procedures involved during both learning and the retrieval of map-related information. Learning requires to create a higher framework of specific graphic features (e.g. map-inherent features or grids). While studying, a map reader first perceptually divides the map into a number of spatial chunks. In this context, the structuring map elements, such as roads, hydrographic features or gridlines, initiates chunking process, thus, helps regionalizing the map and assists learning of map elements and their spatial relations. These structuring elements represent the spatial information of the map content in a hierarchically structured fashion and form fundamental units of cognitive maps, therefore, facilitate the perception and recognition of object locations (Edler et al., 2014).

The first step of retrieval process is the orientation of the participant regarding the task (i.e. establishing a strategy to execute the task from the beginning to the end) and the surroundings (in this case, the drawing environment and its tools). The second step is task execution, in which participants form links between cognitive processes through WM and LTM. In chronological order, the participant first consults WM to check whether there is information about map elements that must be drawn. If the information exists in WM, the participant draws these elements; if not, he must consult LTM, which is responsible for the recalling act. For a participant to draw an element whose information is stored in LTM, this information needs to be transferred to WM. Afterwards, evaluation occurs for editing or redrawing, and then, the participant asks WM once again to finalize the procedure (Ooms et al., 2015). It is important to remember that this procedure is repetitive and continues until the participant is satisfied with the result. During this procedure, the sensory memory captures the image of the sketch map and transfers it to the WM. The memories of this original stimulus, which were previously stored in LTM, need to be recalled. Once the participant retrieves that information, (s)he can compare the sketch map with the original stimulus depending on the location, size, shape, color, etc. The retrieval process for chunks of information requires activation of the related information. This activation involves pointers, schemas and links between schemas stored in LTM. These pointers activate and retrieve the desired chunks of information from LTM and place them in WM (Ooms et al., 2015).

2.2.3. Eye tracking

It is known so far that the early beginnings of perceptual organization is evidenced by the first fixation on a visual stimulus (Edler, et al., 2014). The fixation-related behavior and other eye movement data can be measured via eye tracking which is a widely used quantitative user-testing method. Eye tracking has contributed to human-computer interaction usability studies in numerous disciplines varying from psychology to software engineering, marketing, sports, aviation, navigation and so forth (e.g. Ball, Lucas, Miles & Gale, 2008; Bertrand & Thullier, 2009; Crundall, Underwood & Chapman, 2002; Jacob & Karn, 2003; Poole & Ball, 2006; Schriver et al., 2008; Wedel & Pieters, 2008;). Many cartographers also employed eye tracking in their usability research, especially for the assessment of visual elements (e.g. Çöltekin et al., 2010; Dickmann et al., 2016; Fabrikant, Hespanha & Hegarty, 2010; Ooms, 2012; Ooms et al., 2014b; Ooms & De Maeyer, 2015; Ooms et al., 2017).

As explained earlier in the previous chapter, visual elements in topographic maps assist learning and recognition of location of map elements. Some eye tracking research has revealed how a map user processes those visual elements (e.g. Bestgen et al., 2016; Dickmann et al., 2016; Kuchinke, Dickmann, Edler, Bordewieck & Bestgen, 2016). Eye movement statistics, which can be linked to the cognitive processes when a participant interact with visual stimuli on the screen, consist of a list of pixel coordinates on the screen regarding various positions of the gaze (POR: point of regard). From the raw data, useful metrics such as how long (fixation duration) and how often (fixation count) a person focuses on a specific area of interest, together with his scan-path characteristics (the length and speed of the gaze activity), can be derived (Ooms et al., 2014b). These metrics can also be analyzed for specified regions of the stimulus, called Areas of interest (AoIs). AoIs are subregions of a stimulus that are of high importance for a hypothesis and are created based on the semantic information of the stimulus (Blascheck, et al., 2014).

Our literature study showed that there is a lack of research on the sources of individual differences (e.g. expertise, gender, etc.) and the relationship between the organization of spatial thinking and geographic space. Furthermore, there is a limited empirical evidence on user's cognitive processes involved in map-related tasks, although cartographers hold theoretical knowledge about usability and design issues of maps. Therefore, we aim to evaluate the abovementioned cognitive process on a digital 2D static topographic map to determine the cognitive abilities and/or limitations of map users when they first study the map and retrieve this information later. In this context, we propose collecting data via digital sketch maps, instead of conventional pen and paper method, to be able to link this with ET statistics. Both the ET data and the sketch maps give insights in the users' cognitive processes, but from a different angle. By triangulating the obtained insight, a deeper understanding regarding individual differences of map users can be obtained.

2.3. Methods

2.3.1 Participants

A total of 56 participants took part in the study, with 24 experts and 30 novices. The numbers of female and male participants were 7 and 23, respectively, for novices and 13 and 11, respectively, for experts. The ages of 96% of the participants ranged between 18 and 34, which corresponds to a rather young user group. The novice participants were undergraduate Business and Economy students whose ages varied between 18 and 24 years and who gained credits in return for their participation. The expert group, whose ages ranged between 25 and 34, consisted of participants who had at least a MSc. in Geography, Geomatics Engineering or related areas, and all of them were affiliated with the Department of Geography (Ghent University). The majority of the participants were Belgian (native language=Dutch), and there were six Asian expert participants (native language=Chinese). The experiment itself was designed in English.

While experts work with cartographic products on a daily basis, novices use cartographic products from time to time (e.g. Google maps) and were not trained before the experiment. Eight female and eight male experts had participated in a user experiment with ET previously. Two novice males indicated that they had participated in a user study before. The remaining 38 participants took part in user testing for the first time. All participants unanimously indicated in the post-test questionnaire that the map stimulus was not familiar to them.

2.3.2 Apparatus and recording

The experiment was conducted in the Eye Tracking Laboratory of the Marketing Department of Ghent University. The participants' eye movements were recorded with an SMI RED250 eye tracker mounted to the stimulus monitor. The stimulus was shown on a 22″ color monitor with 1680 x 1050 spatial resolution. We did not use a chin rest and the average distance between the participant and the monitor was 65 cm. Simultaneously with the gaze recording, we performed EEG measurements to estimate the cognitive load (Please see Annex 1 for orientation script). However, this chapter is dedicated to eye tracking and sketch map methods, we will introduce the theoretical background of EEG data acquisition, the synchronization of EEG and ET and related analysis in the upcoming chapters (see Chapter 3-5).

2.3.3 Materials

The stimulus was selected from the Belgian 1:10k topographic map series (Figure 2.2). We paid attention that it was not too complex yet contained some specific main structuring elements. To combat the learning effect, the selected map did not cover a well-known area/city.

Figure 2.2. Original map stimulus shown in memory task (This map stimulus is the same material used by Ooms (2012). This data was produced by Belgian national mapping agency, NGI/IGN (Nationaal Geografisch Instituut/Institut Géographique National)).

2.3.4 Procedure

Participants were instructed to study the map stimulus – for as long as they wanted – to be able to remember the main structural elements (rivers, roads, water bodies, etc.). Once they thought they had studied the map long enough, they pressed a certain key as instructed beforehand and thereby exited the first part of the assignment. Next, they had to draw this map from memory by using MS Paint. This tool was selected because neither experts nor novices would need any prior training. After the execution of the task – in other words, drawing the sketch map – participants used a special key to terminate the task. There was no time limitation for either the studying or the drawing part. While participants studied and drew the map, their eye movements were recorded (please see Annex 2 for full instuctions).

2.3.5 Sketch maps analysis

The first step of sketch map analysis was to quantify the information presented within the maps. Therefore, we determined the structural map elements on the original map/stimulus and then counted and classified them into four main categories: hydrology, land-cover, settlements, and roads. The original map consisted of four hydrographic features, four land-cover features, eight residential areas/settlements, and ten roads (in total, 26 map elements).

The sketch maps were analyzed based on the literature on cognitive processes and sketch map evaluation for cartographic usability (*see* previous section). In this context, two main criteria were identified; (i) drawing order and (ii) the score on drawn elements.

2.3.5.1 Drawing order

Drawing order information was derived from the registered eye tracking video and each participant's data were processed individually. For the assessment of drawing order, the scoring system used by Ooms et al. (2015) was implemented. The scoring was 100, 50, 25, and 5 for the first, second, third, and fourth elements of a certain category, respectively. If a certain element did not exist on the sketch map, it did not receive any point. The rationale behind this scoring algorithm is simply assigning the highest score on the first drawn element and the least to the last drawn one. Among the first three classes (i.e. drawn elements), the weight is halved in value for each consecutive class so that the first drawn element stands out more. The last drawn element (i.e. fourth class) should have the least score, but not zero, since it is drawn on the sketch map. Therefore, its weight equals to the 1/5 of the third class. Finally, the average scores for each map category were calculated separately for expert and novice groups. Higher scores indicated that a certain element belonging to one of the four categories was drawn earlier. Therefore, 100 points would mean that all participants drew this category first. In this way, drawing order analyses contributed to the understanding of the hierarchical construction of the cognitive map.

The visual variables considered for the scoring of the drawn map elements were presence and accuracy (position), size, shape, and color, which corresponded to the qualitative characteristics of the sketch maps. The scoring provided information about how well the sketch map was executed (complete and accurate) and accordingly, how well the cognitive map was constructed.

2.3.5.2 Score on drawn elements

Presence and accuracy

The scoring system as used by Ooms et al. (2015) was implemented to quantify the position of map features. If present and in the correct relative location, an object scored one point. If present and in a considerably wrong relative location, an object scored half a point. Finally, if absent, an object scored zero point. If a person successfully located every map element in the correct location, (s)he scored 26 points (total number of map elements). The results were expressed in percentages with 26 points representing 100%.

Shape, size and color

The shape, size and color characteristics of drawn elements were ranked by employing a system similar to that used by Billinghurst & Weghorst (1995). Their ranking scale was only modified to a 100-point scale; therefore, an incorrect score was 33.3, a partially correct score was 66.7, and a correct score was 100. Here, the participant's drawing ability was neglected, and instead, we focused on how well the sketch map represented the area in the topographic map. For instance, linear objects such as roads and rivers should be illustrated as lines with varying thickness, and when individual roads connect, they should picture the overall road construction. Different logic should be followed for the aggregation of areal objects such that the individual buildings can be grouped and drawn as a single element (i.e. settlement), since the participants were particularly asked to draw the main structural elements. Additionally, only the major shape characteristics of the map elements were taken into consideration for scoring. For instance, both roads and railroads could be drawn as single lines, although they were depicted by double lines in the original map.

2.3.5.3 Aggregation presence & accuracy (1), shape (2), size (3) & color (4)

Presence & accuracy, shape, size and color of drawn elements show "how well" the sketch maps were drawn. Until this point, we have tried to evaluate the influence of each criterion individually. However, the aggregation of all criteria used for scoring the drawn elements can offer a more objective measure to compare the quality of sketch maps. Inherently, the quality of sketch maps reflects the performance of participants. We treated each of the four parameters as if they have equal importance for the overall performance of a participant, and thus, we assigned each parameter the same weight. Overall performance scores were calculated as the average of individual performances for the four different groups (expert females, expert males, novice females and novice males) in a 0-100 scoring scale.

2.3.6 Eye tracking metrics

In addition to extracting the drawing order information from eye tracking data, eye tracking metrics such as the number of fixations per second and the average duration of fixation were analyzed. Similar to Ooms et al. (2015), the number of fixations per second was considered instead of the fixation count because the fixation count is an absolute measure that is related to the length of the trial. Since every participant completes the task in a different time span, the fixation count would be merely a reflection of the trial duration. It is important to note that there is a strong relationship between the number of fixations per second and another widely used metric, average fixation duration. The longer the fixation durations are, the fewer the fixations per second. The fixation duration is also linked to the cognitive processes of the visual stimulus. Longer fixations may indicate that reading the map becomes harder, which causes a rise in the cognitive load (Duchowski, 2007; Ooms et al., 2014b), or that the user finds the map or a certain part of it interesting (Ooms, 2012). People also concentrate their fixations on the most informative parts of the visual stimulus (Henderson & Ferreira, 2004).

These metrics were further complemented with trial durations to study the map on one hand and to draw the associated sketch map on the other hand (results presented separately in 5.1). Although there was no time limitation for both study and drawing parts of the memory task, trial times give insight about motivation and top-down attention. Inherently, longer trial durations for studying the map indicate higher level of interest or difficulty in storing the information in memory.

Furthermore, some ET metrics were analyzed for specific AoIs. These were created on the basis of a previous study of Ooms et al. (2014a) which implemented the same stimuli. This study revealed that, based on a gridded approach of AoI, users tended to focus most on main structuring topographic characteristics in the map stimulus (i.e. major roads, settlements and hydrographic features). In this study, we, thus selected the same object to be included in the AoI. Buffers were created around the linear features similar to what was done by Bargiota, Mitropoulos, Krassanakis & Nakos (2013). Based on the accuracy of eye tracker (0.5°) and the viewing distance (65 cm), buffer size was set to 21 pixels. In this context, the statistics such as how quickly participants notice an element (time to first fixation), how much time the participants spent in the region (dwell time), how many fixations occurred (fixation count, the number of fixations per second) and for how long (average fixation duration) were considered. These metrics were further complemented with trial durations to study the map, on the one hand, and to draw the associated sketch map on the other hand (results presented separately in 5.1).

2.4. Results

2.4.1 Trial durations

Trial durations were assessed in two phases: (i) study time for the map stimulus and (ii) the drawing time for the sketch map. Figure 2.3 illustrates a general overview of the study and drawing performances of experts and novices. The graph clearly shows that drawing took approximately twice – or in some cases more than twice – as much time compared to the study phase.

2.4.1.1. Study time

The average (mean) time for studying the map was 102.7 s (*N= 24, MED= 72.0 s, SD= 61.7 s*) for experts with a minimum of 27.1 s and a maximum of 226.6 s (Figure 2.4a) and 81.5 s (*N= 30, MED= 59.2 s, SD= 57.6 s*) for novices with a minimum of 23.2 s and a maximum of 292.8 s (Figure 2.4b). If we classify the performances of participants regarding to study time, 17% of experts spent 0-50 s; 41%, 50-100 s; 21%, 100-150 s; and 21%, 150 s and more. On the other hand, 35% of novices spent 0-50 s; 46%, 50-100 s; 6%, 100-150 s; and 4%, 150 s and more. The results confirm that experts allocated more time in studying than novices did. No significant interaction effect was observed between expertise and gender ($F_{(1,50)}$= 0.484, p= 0.490).

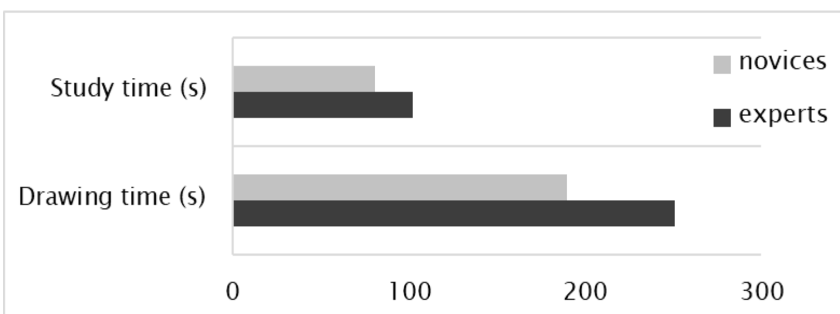


Figure 2.3. Trial durations of experts and novices

2.4.1.2. Drawing time

As for the study part of the memory task, there was no time limitation for the drawing part. The average drawing time for experts was 253.5 s *(N= 24, MED= 175.3 s, SD= 262.9 s*) with a minimum of 76.5 s and a maximum of 356.1 s (Figure 2.5a), whereas the average drawing time was 195.4 s *(N= 30, MED= 196.9 s, SD= 75.6 s*) for novices with a minimum of 50.2 s and a maximum of 1169.4 s (Figure 2.5b). We found no significant interaction effect between expertise and gender ($F_{(1,50)}$= 0.539, p= 0.466).

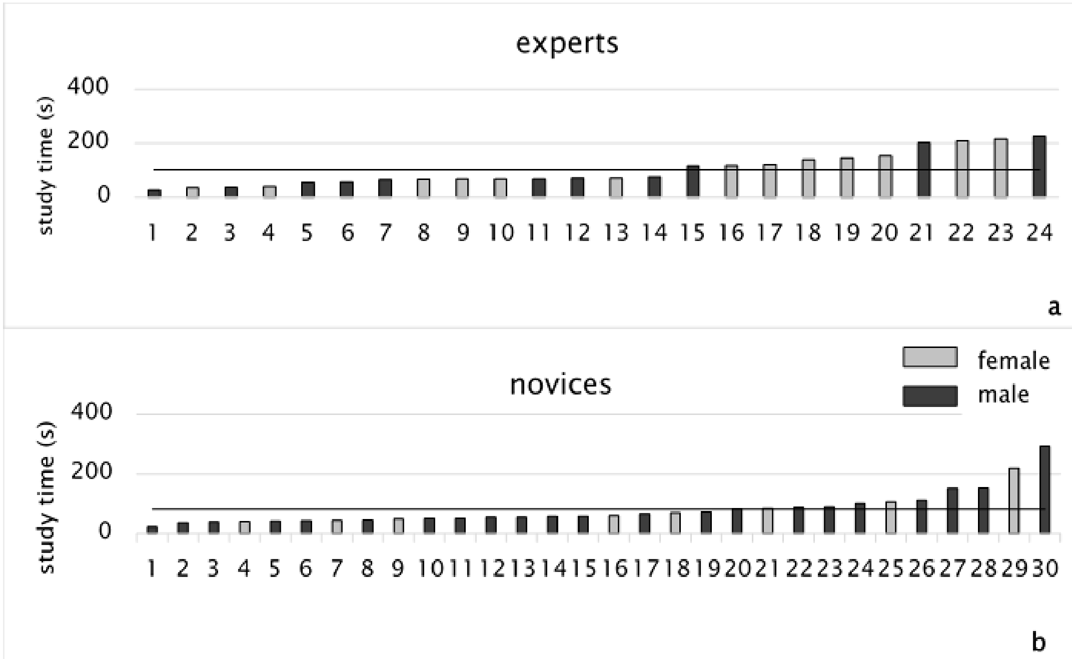


Figure 2.4. Study time of experts (a) and of novices (b) (black line: average)

The time spent on sketching the map might correspond to the richness of detail depicted in the sketch map, the difficulties encountered due to the lack of experience (e.g. unfamiliarity of the task and of the drawing tool), or recall issues. The fact that novices were faster in both studying and drawing may explain that novices were not aware of procedures involved in map production, did not exactly know what to remember. In addition, they are less involved with cartography, thus they might have paid less attention to having good results. Since the average drawing time for experts is greater than that for novices, some experts spent the longest time on the task. The extreme values that occurred in the expert group can be explained by the richness of main structural elements on the sketch maps. These sketch maps were detailed, contained larger numbers of structural elements and scored higher than the average among their group. Unlike in the expert group, there was a more balanced trend among novices (Figure 2.5b). However, the novices who spent the longest time (corresponding to one-third of the time that experts spent) received scores equal to those for experts on their sketch maps.

A Kolmogrov-Smirnov test was used to test of normality on the dependent variables, which are study time and drawing time. For both data, p= 0.000 suggested strong evidence of the data was not normally distributed (D_{study} (54)= 0.209, p < 0.05, and $D_{drawing}$ (54)= 0.258, $p < 0.05$). Since the data did not fit normal distribution, Mann-Whitney U non-parametric method was chosen to test significance of the results. It can be concluded that the differences occurred between novices and experts while both studying (*M= 90.9 s, SD= 59.9 s*) and drawing *(M= 221.2 s, SD= 194.3 s*) were not statistically significant (U_{study}= 275, p= 0.139 and $U_{drawing}$= 320, p= 0.486). Similarly, no significant difference emerged between female and male participants for either studying or drawing (U_{study}= 265, p =0.179 and $U_{drawing}$= 321, p= 0.734).

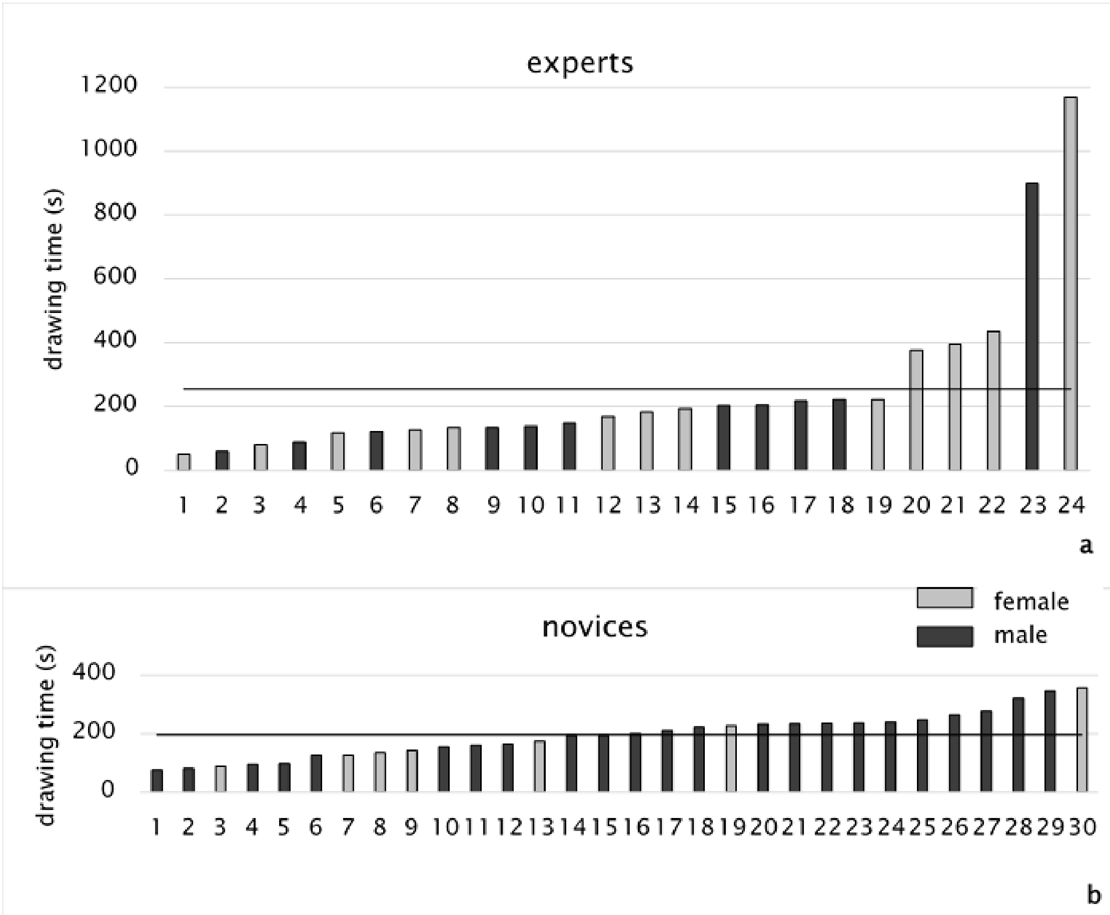


Figure 2.5. Drawing time of experts (a) and of novices (b) (black line: average)

2.4.2 Sketch map analysis

2.4.2.1. Drawing order

Although the spatial distributions of elements on the sketch maps were not properly structured or were even distorted, the drawing order (sequence) was similar to that found by Lynch (1960).

Figure 2.6 depicts the examples of sketch maps drawn by experts and novices for the memory task. According to the average scoring results of all participants, the hydrography (*M= 70.1, MED= 50, SD= 32.5*) and road (*M= 67.7, MED= 50, SD= 33.6)* categories were linked to the highest scores, whereas settlements (*M= 30.5, MED= 25, SD= 21.8*), and land-cover (*M= 9.1, MED= 5.0, SD= 11.7*) were associated with the lowest ones (Figure 2.7). This result means that the majority of participants drew hydrographic objects first. The drawing orders for experts and novices show a slight difference. While experts drew roads first, novices focused more on hydrographic objects such as rivers and water bodies. Hydrography and roads form the main structural elements on the maps. Settlements and land-cover elements (in this case, forest) were drawn third and fourth, respectively, for both user groups.

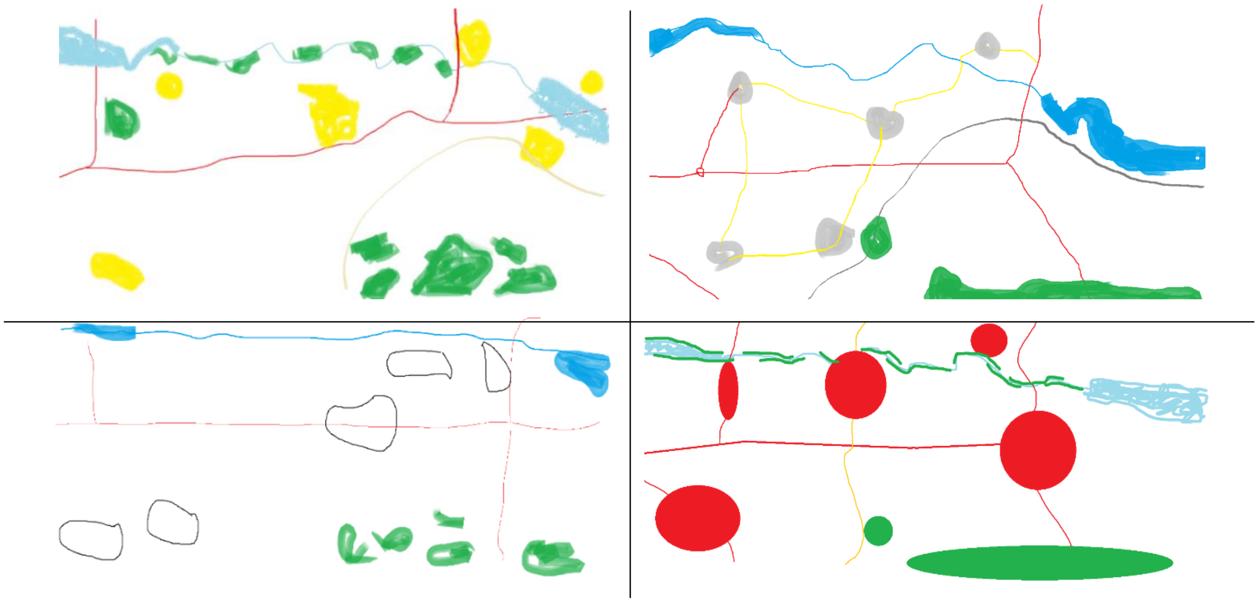


Figure 2.6. Sketch map examples (top and bottom left: novices; top and bottom right: experts)

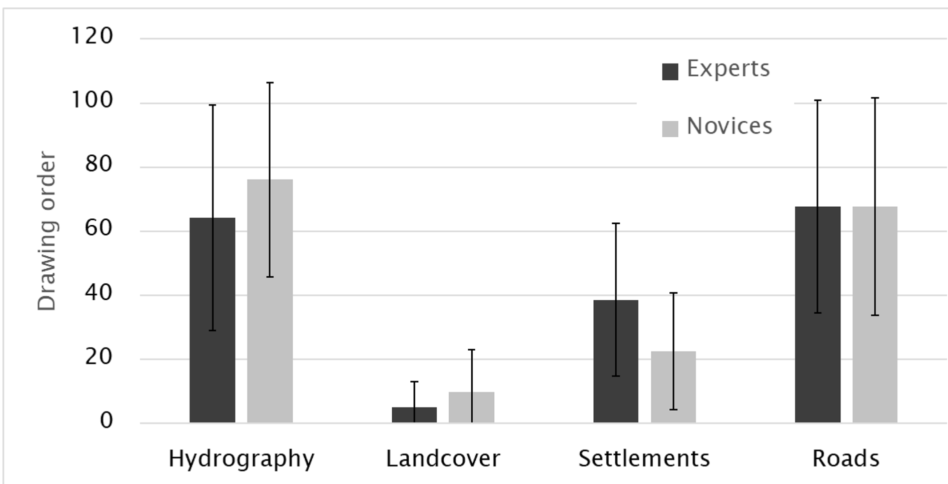


Figure 2.7. Scores for drawing order (Error bars indicate SD)

The fact that both experts and novices drew linear objects (hydrography and roads) first can be explained by the hierarchical structures of schemas in LTM. This fact gives a clear idea that the sketch maps are hierarchically constructed. This finding corresponds to what Huynh and Doherty (2007), Huynh, Hall, Doherty & Smith (2008) and Ooms et al. (2015) found. They discovered that participants start drawing their sketch maps with the main linear structures and continue with other landmarks. Furthermore, female participants started with hydrographic objects, while male participants chose roads in the first place. Accordingly, both females and males drew settlements in the third place, and land-cover objects in the fourth place.

2.4.2.2 Score on drawn elements

The presence and accuracy

A Kolmogorov-Smirnov test was used to test for normality on presence and accuracy, $D(54) = 0.090$, $p= 0.200$ indicated that the data was normally distributed. Based on the average scores of all participants, the average location score was 41.3 (*N=54, MED= 43.3, SD= 14.9*). Experts placed map elements slightly more accurately than the novices did, but according to two-way ANOVA, no significant difference emerged, with $F(1,55)= 0.888$ and $p= 0.350$. The most pronounced performance difference between two groups occurred when placing the settlements (12.0%) (Figure 2.8). The reason for this finding could be explained by the amount, the complexity and the distribution of elements falling into this category. The original stimulus contained eight residential areas, which was the highest number of elements that a category held. Inherently, remembering all of them together with their positions would be harder, especially for novices, compared to other categories having fewer than eight elements. The more isolated the feature was, the more distinctive and easier to remember it became. Hence, the isolated settlements stood out more, and participants tended have higher probabilities of drawing them.

On the other hand, the presence and accuracy results favored females with a 6.3% difference. However, this difference was not statistically significant according to two-way ANOVA ($F(1,55)= 1.672$ and $p= 0.101$).

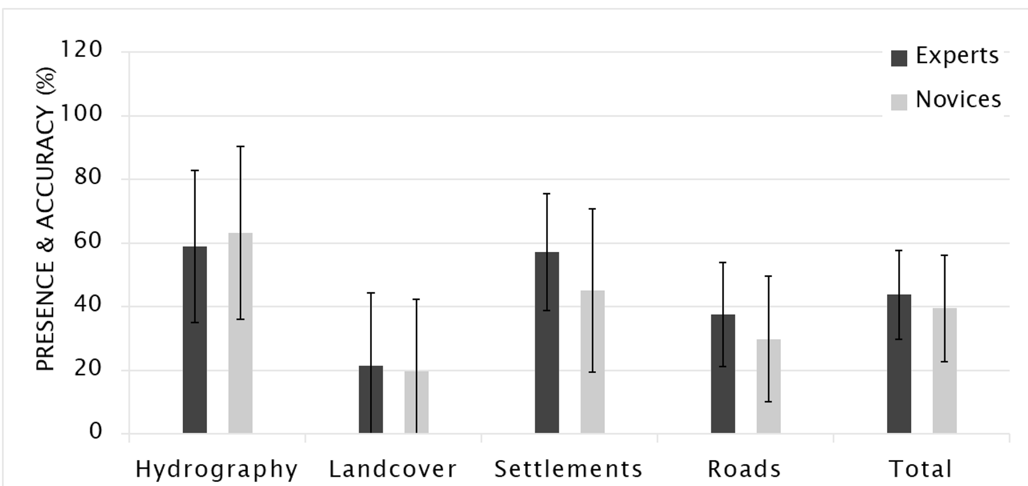


Figure 2.8. Presence and accuracy scores (Error bars indicate SD)

Shape, size and color

Based on average scores all participants, the average shape score was 82.1 (*N=54, MED= 83.3, SD= 12.9*). Figure 2.9 shows the shape scores for experts and novices based on the four main map element categories. A Kolmogorov-Smirnov test was used to test for normality on shape ($D(54) = 0.131$, $p= 0.022$), size ($D(54) = 0.144$, $p= 0.007$) and color ($D(54) = 0.309$, $p= 0.000$). The test results indicated that the data was not normally distributed. Experts illustrated the shape of the map elements 7.5% better than novices did, and Mann-Whitney U test showed that this difference was significant, with $U_{shape}= 247$ and $p= 0.044$. Similar to the results for presence and location, the greatest difference in performances between novices and experts occurred in settlements at 13.8%. On the other hand, female participants outperformed males with a 5.9% difference which was not significant according to Mann-Whitney U test (U_{shape} = 249, p= 0.077).

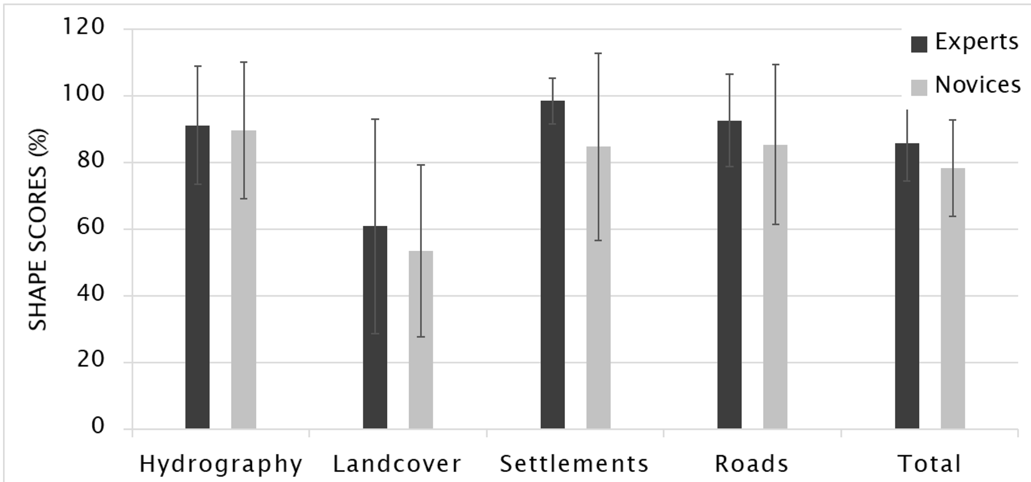


Figure 2.9. Shape scores (Error bars indicate standard deviation)

Size is one of the most effective visual variables in terms of its selectiveness, associativity, and ease of perception as ordered. Larger elements can be perceived immediately compared to smaller ones. To score the size of drawn elements, the relative sizes on the sketch maps were considered. If the size of an element was in line with the size of its surrounding elements, it was accepted as a correct size depiction. Based on the average scores of all participants, the average shape score was 82.7 (*N=54, MED= 83.3, SD= 12.2*). Accordingly, experts drew map elements 7.8% better than novices did considering their size, and based on Mann-Whitney U test, the size scores, with U_{size}= 244.5 and p= 0.040. The greatest difference occurred for settlements (14.3%) (Figure 2.10). A possible explanation could be that the depiction of settlements requires higher-level generalization knowledge. Since individual buildings come together to form a settlement or village, aggregation is needed to define a group of buildings as a settlement. On the other hand, no significant gender difference emerged, according to Mann-Whitney U test (U_{size} = 283.5, p= 0.254).

During the drawing process, participants did not receive any information about using colors. However, the color palette embedded in MS Paint was available to all participants. Other than three novice and five expert participants who chose to use only black, the remaining participants delivered colored sketch maps. Our color assessment criteria regarded the color correspondence of an element drawn on the sketch map with the one on the original map. We also paid attention to whether the elements drawn in the same color represent the same category. Based on the average scores of all participants, the average color score was 75.0 (*N*=54, *MED*= 83.3, *SD*= 31.2). Novices depicted the map elements slightly better using corresponding colors. However, this surprising difference between novices and experts was not statistically significant regarding to Mann-Whitney U test (U_{color}= 342.5 and p= 0.753). The greatest difference in performance was in hydrology (14.9%) (Figure 2.11). This result can be related to missing map elements on the sketch maps (since we assigned a score of zero to absent elements) or to the fact that some experts did not prefer to use color.

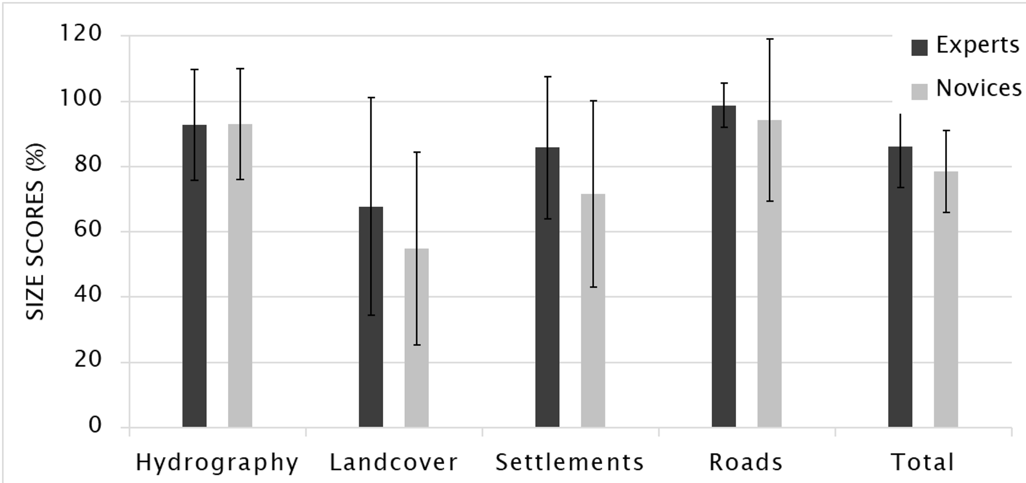


Figure 2.10. Size scores (Error bars indicate standard deviation)

Although women were superior to men for the depiction of colors with 1.7% performance difference, no significant difference occurred among these two groups (U_{color}= 342 and p= 0.934).

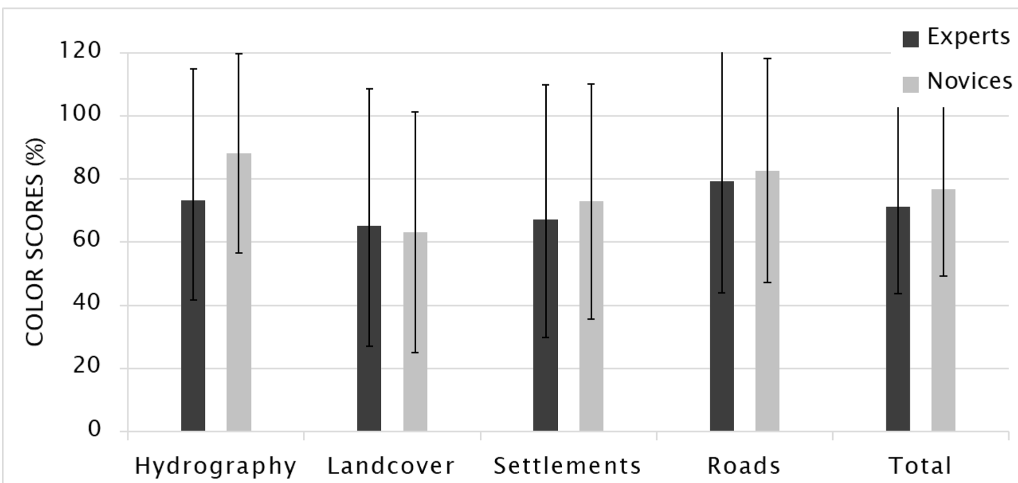


Figure 2.11. Color scores (Error bars indicate standard deviation)

Figure 2.12 shows the performances of experts and novices based on shape, size, color, and presence & location. We clearly see that the lowest overall performances for both groups occurred for presence & location. This result proves that drawing a map element in the correct location was more difficult than describing its shape, size, and color. The sample size was not sufficient to study the differences of four groups; expert males (N= 11), expert females (N= 13), novice males (N= 24) and novice females (N= 7).

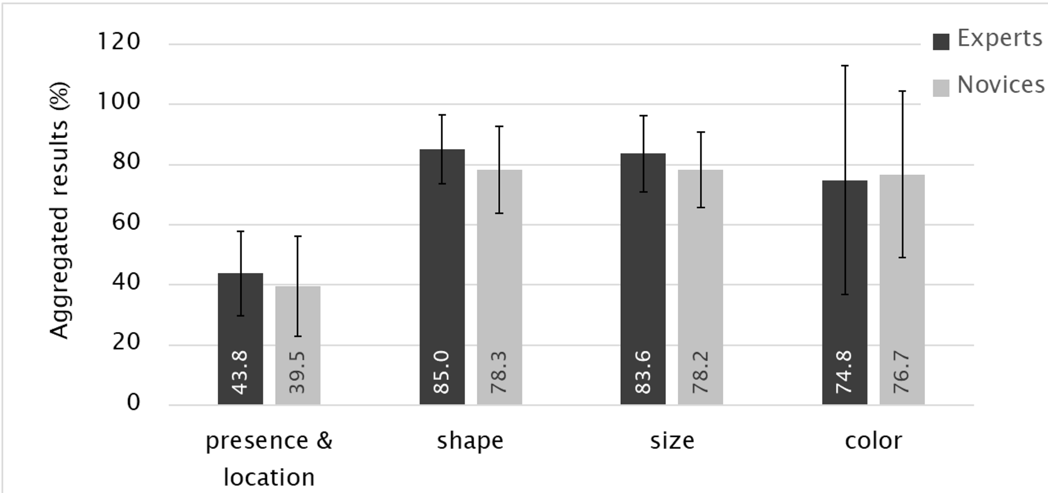


Figure 2.12. Summary of performances (Error bars indicate standard deviation)

2.4.2.3. Aggregation presence & accuracy (1), shape (2), size (3) & color (4).

A Kolmogorov-Smirnov test was used to test for normality on the aggregated scores (*D(54)*= 0.126, *p*= 0.027) and the test results indicated that the data was not normally distributed. According to the aggregated analysis, the average score of experts was 71.8 *(N= 24, MED= 76.8, SD= 19.2*) with a minimum of 39.9 and a maximum of 92.8, whereas it was 68.2 *(N= 30, MED= 68.8, SD= 11.1*) with a minimum of 36.3 and a maximum of 92.2 for novices. The difference of 3.6% on expertise was not statistically significant regarding to Mann-Whitney U test (*U*= 254, *p*= 0.065). The results implied that experts and novices showed no difference in map learning, unless the stimulus required specific map knowledge that only an expert possessed. In terms average scores of drawn elements, we found no significant interaction between expertise and gender (*F (1,50)*= 0.197, *p*= 0.659).

The average score of females was 73.2 (*N=20, MED= 75.7, SD= 14.5*) with a minimum of 39.9 and a maximum of 92.8, whereas it was 68.7 (*N=34, MED= 70.3, SD= 2.2*) with a minimum of 36.3 and a maximum of 92.2 for males. The difference among genders was not statistically significant regarding to Mann-Whitney U test (*U*= 264.5, *p*= 0.146). Although it was not possible to make generalized assumptions or draw conclusions regarding to gender differences between experts and novices as explained earlier, the results showed that both expert and novice females were favored in their groups. Expert females were the most successful group overall with a score of 74.2. Novice females (69.9), then expert males (69.3) and lastly novice males (66.5) followed them.

2.4.3 Eye Tracking

While studying the map, the average duration of the fixations was 230.0 ms *(N= 24, MED= 230.8 ms, SD= 50.1 ms)* for experts and 244.1 ms *(N= 30, MED= 243.0 ms, SD= 48.4 ms)* for novices. These values were 234.0 ms *(N= 20, MED= 239.8 ms, SD= 56.3 ms)* for females, and 240.1 ms *(N= 34, MED= 232.8 ms, SD= 45.3 ms)* for males. A Kolmogorov-Smirnov test was used to test for normality on the average duration of the fixations indicated that the data was normally distributed: *D(54)*= 0.082, *p*= 0.200.

The average duration of fixations for novices was slightly higher than it was for experts, whereas only slight differences emerged between the expert and novice groups and between females and males. However, according to two-way ANOVA, no significant difference was found ($F(1,5)$= 0.074, p= 0.787) between experts and novices, as well as between females and males ($F(1,55)$= 1.001, p= 0.322). Further, Cohen's effect size value (d= 0.09) suggested that the effect was rather small for expertise (d= -0.123) and gender (d= -0.289). No significant interaction effect was observed between expertise and gender ($F(1,50)$= 0.251, p= 0.619).

The average number of fixations per second for the stimulus was 3.5 *(N= 24, MED= 3.7, SD= 1.0)* for experts and 3.6 *(N= 30, MED= 3.6, SD= 0.5)* for novices. These values were 3.4 *(N= 20, MED= 3.4, SD= 1.1)* for females, and 3.7 *(N= 34, MED= 3.7, SD= 0.5)* for males. A Kolmogorov-Smirnov test was used to test for normality on the number of fixations per second indicated that the data did not fit normal distribution: D(54) = 0.145, p= 0.007. The average number of fixation of novices and experts slightly differed, as well as it did for females and males. Regarding to Mann-Whitney U test, the differences emerged neither from expertise, nor from gender were statistically significant ($U_{expertise}$= 338, p= 0.702; U_{gender}= 254, p= 0.123). No significant interaction effect occurred between expertise and gender ($F (1,50)$= 0.286, p= 0.595).

Having visually inspected, we observed that the gaze behaviors of all participants depicted in the focus map clearly reflect the main structural elements of the map stimulus (Figure 2.13). When visually interpreted, the focus map highlighted the main road construction, water bodies and large settlements belonging to the stimulus. The river located in the upper side of the map especially stood out. This result proves why the hydrography was the most remembered category with the highest score in drawing order. Furthermore, forests located on the bottom-right of the map look almost dark, which proves that the participants showed less interest in this part of the map. This finding supports the fact that the land-cover was the least drawn category (see results for drawing order) and also corresponds to what was registered by Ooms et al., 2014a. Therefore, we could use the proposed AoI around the main structuring elements on the map.

The AoIs considered for the further analysis include all three main roads and hydrographic elements, which are aggregated as a single object, four settlements, and one land-cover object as depicted in Figure 2.14. Road 1 with the tilted Y-shape is located in the lower center of the map and forms the longest road feature. The largest settlement is the one located in the upper center of the map (Settlement 1), whereas another fundamental linear feature, the hydrography, covers the upper side of the map.

The time to the first fixation reflects that the larger objects and the objects located in the upper middle of the screen caught a participant's attention earlier than the others did. Both experts and novices gazed at Settlement 1 first (350.7 ms for experts, 49.4 ms for novices), Road 1 second (3463.0 ms for experts, 3162.2 ms for novices) and Hydrography third (3821.1 ms for experts, 4455.2 ms for novices). The longest time to the first fixation was spent for the land-cover object (24976.8 ms for experts, 29863.9 ms for novices) that is located in the bottom-center of the map and has a relatively smaller size.

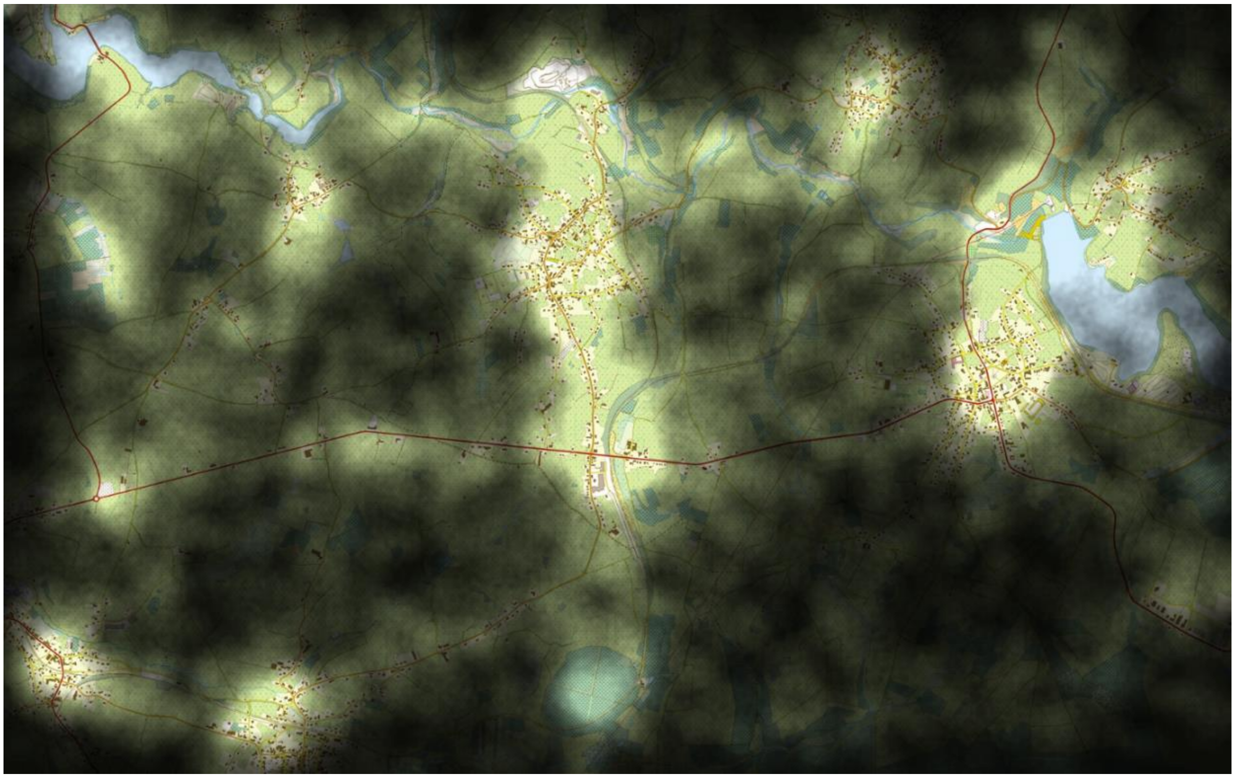


Figure 2.13. Focus map of all 54 participants

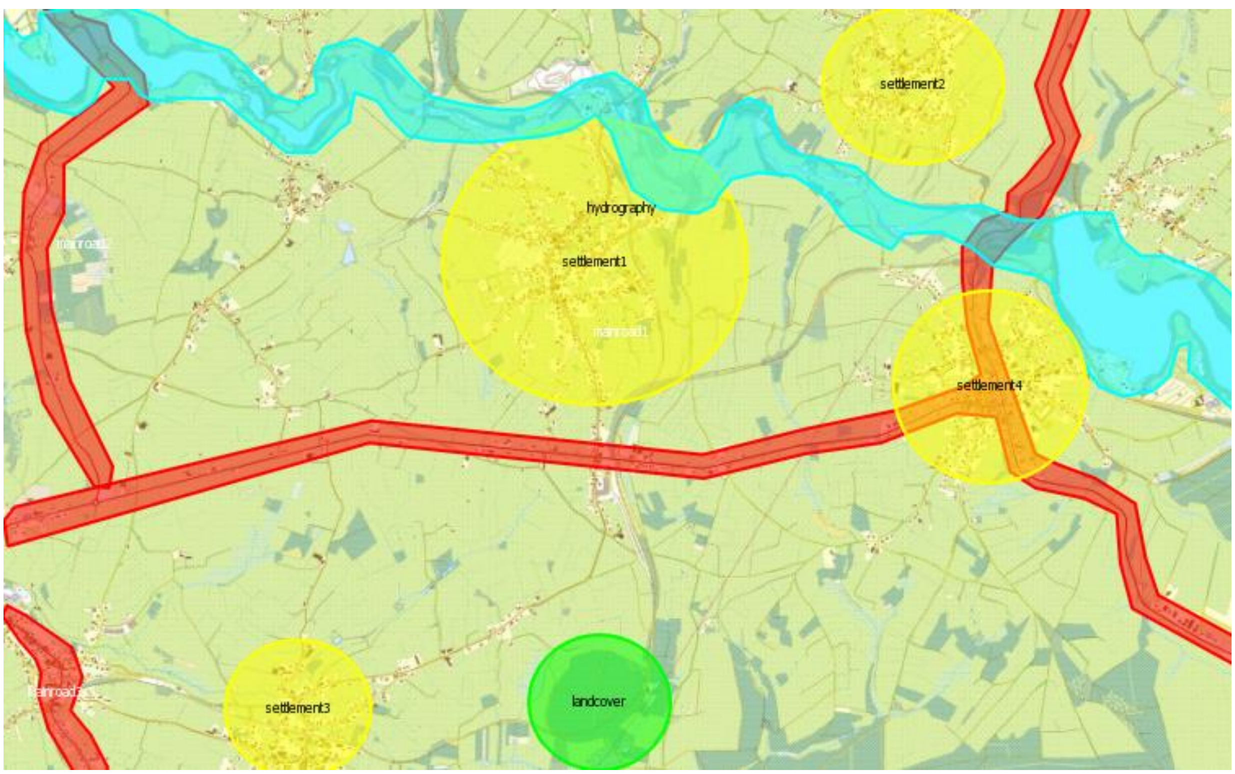


Figure 2.14. Selected AoIs

The dwell times of participants for all AoIs showed that there was similar behavior between experts and novices. The dwell times of experts were higher for Hydrography, whereas novices spent more time for Settlement 1. Both group spent less time for Roads 2 and 3, approximately 1/10 of what they spent for Hydrography and Settlement 1.

On the other hand, the number of fixations within AoIs was slightly higher for experts. Hydrography received the highest fixation counts with 57.4 for experts and 46.5 for novices. The next highest numbers of fixations occurred for Settlement 1 and Road 1 (Figure 2.15). These map elements also resulted in longer dwell times. The fixation count was closely linked to the time a participant spent for a certain region (dwell time). Therefore, the number of fixations per second is a more objective measure to explore differences between experts and novices.

The average fixation durations of participants were higher for all settlements (except Settlement 2) and Road 1 regardless of the expertise. Settlement 3 received the highest average fixation duration, whereas Road 2 received the lowest (Figure 2.16). Although both objects have relatively small sizes, participants seemed to have different reasons why they fixated on those objects for longer or shorter periods. The complexity of the object mostly resulted in higher fixation durations. In this case, the settlement was a more elaborate object compared to the road and required more processing time and thus, more cognitive load. Furthermore, our results proved that the fixation duration and the number of fixations were inversely proportional. The shorter the fixation duration was, the higher the number of fixations per second. For instance, Settlement 3 had the longest average fixation duration (287.0 ms, see Figure 2.16), while it received a lower number of fixations per second (i.e. 3.7, see Figure 2.15) than the other objects did.

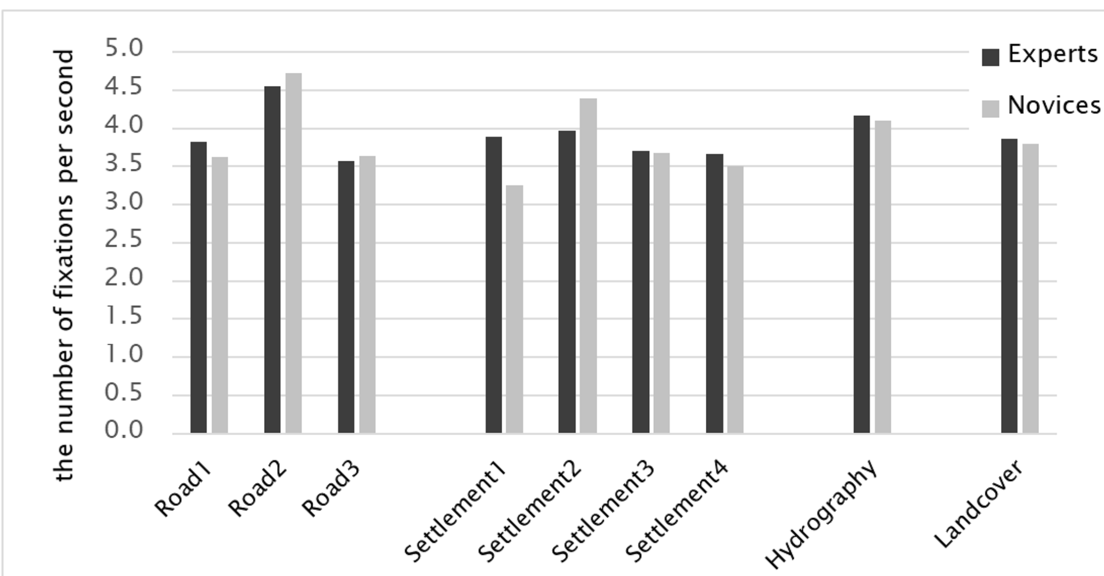


Figure 2.15. The number of fixations per second

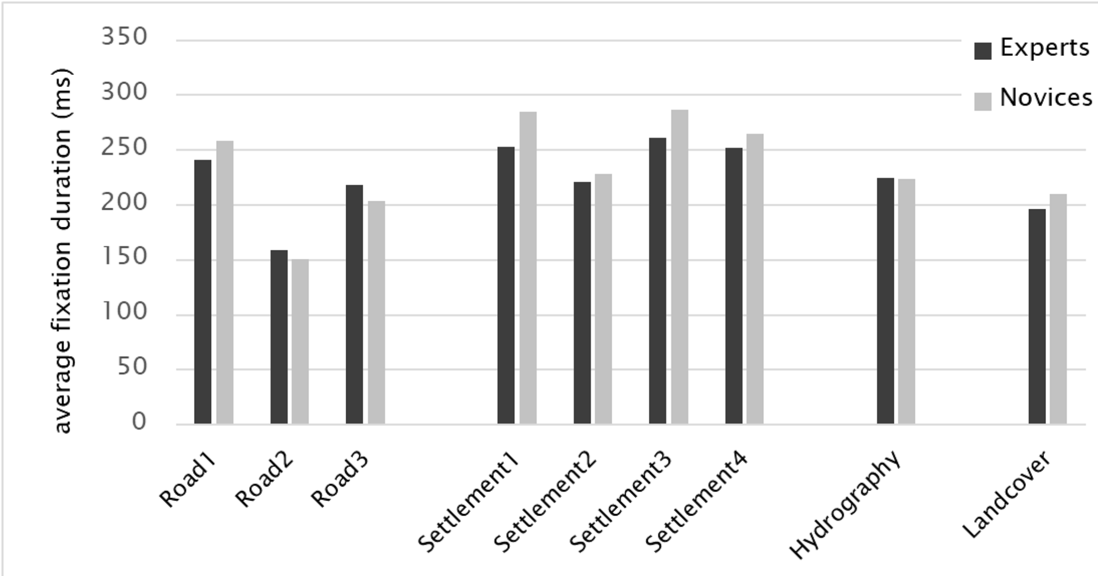


Figure 2.16. Average fixation duration (dark bars: experts, light bars novices)

2.5. Discussion

The results of the study are valid for a specific map stimulus representing only one specific area. However, the map, which was simplified only by removing altitude lines and labels, is a part of a map series covering the whole territory of Belgium. Therefore, the same trends could be observed on all these maps as they are based on the same symbology, although the generalization of results is limited. Although between- subjects design provided some potentially valuable insight, the outcomes may not apply for every condition. The performance of individuals is mainly influenced by the task and stimulus because the cognitive load can be manipulated by the complexity of the visual material and the difficulty of tasks. Therefore, if this study is extended by including other types of map stimulus and tasks, different results might be obtained.

The memory task in the experiment required recalling the main structural elements of a screen map. This retrieval act involved WM-LTM transitions, such as retrieval of spatial information stored in WM through LTM or strategies for constructing hierarchy among map elements.

We regarded visual variables such as location, shape, size and color as though they were equally important for the drawing order which can be influenced by the use of visual variables. Besides other visual variables, color has long been recognized as a preattentive feature (Wolfe, 2000). The order of drawing varied between participants, so that experts drew roads (depicted as red) in first place, whereas novices drew the elements hydrography (depicted as blue). Same situation applies for female and male participants, respectively.

In the original stimulus, roads were linear objects depicted in red, whereas hydrographic objects could be linear (rivers) or areal (water bodies) representations depicted in blue. Our retina includes light -sensitive cells named rods and cones. While rods mediate night vision, cones play role in photopic vision (during daylight) (Hsia & Graham, 1952). The spectral sensitivity of cones follows the order of the visual spectrum. Therefore, our eyes perceive the most in red wavelengths (500-760 nm) and the least on blue wavelengths

(380-550 nm), and green wavelengths (430-673 nm) fall under the red range (Schubert, 2006). To the best of our knowledge, in map design, red tends to focus in the foreground; yellow and green, in the middle; and blue, in the background (NRCan, n.d.). Thus, important objects or the ones to emphasize are shown in red, and blue is a good color for backgrounds. This feature could be the reason why the experts drew the red linear objects (roads) first. On the other hand, having drawn the hydrographic elements first, novices might have found areal objects as important or interesting and thus as memorable as linear objects. We can infer that size is as important as color for the retrieval of an object. Except for one participant, all novices drew water bodies on their sketch maps regardless of the order. Therefore, it is suggested that experts and novices use different strategies in spatial orientation, as well as females and males do. For instance, men tend to refer to environmental geometries or structuring elements, while women rely on landmarks (Sandstrom, Kaufman & Huettel, 1998; Voyer et al., 2007). However, the common characteristic of the first drawn elements by all participants was that they both contained linear objects. This finding referring that the structuring elements guide spatial recognition is in line with what Edler et al. (2014) and Ooms et al. (2015) found. Additionally, the hydrography category included lakes, which were areal representations. Starting with the areal elements instead of linear ones (or in our case, polygons (lakes) and lines (rivers) that were parts of a whole (hydrography)) proves that size of an object also plays an important role when recalling map information.

Based on the assessment of sketch maps considering the aggregated analysis of presence & location, shape, size, and color of drawn elements, we concluded that neither expertise, nor gender differences were influential on the retrieval of spatial information. Our findings related to gender differences corresponds to those by Lloyd & Steinke (1984), Patton & Slocum (1985), Beatty & Bruellman (1987), Golledge, Dougherty & Bell (1995), Lloyd & Bunch (2010) and Edler et al. (2014). On the other hand, our findings on the influence of expertise agree with the earlier research of Thorndyke & Stasz (1980) who focused on experts' and novices' abilities to learn and remember information presented via maps. The fact that novices and experts did not differ in terms of how they learned and remembered map-related information could be explained by the general map knowledge that stepped in when both user groups observed a typical planimetric map stimulus. Hence, various levels of map experience may have resulted in modest differences (Kulhavy & Stock, 1996). The original map shown to participants was a simplified 1:10k topographic map and did not contain any familiar places (or names) to eliminate or minimize the degree of familiarity. Thus, both experts and novices observed the map for the first time, and we presumed that the maplikeness of the stimulus had a great influence on their map learning (study and recall) process. However, the later work (e.g. Gilhooly, Wood, Kinnear & Green, 1988; Ooms et al., 2015) failed to replicate Thorndyke & Stasz's (1980) findings. Instead, they found that experts performed better in recalling schemas in a richer and more detailed fashion. Although our results present that experts and novices do not differ in terms of the amount of information they recall, the learning/recalling strategies of experts and novices may differ. The drawing order results could be evidence that they might use different approaches.

In addition to the maplikeness and the simplicity of the map, the task to be executed was influential on performance. It is important to remember that if the task required domain-specific knowledge about geography or related areas, experienced users would perform better compared to novices (Kulhavy &

Stock, 1996; Thorndyke & Stasz, 1980). Although individual factors other than expertise and gender might have affected the results, the sample size was not sufficient to draw conclusions regarding ethnicity or native language.

While encoding spatial information through maps, structuring elements (e.g. topographic details and grid lines) lead attentional shifts towards “to-be-learned object locations” which improve memory performance. The fact that the first fixation is influenced by experimental manipulations can be seen during recognition and it suggests that the structuring elements are involved in cognitive map production (Kuchinke et al., 2016). Therefore, eye tracking metrics provided valuable insight on how mental representations formed. In this context, average fixation duration and the number of fixations per second revealed that there was no significant difference between the expert and novice groups, as well as between men and women. Although this outcome was different from what was found by Ooms et al. (2014a), it supports our results obtained by digital sketch map assessment.

In addition, the eye tracking metrics (time to first fixation, dwell time, fixation count, the number of fixations per second, average fixation duration) for selected AoIs were explored. The time to first fixation statistics showed that larger AoIs were gazed at earliest and the dwell times for such objects were much longer compared to those for other AoIs. As expected, the majority of participants drew these map elements on their sketch maps. On the other hand, most participants paid less attention (late first fixation and less dwell time) to the relatively small linear (i.e. roads) and areal features (i.e. land cover) within the specified AoIs. However, when comparing the presence and accuracy scores of drawn elements, both groups mostly drew small roads on their sketch maps but not land-cover features. We could infer from this result that the linear features were easier to learn and remember, although the viewer did not pay much attention. Additionally, our results supported the fact that shorter fixation durations resulted in higher numbers of fixations per second. Consequently, longer average fixation durations for a specific AoI indicated that the chances were higher to remember that object. This finding corresponded to the number of objects depicted on the sketch maps; the objects that were absent on the sketch map received the shortest fixation durations during the study phase. However, longer fixation durations may also indicate participants’ difficulty to recognize the information in the observed visual scene.

Although it was beyond the scope of this study, the sequence of visited AoIs can be further explored to analyze how the map elements within specified AoIs are associated to form a sketch map. The sequential order of included elements may vary among individuals who draw sketch maps of the same map stimulus and sequence analysis can provide more insightful outcomes related to how map users encode structure, learn, remember and later use the spatial information presented via maps (e.g. Huynh, et al., 2008). Furthermore, the similarity between sequences can be studied by quantifying and comparing scanpath behaviors of individuals. Scanpath analysis promises rich information regarding to spatial and temporal characteristics of eye movements and contributes to understanding individual differences in a more systematic way (e.g. Anderson, Anderson, Kingstone & Bischof, 2014; Dolezalova & Popelka, 2016).

2.6. Conclusion

This study utilizes digital sketch maps to understand the cognitive abilities and limitations of a group of map users during a memory task via drawing. On one hand, we assessed the quality of sketch maps based on the drawn elements (e.g. the influence of visual variables), which we predicted would reflect the performances of different user groups and might reveal significant insights about their cognitive processes and strategies of retrieving spatial information. On the other hand, we integrated ET statistics to quantify the cognitive processes to advance time-related, gaze activity-related (especially fixations) analyses. We also derived the order in which the sketched objects were drawn from the ET data. The order of drawing offered significant insight into the hierarchical construction of cognitive maps and might have unveiled the differences in the retrieval strategies of experts and novices, if there were any.

Instead of traditionally used pen and paper method, we collected sketch maps digitally to be able to match them with the corresponding eye tracking metrics. Therefore, ET and sketch map were considered as complementary user testing methods providing detailed insight into user behaviors. No significant differences emerged between experts and novices, as well as females and males based on sketch map analyses, and this result was also confirmed by a number of ET statistics. This finding arose from a user experiment that considered a simplified static map for a memory task related to the map elements. However, this research can be extended by considering more rapidly evolving cartographic stimuli (3D visualizations, interactive displays, mobile maps, etc.) and tasks that require different levels of expertise to achieve a better understanding of map users. The more we understand the cognitive limits and abilities of map users, the more we become able to create effective cartographic products.

References

Anderson, N. C., Anderson, F., Kingstone, A., & Bischof, W. F. (2014). A comparison of scanpath comparison methods. Behavior Research Methods, 47(4), 1377-1392.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes1. Psychology of learning and motivation, 2(1), 89-195.

Ball, L. J., Lucas, E. J., Miles, J. N., & Gale, A. G. (2008). Inspection times and the selection task: What do eye-movements reveal about relevance effects?. The Quarterly Journal of Experimental Psychology, 56(6), 1053-1077.

Bargiota, T., Mitropoulos, V., Krassanakis, V., & Nakos, B. (2013). Measuring locations of critical points along cartographic lines with eye movements. In M. F. Buchroithner, N. Prechtel, D. Burghardt, K. Pippig, & B. Schröter (Eds.), Proceedings of the 26th International Cartographic Association Conference, Dresden, Germany.

Beatty, W. W., & Bruellman, J. A. (1987). Absence of gender differences in memory for map learning. Bulletin of the Psychonomic Society, 25(4), 238-239.

Bell, S., & Archibald, J. (2011). Sketch mapping and geographic knowledge: what role for drawing ability. An interdisciplinary approach to understanding and processing sketch maps, 42(1), 1-10.

Bertin, J. (1967). Semiology of graphics, (W. J. Berg Trans.). Madison, WI, USA: University of Wisconsin Press.

Bertrand, C., & Thullier, F. (2009). Effects of player position task complexity in visual exploration behavior in soccer. International Journal of Sport Psychology, 40(2), 306-323.

Bestgen, A. K., Edler, D., Kuchinke, L., & Dickmann, F. (2016). Analyzing the effects of VGI-based landmarks on spatial memory and navigation performance. KI-Künstliche Intelligenz, 31(2), 179-183.

Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., & Ertl, T. (2014, June). State-of-the-art of visualization for eye tracking data. In Proceedings of EuroVis. Swansea, Wales, UK.

Billinghurst, M., & Weghorst, S. (1995, March). The use of sketch maps to measure cognitive maps of virtual environments. Proceedings of Virtual Reality Annual International Symposium, 40-47. 4912223,

Crundall, D., Underwood, G., & Chapman, P. (2002). Attending to the peripheral world while driving. Applied cognitive psychology, 16(4), 459-475.

Cooper, G. (1998). Research into cognitive load theory and instructional design at UNSW. 347-362.

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

Dickmann, F. (2013). City maps versus map-based navigation systems–an empirical approach to building mental representations. The Cartographic Journal, 49(1), 62-69.

Dickmann, F., Edler, D., Bestgen, A. K., & Kuchinke, L. (2016). Exploiting illusory grid lines for object-location memory performance in urban topographic maps. The Cartographic Journal, 54(3), 242-253.

Dolezalova, J., & Popelka, S. (2016). ScanGraph: A novel scanpath comparison method using graph cliques visualization. Journal of Eye Movement Research, 9(4), 13.

Downs, R. M., & Stea, D. (1977). Maps in minds: Reflections on cognitive mapping. New York: Harper & Row.

Duchowski, A. T. (2007). Eye tracking methodology - Theory and practice. London: Springer, 328(614), 2-3.

Edelman, G. M. (1992). Bright Air Briliant Fire: On the Matter of the Mind, London: Penguin Books.

Edler, D., Bestgen, A. K., Kuchinke, L., & Dickmann, F. (2014). Grids in topographic maps reduce distortions in the recall of learned object locations. Plos one, 9(5).

Fabrikant, S. I., Hespanha, S. R., & Hegarty, M. (2010). Cognitively inspired and perceptually salient graphic displays for efficient spatial inference making. Annals of the Association of American Geographers, 100(1), 13-29.

Forbus, K. D., Usher, J., & Chapman, V. (2004). Qualitative spatial reasoning about sketch maps. AI magazine, 25(3), 61-72.

Gilhooly, K. J., Wood, M., Kinnear, P. R., & Green, C. (1988). Skill in map reading and memory for maps. Quarterly Journal of Experimental Psychology, 40(1), 87-1 07.

Golledge, R. G., & Spector, A. N. (1978). Comprehending the urban environment: Theory and practice. Geographical Analysis, 10(4), 403-426.

Golledge, R. G., Dougherty, V., & Bell, S. (1995). Acquiring spatial knowledge: Survey versus route-based knowledge in unfamiliar environments. Annals of the association of American geographers, 85(1), 134-158.

Hafting, T., Fyhn, M., Molden, S., Moser, M. B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. Nature, 436(7052), 801-806.

Haken, H. (1977). Synergetics. Physics Bulletin, 28(9), 412.

Haken, H., & Portugali, J. (1996). Synergetics, inter-representation networks and cognitive maps. The construction of cognitive maps, 32, 45-673.

Haken, H., & Portugali, J. (2016). Information and Self-organisation. Entropy, 19(1), 18.

Henderson, J. M., & Ferreira, F. (2004). Scene Perception for Psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), The interface of language, vision, and action: Eye movements and the visual world, 1-58. New York, NY, US: Psychology Press.

Herbert, G., & Chen, X. (2015). A comparison of usefulness of 2D and 3D representations of urban planning. Cartography and Geographic Information Science, 42(1), 22-32.

Hsia, Y., & Graham, C. H. (1952). Spectral sensitivity of the cones in the dark adapted human eye. Proceedings of the National Academy of Sciences, 38(1), 80-85.

Huynh, N. T., & Doherty, S. T. (2007). 'Digital sketch map drawing as an instrument to collect data about spatial cognition', Cartographica, 42, 285–296.

Huynh, N. T., Hall, G. B., Doherty, S., & Smith, W. W. (2008). Interpreting urban space through cognitive map sketching and sequence analysis. The Canadian Geographer/Le Géographe canadien, 52(2), 222-240.

Intons-Peterson, M. J., & Mcdaniel, M. A. (1991). Symmetries and asymmetries between imagery and perception. Imagery and Cognition, 47-76.

Jacob, R. J., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In: R. Radach, J. Hyona & H. Deubel (Eds.), The mind's eye: Cognitive and applied aspects of eye movement research(573-605). Amsterdam: Elsevier.

Kettunen, P., Putto, K., Gyselinck, V., Krause, C. M., & Sarjakoski, L. T. (2015) Perception and recall of landmarks for personal navigation in nature at night versus day. In J. Brus, A. Vondrakova, V. Vozenilek (Eds.), Modern Trends in Cartography. Lecture Notes in Geoinformation and Cartography, 281-301, Springer, Cham.

Kimerling, A. J., Buckley, A. R., Muehrcke, P. C., & Muehrcke, J. O. (2009). Map use: reading and analysis. Esri Press.

Kuchinke, L., Dickmann, F., Edler, D., Bordewieck, M., & Bestgen, A. K. (2016). The processing and integration of map elements during a recognition memory task is mirrored in eye-movement patterns. Journal of Environmental Psychology, 47, 213-222.

Kulhavy, R. W., Schwartz, N. H.,& Shaha, S. N. (1983). Spatial Representation of Maps. American Journal of Psychology, 96(3), 337-351.

Kulhavy, R. W., Stock, W. A., & Kealy, W. A. (1993a). How geographic maps increase recall of instructional text. Educational Technology Research and Development, 41(4), 47-62.

Kulhavy, R. W., Stock, W. A., Woodard, T. A., & Haygood, R. C. (1993b). Comparing elaboration and dual coding theories: The case of maps and text. American Journal of Psychology, 106(4), 483-498.

Kulhavy, R. W., & Stock, W. A. (1996). How cognitive maps are learned and remembered. Annals of the Association of American Geographers, 86(1), 123-145.

Kveladze, I., Kraak, M., & Van Elzakker, C. P. (2017). The usability of a GeoVisual analytics environment for the exploration and analysis of different datasets. Advances in Cartography and GIScience Lecture Notes in Geoinformation and Cartography, 207-221. Springer, Cham.

Lakoff, G. (1987). Women, fire and dangerous things. Retrieved from https://the-eye.eu/public/Books/qt.vidyagam.es/library/Computational%20Linguistics/Semantics/Lakoff-1987-Women-Fire-and-Dangerous-Things.pdf

Lloyd, R., & Steinke, T. (1984). Recognition of disoriented maps: The cognitive process. Cartographic Journal, 21(1), 55-59.

Lloyd, R. E., & Bunch, R. L. (2010). Learning geographic information from a map and text: learning environment and individual differences. Cartographica: The International Journal for Geographic Information and Geovisualization, 45(3), 169-184.

Lynch, K. (1960). Image of the city. Cambridge, MA: Technology Press/ Harvard University Press.

MacEachren, A. M. (2004). How maps work: representation, visualization, and design. New York, NY: Guilford Press.

Montello, D. R., Sullivan, C. N., & Pick, H. L. (1994). Recall memory for topographic maps and natural terrain: effects of experience and task performance. Cartographica, 31(3), 18–36.

NRCan (Natural Resources Canada). (n.d.). Cartography [lecture notes] Retrieved February 23, 2017 from http://atlas.gc.ca/english/carto

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Research, 34, 171-175.

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. Experimental Neurology, 51(1), 78–109.

O'Keefe, J., & Nadel, L. (1978). The Hippocampus as a Cognitive Map. Retrieved from https://repository.arizona.edu/handle/10150/620894

Ooms, K. (2012). Maps, how do users see them?: an in depth investigation of the map users' cognitive processes. Doctoral dissertation, Ghent University, Ghent. Retrieved from https://biblio.ugent.be/publication/3103344

Ooms, K., De Maeyer, P., & Fack, V. (2014a). Study of the attentional behavior of novice and expert map users using eye tracking. Cartography and Geographic Information Science, 41(1), 37-54.

Ooms, K., Coltekin, A., De Maeyer, P., Dupont, L., Fabrikant, S., Incoul, A., & Van der Haegen, L. (2014b). Combining user logging with eye tracking for interactive and dynamic applications. Behavior Research Methods, 47(4), 977-993.

Ooms, K., & De Maeyer, P. (2015, August). Georeferencing Eye tracking Data on Interactive Cartographic Products. In C. R. Sluter, C. B. M. Cruz, S. P. Camboim, L. S. Delazari, M. D. C. Fernandes, R. S. D. Barros, H. Firkowski, & P. M. L., De Menezes (Eds.), Proceedings of 27th International Cartographic Conference, Rio de Janeiro, Brazil. Retrieved from https://icaci.org/files/documents/ICC_proceedings/ICC2015/papers/13/fullpaper/T13-35_1427704736.pdf

Ooms, K., De Maeyer, P., & Fack, V. (2015). Listen to the map user: cognition, memory, and expertise. The Cartographic Journal, 52(1), 3-19,

Ooms, K., De Maeyer, P., Dupont, L., Van Der Veken, N., Van de Weghe, N., & Verplaetse, S. (2016). Education in cartography: What is the status of young people's map-reading skills?. Cartography and Geographic Information Science, 43(2), 134-153.

Ooms, K. (2016, June). Cartographic User Research in the 21st Century: Mixing and Interacting. In T. Bandrov, & M. Konecny, (Eds.), 6th International Conference on Cartography and GIS Proceedings (p. 367). Albena, Bulgaria:.Bulgarian Cartographic Association. Retrieved from http://iccgis2016.cartography-gis.com/6iccgis-2016-proceedings/

Ooms, K., Dupont, L., & Lapon, L. (2017). Mixing methods and triangulating results to study the influence of panning on map users' attentive behaviour. The Cartographic Journal, 54(3), 1-18.

Patton, J. C., & Slocum, T. A. (1985). Spatial pattern recall/an analysis of the aesthetic use of color. Cartographica 22(3), 70-87.

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. Encyclopedia of human computer interaction, 211-219.

Popelka, S., Stachoň, Z., Šašinka, Č., & Doležalová, J. (2016). EyeTribe tracker data accuracy evaluation and its interconnection with hypothesis software for cartographic purposes. Computational intelligence and neuroscience, 2016, 20.

Portugali, J. (1996). Inter-representation networks and cognitive maps. In J. Portugali (Eds.), The Construction of Cognitive Maps. GeoJournal Library, 32, 11-43.

Roth, R. E., Çöltekin, A., Delazari, L., Filho, H. F., Griffin, A., Hall, A., … & van Elzakker, C. P. (2017). User studies in cartography: opportunities for empirical research on interactive maps and visualizations. International Journal of Cartography, 3(sup1), 61-89.

Sandstrom, N. J., Kaufman, J., & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task[1]. Cognitive brain research, 6(4), 351-360.

Schriver, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. Human Factors, 50(6), 864-878.

Schubert, E. F. (2006). Human eye sensitivity and photometric quantities. Light-emitting diodes, 275-291. doi:10.1017/cbo9780511790546.017

Shettleworth, S. J. (2010). Cognition, evolution, and behavior. Oxford: Oxford University Press.

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

Simon, H. A. (1989). Models of thought. New Haven (Conn.): Yale University Press.

Stachoň, Z., Šašinka, Č., Štěrba, Z., Zbořil, J., Březinová, Š., & Švancara, J. (2013). Influence of graphic design of cartographic symbols on perception structure. Kartographische Nachrichten, 63(4), 216-220.

Thorndyke, P. W., & Stasz, C. (1980). Individual Differences in Procedures for Knowledge Acquisition from Maps. Cognitive Psychology, 12(1), 137- 175.

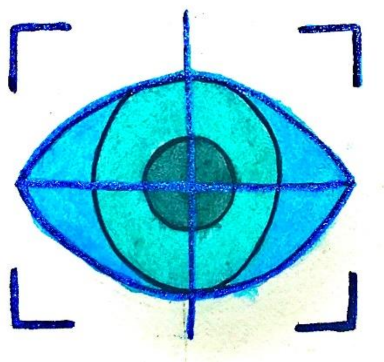
Tversky, B. (1992). Distortions in cognitive maps. Geoforum. 23(2), 131-138. doi:10.1016/0016-7185(92)90011-R

Voyer, D., Postma, A., Brake, B., & Imperato-McGinley, J. (2007). Gender differences in object location memory: A meta-analysis. Psychonomic bulletin & review, 14(1), 23-38.

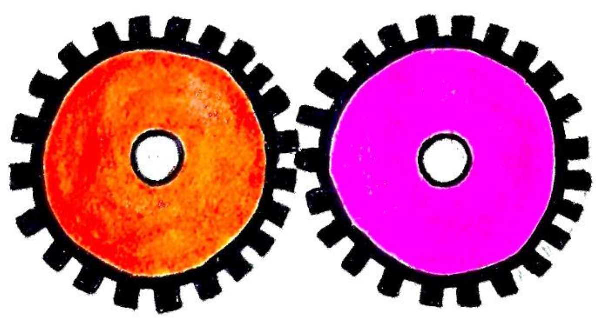
Wedel, M., & Pieters, R. (2008). Eye tracking for visual marketing. Foundations and Trends® in Marketing, 1(4), 231-320.

Wolfe, J. M. (2000). Chapter 8 - Visual Attention, In Karen K. De Valois (Eds.), Handbook of Perception and Cognition, Seeing, 335-386. San Diego, CA: Academic Press.

Chapter 3:
The Design of Experiment 2



"



"A well-designed experiment is likely to produce meaningful and interpretable results that have implications for theories and may inspire new research, even if the data are noisy and only basic analyses are performed"

Mike X Cohen, 2014

Abstract. *Designing a user experiment with ET and EEG is not straightforward, because of the methodological and technical issues this integration causes. The methodological issues are related to the experiment design, including research goals, participants, task and stimuli, psychological measures to use, evaluation methods and possible analyses for collected data, whereas the technical concerns refer to the synchronization of ET and EEG recording systems, their accuracy and quality, and numerous data processing steps. In this chapter, both issues will be explained and discussed within the frame of cartographic user experiment design. The issues mentioned in this chapter and lessons learned in the first experiment were incorporated in the design of the second experiment.*

Author Contributions*: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing-Original Draft Preparation: Merve Keskin; Resources: Kristien Ooms, Philippe De Maeyer*

3.1. Methodological issues while designing an EEG & ET user experiment in a cartographic context

Although there exists quite good amount of literature on EEG and ET experiments in psychology or related domains (e.g. Dimigen, Sommer, Hohlfeld, Jacobs & Kliegl, 2011; Nikolaev, Meghanathan & van Leeuwen, 2016), EEG is quite new for cartographic user research and besides psychological experiment design rules, there are many other factors that should be taken into account when maps come into play. Therefore, it is essential to mention methodological issues, substantially based on hands-on experience, for researchers who aim to study the individual differences of map users by employing ET and EEG. In this context, we would like to present how an ET-EEG experiment set up should be done and what type of analysis is possible relevant to the experiment design with a special focus on a spatial memory task related to digital 2D static maps. The referred task is designed to evaluate the cognitive strategies of expert and novice participants when they are asked to memorize and then remember a (part of) map content. This chapter aims to explain and discuss the methodology to be used for designing an ET & EEG experiment, rather than answering the research questions, which are covered in the following chapters within the actual experiments, such as:

- How do cognitive strategies differ for experts and novices for memorizing map-related information? How does cognitive load vary across those two groups during this process?
- How does the complexity of task influence cognitive load while memorizing map content?

The methodological issues are situated in many aspects of the experimental design and its set-up, which includes identifying the research goals, participants, task and stimuli, psychological measures to use, evaluation methods and possible analyses of the collected data. Within in this section, a theoretical overview covering the fundamentals of the experimental design is presented regarding to the best practices in the existing literature.

3.1.1 Theoretical background

A well-designed user experiment is essential to obtain meaningful and interpretable results of complex behavioral and cognitive processes, regardless of the quality of the collected EEG or ET data (Cohen, 2014). Even though the purpose of the experiment is established in a cartographic context to assess the cartographic content and the users trying to interpret it, the experimenter should primarily consider the fundamental design principles of psychological experiments (e.g. empiricism, determinism) for the study to be reliable (Gregory, 2004; Urbina, 2014). A psychological test is conducted "to evaluate individual differences or variations among individuals" (Kaplan & Saccuzzo, 2017, p.9). The most crucial step in psychological testing is to determine a clear research goal and related hypotheses and this in turn needs to be investigated with appropriate tools for data collection considering "their administration, scoring, interpretation, and the judicious use of the data collected to make inferences about the question at hand" (Urbina, 2014, p.25). In this context, it is important to verify beforehand that the identified research questions can indeed be answered by implementing (ET and) EEG in the experiment design. Due to its high temporal resolution and non-invasive application, EEG is more likely to explore the timing of cognitive processes and it is quite sensitive to capture the brain organization related to perception, attention, and memory (e.g. cognitive load). Therefore, the most common application of EEG is to study stimulus or event related response of the brain (for detailed overview see Teplan, 2002). Accordingly, with EEG, it is possible to answer our research questions focusing on the cognitive differences of experts and novices while they study map stimuli to remember their main structuring elements.

3.2.1.1. Within vs. between user study design

The research question underlies the entire experimental design, especially to decide whether within-subjects (i.e. repeated measures) or between-subjects design, will be employed. If the influence of an independent variable, in a single group is explored, within-subject design is preferred. Cognitive load can be manipulated in different ways, such as assigning a difficulty level (i.e. easy, moderate, hard) to tasks, increasing/decreasing the presentation time of the stimulus (i.e. the duration of the map stimulus on the screen) and varying the complexity of maps (e.g. each (group of stimuli or) stimulus including a different number of the map elements). The increasing number of map elements and the decreasing presentation time of the map stimulus correspond to increasing cognitive load. All of these variables affecting cognitive load can be implemented as independent variables in the study design. However, to study the influence of an independent variable among two or more groups, between-subject design is appropriate. In between-subject experiments, it is important to recruit the groups of novices and experts whose age and gender are counterbalanced, which act as possible confounders to data analysis because EEG is very sensitive to those factors. On the contrary, the individual differences of participants (i.e. age, gender) have no impact on within-subject design, since the outcome of the experiment relies on a large number of trials. Note that between- and within-subject design can perfectly be combined in one experiment.

3.2.1.2. Stimuli and trials

Having identified the research question, one should build the design of the task and stimuli around it. First, regardless of the type of the task, a large number of trials are typically necessary for the reliable EEG analysis (Cohen, 2014). In this context, trials refer to the repetition of the same stimulus condition. Theoretically, the recorded EEG contains two types of signals; one is stemmed from brain-related activity and the other is the noise elicited from external sources (e.g. eyes, muscles, recording device) for any given trial. If an infinite number of trials were possible, the effect of the noise part would be canceled out and the remaining signal part would be an indicator of the brain-related activity (De Haan, 2007; Talsma & Woldorff, 2005). This assumption is valid for an ideal situation, but one could question the influence of the fatigue of the users on the measurements if the number of trials increases – and thus the experiment duration is very long (more than one hour). However, increasing the number of trials is a common approach to eliminate the noise and reveal the EEG signal occurred due to the brain activity. For instance, many trials of the same stimulus are averaged to disentangle the event-related potentials (ERP), which are the time-locked electrical activity derived from the EEG signal in response to a discrete stimulus or event (Handy, 2005). ERPs can be considered any measured brain responses that are directly the result of a thought or perception; therefore, they can identify distinct phases of cortical processing in response to stimulus presentation such as cognitive load (Winslow et al., 2013). Despite that there is no recommended number for trials in the literature, most ERP studies include approximately 100 trials (Delorme, Miyakoshi, Jung & Makeig, 2015; Handy, 2005; Ouyang, Sommer & Zhou, 2016; Talsma & Woldorff, 2005), which correspond to a significant number of stimuli compared to those used in other cartographic user studies (e.g. ET, think-aloud, sketch map). However, the number of trials highly depends on the component the experimenter looking at. For instance, A P300- component could be observed with 30 trials, whereas for an early attentional component like the C1, one would need at least 300 trials to have a decent signal.

Besides their quantity, the order in which the trials and stimuli are presented in the experiment should be well planned. For instance, the duration of the stimulus presentation (i.e. presentation time) should be identified and limited based on the existing literature or pilot tests. The difficulty levels of the tasks should also be assigned beforehand.

3.2.1.3. Psychological Measures

Another essential issue of the experimental design is choosing which psychological measures will be utilized to extract the cognitive load. Achieving meaningful outcomes from the integration of EEG and ET methods is only possible with an appropriate experiment design. The preparation and presentation of the stimuli and the task should allow collecting data required for cognitive load assessment. In the literature, there are two widely used ET metrics as indicators of cognitive load: fixations and pupil size. Fixation duration indicates attention paid to a fixated location because the eye fixates to a certain point as long as the information is being processed (Henderson & Ferreira, 2013). Fixation duration increases under processing load, i.e. when visual processing requires more effort (Duchowski, 2007; Holmqvist, et al., 2011; Meghanathan, van Leeuwen & Nikolaev, 2015). Besides, task-invoked pupillary response directly reflects the cognitive load on working memory. Greater pupil dilation is found to be associated with high cognitive load (Beatty, 1982; Fehrenbacher & Djamasbi, 2017; Granholm, Asarnow, Sarkin & Dykes, 1996) and pupil size is particularly sensitive to cognitive load when it is above the limit of memory capacity (Meghanathan, et al., 2015). Pupil metrics are only useful when eye movements are restricted by the experiment design, e.g. by using fixated targets to direct gaze activity of the participant to a certain portion of the screen where the stimulus is shown.

Cognitive load can be measured using EEG activity power spectrum and several researchers have repeatedly proved that the spectral power changes under alpha and theta frequency band are related to task difficulty and therefore good predictors of cognitive load in a variety of working memory task demands (Gevins & Smith, 2000; Witvoet, 2013). These studies have found that alpha activity (particularly over the parietal and occipital areas) decreases with growing task demands that inherently cause working memory performance to decrease, whereas theta power increases (over frontal midline areas) when there is a cognitive activity and encoding new information (Klimesch, 1999; Kumar & Kumar, 2016; Pfurtscheller & da Silva, 1999; Witvoet, 2013). Event-related power changes in EEG signal can be quantified in a specified frequency band during a cognitive task. Event-related power decreases that cause a reduction of amplitude in response to a stimulus, are called as event-related desynchronization (ERD), whereas power increases resulting an increment of amplitude with the stimulus presentation are referred to event-related synchronization (ERS) (Pfurtscheller & Da Silva, 1999; Witvoet, 2013). Alpha desynchronization and theta synchronization are fundamental EEG phenomena that have been used in multiple studies on cognitive load and task difficulty (Anderson, et al., 2011; Gevins, Smith, McEvoy & Yu, 1997; Gevins & Smith, 2003; McEvoy, Smith & Gevins, 1998; Sauseng et al., 2005). ERD/ERS of the alpha band has been found to be especially sensitive to cognitive task performance and higher cognitive abilities (e.g., Fink, Grabner, Benedek & Neubauer, 2006; Neubauer and Fink, 2009). For instance, Gevins et al. (1997) examined changes in cortical activity during spatial and verbal working memory tasks and observed lower alpha activity in the difficult tasks compared to easy ones and theta activity increased in magnitude with higher task difficulty. These results suggest that alpha and theta oscillations are differently related to the task difficulty. As task difficulty increases, alpha activity decreases (desynchronize), whereas theta activity increases (synchronize) (Witvoet, 2013). In this context, the event-related changes in EEG power spectral density (PSD) can be calculated for alpha and theta frequency bands. PSD refers to the measure of signal's power content versus frequency (see Keskin et al., 2020). I will explain further with the formulas to calculate ERS and ERD in Chapter 4.2.2.4.

Alternatively, one of the most widely used EEG components is ERP referring to a set of voltage changes contained within EEG epochs that are time-locked to the stimulus (Coles & Rugg, 1995). In most cases to extract the ERP signal, a number of EEG epochs are averaged through the repetition of the same event. Cognitive psychologists have long studied ERP, and it has contributed to sensory, perceptual and cognitive research (see for review: Coles & Rugg, 1995; Jiang, 2014). It is also possible to track cognitive processes that occur during a single eye fixation by using eye-fixation-related-potentials (EFRP) which combine eye movements and ERPs for studying cognitive behavior. As Baccino & Manunta (2005) stated, both eye movements and ERPs suggest temporal measurements, but eye movements (i.e. fixation duration) describe a summation of all the cognitive processing occurring during identification while ERPs show the sequence of the processes. The major advantage of EFRP is to couple accurate time measures from ERPs and the location of the eye on the stimulus, so it can be used to disentangle perceptual/attentional/cognitive factors affecting retrieval process. Furthermore, with the EFRP technique, experimental settings allowing for a strong ecological validity can be used since the technique allows one to move the gaze freely onto any complex stimuli (text, visual scenes, etc.) (Baccino, 2011). However, EFRP method requires high precision in terms of time synchronization of ET and EEG systems.

Besides ET and EEG data, behavioral data (e.g. reaction time and correct answers) is a good indicator of cognitive load. In psychological experiments, tasks are often presented with strong time constraints (e.g. to respond as quickly as possible to which option is true by pressing a button) in order to study the effect of some variables (e.g. task difficulty) on the mean response time of a given number of subjects. In this context, it is assumed that the longer the response time, the more complex the cognitive process leading to that choice and inherently the harder the decision to make. When evaluating the reaction time, it is common to exclude the trials with incorrect answers, since only a correct answer is an evidence that the subject completely and correctly processed the task. This approach is also valid for the assessment of EEG and ET data due to the necessity of having the same trials contributing the analyses of behavioral, gaze and neurophysiological data, and the fact that wrong answers contain error related negativity, which may affect the correct EEG pattern. Accordingly, the inverse efficiency score (IES) is the oldest and the most frequently used measure to consider the corrected reaction times based on the amount of errors (Townsend & Ashby, 1978) (Equation 3.1),

$$\mathrm{IES} = \mathrm{RT} / (1 - \mathrm{PE}) \tag{3.1}$$

where RT is the subject's average (correct) RT of the condition, and PE is the subject's proportion of errors in the condition. Table 3.1 summarizes all the methodological aspects of the experiment design explained until here, covering research question & goals, participants, materials, procedures and analysis.

Table 3.1. Summary of the methodological issues in our experiment design

Issue	Summary
Research Question & Goals	- How does cognitive load vary between experts and novices while memorizing map content and how does the complexity of task influence the cognitive load? *Hypothesis*: We expect that memory task will cause higher cognitive load in novices *Goal*: to test the effect of task difficulty on behavior, which is the retrieval of maps and map features.
Participants Stimulus & Task	Novice and expert participants matching age and gender with normal or corrected-to-normal vision - 10-15 participants for each group (age range: 25-35) The map stimuli: Screenshots of Google's road maps at zoom level 15 with 1km scale bar. The sketch maps in the answer screen: by digitizing the main structuring map elements (i.e. major roads, hydrography, green areas) Randomized block design: Seven blocks representing seven difficulty types included in the experiment. Each block includes 50 map stimuli and 50 trials (i.e. one for each stimulus) focusing on the similarity of one of the below listed: Block 1: The whole sketch map Block 2: Roads and hydrography Block 3: Roads and green areas Block 4: Green areas and hydrography Block 5: Green areas Block 6: Hydrography Block 7: Roads
Procedures	The order of blocks and trials within blocks are randomized across participants to combat learning effects. • Preparation stage (e.g. placement of EEG cap, instructions for the experiment, ET calibration, presentation of the training task) • Main task (presentation of fixation cross, encoding display (map stimulus), search display (graphical response screens)) • Post-test questionnaire *2.5 hours (with two three short breaks because fatigue starts at ~30 minutes)*
Analysis	Interpretation of results based on statistics applied to the collected data. Metrics to estimate cognitive load: behavioral: accuracy/correctness, reaction time, neurophysiological: EEG (ERD – ERS, spectral power analysis), gaze: ET (fixation duration, number of fixations) Report both descriptive and inferential statistics.

3.3. Technical issues while designing an EEG & ET user experiment in a cartographic context

Technical concerns refer to the (i) synchronization of ET and EEG recording systems, their accuracy and quality, and (ii) numerous processing steps (i.e. preprocessing, the alignment of collected ET and EEG data, removal of non-cerebral activities from EEG data, segmentation and re-referencing). In this part, we present an overview of the theory behind each technical issue and provide information about how we dealt with it in our experiments (see Table 3.2 for a summary of this section).

3.3.1. Synchronization of EEG and ET recording systems, their accuracy and quality

As Xue, Quan, Li, Yue & Zhang (2017) explained, it would be ideal to sample electrophysiological and behavioral data with the same device and sampling rate. However, due to different hardware and software used for data acquisition in practice, we have to implement hardware and software based synchronization methods. Because both hardware and software have their own internal latencies, the temporal accuracy of the co-registration of an ET and EEG has to be identified. To achieve synchronized ET and EEG data collection, there are two most widely acknowledged hardware methods based on the signal characteristics of the events; TCP/IP (Transmission Control Protocol/Internet Protocol, i.e. network) and TTL (Transistor-transistor logic). In network or TCP/IP method, the events are sent through a LAN network where both the sender and the receiver are the recording computers sharing a common TCP/IP communication protocol (e.g. Xue et al., 2017).

Using TTL pulses to trigger the sampling onset is known to be the most reliable ways to send information with accurate time, especially if they are generated by a dedicated hardware. The events are based on standardized electrical signals levels sent through the different pins of a TTL output port. Depending on the number of pins of the TTL output and input port, the number of different events that can be sent varies in a binary format (e.g. 8 pins: up to 255 events). (e.g. Baccino, 2011; Dimigen, et al., 2011; Plöchl, Ossandón & König, 2012; Savage, Potter & Tatler, 2013).

Simultaneous data acquisition of ET and EEG can be accomplished with a single or dual PC set up (for further information, please read SMI, 2016). One crucial technical issue in this synchronization is the exact timing of all subparts of the system and the accuracy of this, which is referred as software methods including the timing synchronization protocols (Xue et al., 2017). If the experiment relies on the precise timing, the timings need to be specified in the order of milliseconds. The deviations in the timings can be jittered or fixed delays. For EEG, ET or Reaction Time measurements, effects of jitters can be neglected because the subject variance is high, or in other words: they level out in a large group of participants and stimuli. Nevertheless, the fixed delays must be identified precisely and taken into account either in the experiment or in the analysis stage. Spatio-temporal resolution of the eye-tracker 0.01◦ / 2 kHz and the TTL itself is within one millisecond accurate, which corresponds to the delay between when the TTL pulse is sent – the stimulus PC issues a command to display a stimulus – and when arrives at the EEG isolator. However, the big unknown within this setup is how long it takes an operating system to actually execute the call to display a stimulus and next how long it takes, physically, to display this on the monitor. The monitors are usually the largest source of fixed delays, which can be measured by a photo sensor. What is measured with a photo sensor is not the delay between the ET and EEG system, but the delay between when a stimulus is presented and when it is actually displayed (i.e. how long it takes for monitor to load the stimulus). Furthermore, monitors have a certain refresh rate which brings certain limitations to the synchronization. For instance, if the monitor used in the experiment has a 60Hz refresh rate, the system cannot do any better than 16.67 millisecond in terms of synchronization with the eye tracker and EEG data. Therefore, TTL offset and monitor offset must be taken into an account while synchronizing ET and EEG data (Dimigen et al., 2011).

The quality of EEG recording relies on the quality of EEG system and experimenter's skills in electrode placement. To maintain a sufficient data recording quality (i.e. to maximize signal/noise), impedance at every electrode contact with the scalp must be checked using an EEG impedance meter at the beginning of every experiment. Measuring the impedance between the highly conductive living skin overlying the skull and the electrode allows verifying if skin-electrode interface is good enough in advance of recording (Kappenman & Luck, 2010; Casal & La Mura, 2016). Since modern amplifiers have high input impedance (e.g. the one we use, EEG100C, has 2 M Ohm differential and 1000 M Ohm common mode input impedance), electrode impedances of the whole circuit (i.e. ground, active EEG, reference electrodes) less than 10 K ohms are acceptable (Herman, et al., 2015; Teplan, 2002). Although impedance fluctuation is one of the patient related artifact sources, decreasing the impedance helps to reduce technical artifacts such as AC power line noise.

3.3.2. Data processing steps

3.3.2.1. Preprocessing

Except filtering out the unwanted portions, preprocessing of ET data mostly includes data management (e.g. data export and conversions, restructuring the data content) due to the necessity of converting data into compatible EEGLAB compatible format by following Dimigen & Reinacher's (2013) tutorial of MATLAB Toolbox for Simultaneous Eye tracking & EEG.

While the preparation of ET data can be handled with a small scripting (e.g. written with Python), preprocessing of EEG data requires an extensive workflow with several steps to follow. Although it is not a scientific or an official publication of EEGLAB, Makoto's preprocessing pipeline is recommended as a very useful guideline on how to perform these steps in EEGLAB, and consulted by many researchers dealing with EEG data (Miyakoshi, 2018). It covers all the preprocessing stages from importing data to filtering it regarding on the desired psychological measures to study. Therefore, it might be useful to consult this online and regularly updated documentation, if needed.

If EEG data are imported correctly, first step prior to further analysis is filtering which corresponds to eliminate portions of the EEG record that are contaminated by gross motor movements or eye blinks (i.e. biological artifacts). These artifacts are electrical activities originate from non-cerebral sources like muscles and EOG (Electrooculography). EOG signal should be removed to reveal clean EEG reflecting neural activity. Blind Source Separation (BSS) or Independent Component Analysis (ICA) is commonly applied for EOG removal (see Chapter 3.3.2.3 for detailed information on ICA). Besides biological artifacts, one of the biggest sources for external noise (i.e. environmental artifacts) is power-line interference, which is often characterized by high power at high frequencies or spikes in the power spectrum at some characteristic frequencies (Repovš, 2010; Samadi & Cooke, 2014). AC power line noise, which can be either 50 Hz or 60 Hz, causes a directional effect on channels, making the raw EEG data impossible to analyze. Therefore, it should be removed either on the fly with a built-in filter or with a relevant notch filter afterwards.

Due to the factors such as sweating, lost data during recording, drifts in electrode impedance leading to changes in voltage and the saturation of the amplifier, it is recommended to filter out the frequencies below 0.01 Hz. A low-pass filter at 100 Hz is applied for filtering out the noise at the other end of the spectrum of frequencies that are of interest. On account of the of the contraction of muscles that causes a strong signal with frequencies above 100 Hz, those frequencies should be suppressed to eliminate movement artefacts in the EEG signal (Luck, 2005; Repovš, 2010; Teplan, 2002).

After filtering, bad EEG channels, if there are still any, should be removed prior to the detection of eye movements and the rejection of bad eye-tracking data. Besides ET and EEG data, channel location information of EEG electrodes (provided in polar, spherical or Cartesian coordinates) is necessary to plot EEG scalp maps in either 2-D or 3-D format, or to estimate source locations of the recording electrodes in EEGLAB. A channel location file is simply a text file, which can be created with any text editor and load manually to EEGLAB, or can automatically be detected if the data is compatible with the existing sample location files available in the channel location library of EEGLAB.

3.3.2.2. Alignment of the EEG and ET data

Verifying that the recordings are of sufficient quality, preprocessed ET and EEG data should be aligned. This process is also called offline synchronization and Dimigen et al. (2011) proposed three possibilities to synchronize EEG and ET systems in EEGLAB; (i) message + triggers, (ii) analogue output, (iii) shared triggers. In the first method, while triggers are still sent to the EEG, messages (i.e. short text strings inserted in ET data) are used as the corresponding events for the ET. Here, the ET computer is given a command to insert an ASCII text message (containing a keyword and the value of the corresponding EEG trigger) into the ET data. In the stimulation software, the commands to send a trigger (to the EEG) and a message (to the ET) are given in immediate succession. The latter method is that copy of the eye track is fed directly into the EEG. A digital-to-analogue converter card in the ET outputs (some of) the data as an analogue signal (Dimigen & Reinacher, 2013). The last method is shared triggers which involves sending common trigger pulses ("triggers") frequently from the stimulation computer to both ET computer and EEG recording computer. This is achieved via a Y-shaped cable that is attached to the parallel port of the stimulation computer and splits up the pulse so it is looped through to EEG and ET.

To align the collected EEG and ET data, we chose “message + triggers/events” method as explained by Dimigen & Reinacher (2013). The EYEEEG plugin requires that there are at least two shared events present in the ET and EEG; start-event and end-event. This can be achieved by using a unique event value (e.g. "100") to mark the start-event and other unique value (e.g., "200") for the end-event. Eye-tracking data in between the start-event and end-event are then linearly interpolated to match the sampling frequency of the EEG. In messages + triggers method, synchronization messages sent to the eye tracker need to have a specified format.

On the other hand, the event text file should be composed of three columns; the first containing the latency of the event (in seconds), the second the type of the event, and the third a parameter describing the event (for example, the position of the stimulus). Thus, we wrote a Python script that enables us to remove unwanted columns (channel, label), convert the time unit in seconds and minutes into milliseconds and change the column names into "latency, type, position" as suggested in the EEGLAB tutorial (Delorme & Makeig, 2012). Since EEGLAB cannot read string, each stimulus label was expressed as an integer (i.e. 101, 102, 103, etc.). In Raw ET data, we also assigned every stimulus the same integer code in order to synchronize ET and EEG data. For instance, in raw ET data (text) file, '# Message: NoImage' expression in the stimulus column was replaced with '# Message: SYNC 100'. The SYNC codes created here is used while parsing ET data. Additionally, other string data in ET file should be replaced with integer values. Therefore, we assigned "0" to blinks, "1" to saccades, "2" to fixations and "3" to "-" (see Annex 4 for the Python scripts).

The next step is the detection of fixation and saccades and the rejection of bad eye-tracking data. Before extracting fixation and saccades, eye movement data should be filtered considering the saccade size and fixation durations that are meaningful for further analysis. After detecting eye movements, we filtered out saccades exceeding 20 degrees, and their corresponding following fixations. Fixations whose durations were outside the range of 50-1000 ms were excluded together with their corresponding preceding saccades (Figure 3.1). ET data exceeding the screen coordinates of PC (1680 x 1050) were also rejected (see Annex 5 for the Python scripts).

Once we filtered EEG and ET data, and obtain the channel locations for EEG electrodes are obtained, the alignment of ET and EEG recordings (synchronization) can be established through triggers/events (Figure 3.2). Events represent time stamps when the stimulus is shown to a participant. These time points can be derived from raw ET data with scripting, therefore Python scripts were used to extract events and organize data in a compatible EEGLAB format). Figure 3.3 illustrates a portion of an EEG recording merged and aligned with eye movement data (saccades and fixations). The horizontal axis represents time and the vertical axis shows the amplitude (μV), i.e. the amount of energy in frequency bands listed on the left-hand side of the graph. Amplitude scale was adjusted in a way that the EEG waves are clearly visible but not overlap (Keskin & Ooms, 2018).

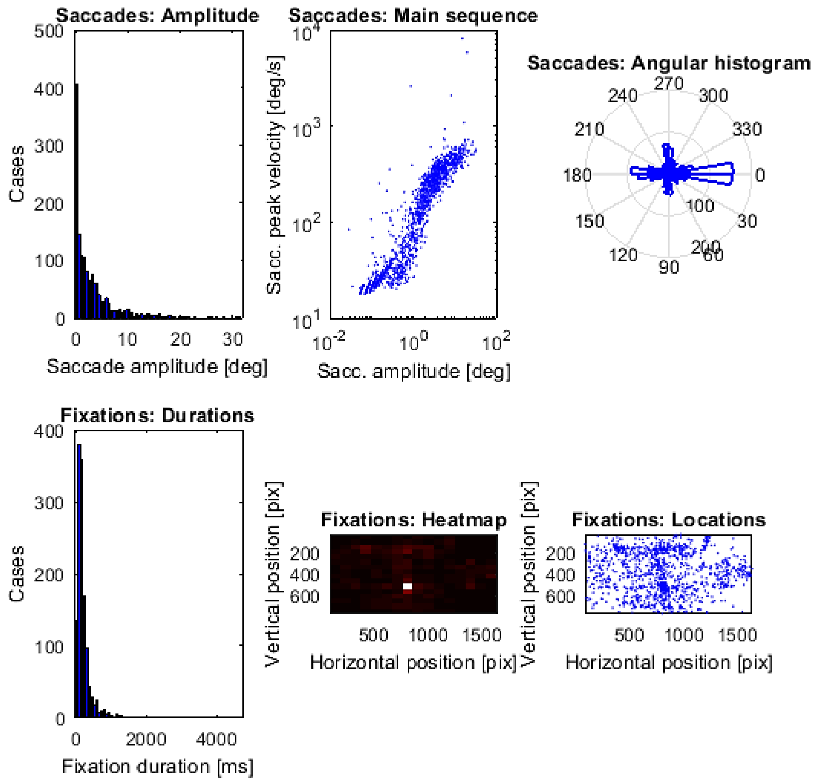


Figure 3.1. Properties of saccades and fixations

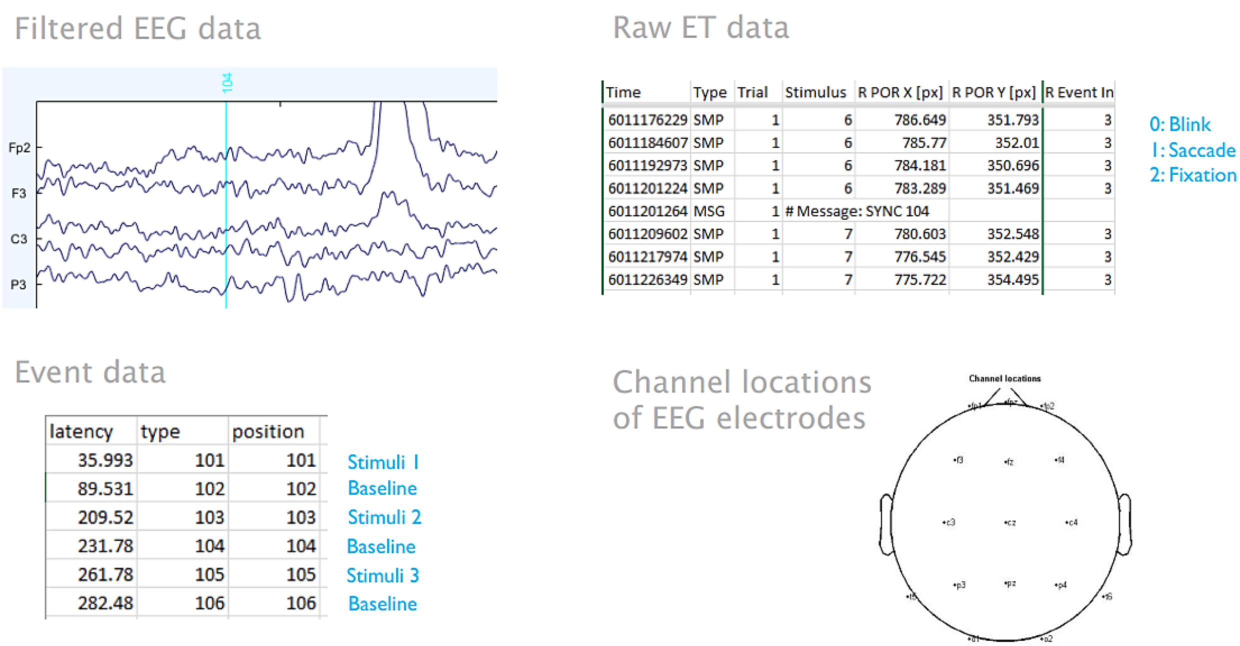


Figure 3.2. Integration of the collected data

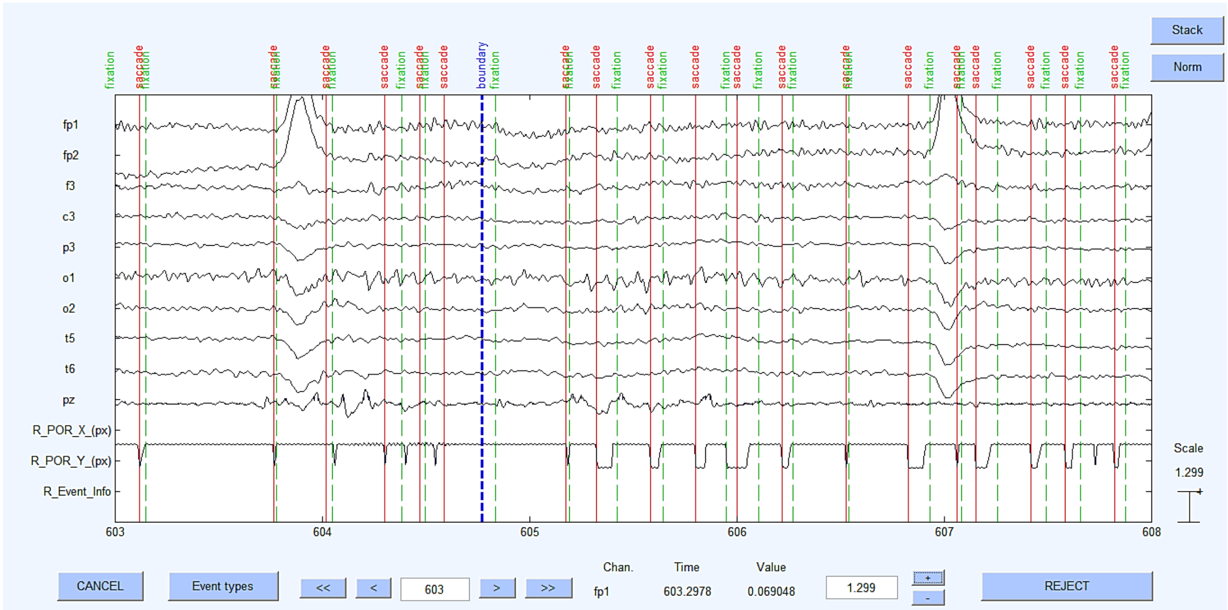


Figure 3.3. Synchronized ET and EEG data

3.3.2.3. Removal of non-cerebral activities from EEG data

EEG data include diverse source information related to either cortical or non-cortical activities such as artifacts that are usually independent of each other. Independent Component Analysis (ICA) is an established and a plausible method to separate these independent EEG activities in each participant's data assuming that the observed EEG signals from electrode channels are a linear mixture of independent source signals (or components). As a result of ICA, the EEG channels are broken down into components reflecting the activity of one source domain that is locally coherent and includes its projections to all the scalp channels, while the activities of unrelated EEG sources will be rejected from this independent component (IC) and isolated into other ICs (Onton & Makeig, 2006; Samadi & Cooke, 2014; Zeng & Song, 2014). For instance, it is very useful to separate out non-cerebral activity (e.g. blinks) from EEG data. However, interpretation of ICA results is mostly carried out subjectively and the classification criteria are rarely reported in the publications, thus, to decide which IC consists non-cortical artifact source requires an expertise (Dimigen, 2017).

There are several different ICA algorithms, however infomax ICA, which is freely available in EEGLAB, in particular, provides reliable results for data of sufficient quantity and quality having almost any number of channels as claimed by Onton & Makeig (2006).

3.3.2.4. Segmentation (epoching) & re-referencing

The next step is the segmentation of the usable data relative to the stimulus onset. Segmentation should be performed in a way to extract the cognitive load through the event-related changes in EEG power spectral density (PSD) for alpha and theta frequency bands and therefore, to calculate the event-related synchronization/desynchronization (ERS/ERD) in alpha and theta, which were explained in Chapter 3.2.1.3.

After epoching the data based on successful trials, EEG electrodes must be re-referenced before EEG power spectrum analysis. Theoretically, a reference could be anywhere but has to be selected carefully, because any activity in the reference electrode will be reflected at other electrodes. There is no best common reference site in the literature, however there are some widely utilized approaches for re-referencing; by using mastoids (i.e. physical references), by using a fixed electrode or by averaging the activity at all EEG electrodes (i.e. reference-free). Re-referencing is essential for the comparison of results (i.e. ERD/ERS) among different participants, since in EEG data voltages recorded at each electrode are relative to voltages recorded at other electrodes and each individual's brain waves are unique (Teplan, 2002) (URL 2). Table 3.2 demonstrates a summary of the technical issues covered in the experiment design.

3.4. Discussion

The motivation of this chapter is to introduce/explain how an EEG&ET user study should be designed in a cartographic context based on the hands-on experience and the relevant literature. I would like to emphasize that there exists no best solution available when it comes to EEG&ET user experiment design, and this is another important reason of why we attempt to list issues related to experimental design together with possible solutions. Consequently, I present a set of rules to follow to achieve a good experiment design within the frame of research objectives. On the one hand, as methodological decisions are highly dependent on the research questions and hypotheses regarding to them, it is important to describe a solid objective for the user study with the psychological design principles in mind and identify relevant metrics answering the research questions specified at the beginning. In this context, I tried to focus on the data we would like to acquire and what EEG can answer.

On the other hand, although the experiment is limited to a spatial memory task, and the methodological design of the other experiments may vary on a large scale, the technical issues to overcome and the preprocessing steps of the collected data are valid for almost all ET&EEG experiments. Technical and data-analytical issues, which unfold in two parts; (i) synchronization of ET and EEG recording systems, and (ii) processing of the collected data, play determinant roles in methodological and the experimental design. Recording EEG and ET data in free viewing tasks has been a challenge and rarely applied especially due to the precise co-registration of gaze position. The technical problems I mentioned throughout the chapter are similar to what was mentioned by Dimigen et al. (2011), such as muscle artifacts stemmed from unnatural sitting positions, electromagnetic artifacts resulting from other electric devices affecting EEG sensors, proper synchronization of EEG and ET records. To minimize the muscle artifacts, using chin rest and adapting the position of the participant is crucial; besides making sure that the participant had enough rest between blocks so that they move as little as possible during the experiment. Electromagnetic artifacts, which is introduced as line noise in EEG data, should be identified and filtered out. For accurate synchronization of both EEG and ET data records, TTL triggers are preferred as it is the most straightforward and reliable method (e.g. Dimigen & Reinacher, 2013; Nikolaev et al., 2016). Although the proper synchronization can be achieved with TTL trigger method, in our experiment, the monitor offset value restricts studying the eye-fixation related potentials requiring high temporal resolution in terms of synchronization of EEG and ET.

Table 3.2. Summary of the technical issues covered in the experiment design

			Theory	Our experiment
System architecture	**Recording Equipment**	For EEG	BIOPAC, BrainVision, Biometric, Biosemi, ObseverXT, iMotions, EGI's Netstation	BIOPAC Acqknowledge hardware & software
		For ET	SMI, Tobii, EyeLink, iMotions	SMI RED 250 Eye tracker, SMI iViewX software
		For stimulus presentation	E-prime, Superlab, OpenSesami, NBS presentation	SMI Experiment Center
Synchronization of EEG & ET recording system	**Synchronization method**	(i) through TTL events (ii) through TCP/IP events	dual PC setup with TTL triggers	
Processing	**Software**	For EEG	Open source: EEGLAB, Brainstorm, Okazolab, etc. Commercial: BIOPAC, Biosemi, EGI's Netstation, iMotions, etc.	EEGLAB: MATLAB tool with EYEEEG plugin to handle EEG & ET data together
		For ET	Opensource: PyGaze, Ogama, etc.	
		For ET & EEG	Open source: EEGLAB Commercial: BIOPAC Acqknowledge	
	Data management	For EEGLAB	converting ET & EEG & event data into EEGLAB compatible format & rearranging ET and events with Python scripting	

		Theory		Our experiment
	Filtering EEG data	(i) Low pass filter ((ii) High pass filter (iii) Notch filter		Low pass filter at 100 Hz High pass filter at 0.01 Hz Notch filter at 50 Hz
	EEG channel location information	polar, spherical or Cartesian coordinates		Creating a text file including polar, spherical or Cartesian coordinates for EEG channel locations
	Bad channel removal	For EEG; due to bad contact or noise		not more 5 channels per participant because re-referencing might be performed by averaging the EEG channels.
Alignment of ET & EEG data	**Method**	(i) shared events (ii) message + triggers (iii) analogue output		message + triggers
Detection of eye movements & Rejection of bad eye-tracking data	**Criteria**	For ET; data exceeding screen coordinates, all blinks, and a part of saccade and fixations		ET data exceeding the screen coordinates of PC (1680 x 1050) saccades > 20 degree and also their corresponding following fixations fixations outside 50-1000 ms interval together with their corresponding preceding saccades
Removal of non-cerebral activities from EEG data	**Method**	For EEG	ICA BSS	ICA
Segmentation & Re-referencing	**Method**	For EEG	Linked mastoids, Fixed electrode Average at all electrodes	Fixed electrode

However, it allows studying EEG activity power spectrum to estimate the cognitive load and ET data can still be synchronized offline and ET metrics can be correlated with EEG data on a trial basis. Therefore, the feasibility of the methodology should always be verified in advance, due to the possible technical constraints stemmed from the recording equipment.

Although some procedures such as data management (e.g. converting data into a compatible format with EEGLAB) and noise filtering (e.g. applying high- and low-pass filters on the fly) can be automatized, many other steps such as bad channel removal, which is mostly carried out by visual inspection, are performed manually. In addition, preprocessing and analyzing the data are inherently the most labor-intensive and complicated part of the study. Since each participant data consists of a number of trials and should be handled individually, processing stage is overall very time-consuming.

3.5. Conclusion

This chapter aimed to provide a methodological overview of what is possible with co-registration of EEG and ET to investigate the temporal characteristics of cognitive processes in free viewing condition, only within the frame of the specific spatial memory study described throughout the chapter. Combining EEG and ET is not straightforward since there are numerous methodological and technical problems to overcome, yet it is indeed a very valuable technique to explore the individual differences and similarities of map users through perceptual and cognitive procedures. If we continue staying engaged with experimental psychology and cognitive science research, it will contribute to the future progress of scientific cartography. The more we know about the limitations and capabilities of visual perception and cognition of different map users, the higher the possibilities to design cartographic products in a more efficient, understandable and usable way.

References

Anderson, E. W., Potter, K. C., Matzen, L. E., Shepherd, J. F., Preston, G. A., & Silva, C. T. (2011, June). A user study of visualization effectiveness using EEG and cognitive load. In Computer Graphics Forum (Vol. 30, No. 3, 791-800). Blackwell Publishing Ltd.

Baccino, T., & Manunta, Y. (2005). Eye-fixation-related potentials: Insight into parafoveal processing. Journal of Psychophysiology, 19(3), 204.

Baccino, T. (2011). Eye movements and concurrent event-related potentials': eye fixation-related potential investigations in reading. Eye, 2, 1.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychological bulletin, 91(2), 276.

Casal, L., & La Mura, G. (2016, April). Skin-electrode impedance measurement during ECG acquisition: Method's validation. In Journal of Physics: Conference Series (Vol. 705, No. 1, p. 012006). IOP Publishing.

Cohen, M. X. (2014). Analyzing neural time series data: theory and practice. MIT press.

Coles, M. G. H., & Rugg, M. (1995). Electrophysiology of mind: Event-related potentials and cognition.

De Haan, M. (2007). Visual attention and recognition memory in infancy. In M. de Haan (Ed.), Studies in developmental psychology. Infant EEG and event-related potentials (101-143). New York, NY, US: Psychology Press.

Delorme, A., Miyakoshi, M., Jung, T. P., & Makeig, S. (2015). Grand average ERP-image plotting and statistics: A method for comparing variability in event-related single-trial EEG activities across subjects and conditions. Journal of neuroscience methods, 250, 3-6.

Delorme, A., & Makeig, S. (2012). EEGLAB tutorial outline. Retrieved from: http://cognitrn.psych.indiana.edu/busey/temp/eeglabtutorial4.301/eeglabtut.html Last accessed: 17.01.2019.

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: analyses and review. Journal of experimental psychology: General, 140(4), 552-572.

Dimigen, O., & Reinacher, U. (2013, January). A MATLAB Tool for Simultaneous Eye Tracking and EEG. In Journal of Cognitive Neuroscience, 143-144.

Dimigen, O. (2017, July) EYE-EEG: Tutorial. Retrieved from: http://www2.hu-berlin.de/eye tracking-eeg/tutorial.html, Last accessed: 17.01.2019.

Duchowski, A. T. (2007). Eye tracking methodology. Theory and practice, 328.

Fehrenbacher, D. D., & Djamasbi, S. (2017). Information systems and task demand: An exploratory pupillometry study of computerized decision making. Decision Support Systems, 97, 1-11.

Fink, A., Grabner, R. H., Benedek, M., & Neubauer, A. C. (2006). Divergent thinking training is related to frontal electroencephalogram alpha synchronization. European Journal of Neuroscience, 23(8), 2241-2246.

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cerebral cortex (New York, NY: 1991), 7(4), 374-385.

Gevins, A., & Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. Cerebral cortex, 10(9), 829-839.

Gevins, A., & Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. Theoretical Issues in Ergonomics Science, 4(1-2),113-131.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. Psychophysiology, 33(4), 457-461.

Gregory, R. J. (2004). Psychological Testing: History, Principles, and Applications. Allyn & Bacon.

Handy, T. C. (Ed.). (2005). Event-related potentials: A methods handbook. MIT press.

Henderson, J., & Ferreira, F. (2013). The interface of language, vision, and action: Eye movements and the visual world. Psychology Press.

Herman, S. T., Abend, N. S., Bleck, T. P., Chapman, K. E., Drislane, F. W., Emerson, R. G., . . . Hirsch, L. J. (2015). Consensus Statement on Continuous EEG in Critically Ill Adults and Children, Part II: Personnel, Technical Specifications and Clinical Practice. Journal of Clinical Neurophysiology : Official Publication of the American Electroencephalographic Society, 32(2), 96–108.

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

Kaplan, R. M., & Saccuzzo, D. P. (2017). Psychological testing: Principles, applications, and issues. Nelson Education.

Kappenman, E. S., & Luck, S. J. (2010). The effects of electrode impedance on data quality and statistical significance in ERP recordings. Psychophysiology, 47(5), 888-904.

Keskin, M., & Ooms, K. (2018, January). Possibilities of eye tracking and EEG integration for visual search on 2D maps. In Eye tracking for Spatial Research, Proceedings of the 3rd International Workshop. ETH Zurich.

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

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain research reviews, 29(2-3), 169-195.

Kumar, N., & Kumar, J. (2016). Measurement of Cognitive Load in HCI Systems Using EEG Power Spectrum: An Experimental Study. Procedia Computer Sci. 84, 70-78.

Luck, S. J. (2005). Ten simple rules for designing ERP experiments. Event-related potentials: A methods handbook, 262083337.

Miyakoshi, M. (2018). Makoto's preprocessing pipeline. Retrieved from: https://sccn.ucsd.edu/wiki/Makoto's_preprocessing_pipeline last accesed: 19.03.2018.

Meghanathan, R. N., van Leeuwen, C., & Nikolaev, A. R. (2015). Fixation duration surpasses pupil size as a measure of memory load in free viewing. Frontiers in human neuroscience, 8, 1063.

McEvoy, L. K., Smith, M. E., & Gevins, A. (1998). Dynamic cortical networks of verbal and spatial working memory: effects of memory load and task practice. Cerebral cortex (New York, NY: 1991), 8(7), 563-574.

Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. Neuroscience & Biobehavioral Reviews, 33(7), 1004-1023.

Nikolaev, A. R., Meghanathan, R. N., & van Leeuwen, C. (2016). Combining EEG and eye movement recording in free viewing: Pitfalls and possibilities. Brain and cognition, 107, 55-83.

Onton, J., & Makeig, S. (2006). Information-based modeling of event-related brain dynamics. Progress in brain research, 159, 99-120.

Ouyang, G., Sommer, W., & Zhou, C. (2016). Reconstructing ERP amplitude effects after compensating for trial-to-trial latency jitter: A solution based on a novel application of residue iteration decomposition. International Journal of Psychophysiology, 109, 9-20.

Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. Clinical neurophysiology, 110(11), 1842-1857.

Plöchl, M., Ossandón, J. P., & König, P. (2012). Combining EEG and eye-tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data. Frontiers in human neuroscience, 6, 278.

Repovš, G. (2010). Dealing with noise in eeg recording and data analysis, Informatica Medica Slovenica. 15(1), 18–25.

Samadi, M. R. H., & Cooke, N. (2014, September). EEG signal processing for eye-tracking. In Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European. 2030-2034.

Savage, S. W., Potter, D. D., & Tatler, B. W. (2013). Does preoccupation impair hazard perception? A simultaneous EEG and eye-tracking study. Transportation research part F: traffic psychology and behaviour, 17, 52-62.

Sauseng, P., Klimesch, W., Schabus, M., & Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. International Journal of Psychophysiology, 57(2), 97–103.

SMI (SensoMotoric Instruments). (March 2016). SMI BIOPAC Remote Integration Mini Manual V1.0.

Talsma, D., & Woldorff, M. G. (2005). Methods for the estimation and removal of artifacts and overlap. Event-related potentials: a methods handbook. MIT Press, Cambridge, MA, 115-148.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. Cognitive theory, 3, 199-139.

Teplan, M. (2002). Fundamentals of EEG measurement. Measurement Science Review, 2(2), 1–11.

Urbina, S. (2014). Essentials of psychological testing. John Wiley & Sons.Verhulst, N., Slabbinck, H., Vermeir, I., & Larivière, B. (2018). A Multi-Method Exploration of Emotions During Service Delivery through Self-Service Technology or by an Employee. In La Londe Conference.

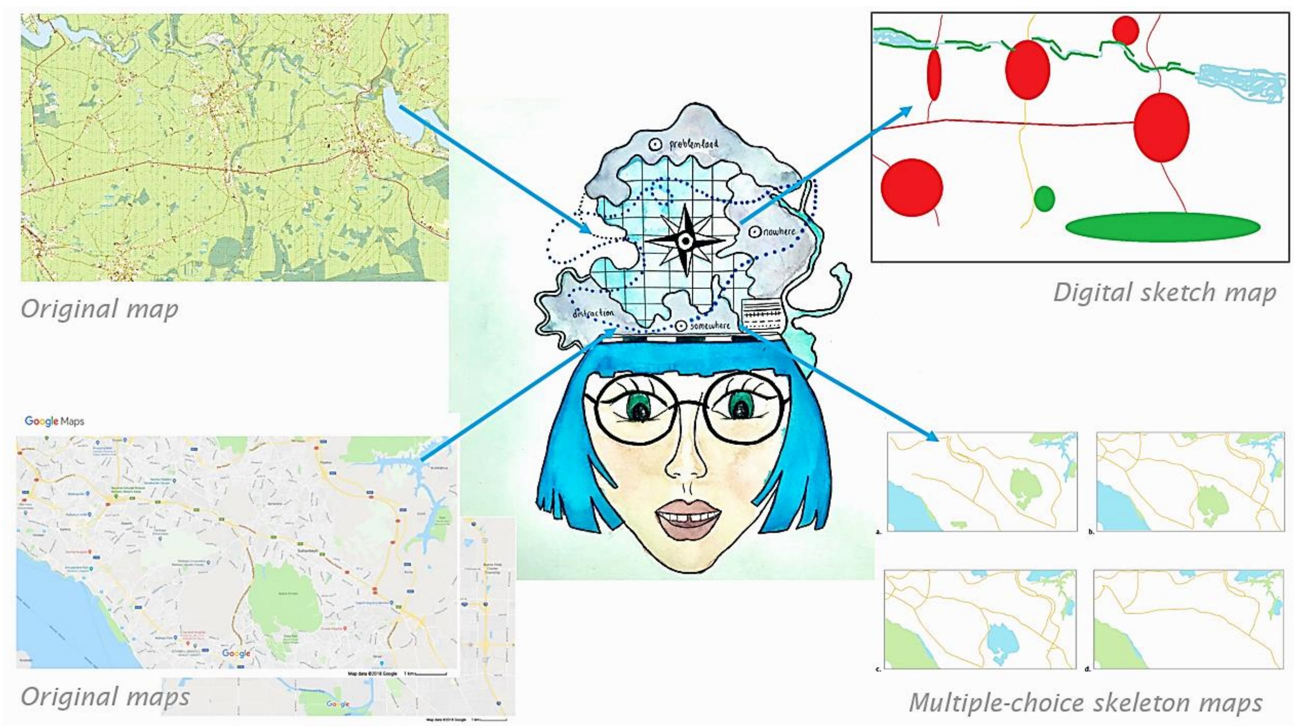
Winslow, B., Carpenter, A., Flint, J., Wang, X., Tomasetti, D., Johnston, M., & Hale, K. (2013). Combining EEG and Eye Tracking: Using Fixation-Locked Potentials in Visual Search. Journal of Eye Movement Research, 6(4), 1-11.

Witvoet, J. (2013). Does cognitive load influence performance in a game-based learning task? Bachelor's thesis, University of Twente, Enschede.

Xue, J., Quan, C., Li, C., Yue, J., & Zhang, C. (2017). A crucial temporal accuracy test of combining EEG and Tobii eye tracker. Medicine, 96(13).

Zeng, H., & Song, A. (2014). Removal of EOG artifacts from EEG recordings using stationary subspace analysis. Sci. World J. 2014, 2014, 259121.

Chapter 4:
Experiment 1 vs Experiment 2:
Eye tracking & EEG



"The brain internally simulates what will happen
if you were to perform some action under specific conditions.
Internal models not only play a role in motor acts (such as catching or dodging)
but also underlie conscious perception."

David Eagleman, Incognito: The Secret Lives of the Brain, 2011

Abstract. *The aim of this chapter is to evaluate the use of ET and EEG for studying the cognitive processes of a group of expert and novice map users and to explore these processes by comparing two types of spatial memory experiments through cognitive load measurements. The first experiment consisted of single trials and participants were instructed to study a map stimulus without any time constraints in order to draw a sketch map afterwards. According to the ET metrics (i.e., average fixation duration and the number of fixations per second), no statistically significant differences emerged between experts and novices. A similar result was also obtained with EEG Frontal Alpha Asymmetry calculations. On the contrary, in terms of alpha power across all electrodes, novices exhibited significantly lower alpha power, indicating a higher cognitive load. In the second experiment, a larger number of stimuli were used to study the effect of task difficulty. The same ET metrics used in the first experiment indicated that the difference between these user groups was not statistically significant. The cognitive load was also extracted using EEG event-related spectral power changes at alpha and theta frequency bands. Preliminary data exploration mostly suggested an increase in theta power and a decrease in alpha power.*

__Author Contributions:__ Conceptualization: Merve Keskin, Kristien Ooms; Methodology: Kristien Ooms, Merve Keskin; Software: Merve Keskin; Validation: Ahmet Ozgur Dogru; Formal Analysis: Merve Keskin; Investigation, Kristien Ooms, Ahmet Ozgur Dogru; Resources: Philippe De Maeyer; Data Curation: Merve Keskin; Writing-Original Draft Preparation: Merve Keskin; Writing-Review & Editing: Merve Keskin, Ahmet Ozgur Dogru, Kristien Ooms, Philippe De Maeyer.

This chapter is modified from:

Keskin, M., Ooms, K., Dogru, A.O., De Maeyer, P. (2019). EEG & Eye Tracking User Experiments for Spatial Memory Task on Maps. ISPRS Int. J. Geo-Inf., 8, 546

4.1. Introduction

Developments in medical research allow scientists to observe neurons in the brain with a high spatial and temporal resolution. As "scientific cartography" emerged in the early 1900s, it became possible to borrow theories and methods from experimental psychology to study how map design influences map use in a formal, systematic and empirical way rather than the trial and error method (Griffin, 2017). In this respect, scientific cartography has long dealt with cognitive issues of maps and map use, and Eckert applied experimental psychology principles to establish the laws of map logic (Griffin, 2017). According to him, map logic complies with the map creation laws, which strongly influence cartographic perception (Eckert, 1977).

To be able to understand map users' behaviors, it is important to identify the cognitive procedures. Human memory functions within a sequence of three stages; sensory memory, working (short-term) memory, and long-term memory (LTM) (Atkinson & Shiffrin, 1968). Cognitive processes and strategies particularly occur during circumstances such as being aware of where to look at in a map or the interpretation of map-related information regarding other knowledge stored in LTM. Since different map users have different information stored in their memory, they are expected to have different strategies while reading maps (Griffin, 2017). Therefore, expertise is one of the major individual differences across map users. During a map-related task, all map users, especially novices, rely on the general map knowledge (e.g., knowing that the contour lines represent the elevation) whereas experts mostly consult their specific map knowledge (e.g., knowing the direction of slope by interpreting the contour lines). Specific map knowledge enables experts to establish spatial relationships in a more structured and systematical way compared to non-experts (Keskin, Ooms, Dogru & De Maeyer, 2018). With the understanding of map knowledge of users, cartographers can focus on effective map designs that ideally do not cause a high cognitive load.

As fundamental units of cartographic design, Bertin's (1967) visual variables (i.e., position, size, shape, value, color hue, orientation, and texture) maintain the visual hierarchy (as described in Gestalt theory) which is essential to improve map logic by distinguishing and grouping map symbols and encoding map-related information (Griffin, 2017; Keskin et al., 2018). That is why map perception is, in a way, dependent on the decision of the visual variables used in its design. As important as design elements, the map task is fundamental in cognition because cognitive procedures generally occur instantaneously and within a specific task or context.

To measure real time cognitive responses of map users, cartographers have hitherto implemented many methods in cartographic usability research such as eye tracking (ET), sketch maps, thinking aloud, interviews or questionnaires (e.g. Herbert & Chen, 2015; Kveladze, Kraak & van Elzakker, 2017; Ooms, De Maeyer, Dupont, Veken, de Weghe & Verplaetse, 2016). The cartographic eye-tracking research has focused on the interpretation of visual information while performing a complex visual and cognitive task (e.g. Duchowski, 2007; Jacob & Karn, 2003), visual interaction with highly interactive interfaces (e.g. Çöltekin, Fabrikant & Lacayo, 2010), cognitive processes linked with visual search in maps (e.g. Ooms, 2012), and learning and remembering the information presented via maps (e.g. Keskin et al., 2018). The insights provided by eye tracking studies are promising for understanding how the human brain handles map-related cognitive tasks, yet there is still much research to do to elaborate how visual elements affect map use (e.g., perception, memory, cognition, etc.) and how to leverage visual variables to facilitate map use with less cognitive load. There is also a lack of empirical evidence on the users' cognitive processes involved in map tasks, especially on the sources of individual differences (i.e., expertise, gender, etc.) and the relationship between the organization of spatial thinking and geographic space (e.g. Kimerling et al., 2009; MacEachren, 2004; Ooms, de Maeyer, and Fack, 2015).

Authors propose that non-invasive brain-imagining techniques (e.g., EEG) can benefit cartographic user studies by providing direct measures of brain activity during cognitive processes. EEG, which is used to monitor the electrical activity in the brain, can be combined with other quantitative methods, such as eye tracking, to gain a better understanding of cognitive abilities and limitations of different groups of map users and how visual elements influence map use. The insights that particularly arose from the differences due to expertise will henceforth contribute to creating effective cartographic products.

Although there exists not much research showing the relationship between ET and EEG within the cartographic context, the outcomes of these two methods might provide different outcomes in terms of cognitive load. For instance, Gedminas (2011) explored how hurricane advisory maps are perceived by comparing the existing maps and their alternative map designs and found out that fixation durations and the number of fixations that these maps received do not differ significantly, while frontal EEG analysis indicated that the alternative maps had a larger and more positive effect on the user.

This study deals with the cartographic user experiments employing ET and EEG as simultaneous and synchronized data collection methods in collaboration for spatial memory tasks on maps (see Figure 4.1). The goal of the study is twofold: (i) studying the cognitive processes of participants, and (ii) evaluating the use of EEG for these processes by comparing two types of experiments, also allowing to triangulate ET and EEG findings and draw conclusion on the suitability of the methods, especially the contribution of EEG.

In this context, we introduce two user experiments both aiming to explore the (cognitive) strategies of experts and novice participants through cognitive load measurements when they are asked to memorize and then remember a (part of) map content with varying levels of complexity. Due to the methodological differences in the experiment designs, in the first experiment, we used simple and exploratory measurements for cognitive load extraction. However, the findings of the first experiment contributed to the experimental design of the second one. They were utilized as inputs for hypothesizing the second experiment as some of the findings were fundamental for the motivation of the second experiment. Therefore, the second experiment was designed in a more complex way of addressing in-depth investigation of cognitive load. In other words, it became possible to identify how cognitive load affects the recalling performance for both map user groups, and whether some features are recalled independently of task difficulty. If so, we can identify which features are recalled easily/primarily with respect to other features recalled within the task, especially when the task demands higher cognitive load. Moreover, we hope to contribute to cartographic usability research by introducing a brief overview on the methodology of ET and EEG experiments, because it enables us to explore the behavioral and neurophysiological responses of map users and helps with understanding the influence of cartographic design and task on individual map users. It has rarely been applied in a cartographic setting before, especially for map reading instead of map usability (e.g. Gedminas, 2011).

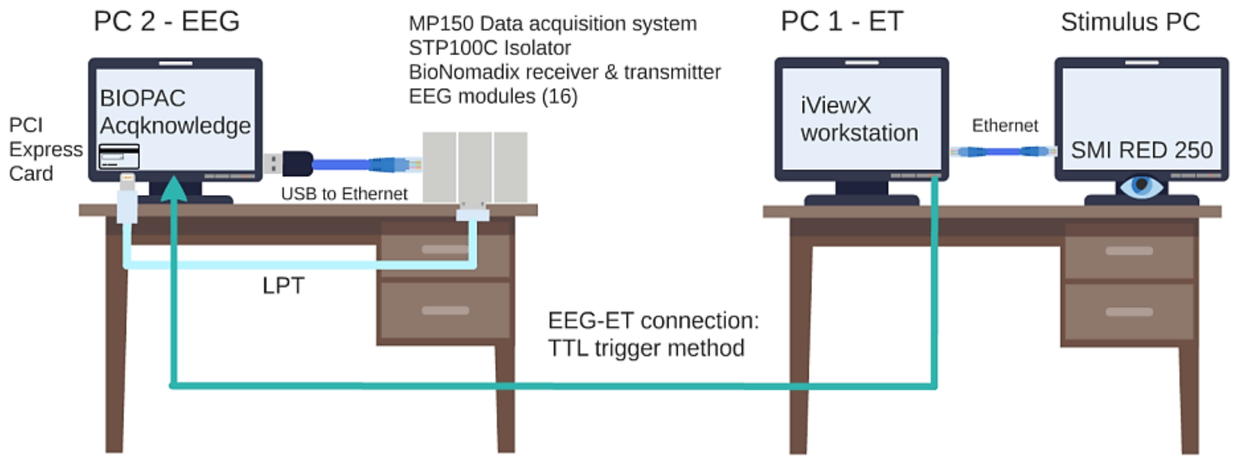


Figure 4.1. Synoptic diagram of our dual PC hardware set-up

4.2. Methodology

The (spatial) memory task in both user experiments focuses on the study process of the main structuring map elements (i.e., roads, green areas, and hydrography) of a map stimulus to be retrieved later. Accordingly, visual variables (e.g., shape, size, color, etc.) used for depicting those elements play an important role in the experimental design because we utilize them to design maps to be used either with less or more cognitive load. While roads contain only linear, and green areas contain only polygon features, hydrography contains both linear and polygon features. Inherently, recalling one or a combination of those can be linked to the different levels of task difficulty; hence, each different (or a group of) task is assumed to cause different cognitive loads. For instance, linear features are easier to learn and remember regardless of paying too much attention, and besides the color, the shape and size of map elements have an equally important impact on visuospatial memory (Keskin et al., 2018). Table 1.4 in Chapter 1 summarizes all the aspects of the experiment design for the first and second experiment. Please also see Annex 1 for the orientation scripts and Annex 2 for the full instructions of both experiment.

4.2.1. Experiment 1

In the first experiment, participants were asked to study a map stimulus as long as they would like in order to draw a sketch map of what they had studied. The map stimulus used in this experiment is a simplified topographic map that was produced by Belgian National Mapping Agency, NGI/IGN (Nationaal Geografisch Instituut/Institut GéographiqueNational), and was also used by Ooms (2012) and Keskin et al. (2018). According to the results of these studies, we hypothesized that the spatial memory task will cause higher cognitive load in novices. To explore participants' recalling strategies, we evaluated the drawn elements in the sketch maps and analyzed fixation related and AOI-based eye tracking metrics (for more detail about the experimental settings and results, please read Keskin et al. (2018).

This experiment resulted as single trials of one spatial memory task, but of long ET and EEG recordings due to the absence of time constraints. The ET metrics used as indicators of cognitive load were (i) average fixation duration and (ii) the number of fixations per second, which are commonly used metrics by many researchers when studying individual differences (e.g. Gedminas, 2011; Ooms, 2012; Togami, 1984). Average fixation duration is useful to study attentional procedures to one specific stimulus, whereas the number of fixations per second reveals the speed of attention (Bigne, Llinares and Torrecilla, 2016).

To extract the cognitive load from EEG data, we first averaged alpha power for all recording EEG channels and calculated Frontal Alpha Asymmetry (FAA) using frontal channels. Cognitive load can be measured using EEG activity power spectrum, and several researchers have repeatedly proved that the spectral power changes under alpha and theta frequency bands are related to task difficulty and therefore good predictors of cognitive load in a variety of working memory task demands (Gevins and Smith, 2000; Witvoet, 2019). These studies have found that alpha activity (particularly over the parietal and occipital areas) decreases with growing task demands that inherently cause working memory performance to decrease, whereas theta activity increases (especially over frontal midline areas) when encoding new information (Klimesch, 1999; Kumar and Kumar, 2016; Pfurtscheller and Da Silva,1999; Witvoet, 2019). A decrease in alpha power is a sign of attentional demands or comparatively high neuronal excitability (i.e., processing visual information or responding to internal events, e.g., mental activation or cognitive effort), while an increase in power reflects inhibition or cortical deactivation (Guay, De Beaumont, Drisdelle, Lina, Jolicoeur, 2018).

FAA is a commonly used measure for motivation, emotion, and cognitive control (e.g. Coan and Allen, 2003; Harmon-Jones, 2003). Greater relative left frontal activity is associated with increased memory & attentional performance and more-focused task performance (Lanini-Maggi, 2017). FAA is the average hemispheric difference in EEG alpha power between the left and right frontal regions of the brain during EEG recording (Adolph, Glischinski, Wannemüller, and Margraf, 2017; Quaedflieg, Smulders, Meyer, Peeters, Merckelbach, and Smeets, 2016; Smith, Reznik, Stewart, and Allen, 2017). We computed the alpha asymmetry using the left (F3) and right (F4) frontal channels with the following formula (Davidson, 1984) (Equation 4.1):

$$\text{FAA} = \log(\text{alpha F4}) - \log(\text{alpha F3}) \tag{4.1}$$

Since EEG power is inversely correlated with the activation, the negative alpha asymmetry scores correspond to greater relative right frontal activation, whereas positive ones indicate greater relative left frontal activity (Coan, Allen, McKnight, 2006; Davidson, 1984, 1998; Tomarken, Davidson, Wheeler, 1992). More activity in the left-frontal hemisphere indicates approach and motivation, whereas greater relative right activation refers to withdrawal and avoidance (Lanini-Maggi, 2017).

4.2.2. Experiment 2

The goal of the second experiment is fundamentally the same as the first one, but more complex, and there were some important modifications in terms of experimental design including participants, task and stimuli, procedures, and psychological measures to extract cognitive load. In this chapter, we will touch on all these aspects in detail.

The theoretical background for formulating the hypotheses of the second experiment was based on the observations in Experiment 1 (Keskin et al., 2018), whose outcomes will be presented in the results section, and the existing literature (Edler et al., 2014; Ooms et al., 2015). In this context, linear features are primary to construct the whole map, and therefore, they are easily accessible in working memory. We additionally expect that experts would perform better at tasks demanding higher cognitive load.

4.2.2.1. Participants

Since we intend to explore the influence of the task on cognitive load between expert and novice participants, within and between designs should be combined. In this context, both experts (5 females, 6 males), novices (6 females, 5 males), whose age and gender match (N= 22, MED= 27.5, SD= 3.9), performed the same experiment under the same conditions. The dataset used for the combined eye tracking and EEG analysis is a subset of a larger dataset and the user characteristics of the recruited participants are given in Table 1.4 and Annex 6.

4.2.2.2. Task and Stimuli

The spatial memory task in this experiment focuses on the retrieval of the main structuring elements with varying difficulty levels. Compared to the first experiment, we increased the number of stimuli of interest and the reference stimuli (e.g., fixation cross), presented them as randomized blocks, added more levels of task difficulty, put time constraints in place, and allowed participants to select from multiple choices instead of drawing the sketch maps themselves as applied in the first experiment.

Next to the fixation crosses used as a pre-stimulus reference, the experiment included two types of visual materials: (i) original map stimuli to be studied and (ii) the corresponding skeleton maps displayed on the graphical response screens. The original map stimuli were acquired from Google maps at zoom level 15 with 1 km scale bar (since the resolution of a map with the Mercator projection is dependent on the latitude, the scale of the maps (collected from regions all around the world) varies slightly but is approximately 1:40000). The skeleton maps are the simplified representations of map stimuli indicating the main structuring map elements of interest for that specific task and were prepared by digitizing the main structuring map elements on the original stimuli using a GIS software. Throughout the design of the skeleton maps used in the experiment, we paid attention to depict each map feature class with a unique color and to make sure that these colors remained true to the ones used in the original stimuli. Accordingly, the main roads were assigned to yellow, major hydrographic features to light blue and the green areas to light green. The maps (1344 × 768 pixels, 14′ × 8′) and the graphical response screens including four panels (576 × 326 pixels, 6′ × 3.4′) were shown on a 22″ color monitor with 1680 × 1050 spatial resolution.

Figure 4.2. Example stimulus and experiment blocks a. Original stimulus, b. Block 1: The whole map, c. Block 2: Roads and hydrography, d. Block 3: Roads and green areas, e. Block 4: Green areas and hydrography, f. Block 5: Green areas, g. Block 6: Hydrography, h. Block 7: Roads.

Tasks including the same number of trials related to the same map element were classified as blocks. For the randomization of stimuli used in trials, randomized block design was used and in total seven blocks of trials were designed. Each block consisted of one trial for each stimulus (i.e., 50 trials within a block) focusing on the similarity of one of the criteria listed in Figure 4.2: the main structuring elements of (b) the whole map, (c) roads and hydrography, (d) roads and green areas, (e) green areas and hydrography, (f) green areas, (g) hydrography, and (h) roads. The trials in Block 1 were designed to study the recalling performance related to the entire map stimulus; therefore, the skeleton maps were prepared by digitizing all the main roads, all the major hydrographic features and the green areas on the original map stimuli. The trials included in Block 2, 3 and 4 were dedicated to the retrieval of the combination of two map feature classes. In this case, Block 2 refers to the main roads and major hydrographic features, whereas Block 3 addresses the main roads and green areas, and Block 4 involves major hydrographic features and green areas. The trials belonging to Block 5, 6 and 7 deal with a single map feature class; either green areas, hydrographic or road features, respectively, and each of them were digitized individually on the original stimuli.

One important concern about the design is that the task difficulty may not be predicted easily in advance, because it depends on many factors rather than only the number of object classes to remember. According to the average reaction time of the correct answers (i.e. inverse efficiency score) provided by all participants, we observed clustering among some blocks and natural breaks between those clusters (as explained in Chapter 3.2.1.3). Subsequently, Block 1 and 2 were designated as hard; Block 3 and 4 as moderate; and the rest were assigned to easy level (Figure 4.3). By this way, the blocks falling into the same category can be treated similarly when analyzing and interpreting the gaze and neurophysiological data (i.e., eye tracking, EEG) collected during the entire experiment.

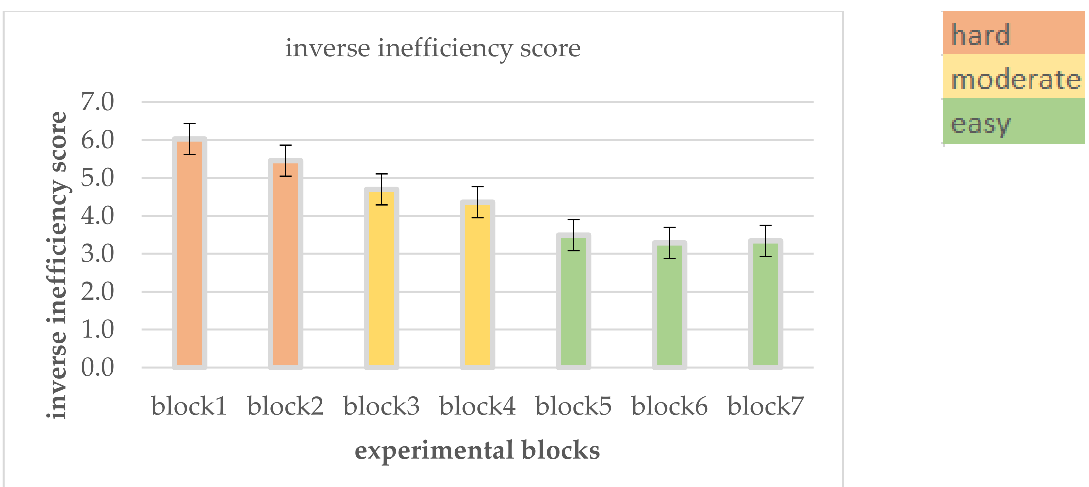


Figure 4.3. Experiment blocks and difficulty levels

4.2.2.3. Procedures

Measuring the cognitive load is linked to how a participant indicates a correct answer on the response screen presented to her/him, and reaction times of key presses are a simple and rather reliable way to measure it. During the trial, participants were first asked to study a map stimulus and during the stimulus presentation, they were free to shift their gaze across the display. The response screen appeared with four graphical response panels that shows skeleton maps indicating specific main structuring map elements (Figure 4). Only one of the panels corresponded to the map that a participant just saw (a correct response).

Participants were instructed to press the space bar immediately when they found the panel with the correct skeleton map and to remember the corresponding letter.

Pressing the space bar indicated that the search was complete by allowing participants to move to the second response screen where they would see only the letters (i.e., no pictures) (Figure 4.4). They should click on the letter, which they were keeping in memory to complete the task. If multiple features were needed to be remembered (e.g., roads and hydrography), a participant might remember only one type (e.g., hydrography), and then find a correct skeleton map based only on this type of information. Thus, the options in the graphical response panels assured that a response based on partial information was impossible. Additionally, the possible answers (correct skeleton maps) appeared at different locations between each consecutive trial and the block orders were counter-balanced across participants. Overall, each participant had to complete all seven blocks (please see Annex 2 & 3).

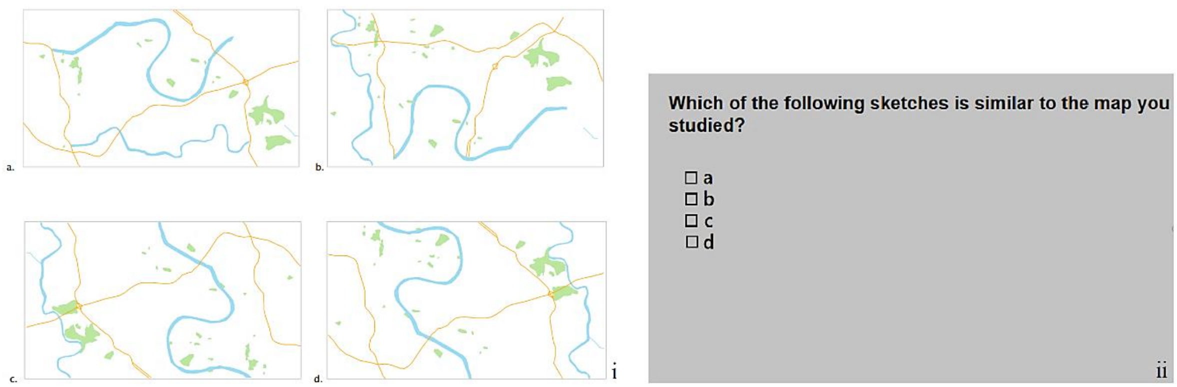


Figure 4.4. i. Graphical response screen, ii. Response screen with letters

4.2.2.4. Psychological Measures to Use: ET & EEG Metrics

We used the same eye-tracking metrics employed in the first experiment to extract the cognitive load: average number of fixations per second and fixation durations for each trial.

Events refer to the time points where the stimuli of interest are presented to the participants. During a cognitive task, event-related power changes in EEG bands can be calculated in a specific frequency band. As explained in Chapter 3, if the event-related power decreases, it causes a reduction of amplitude in response to a stimulus, and therefore is called event-related desynchronization (ERD), whereas power increases result in an increment of amplitude with stimulus presentation, and hence, are referred to as event-related synchronization (ERS) (Gevins and Smith, 2000; Pfurtscheller and Da Silva, 1999). Many studies suggest that alpha activity decreases (i.e., desynchronizes) and theta activity increases (i.e., synchronizes) as task difficulty increases (e.g. Anderson et al., 2011; Gevins et al., 1997; Gevins and Smith, 2003; Sauseng et al, 2005).

To be able to extract the alpha and theta spectral powers, the EEG data went through a series of preprocessing steps. For handling EEG and ET data together, we decided to use EEGLAB, an open source and interactive MATLAB toolbox (Delorme and Makeig, 2012), with the EYEEEG extension (Dimigen, Sommer, Hohlfeld, Jacobs, and Kliegl, 2011). EEGLAB processes continuous and event-related EEG and other electrophysiological data (supports data from most of the commercially available software), and performs time/frequency analysis, artifact rejection, event-related statistics, and visualization of averaged or single-trial EEG data. Figure 4.5 demonstrates the pre-processed (i.e., filtered, bad channels removed, events added and modified based on correct responses, re-referenced, and segmented) EEG recordings belonging to an expert female participant. The vertical axis shows the amplitude (μV), i.e., the amount of energy in artifact-free EEG frequency bands listed on the left-hand side of the graph, whereas the horizontal axis represents time in seconds. The vertical lines on the graph labeled with vertical lettering (e.g., 148, 149, 150) are the event codes, and the intervals represented between the blue vertical lines and numbers above the upper part border of the graph (e.g., 16, 17, 18) indicate the epochs.

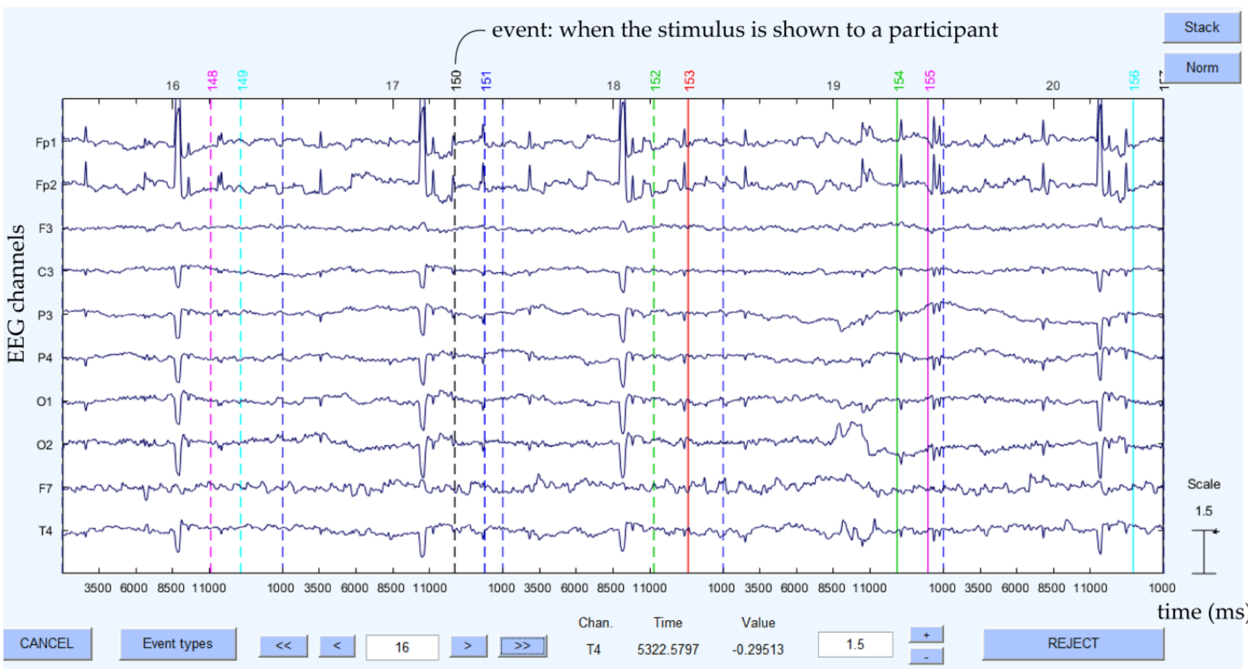


Figure 4.5. Preprocessed EEG data

Once EEG data had gone through preprocessing steps, we segmented it based on trials. Figure 4.6 demonstrates the trial sequence of the experiment. To be able to calculate event-related power change at an electrode, we created epochs from the events of our interest based on two different intervals:

- [0 2] s for the events in the reference interval - fixation crosses
- [2 9] s for the events in the activation interval - map stimuli

Bad epochs containing blink or muscle artifacts were rejected based on visual inspection and collected eye-tracking data. Prior to epoching, we synchronized the EEG recording with its corresponding ET recording through shared events present in the ET and EEG: start-event and end-event. Although the time synchronization accuracy of our system was not sufficient for studying eye-fixation-related potentials, fixation and saccade detection on EEG helps explaining the EEG spikes elicited from the eye movements.

Therefore, we think offline synchronization of ET data is still useful for artifact rejection and excluding the epochs contaminated with blinks (Figure 4.7).

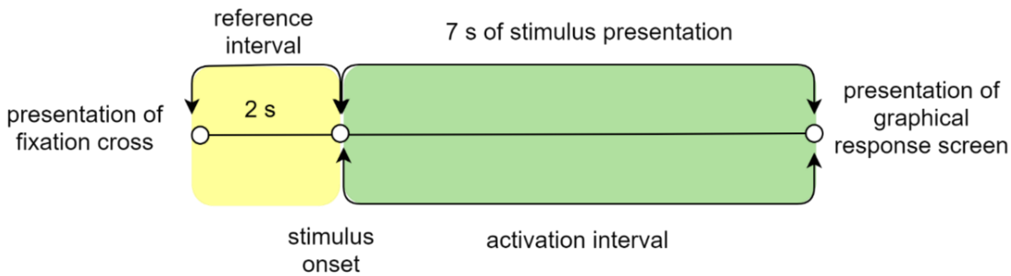


Figure 4.6. Trial sequence. The fixation cross was followed by the stimulus presentation; the stimulus remained visible throughout the study time. Activation interval ended with the presentation of a graphical response screen.

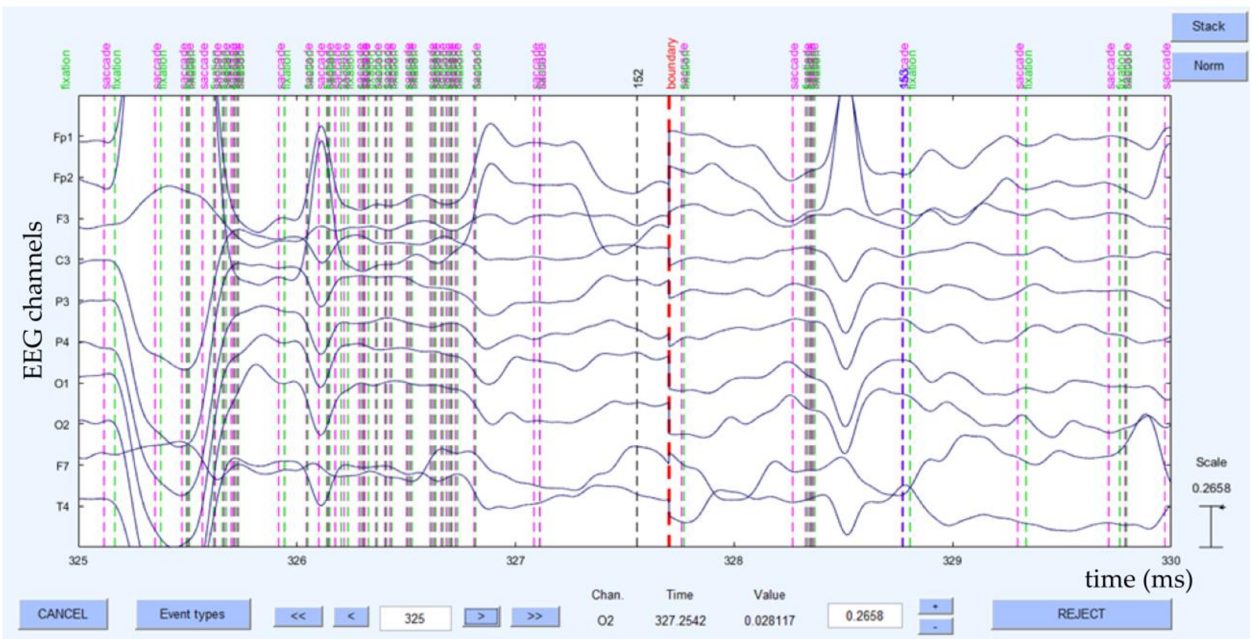


Figure 4.7. EEG and eye movement data synchronized through shared events.

For the computations of the changes in EEG power spectral density (PSD), first, the band power of the EEG signal was computed by means of a time–frequency analysis that employs a standard Fast Fourier Transform (FFT). FFT transforms the EEG signal from the time domain into the frequency domain. Therefore, any time-dependent signal can be broken down into a collection of sinusoids, and EEG recordings can be plotted in a frequency power-spectrum. After the transformation, we averaged the spectral power of alpha (8.5–12.5 Hz) and theta (4.5–6.5 Hz) bands for our 7-seconds-long EEG recording (i.e., duration of the stimulus on the screen, activation period) of valid trials in each block.

We first computed EEG power spectral density (PSD) for alpha and theta frequency bands and to extract the cognitive load, event-related power changes can be quantified by contrasting the power in a specified frequency band during a cognitive task (e.g., spatial memory) with a preceding reference interval (i.e., ERD

& ERS) (for detailed information please read Pfurtscheller, and Da Silva, 1999). In this context, the baseline (pre-stimulus) period of EEG was used to compare with the event-related EEG power dynamics during the activation intervals in each epoch (Chuang, Cao, King, Wu, Wang, and Lin, 2018). Event-related power change (ERP) at an electrode was obtained by subtracting the log-transformed power during pre-stimulus reference intervals from the log-transformed power during the activation intervals according to the following formula (Benedek, Schickel, Jauk, Fink & Neubauer, 2014) (Equation 4.2).

$$ERP(i) = \log(Powi, activation) - \log(Powi, reference), \quad (4.2)$$

Note that this ERP should not be confused with the commonly used abbreviation for event-related potentials in EEG domain. After computing ERPs at alpha and theta frequency bands for all task difficulty levels considering expertise, the powers were compared to study the differences between expert and novices particularly based on low and high levels of complexity of tasks.

4.3. Results

4.3.1. Experiment 1

The average duration of the fixations was 230.0 ms (*N*= 24, *MED*= 230.8 ms, *SD*= 50.1 ms) for experts and 244.1 ms (*N*= 30, *MED*= 243.0 ms, *SD*= 48.4 ms) for novices. Two-way ANOVA test suggested that no significant difference occurred between experts and novices ($F(1,55)$= 0.074, *p*= 0.787). The average number of fixations per second for the map stimulus was 3.5 (*N*= 24, *MED*= 3.7, *SD*= 1.0) for experts and 3.6 (*N*= 30, *MED*= 3.6, *SD*= 0.5) for novices. Regarding the Mann-Whitney U test, the difference between these two user groups was not statistically significant ($U_{expertise}$= 338, *p*= 0.702) (for detailed results, please read Keskin et al., 2018).

The average alpha power across all usable common EEG electrodes (i.e., C3, F3, F4, O1, P3, T5, T6) for all participants (usable data: 6 novices, 6 experts) was 0.000939 (*SD*= 0.000051, *range*= 0.000225 - 0.002218). Shapiro–Wilk test was used to test the normality of the distribution of the data since our dataset is smaller than 2000 samples (*N*= 84). p= 0.000 suggested strong evidence that the data was not normally distributed ($D(84)$= 0.930, $p < 0.05$). The difference of 0.000171 in alpha power between experts (M = 0.001282, SD = 0.000064) and novices (*M*= 0.000853, *SD*= 0.0000777) was statistically significant according to non-parametric Man–Whitney U test (*p*= 0.024 < 0.05). The greater alpha power is associated with the lesser cognitive load, therefore, the results indicate that experts spent considerably lesser cognitive load on this memory task compared to novices. This outcome was important because while sketch map evaluation and ET metrics claimed the other way, EEG alpha power provided an additional insight referring to a significant difference in the spatial memory performance of experts and novices.

For the memory task, average FAA score across participants (usable data: 7 novices, 10 experts) was −0.149 (*SD*= 0.275, *range*= −0.810 to 0.160). According to the Shapiro–Wilk test, p= 0.006 showed that the data was normally distributed ($D(17)$= 0.870, $p > 0.05$), therefore, we applied two-way ANOVA to explore whether the difference between expert and novice groups was statistically significant. Novices (*M*= −0.054, *SD*= 0.252) and experts (*M*= −0.216, *SD*= 0.283) showed no significant difference in FAA scores ($F(1,15)$= 0.199, *p*= 0.245). However, 70% of experts had negative scores on this metric, which reflects greater relative right activation, suggesting withdrawal-related motivation. Although the average FAA scores were negative for novices, 57% of them exhibited larger left-hemispheric activation, which is an indicator of approach-oriented motivation and positive affective states.

4.3.2. Experiment 2

ET results are shown in Figure 4.8. Fixation durations of novices were longer, and the difference between experts and novices increased as the difficulty increased. For the hard tasks, this difference was the highest. On the contrary, the number of fixations (per second) of experts was higher, and the difference increased as the difficulty decreased. Therefore, these two groups differed the most for the easy tasks. The eye movement data for both metrics fit normal distribution (Shapiro–Wilk test) for easy and moderate tasks. For these two categories of task difficulty, no statistically significant difference emerged between experts and novices in terms of the average fixation duration (F_{easy}= 0.261, p= 0.232; $F_{moderate}$= 0.174, p= 0.514). Additionally, no significant interaction effect was observed between expertise and task difficulty ($F_{(2,22)}$= 0.208, p= 0.814).

The difference in the number of fixations per second was not significant ($F_{moderate}$= 1.861, p= 0.165) for moderate tasks, whereas it was significant for easy tasks (F_{easy}= 0.006, p = 0.019). For the hard tasks, the average fixation durations were not normally distributed across participants (Shapiro–Wilk test) and we observed no statistically significant difference between expert and novice groups (Mann–Whitney U, p = 0.886). The data for the number of fixations per second fit the normal distribution (Shapiro–Wilk test), and no significant difference occurred between the two groups based on two-way ANOVA test (F_{hard}= 0.064, p= 0.983). Moreover, the interaction between expertise and task difficulty was not significant ($F_{(2,22)}$= 1.616, p= 0.221).

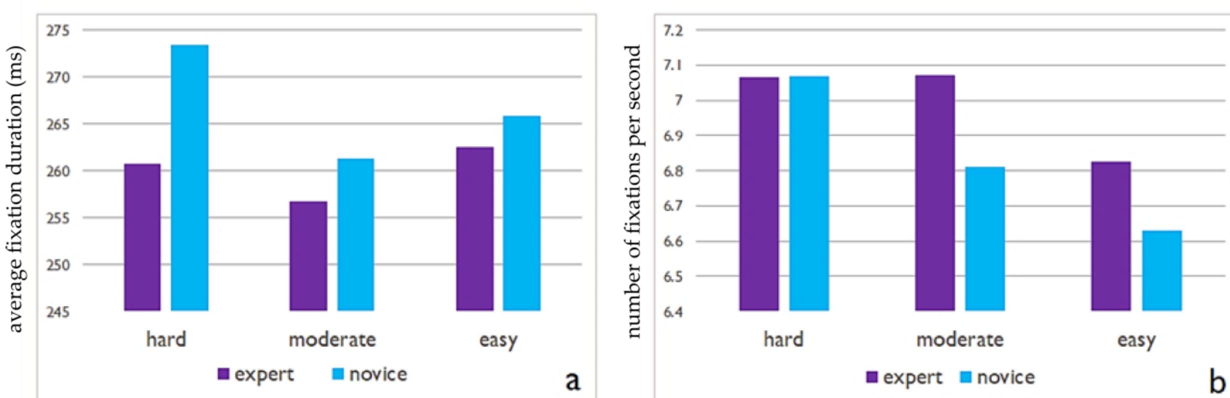


Figure 4.8. ET metrics belonging to Experiment 2 (a. average fixation duration (ms), b. the number fixations per second)

Figure 4.9 depicts the ERPs in theta and alpha averaged for a novice male for Block 1 and Block 2. Here, we would like to show how individual data might include inconsistencies, although we observed negative alpha power and on the contrary, positive and relatively higher powers in theta frequency in most EEG channels (see Table 4.2). Frontal channels (e.g., Fp1, F3, F7, and F8) might not be trusted because they might still contain small blink artifacts acting as confounding effects, however, except for that, we usually observed ERD in alpha and ERS in theta power. Obviously, the cognitive load cannot be interpreted based on one or two participant data for a single block. The overall results will be of aggregating blocks based on task difficulty (i.e., easy, moderate, hard) and averaging many trials of many participants for every difficulty level. However, the preliminary data analysis seems promising for further analysis of the EEG power spectrum. With this study, we attempted to verify the proposed methodology and prove that with our experiment design and hardware & software set up, it is possible to synchronize ET and EEG data to obtain more detailed insight on user behaviors and observe the EEG metrics, alpha and theta power.

Table 4.1. The event-related changes in EEG power spectral density (PSD) for alpha and theta frequency bands between activation and reference intervals.

EEG channels	ERP theta (4-8 Hz)	ERP alpha (8-13 Hz)	ERP theta (4-8 Hz)	ERP alpha (8-13 Hz)
	Block 1 ($\mu V^2/Hz$)		Block 2 ($\mu V^2/Hz$)	
C3	$4.76 x 10^{-5}$	$-5.83 x 10^{-6}$	$8.57 * 10^{-1}$	$-4.92 * 10^{-2}$
F3	$-1.37 x 10^{-5}$	$-3.87 x 10^{-6}$	$4.58 * 10^{-1}$	$-1.15 * 10^{-2}$
F7	$2.58 x 10^{-4}$	$2.59 x 10^{-7}$	3.15	$9.89 * 10^{-2}$
F8	$3.25 x 10^{-4}$	$8.36 x 10^{-6}$	5.23	$-2.18 * 10^{-3}$
Fp1	$5.42 x 10^{-4}$	$-1.54 x 10^{-6}$	6.15	$-1.7 * 10^{-2}$
Fp2	$7.15 x 10^{-4}$	$3.22 x 10^{-6}$	7.13	$3.49 * 10^{-3}$
O1	$3.29 x 10^{-5}$	$-2.76 x 10^{-6}$	$9.32 * 10^{-1}$	$-6.22 * 10^{-2}$
O2	$3.11 x 10^{-5}$	$-1.23 x 10^{-5}$	$5.41 * 10^{-1}$	$-7.39 * 10^{-2}$
P3	$1.11 x 10^{-4}$	$-3.49 * 10^{-6}$	1.75	$-3.83 * 10^{-2}$
P4	$6.03 x 10^{-5}$	$-5.18 x 10^{-6}$	$3.72 * 10^{-1}$	$-2.32 * 10^{-1}$
T4	$9.72 x 10^{-5}$	$-5.30 x 10^{-6}$	1.05	$-1.54 * 10^{-2}$

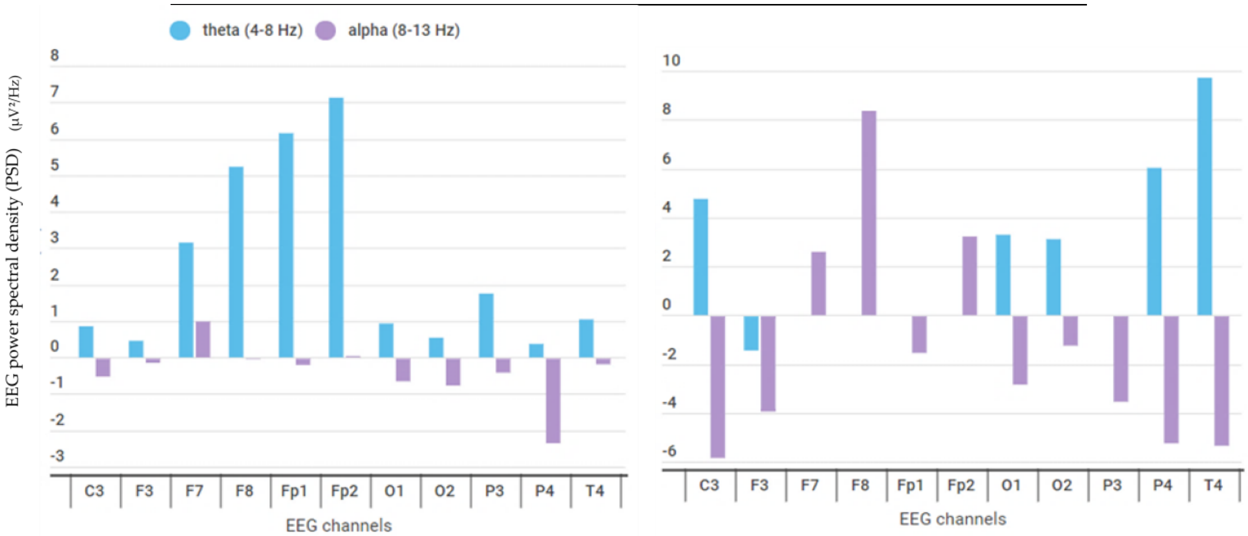


Figure 4.9. The event-related changes in EEG power spectral density (PSD) for alpha and theta frequency bands (left figure for Block 1, right figure for Block 2) ($\mu V^2/Hz$).

4.4. Discussion

ET metrics in the first experiment showed that there was no significant difference between expert and novice participants, similar to what was found in the second experiment, except for novices exhibiting a lesser number of fixations for easy tasks. This finding is interesting considering our hypotheses and it can be evidence that expert and novices use similar strategies for moderate and hard tasks; however, novices might think about easy tasks more deeply or they find even easy tasks more confusing. On the other hand, we observed more fixations and longer fixation durations for novices in hard tasks, and a similar situation applies for experts in easy tasks. To be able to interpret this outcome, we can look into saccade-related metrics because of longer search times, more fixations, shorter saccades, and longer fixation durations with increasing crowding and decreasing span size (Vlaskamp, and Hoog, 2006). Triangulating ET data with EEG data might also contribute to judging this result better, therefore, there is still a lot of work to do in terms of further analysis. For instance, while ET metrics do not differ across different conditions, EEG metrics argue otherwise (e.g. Gedminas, 2011).

Studying the EEG metrics indicating the cognitive load suggested an important insight on map users and seems assuring to be integrated as a complementary methodology and a way of assuring the validity of research. However, EEG requires a quite extensive experience to acquire, analyze and interpret the data, and one of the motivations of this research was to emphasize the importance of the experiment design, especially when EEG comes into play.

On the one hand, as methodological decisions are highly dependent on the research questions and hypotheses regarding them, it is important to describe a solid objective for the user study with psychological design principles in mind and to identify the key metrics answering the research questions. On the other hand, although the experiment within this study is limited to a spatial memory task and the methodological design of the other experiments may vary on a large scale, the technical issues to overcome and the preprocessing steps of the collected data are valid for almost all ET&EEG experiments. Recording EEG and ET data in free-viewing tasks has been a challenge and rarely applied, especially due to the precise co-registration of gaze position. To minimize the muscle artifacts due to unnatural sitting positions, using a chin rest and adapting the position of the participant is crucial; besides, this makes sure that the participant has enough rest between blocks so that they do not exhibit fatigue and move as little as possible during the experiment. Electromagnetic artifacts that are elicited from other electrical devices and introduced as line noise in EEG data should be identified and filtered out. For accurate synchronization of both EEG and ET data records, Transistor-transistor logic (TTL) triggers is preferred as it is the most straightforward and reliable method (e.g. Dimigen, and Reinacher, 2013; Nikolaev, Meghanathan, and van Leeuwen, 2016). Although proper synchronization can be achieved with the TTL trigger method, in our experiment, the monitor offset value limited studying the eye-fixation-related potentials requiring high temporal resolution in terms of synchronization of EEG and ET. However, our experiment setting allows for studying the EEG activity power spectrum, and ET data can still be synchronized offline and ET metrics can be correlated with EEG data on a trial basis. Therefore, the feasibility of the methodology should always be verified in advance considering the possible technical constraints related to the recording equipment.

Although some procedures such as data management (e.g., converting data into a compatible format with EEGLAB) and noise filtering (e.g., applying high- and low-pass filters on the fly) can be automatized, many other steps such as bad channel removal, which is mostly carried out by visual inspection, are performed manually. In addition, preprocessing and analyzing the data is inherently the most labor-intensive and complicated part of the study. Since each participant's data consists of a number of trials and should be handled individually, the processing stage is overall very time-consuming.

4.5. Conclusions

We presented two cartographic user experiments first to demonstrate what is possible with the co-registration of EEG and ET and to investigate the spectral characteristics of cognitive processes in free viewing conditions, only within the frame of the specific spatial memory task described for this study. Our results showed that EEG can be employed as a complementary technique to get a detailed insight about user actions and behaviors and reveal the information that we did not observe with eye tracking. While eye tracking metrics in the first experiment demonstrated that the difference between experts and novices are not significant, the EEG alpha power analysis suggested that this difference was significant, indicating that this specific spatial memory task caused more cognitive load in novices. Therefore, triangulating EEG and ET data seems useful to be able draw conclusions on user's behavior and also shows that the data require more investigation.

Although the analysis of the second experiment is still in progress, preliminary results of event-related power changes in alpha and theta allowed us to estimate the variations in the cognitive load that a certain task demands. The future work will focus on alpha & theta power computations considering both user groups and varying task difficulties. In this respect, alpha and theta power changes will be averaged for easy, moderate and hard tasks considering experts and novices to explore the influence of expertise on the cognitive load. By this way, we will be able to tell whether there is a difference across participants, and if so, how much this difference is and how significant it is. Having ET metrics calculated, we will then link and correlate them with EEG metrics to estimate the overall cognitive load.

Combining EEG and ET is not straightforward since there are numerous methodological and technical problems to overcome, yet it is indeed a very valuable technique to explore the individual differences and similarities of map users through perceptual and cognitive procedures. If we continue staying engaged with experimental psychology and cognitive science research, it will contribute to the future progress of scientific cartography. The more we know about the limitations and capabilities of visual perception and cognition of different map users, the higher the possibilities to design cartographic products in a more efficient, understandable and effective way.

References

Adolph, D., von Glischinski, M., Wannemüller, A., & Margraf, J. (2017). The influence of frontal alpha-asymmetry on the processing of approach-and withdrawal-related stimuli—A multichannel psychophysiology study. Psychophysiology, 54(9), 1295-1310.

Anderson, E. W., Potter, K. C., Matzen, L. E., Shepherd, J. F., Preston, G. A., & Silva, C. T. (2011, June). A user study of visualization effectiveness using EEG and cognitive load. In Computer Graphics Forum (Vol. 30, No. 3, 791-800). Blackwell Publishing Ltd.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes1. Psychology of learning and motivation, 2(1), 89-195).

Benedek, M., Schickel, R. J., Jauk, E., Fink, A., & Neubauer, A. C. (2014). Alpha power increases in right parietal cortex reflects focused internal attention. Neuropsychologia, 56(100), 393–400.

Bertin, J. (1967). Semiology of graphics, (W. J. Berg Trans.). Madison, WI, USA: University of Wisconsin Press.

Chuang, C. H., Cao, Z., King, J. T., Wu, B. S., Wang, Y. K., & Lin, C. T. (2018). Brain electrodynamic and hemodynamic signatures against fatigue during driving. Frontiers in neuroscience, 12, 181.

Bigne, E., Llinares, C., & Torrecilla, C. (2016). Elapsed time on first buying triggers brand choices within a category: A virtual reality-based study. Journal of Business Research, 69(4), 1423-1427.

Coan, J. A., & Allen, J. J. (2003). State and trait of frontal EEG asymmetry in emotion. The asymmetrical brain, 566-615.

Coan, J. A., Allen, J. J. B., & McKnight, P. E. (2006). A capability model of individual differences in frontal EEG asymmetry. Biological Psychology, 72, 198–207.

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

Davidson, R.J. (1984). Affect, cognition, and hemispheric specialization. In Emotion, Cognition, and Behavior; Izard, C.E., Kagan, J., Zajonc, R.B. (Eds.), Cambridge University Press: Cambridge, UK, 320–365.

Davidson, R. J. (1998). Affective style and affective disorders: Perspectives from affective neuroscience. Cognition & Emotion, 12(3), 307-330.

Dimigen, O., & Reinacher, U. (2013, January). A MATLAB Tool for Simultaneous Eye Tracking and EEG. In Journal of Cognitive Neuroscience, 143-144.

Delorme, A., & Makeig, S. (2012). EEGLAB tutorial outline, Retrieved from:
http://cognitrn.psych.indiana.edu/busey/temp/eeglabtutorial4.301/eeglabtut.html Last accessed: 17.01.2019.

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: analyses and review. Journal of Experimental Psychology: General, 140(4), 552–572.

Duchowski, A. T. (2007). Eye tracking methodology - Theory and practice. London: Springer: 328(614), 2-3.

Eckert, M. (1908). (trans. By W. Jeorg). On the Nature of Maps and Map Logic. Bulletin of the American Geographical Society, 40(6), 344-351.

Edler, D., Bestgen, A. K., Kuchinke, L., & Dickmann, F. (2014). Grids in topographic maps reduce distortions in the recall of learned object locations. PloS one, 9(5), e98148.

Fink, A., Grabner, R. H., Benedek, M., & Neubauer, A. C. (2006). Divergent thinking training is related to frontal electroencephalogram alpha synchronization. European Journal of Neuroscience, 23(8), 2241-2246.

Gedminas, L. (2011). Evaluating hurricane advisories using eye-tracking and biometric data. Master's Thesis, East Carolina University: North Carolina.

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cerebral cortex (New York, NY: 1991), 7(4), 374-385.

Gevins, A., & Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. Cerebral cortex, 10(9), 829-839.

Gevins, A., & Smith, M. E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. Theoretical Issues in Ergonomics Science, 4(1-2), 113-131.

Griffin, A. L. (2017). Cartography, visual perception and cognitive psychology. The Routledge handbook of mapping and cartography, 44-54.

Guay, S., De Beaumont, L., Drisdelle, B. L., Lina, J. M., & Jolicoeur, P. (2018). Electrophysiological impact of multiple concussions in asymptomatic athletes: a re-analysis based on alpha activity during a visual-spatial attention task. Neuropsychologia, 108, 42-49.

Harmon-Jones, E. (2003). Clarifying the emotive functions of asymmetrical frontal cortical activity. Psychophysiology, 40(6), 838-848.

Herbert, G., & Chen, X. (2015). A comparison of usefulness of 2D and 3D representations of urban planning. Cartography and Geographic Information Science, 42(1), 22-32.

Jacob, R., & Karn, K. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In: R. Radach, J. Hyona & H. Deubel (Eds.), The mind's eye: Cognitive and applied aspects of eye movement research(573-605). Amsterdam: Elsevier.

Keskin, M., Ooms, K., Dogru, A. O., & De Maeyer, P. (2018). Digital sketch maps and eye tracking statistics as instruments to obtain insights into spatial cognition. Journal of Eye Movement Research, 11(3).

Kimerling, A. J., Buckley, A. R., Muehrcke, P. C., & Muehrcke, J. O. (2009). Map use: reading and analysis. Esri Press

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain research reviews, 29(2-3), 169-195.

Kumar, N., & Kumar, J. (2016). Measurement of Cognitive Load in HCI Systems Using EEG Power Spectrum: An Experimental Study. Procedia Computer Sci. 84, 70-78.

Kveladze, I., Kraak, M. J., & van Elzakker, C. P. (2017, July). The Usability of a GeoVisual Analytics Environment for the Exploration and Analysis of Different Datasets. In International Cartographic Conference (p. 207-221). Springer, Cham.

Lanini-Maggi, S. (2017). Depicting movement data with animations for embodied and real-time decision-making: a user study with air traffic control displays and real-time movement data. Doctoral dissertation, University of Zurich, Zurich.

MacEachren, A. M. (2004). How maps work: representation, visualization, and design. New York, NY: Guilford Press.

Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency. Neuroscience & Biobehavioral Reviews, 33(7), 1004-1023.

Nikolaev, A. R., Meghanathan, R. N., & van Leeuwen, C. (2016). Combining EEG and eye movement recording in free viewing: Pitfalls and possibilities. Brain and cognition, 107, 55-83.

Ooms, K. (2012). Maps, how do users see them?: an in depth investigation of the map users' cognitive processes. Doctoral dissertation, Ghent University, Ghent.

Ooms, K., De Maeyer, P., & Fack, V. (2015). Listen to the map user: cognition, memory, and expertise. The Cartographic Journal, 52(1), 3-19.

Ooms, K., De Maeyer, P., Dupont, L., Van Der Veken, N., Van de Weghe, N., & Verplaetse, S. (2016). Education in cartography: what is the status of young people’s map-reading skills?. Cartography and Geographic Information Science, 43(2), 134-153.

Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. Clinical neurophysiology, 110(11), 1842-1857.

Quaedflieg, C. W., Smulders, F. T., Meyer, T., Peeters, F. P. M. L., Merckelbach, H. L. G. J., & Smeets, T. (2015). The validity of individual frontal alpha asymmetry EEG neurofeedback. Social cognitive and affective neuroscience, 11(1), 33-43.

Sauseng, P., Klimesch, W., Schabus, M., & Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. International Journal of Psychophysiology, 57(2), 97–103.

Smith, E. E., Reznik, S. J., Stewart, J. L., & Allen, J. J. (2017). Assessing and conceptualizing frontal EEG asymmetry: An updated primer on recording, processing, analyzing, and interpreting frontal alpha asymmetry. International Journal of Psychophysiology, 111, 98-114.

Togami, H. (1984). Affects on visual search performance of individual differences in fixation time and number of fixations. Ergonomics, 27(7), 789-799.

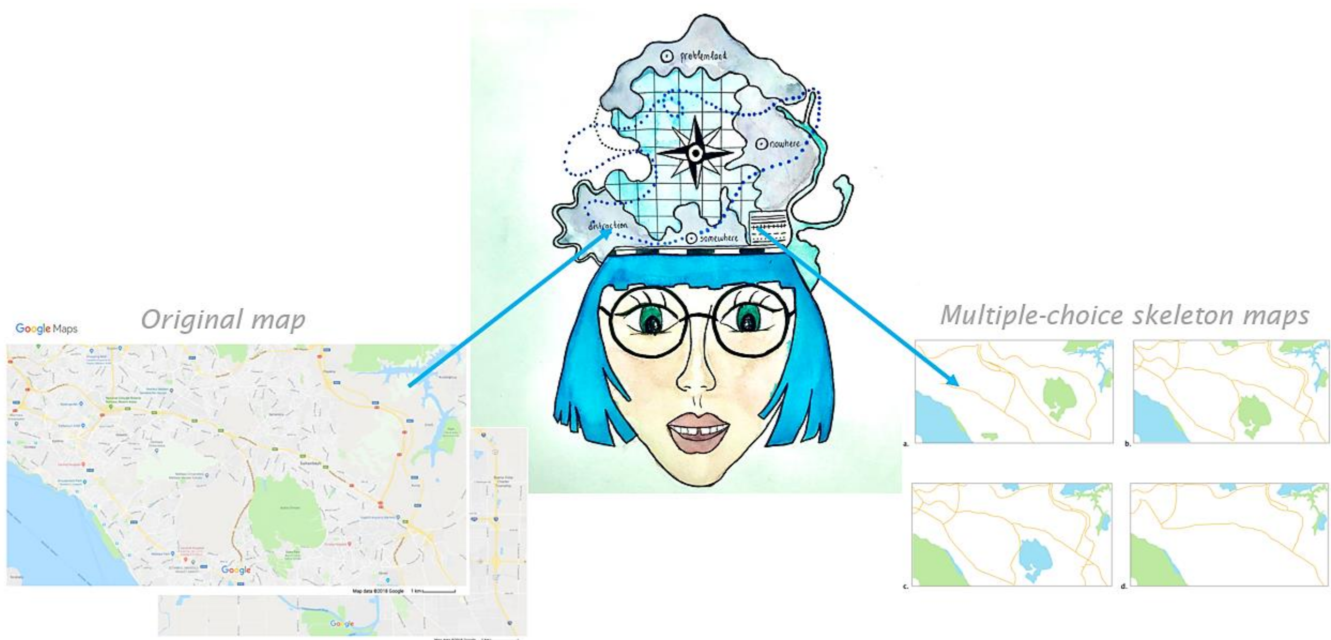
Tomarken, A. J., Davidson, R. J., Wheeler, R. E., & Doss, R. C. (1992). Individual differences in anterior brain asymmetry and fundamental dimensions of emotion. Journal of Personality and Social Psychology, 62, 676–87

Vlaskamp, B. N., & Hooge, I. T. C. (2006). Crowding degrades saccadic search performance. Vision research, 46(3), 417-425.

Winslow, B., Carpenter, A., Flint, J., Wang, X., Tomasetti, D., Johnston, M., & Hale, K. (2013). Combining EEG and Eye Tracking: Using Fixation-Locked Potentials in Visual Search. Journal of Eye Movement Research, 6(4), 1-11.

Witvoet, J. (2013). Does cognitive load influence performance in a game-based learning task? Bachelor's thesis, University of Twente, Enschede.

Chapter 5:
Experiment 2: Eye tracking & EEG



"One of the most pervasive mistakes is to believe that
our visual system gives a faithful representation of what is "out there"
in the same way that a movie camera would.
…
You are not perceiving what is out there.
You are perceiving whatever your brain tells you."

David Eagleman, Incognito: The Secret Lives of the Brain, 2011

Abstract. *The main objective of this chapter is to explore the cognitive processes of a group of expert and novice map users during the retrieval of map-related information within varying difficulty levels (i.e. easy, moderate, hard) by using eye tracking and EEG. In this context, we present a spatial memory experiment consisting of a large number of stimuli to study the effect of task difficulty on map users' behavior through cognitive load measurements. Next to the reaction time and success rate, we used fixation and saccade related eye tracking metrics (i.e., average fixation duration, the number of fixations per second, saccade amplitude and saccade velocity), and EEG power spectrum (i.e. event-related power changes in alpha and theta frequency bands) to identify the cognitive load. While fixation metrics indicated no statistically significant difference between experts and novices, saccade metrics proved the otherwise. EEG power spectral analysis, on the other side, suggested an increase in theta power (i.e. event-related synchronization) and a decrease in alpha power (except moderate tasks) (i.e. event-related desynchronization) at all difficulty levels of the task for both experts and novices, which is an indicator of cognitive load. Although no significant difference emerged between two groups, we found a significant difference in their overall performances when the participants were classified as good and relatively bad learners. Triangulating EEG results with the recorded eye tracking data and the qualitative analysis of focus maps indeed provided a detailed insight on the differences of the individuals' cognitive processes during this spatial memory task.*

Author Contributions: *Conceptualization: Merve Keskin, Kristien Ooms; Methodology: Merve Keskin; Software: Merve Keskin; Validation: Ahmet Ozgur Dogru; Formal Analysis: Merve Keskin; Investigation: Kristien Ooms, Ahmet Ozgur Dogru; Resources: Philippe De Maeyer; Data Curation: Merve Keskin; Writing-Original Draft Preparation: Merve Keskin; Writing-Review & Editing: Merve Keskin, Ahmet Ozgur Dogru, Kristien Ooms, Philippe De Maeyer*

This chapter is moodified from:

Keskin, M., Ooms, K., Dogru, A.O., De Maeyer, P. (2020). Exploring the Cognitive Load of Expert and Novice Map Users Using EEG & Eye Tracking. ISPRS Int. J. Geo-Inf. 9(7), 429.

5.1. Introduction

Cognitive processes emerge from both overt (externally detectable) and covert attention (internally detectable), and attention is a fundamental cognitive function that controls all the other cognitive processes such as perception, memory and learning. Attention can be driven unintentionally by external events (i.e. bottom-up) or deliberately by internal expectations requiring cognitive control (i.e. top-down). Top-down attention influences the selection of visual stimuli based on previous experience and current goals, while filtering out distracting objects/visuals. The working memory, whose performance depends on the cognitive demands of the task, plays a critical role in guiding these top-down attentional processes by keeping present goals in mind (Pratt et al, 2011). Map learning involves complex cognitive processes and is different from other learning concepts, in the sense that it requires understanding and memorizing the information presented in map format, and this information is presented at once (Thorndyke & Stasz, 1980). When people need to perform a spatial memory task, they tend to memorize the location, color, shape, and size of the objects (i.e. visual variables) together with their spatial relationships between each other (Keskin et al., 2018). They also adapt themselves to when, how and in which order they select and focus on a map object of their interest. Therefore, each map user can develop their own strategy to approach the spatial information on maps. Being an example of top-down attentional tasks, map learning causes a cognitive load that varies depending on the task difficulty and the individual characteristics. Cognitive load refers to the used amount of working memory resources. It has been used to explain how humans deal with increasing cognitive demands associated with the increased task difficulty in actions where the cognitive skills are more important than the physical ones. Even if the task's difficulty is one of the most essential factors affecting performance, cognitive load is used to describe the mental cost of accomplishing task demands. Fluctuations of attentional state are also modulated by cognitive load in a sense that an increase in cognitive load involves increased attentional processing (Di Stasi et al., 2011). In this context, map design and the level of complexity of maps might have an impact on cognitive load and even influence how difficult a particular task can be.

Performance is generally defined by the reaction times and accuracy/success rates. It is worth mentioning that the reaction time can be related as a metric to indicate the fatigue of participants. More difficult tasks would require more cognitive effort from a user and may result in longer response times. Cognitive load can be a complementary measure to distinguish between users who perform a task with equal reaction time and accuracy rates but with different levels of cognitive effort, helping to develop interfaces that require less cognitive capacities.

Cognitive load can be extracted by using both fixation and saccade related eye tracking metrics, On the one hand, fixations are stable point of regards (PORs) during a certain time span (at least 80 to 100 ms) and indicate the users' content interpretation at that location (Ooms, 2012). For instance, average fixation duration is associated with the attentional procedures, while the number of fixations per second indicates the speed of attention. Fixation duration and the number of fixations are generally inversely correlated and higher fixation durations indicate higher processing load (Keskin et al 2018, 2019). On the other hand, saccades are short (e.g. typically 30-80 ms) and voluntary eye movements between two fixations and can be visualized as scan paths. Saccadic eye movements are identified with their amplitude, duration, and velocity, and the relationship between these three parameters is called the 'main sequence'. The measurements of saccade velocity and amplitude helps observing the pattern of a scan path and exploring the cognitive effort. Saccade amplitude (length) and saccade velocity are highly correlated to each other and discriminatory parameters in terms of cognitive performance (Di Stasi et al., 2011). Saccade velocity (°/s) is the average saccade speed in degrees per second, whereas saccade amplitude (°) is the size of the saccade in degrees. Higher saccade velocity average indicates higher stress and task complexity and lower concentration while doing the task. The higher the cognitive load, the shorter the saccades, and the higher the saccade velocity (Behroozi et al., 2018).

Based on the existing eye tracking literature on the differences between expert and novice map users, we know that experts have better defined eye-scanning patterns, mostly have shorter reaction times and fixations and more fixations per second (e.g. Li et al., 2013, Ooms et al., 2012), and also fewer saccades (e.g. Dong et al., 2018) of which all are correlated with a low cognitive load. Regardless of their expertise, users' eye movements reflect the main elements on map stimuli and their attention is influenced by deviating colors on maps (Ooms et al., 2014). The cognitive strategies of experts and novices might differ as well, regardless of the type of the visual stimuli. In the context of solving a physics problem, correct answers are associated with the fact that the participants look at thematically relevant areas, unlike wrong answers being correlated with their focus on perceptually salient areas of the visual stimulus (Carmichael, et al., 2010). Similarly, while solving a thematic map problem, unsuccessful participants were not able to use of the thematic legend properly, focus on the relevant map layout elements and adequate map content (Havelková & Gołębiowska, 2020).

EEG is another non-invasive and direct method to measure cognitive processes in the brain. The EEG signal represents oscillations observed across a wide range of frequencies which are commonly divided into distinct frequency bands (i.e. delta band: <4 Hz, theta band: 4-8 Hz, alpha band: 8–12 Hz, beta band: 13–30 Hz, gamma band: >30Hz) (Fink and Benedek, 2013). Spectral analyses of the EEG (i.e. power spectral density (PSD)) can be used to compute the band-specific frequency power for given periods of time, i.e. during a task/trial. Event-related power decreases from a reference to an activation interval are good indicators of cognitive load. As explained in Chapter 3 and 4, alpha and theta power are associated with the cognitive load in a sense that alpha decreases and theta increases as cognitive processing increases (Pfurtscheller & Da Silva, 1999). The power deccreases are commonly referred to as event-related desynchronization (ERD), on the contraray, the power increases are referred to as event-related synchronization (ERS).

To study the cognitive procedures of individuals during a map learning task, eye tracking and EEG technologies can be combined (e.g. Al-Samarraie, 2019; Gedminas, 2011; Keskin et al., 2018; 2019; Lanini-Maggi, 2017). Since eye movements and attentive cognition are linked, it is possible to detect users' cognitive states in situ via eye trackers. Once these cognitive states are understood, effective spatial visualizations that adapt themselves to their users' current cognitive capacities (e.g. cognitive load) can be developed (Zagermann et al., 2016). While eye tracking is used to detect overt attention through gaze movements, EEG, which is sensitive to the instantaneous changes in the brain, is more likely to detect covert attention through direct measures of the electrical activity along the scalp. As well as eye tracking, EEG requires a statistical and visual analysis of cognitive processes. Furthermore, it has commonly been applied in cognitive and experimental psychology to study how the human brain responds to any kind of external stimuli (e.g. Bombeke, 2017; Meghanathan et al.,2019; Verhulst, 2018). Therefore, the co-registration of eye movements and EEG rhythms is promising for cartographic usability research especially when studying the behavior of different map user groups (e.g. experts, novices) as the insights that particularly arose from the personal differences contribute to creating effective and user-friendly cartographic products for those user groups.

Our main research objective is to explore the cognitive processes of expert and novice map users during the retrieval of map-related information contained in a map stimulus and within varying difficulty levels. Therefore, we aim to test the effect of the task difficulty on map users' attentional behavior through cognitive load measurements. We are interested to explore whether the cognitive procedures used by experts and those used by novices differ for basic spatial memory tasks. We expect that experts might apply more structured strategies that are particular for map use and might execute the tasks faster and in a more efficient way due to their specific map knowledge.

Previously, we investigated the spatial memory (i.e. map learning) abilities of map users through two user experiments employing mixed methods of (i) sketch maps and eye tracking (Chapter 2) and (ii) eye tracking and EEG (Chapter 4) by emphasizing the importance of cartographic/psychological experimental design. While, in Chapter 4, we mostly focused on the experimental set-up of the user study presented in this chapter, we now present the results of the EEG analysis in detail with the aim of triangulating them with the recorded eye tracking data. With this approach, we will be able to interpret cognitive processes occurring during this spatial memory task in a more holistic way. In this context, we introduce the behavioral data (i.e. reaction time, response accuracy), saccade-related metrics such as saccade velocity and saccade amplitude, their relationship with the previously obtained fixation related metrics and their impact on understanding the cognitive strategies of expert and novice map users. Additionally, attentional behaviors of two groups are further explored with the qualitative analysis of focus maps. We also provide event-related EEG analysis (i.e. PSD) of two user groups for different difficulty levels of the spatial memory task. Alternatively, the recruited participants were classified as good learners and relatively poor learners based on their overall task success rates, and we present the results of the EEG analysis conducted with respect to this classification as well.

5.2. Methodology

The methodology used to process the collected eye tracking and EEG data is not straightforward, and the algorithms used to detect fixation, and saccades or EEG rhythms could influence the results. We used the same methodology for the experimental design and a subset of the same dataset for data analysis of the collected data as in Chapter 4. Table 5.1 summarizes the experimental design elements. Preprocessing of the EEG recordings including the steps such as noise filtering, bad channel removal, channel interpolation, re-referencing and segmentation was handled in EEGLAB open-source MATLAB toolbox by following Makoto's preprocessing pipeline (Delorme & Makeig, 2012). Event-related changes in the spectral power density (PSD) with respect to alpha and theta frequency bands were calculated as explained in Klimesch (1999) and Pfurtscheller (1999). We calculated the EEG metrics not only for experts and novices (i.e. classification based on expertise), but also for good learners and relatively poor learners (i.e. classification based on success rates). Theoretically similar to what was done by Thorndyke and Stazs (1980), we defined good learners as those who performed better than the average did and the rest would be relatively poor map learners, regardless of their expertise.

Table 5.1. Experimental design elements (modified from Table 1.4).

	Experiment 2
Participants	20 participants 10 experts (5 females, 5 males) (average age: 28) 10 novices (5 females, 5 males) (average age: 26.4) Average age: 27.2, SD: 3.9
Task procedures & Stimuli	Randomized block design: Seven blocks representing seven difficulty types. Each block includes 50 trials (i.e., one for each stimulus) focusing on the similarity of: Block 1: All map elements (the whole map) Block 2: Roads and hydrography Block 3: Roads and green areas Block 4: Green areas and hydrography Block 5: Green areas Block 6: Hydrography Block 7: Roads
Independent variables	7 task difficulty levels (i.e., classified as easy, moderate, hard) 2 expertise levels (i.e., experts vs. novices)
Dependent variables	Response time, success rate (i.e. correct answers), eye tracking metrics (average fixation duration, the number of fixations per second, saccade velocity, saccade amplitude, EEG metrics (alpha and theta power changes (ERD/ERS))

5.2.1. Apparatus

A dual PC set-up was established for EEG and eye tracking to simultaneously capture participants' psychological data (see Keskin et al., 2019). EEG was recorded using BIOPAC Acqknowledge software and hardware, and an International 10-20 System ECI electrode cap (i.e. recording electrodes: Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6; the linked mastoids as reference and ground) with a sampling rate of 500 Hz. The SMI RED 250 eye tracker was synchronized with EEG to capture the gaze activities simultaneously and to monitor the possible eye movement artifacts on EEG. To ensure a good enough skin-electrode interface, the impedance was measured using BIOPAC EL-CHECK in advance of recording for every participant. We paid attention to keep the electrode impedances of the whole circuit (i.e. ground, active EEG, reference electrodes) less than 10 K ohms as suggested by Herman et al. (2015) and Teplan

(2002). To ensure that participants remained at a fixed distance from the screen and to avoid head movements, we established a chin rest, which was positioned at 70 cm from the screen. The horizontal and vertical eye positions for both eyes were recorded at a rate of 250 Hz.

5.2.2. Participants

This research got the approval of the Ethics Committee of the Faculty of Business and Economics of Ghent University where we conducted our experiments. We considered experts who hold at least an MSc degree in geomatics and other geo-related domains. The novices were selected from the volunteers who had no professional experience with maps. As a rule of psychological experiments, they were at least above 23-years-old to be able match their age with the average age of experts with a low standard deviation. In total, both experts and novices, whose age (N = 20, MED = 27.2, SD = 3.9) and gender (10 F, 10 M) match, performed the same experiment under the same conditions (see Table 5.1 for more detail).

5.2.3. Task and Stimuli

The spatial memory task related to the retrieval of the main structuring elements of maps varied in difficulty; hard, moderate, and easy. Hard tasks focused on (1) all elements, and (2) roads & hydrography; moderate ones focused on (3) roads & green areas, and (4) green areas & hydrography; and easy ones focused on (5) green areas, (6) hydrography, and (7) roads. Accordingly, the stimuli included in tasks were presented as seven randomized blocks; each including 50 trials of the same type of map elements (i.e. in total 350 trials) (Figure 5.1). For the classification of the task difficulty, we considered the average reaction times of all participants corrected based on the amount of errors committed, i.e. the inverse efficiency score (Townsend & Ashby, 1978).

5.2.4. Procedures

Before each trial, participants were shown a fixation cross in the middle of the screen for a duration of two seconds. This is called baseline period (i.e. reference interval) and refers to the pre-stimulus duration without any task demands except for concentrating on a displayed cross. During the trial, participants were asked to study a map stimulus in a free-viewing condition for seven seconds. This is called the task period (i.e. activation interval), where the participants were required to perform the experimental task. After studying the map stimulus, four graphical response panels appeared. The panels included skeleton maps in which only the main structuring element(s) of interest were drawn and the participants were instructed to select the panel with the correct skeleton map corresponding to the stimulus they had just studied. By the time they decided on their answers with a simple key command, the trial was terminated and they were automatically presented with the fixation cross for two seconds and then the next map stimulus/trial was initiated. With the preparation of the participant (i.e. reading instructions, signing (please see Annex 2 for full instuctions and Annex 3 for the task structure), wearing the EEG cap, impedance check, and calibration of the eye tracker) and small breaks between blocks to combat the fatigue, each participant averagely required 2.5 hours to complete the experiment consisting of seven blocks.

Figure 5.1. Example stimulus and experiment blocks: (a) Original stimulus; (b) Block 1: All map elements; (c) Block 2: Roads and hydrography; (d) Block 3: Roads and green areas; (e) Block 4: Green areas and hydrography; (f) Block 5: Green areas; (g) Block 6: Hydrography; (h) Block 7: Roads (modified from Keskin et al., 2019).

5.2.5. Psychological measures to estimate cognitive load

Next to the average fixation duration and the number of fixation per second, which were previously published in (Keskin et al., 2019), we explored saccadic eye movements as measures of cognitive processing demands, i.e. cognitive load, because they are highly distinctive, task-dependent and can be correlated with fixation duration to interpret the overall cognitive load (Zagermann et al., 2016). There is a strong evidence that longer fixation duration and shorter saccades are related to higher cognitive load (Holmqvist et al., 2011) and indicate that more attentional resources were required (Debue & Leemput, 2014). Consequently, it is possible to formulate that fixation duration and saccade velocity increase but saccade length decreases when information processing rises.

In addition to the above-mentioned quantitative analysis of the collected eye fixations and saccades, we randomly selected 10 stimuli used in the experiments and conducted a qualitative analysis (e.g. visual inspection of eye fixations) focusing on the attentional behavior of the participants to the map elements of interest using focus/heat maps.

For EEG, we focused on the specific group of electrodes due to the fact that alpha reduction (alpha ERD) is generally observed at parietal regions e.g. (Keil et al 2006; Klimesch, Doppelmayr, Roehm, Pöllhuber & Stadler, 2000) and theta increase (theta ERS) is most profound over frontal electrode locations (Brouwer, 2012). In this context, we focused on the Fp1, Fp2, F3, F4, F7 and F8 frontal channels for theta power and the P3 and P4 parietal channels for alpha power e.g. (Jensen & Tesche, 2002; Missonnier, et al., 2006; Sauseng et al., 2005) (Figure 5.2). After averaging all usable trials within the reference and activation periods, event-related power changes at an electrode were calculated by subtracting the log-transformed the power during activation intervals from the log-transformed the power during the reference intervals (Keskin et al., 2019) (see Chapter 4). Consequently, we grouped the trials based on the task difficulty and averaged the theta spectral power at frontal channels and alpha spectral power at parietal channels for novices and experts separately.

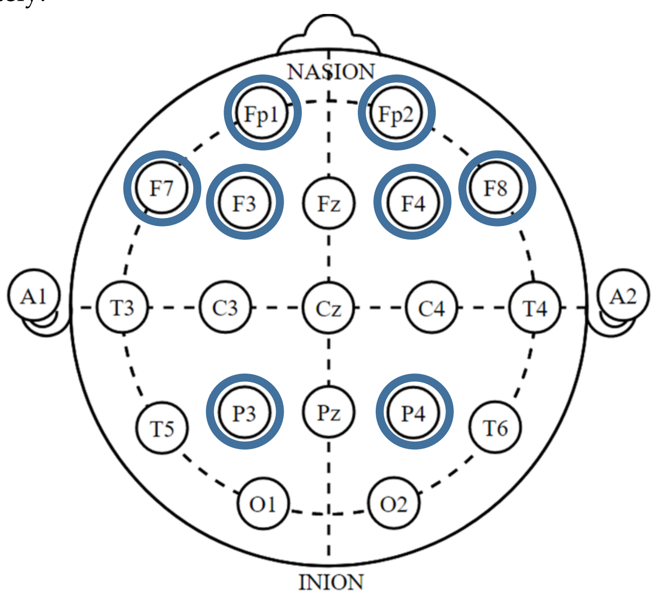


Figure 5.2. Selected electrodes for the EEG analysis (International 10-20 system).

5.3. Results

5.3.1. Behavioral measures

The overall average reaction time was 5.1 s (*SD*= 1.1) for experts and 3.7 s (*SD*= 0.5) for novices, consequently, experts spent more time than novices did for all tasks. For hard tasks, experts spent 6.9 s (*SD*= 1.4) whereas novices took 5.0 s (*SD*= 0.6); for moderate tasks, experts completed them in 5.4 s (*SD*= 1.4) while novices took in 2.9 s (*SD*= 0.6); and lastly for easy tasks, experts responded in 3.7 s (*SD*= 0.9) and novices took in 2.8 s (*SD*= 0.4) (Figure 5.3a). These differences between the two groups are statistically significant across the hard and moderate tasks (Mann–Whitney U test: U_{hard}= 106.000, p= 0.022; $U_{moderate}$= 114.000, p= 0.035), however, it was not the case for easy tasks (Mann–Whitney U test: U_{easy}= 128.000, p= 0.138).

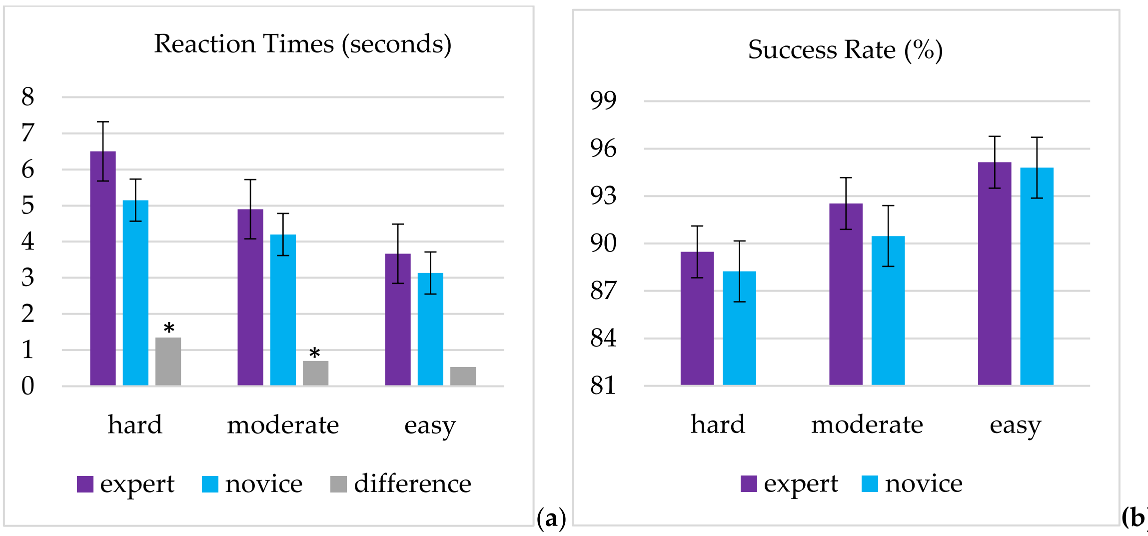


Figure 5.3. (a) Reaction times; (b) Success rates (*significant difference).

The overall average success rate (i.e. correct answers in %) was quite high for both group of participants, averaged for 350 trials per participant in total, $M_{overall}$= 91.8% (*SD*= 4.7, *range*= 78.3%-98.3%). For hard tasks, experts scored 90.6% whereas novices scored 86.8%; for moderate tasks, experts achieved the score of 93.5%, while novices scored 88.8%; and lastly for easy tasks, experts responded the trials with a 95.5% success rate and novices scored a 93.3% success rate (Figure 5.3b). The success rate did not significantly differ across the categories of expertise for any difficulty level ($M_{experts}$= 93.2%, $M_{novices}$= 89.6%; Mann–Whitney U test: U_{easy}= 165.000, p= 0691, $U_{moderate}$= 150.000, p= 0.400, U_{hard}= 178.500, p= 1.000). The reason for high success rates underlies the design of the experiment because we intended to collect the data with as many correct answers as possible for EEG analysis. When accomplishing a task or failing it, different cognitive processes occur in the brain, hence, it is appropriate to consider correct and wrong answers separately. If approximately equal in number, correct and wrong answers could be compared in terms of EEG-related analysis. However, we were interested in the cognitive processes occurring during the accomplishment of the task, in other words, correct answers. In this context, the experiment was designed as a rather easy one so that when we exclude the wrong answers, which are much less in number, there would still remain enough trials to average for the EEG power spectrum analysis.

Although we observed slight differences in performances, it might be interesting to explore success rates in terms of good and relatively poor map learners instead of experts and novices. Theoretically similar to what was done by Thorndyke & Stazs (1980), we defined good learners who performed better than the average did and the rest would be relatively poor map learners regardless of their expertise. Out of 20, 15 participants were good learners with an average score of 94.6% overall and consisted of nine experts (4F, 5M) and six novices (3F, 3M); the remaining five were relatively poor learners with an average score of

85.8% and this involved one expert (F) and four novices (2M, 2F). This difference was statistically significant (Mann–Whitney U test: U= 67.500, p= 0.000) and showed that good map learners remembered more map elements compared to the relatively poor learners.

The overall average success rate for hard tasks were 88.7% with the lowest score of 66.0% and the highest being 98.0%. Good learners averagely scored 92.5%, whereas the other group scored 77.4%. The overall average success rate for moderate tasks was 91.2% with the lowest score of 74% and the highest being 98.0%. The average score of good learners was 93.7% and the relatively poor learners scored 83.4%. The overall average success rate for easy tasks was 94.4% with the lowest score of 80.0% and the highest being 99.3%. Good learners resulted as 96.1% while the remaining group as 89.2%. The difference occurred between good learners and the relatively poor learners for easy tasks was 6.9%; for moderate ones, 10.3%; and for hard ones, 15.1%. It is also important to mention that we observed an increase in terms of the performance differences between good learners and relatively poor learners as the task difficulty increases (Figure 5.4).

One interesting finding to note is that the reaction times longer than the average (i.e. between 4.5 s - 6.6 s) all belonged to the good learners, which consisted of five experts and one novice. Additionally, the top three shortest overall reaction times (i.e. 3.0s, 3.1s, 3.5 s.) all belonged to the relatively poor learners (all novices) with the top three lowest overall success rates (i.e. 78.3%, 80.0%, 85.1%). This shows that spending more time on tasks helped experts achieving higher accomplishment rates whereas the fast responses of novices resulted in a lower number of correct answers.

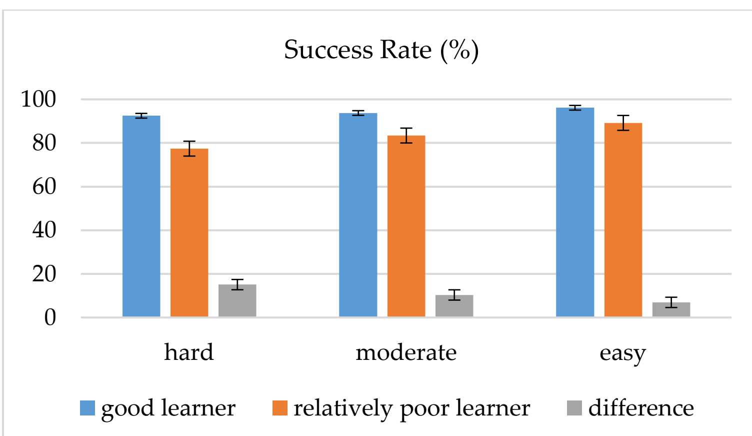


Figure 5.4. Success rates of good and relatively poor learners

5.3.2. Psychological measures

Although novices had longer fixation durations compared to the experts did for all tasks (Figure 5.5a),this difference for fixation duration was not considered as significant as a result of applied statistical tests (F_{easy}= 0.261, p= 0.232; $F_{moderate}$= 0.174, p= 0.514, Mann–Whitney U test: U_{hard}= 1812391.500 p= 0.886). Experts mostly exhibited a higher number of fixations per second for all difficulties (Figure 5.5b). However, the difference between experts and novices was not significant for hard and moderate tasks ($F_{moderate}$= 1.861, p= 0.165, F_{hard}=

0.064, p= 0.983), whereas it was significant for easy tasks (F_{easy}= 0.006, p= 0.019) (see Chapter 4 for more detail).

In Chapter 4, we suggested that it would be useful to investigate saccade related metrics to interpret the cognitive load further. Figure 5.5c and 5.5d show saccade amplitude and velocity varied for experts and novices for the easy, moderate and hard tasks. As the task becomes harder, we observed that the saccade amplitude becomes smaller; hence, the saccades become shorter which indicates a higher cognitive load. Regarding to saccade velocity, a contradicting trend is seen between experts and novices. Novices exhibited the highest velocity with the easy category, which is linked with the highest amplitude, and they demonstrated the lowest with the hard category, which is linked with the lowest amplitude. This finding is in line with the previous research, e.g. (Behroozi et al., 2018), however, the expert group showed the opposite result, in the sense that they had the highest velocity with the most difficult category and thus the lowest amplitudes.

None of the saccade related metrics for all types of tasks (i.e. easy, moderate, hard) fits the normal distribution. (Shapiro-Wilk test for saccade amplitude: W= 0.933, p= 0.000 < 0.05; for saccade velocity: W= 0.970, p= 0.000 < 0.05). Therefore, we applied Mann-Whitney U non-parametric test to measure the significance of the differences between two groups, and as a result, saccade amplitude (U_{hard}= 1554376.500, p_{hard}= 0.000 < 0.05; $U_{moderate}$= 1363219.500, $p_{moderate}$= 0.000 < 0.05; U_{easy}= 3061036.000, p= 0.000 < 0.05) and saccade velocity (U_{hard}= 1750439.000, p_{hard}= 0.000 < 0.005; $U_{moderate}$= 1536918.500, $p_{moderate}$= 0.000 < 0.05; U_{easy}= 3368847.500, p_{easy}= 0.000 < 0.05) of expert and novice participants were all significantly different for all types of tasks. For both metrcis, we observed no significant interaction between expertise and task difficulty (i.e. for saccade amplitude: $F\ (2,22)$= 0.310, p= 0.736; for saccade velocity: $F(2,22)$= 0.281, p= 0.757).

Compared to experts, novices exhibited larger saccades at all difficulty levels, and the difference in saccade amplitude between experts and novices increased as the task difficulty decreased (Figure 5.5c). The easy tasks received the larger saccades and the hard tasks received shorter saccades as expected. Due to the higher number of elements to pay attention to in hard tasks, the participants had to jump from one object to another in a short amount of time; therefore, they exhibited shorter saccades. Shorter saccades demonstrated a higher cognitive load for experts at all difficulty levels, however, the saccade velocity data claimed slightly differently. The novices accomplished moderate and easy tasks with a faster saccade velocity, whereas experts had higher saccade velocity in hard tasks (Figure 5.5d). These findings show that experts manifested more cognitive load in hard tasks according to their shorter saccade amplitudes and higher saccade velocity, although their fixation durations were shorter compared to the novices but not significantly. Accordingly, experts did not accomplish the tasks with a lesser cognitive load but they scan the map faster and in a more effective way (with higher success rates) when it comes to hard tasks.

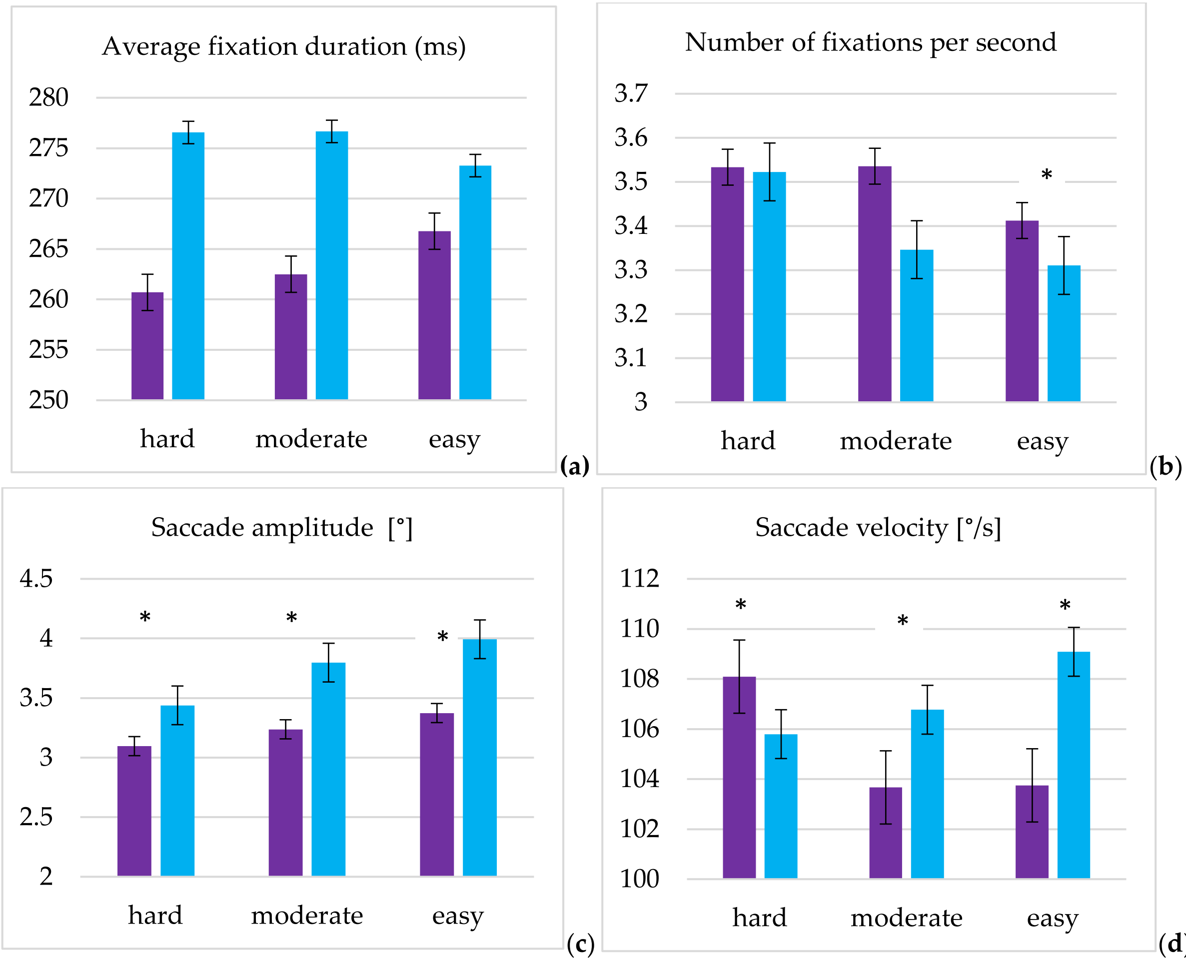


Figure 5.5. (a) fixation duration; (b) number of fixations per second (also published in Keskin et al., 2019); (c) saccade amplitude, (d) saccade velocity (purple bars: experts, blue bars: novices, *: significant difference, error bars indicate standard deviation).

We observed several common characteristics between expert and novices when heat/focus maps of randomly selected ten stimuli were visually evaluated. Some fundamental remarks are listed as follows (see Figure 5.6):

- Block 1 – All map elements: the road junctions, especially in the center or close to the center of the map, are where all the participants inherently focused the most. Both experts and novices generally paid most attention to the green areas that are large and isolated. These isolated and large green areas received more and longer fixations in comparison to the water bodies. Hydrographic features received lesser fixations compared to others. The labels/texts on the map also received much attention from both groups. This outcome might be due to the unfamiliar language used for labels or the size and position of the labels; therefore, it was a distraction, yet it could be a useful input in map design.
- Block 2 – Roads & Hydrography: Independently from what the spatial memory task demands, green areas received as many fixations as roads and hydrography did from both groups. In some cases, participants drew their attention even to the smaller green areas.

- Block 3 – Roads & Green areas: Similar situation as in Block 2 occurred for Block 3, and in this case, the hydrographic elements received as many fixations as the roads and green areas did.
- Block 4 – Hydrography & Green areas: Large green areas and road junctions received the most fixations. In this case, the relatively larger hydrographic areas did not receive as many fixations as the smaller ones did.
- Block 5 –Green areas, Block 6 - Hydrography, Block 7 – Roads: Both expert and novice participants majorly focused only on what the task demanded, therefore, in their focus maps only the map elements of interest stood out. This shows that it was easier to maintain an undivided attention when participants needed to focus on only one map element class.

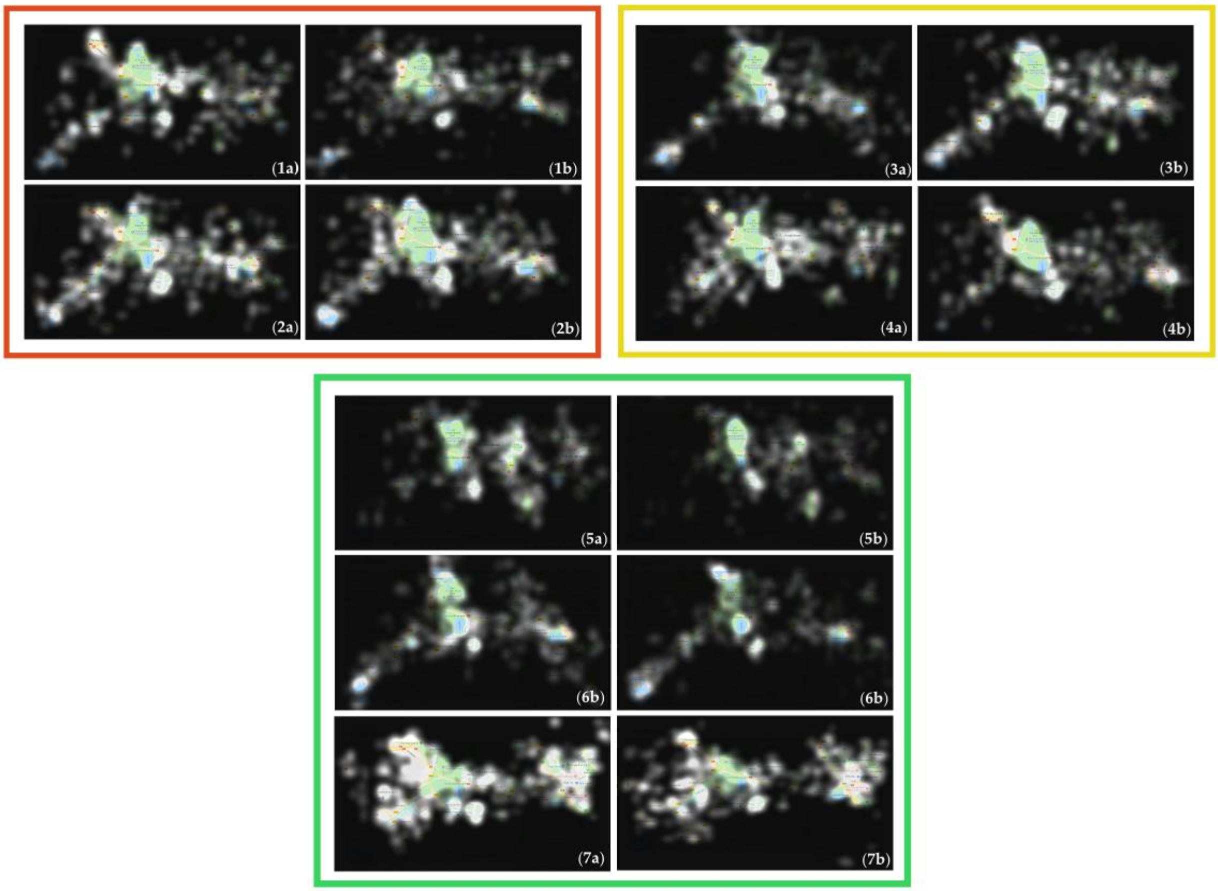


Figure 5.6. Focus maps belonging to the example stimulus in Figure 1. The maps in the red rectangle represent hard tasks: (1a) Block 1 - experts, (1b) Block 1 - novices, (2a) Block 2 - experts, (2b) Block 2 - novices. The maps in the yellow rectangle represent moderate tasks: (3a) Block 3 - experts, (3b) Block 3 - novices, (4a) Block 4 - experts, (4b) Block 4 - novices. The maps in the green rectangle represent easy tasks: (5a) Block 5 - experts, (5b) Block 5 - novices, (6a) Block 6 - experts, (6b) Block 6 - novices, (7a) Block 7 - experts, (7b) Block 7 - novices.

Based on the PSD analysis of alpha and theta, we observed an ERD alpha and ERS in theta for easy and hard tasks. This finding is in line with the literature on the frontal theta activity increasing e.g., (Antonenko et. al, 2010; Jensen & Tesche, 2002; Missonnier et al., 2006; Mussel, 2016), and parietal alpha decreasing e.g., (Keil et al., 2006; Klimesch et al., 2000; Morton et al., 2019) with the cognitive load in a working memory task. For moderate tasks, alpha power was observed to be increasing as well as theta was. Although the changes in alpha power seem very small, the theta effects seem stronger and confirm that the experts and novices have a different experience in this spatial memory task in a sense that experts exhibited more theta in moderate and hard tasks whereas novices did more in easy tasks. The increase in alpha during moderate tasks might be due to this spectral power feature possibly not being sensitive enough to discriminate on an aggregated level. A great deal of information is lost considering the values for alpha power activity are averaged for the whole duration of the condition (Morton et al., 2019). However, the results indicate an interaction with the participants for easy and hard tasks (see Figure 5.7, Table 5.2).

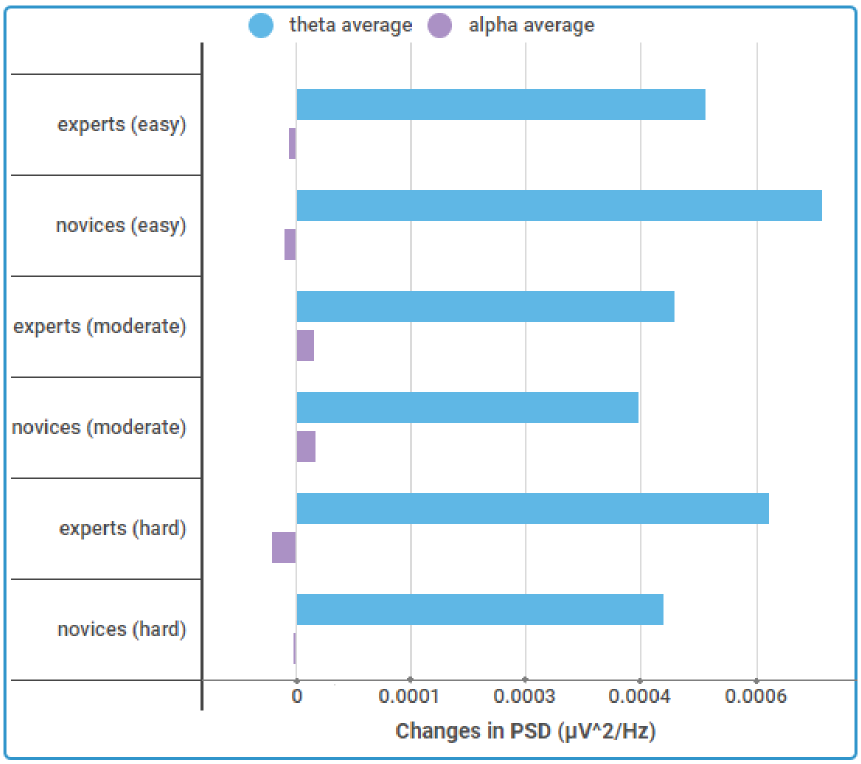


Figure 5.7. Changes in PSD of alpha and theta for experts and novices (μV²/Hz) (Alpha values are multiplied by 10 for visualization purposes).

Table 5.2. Changes in PSD of theta and alpha averaged for all tasks.

Changes is PSD ($\mu V^2/Hz$)	**Easy**		**Moderate**		**Hard**	
	Experts	Novices	Experts	Novices	Experts	Novices
theta (average of frontal channels)	0.0005319	0.0006844	0.0004911	0.0004449	0.0006153	0.0004781
alpha (average of parietal channels)	-0.0000008	-0.0000014	0.0000023	0.0000026	-0.0000029	-0.0000002

To compare the cognitive load based on the task difficulty, we focused on theta power since a very small alpha effect was observed. The difference of theta power changes between experts and novices was

0.0001525 for easy tasks; 0.0000462 for moderate tasks and 0.0001372 for hard tasks. For none of the difficulty levels, theta values fit the normal distribution (Shapiro Wilk test: W_{hard}= 0.875, p= 0.001; $W_{moderate}$= 0.773, p= 0.000 < 0.05; W_{easy}= 0.922, p= 0.002), accordingly, we applied Mann-Whitney U test for assessing the significance of the differences. The distribution of the theta power change was the same across categories of expertise and difficulties (U_{hard}= 77.000, p_{hard}= 0.519 > 0.000; $U_{moderate}$= 124.000 $p_{moderate}$= 0.367; U_{easy}= 262.000 p_{easy}= 0.766), which shows that the difference between two user groups was not statistically significant. This finding suggests theta power may not be as sensitive for average cognitive load and that it may be developed into a valid objective measure of average cognitive load although its true potential lies in the possibility to measure online fluctuations in cognitive load or instantaneous cognitive load (Castro-Meneses et al., 2020). We found no significant interaction between expertise and task difficulty ($F(2,103)$= 0.443, p= 0.644)

Although we expected that there would be a greater effect on theta power for the hard tasks compared to the others, we observed the greatest difference for easy tasks. This could be due to the hard tasks being too overwhelming, it being hard to stay motivated, and also the participants' tendency to give up and not to invest mental effort and resources anymore e.g. (Morton et al., 2019). Participants confirmed in their post-test questionnaires that they find the task hard to focus on and tiring.

Alternatively, we calculated event-related theta and alpha power changes in EEG power spectrum for good and relatively poor map learners (Figure 5.8, Table 5.3). Good learners exhibited slightly higher cognitive load at all the levels of difficulty. Regarding the overall performances, only small and non-significant power changes occurred in alpha (Mann–Whitney U test: U_{alpha}= 846.000, p= 0.501 > 0.05), whereas the theta power seemed higher for good learners in all tasks and the difference that emerged between good and relatively poor learners was significant (Mann–Whitney U test: U_{theta}= 753.000, p= 0.020 < 0.05). It shows that the good learners exhibited higher cognitive load, regardless of the task difficulty. The biggest difference (0.000377) in terms of theta power change between good learners and relatively poor learners was observed for easy tasks. However, the differences in theta power among none of the difficulty levels was statistically significant ((Mann–Whitney U test: U_{hard}= 53.000, p_{hard}= 0.589 > 0.000; $U_{moderate}$= 90.000 $p_{moderate}$= 0.323; U_{easy}= 125.000 p_{easy}= 0.068).

Table 5.3. Changes in PSD of theta and alpha averaged for all tasks.

Changes in PSD ($\mu V^2/Hz$)	**Hard**			**Moderate**			**Easy**		
	GL[1]	**RPL**[2]	**ΔP**[3]	**GL**[1]	**RPL**[2]	**ΔP**[3]	**GL**[1]	**RPL**[2]	**ΔP**[3]
theta	0.000594	0.000391	0.000203	0.000511	0.000351	0.000160	0.000695	0.000318	0.000377
alpha	-0.000002	0.000000	-	0.000003	-0.000001	-	-0.000001	-0.000005	-

[1] GL= Good Learners, [2] RPL= Relatively Poor Learners, [3] ΔP= difference in power change.

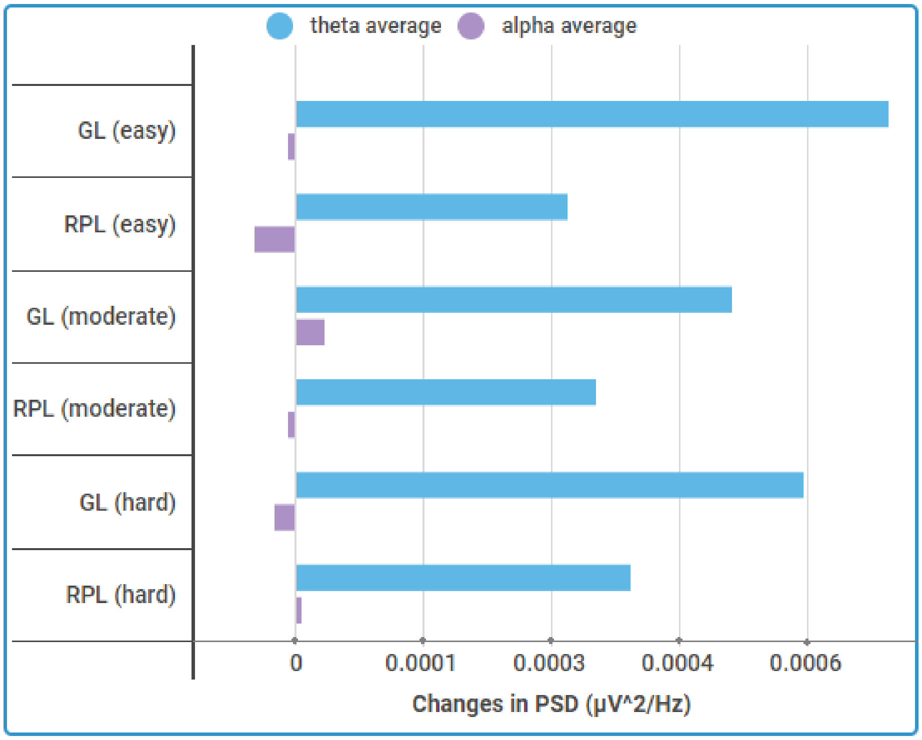


Figure 5.8. Changes in PSD of alpha and theta for good learners and relatively poor learners ($\mu V^2/Hz$) (Alpha values are multiplied by 10 for visualization purposes).

5.4. Discussion and Conclusion

In this chapter, we investigated the spatial memory abilities of a group of expert and non-expert map users within a simple map-learning task using eye tracking and EEG and triangulated the behavioral and psychological data to indicate the cognitive load caused by the task. Some highlights of the findings are listed as follows:

- Experts had longer reaction times (significantly longer for moderate and hard tasks), but higher success rates. They might be a bit more ambitious and driven to accomplish the task compared to novices, and have saved an extra time to review or verify their response before submitting it. It seems the fact that experts exhibiting more cognitive load paid off with higher success rates.
- Novices were observed to have longer fixation durations, mostly lower number of fixations per second and higher saccade velocity (except for hard tasks) which indicate a higher cognitive load for novices. Additionally, the saccade amplitudes of novices were longer. In longer saccades (i.e. larger amplitude), the search goes all across the image and is thus less organized. Experts exhibiting shorter saccades means a more targeted search from one focal point to the next, which are close to each other in the map, therefore, a less chaotic search pattern. The shorter fixations of experts also show that they needed less time to interpret what they saw.
- Although not significant, experts demonstrated higher theta power (except for easy tasks) which can be associated with a higher cognitive load.
- Qualitative analysis of the eye movements shows that both groups showed similar attentional behavior in terms of the map area covered and the map elements on which they focused.

Based on the findings, it is difficult to favor one user group in terms of their performance while retrieving the map-related information. The map-learning and recalling strategies of experts and novices and their approach to the task might not be similar, however, the overall performances of them did not differ much. In fact, novices, in some cases, outperformed experts. This outcome might seem to contradict the results within the expert-novice research paradigm e.g. (Dong et al., 2018; Ooms et al., 2012); however, it is in line with the findings in the field of geography e.g. and in map learning domains e.g. (Thorndyke & Stasz, 1980).

On the one hand, the reason why we did not find significant differences between novices and experts might be that they pay attention to the different aspects of a task. This affects both their perceptions of task complexity (i.e., task analyzability and variability) and their performance on the task. Superior performance by experts depends on the match between the experts' cognition and the demands of the task (Haerem & Rau, 2007). The fact that novices sometimes perform better than experts' would be an evidence that they use different learning strategies. As explained by Postigo and Pozo (1998) *'the subject lacking domain-specific knowledge tends to construct a visual-spatial mental representation, as opposed to the semantic representation of the expert. Experts represented given information in a domain-specific manner that was concerned with the deep semantic structure of that information, whereas the novices mentally represented focused-upon superficial domain-general aspects'* (p. 77-78)**.**

On the other hand, the reason why expertise is not as influential as we think especially for simple map-learning tasks is due to the effect of other individual differences. According to Hunt (1978), those differences originate from *'the use of simple processing procedures, knowledge related to the task and the ability to perform the low-level mechanics of information processing*'. Experts did not always outperform novices, which could explain that domain knowledge was not that relevant to the task, instead general education and the ability to perform basic operations such as decoding, visualization, selective filtering, memory retrieval, and memory comparison, played an influential role in high-level procedure and strategy choices. Therefore, the variation in performance and strategy choice might arise from the differences in basic visual or spatial ability. Additionally, the competence of the expert group does not only rely on their extensive knowledge, but also the organization of this knowledge that forms their cognitive representations and characterizes them (Postigo & Pozo, 1998; Thorndyke & Stasz, 1980).

We alternatively grouped the participants as good learners and relatively poor learners based on their success rates, and we calculated the changes in EEG power spectral density with respect to this classification. We observed that good learners exhibited significantly higher theta ERS considering their overall performance. Although the cognitive load of these groups did not differ based on the task difficulty within the frame of this study, classifying participants based on their spatial memory performances provided different insights on map user' cognitive processes. Similar to what was found by Havelková and Gołębiowska (2020), unsuccessful participants differed in the general problem-solving approach, in a way that they tended to choose fast, less cautious strategies and lacked motivation.

This study also showed that high cognitive load is not necessarily associated with the low task performance, in fact, for most cases, it was an indicator of more elaborate, structured and efficient cognitive strategies especially demonstrated by experts. Therefore, it is useful to triangulate data collected via difference sources (i.e. quantitative and qualitative methods) to interpret the cognitive load and to understand the underlying behavior of the participants.

Within this chapter, we analyzed the influence of the independent variables such as task difficulty and expertise level on the cognitive strategies of map users. As well as task and expertise, map design characteristics play an important role in users' cognitive load and learning performance, hence, should be evaluated in order to contribute to enhancing the design and usability of cartographic products e.g. (Al-Samarraie, 2019). We used screenshots of Google's road maps, which is designed for everyone (i.e. regardless of the users' individual differences of spatial cognition) as stimuli in our experiments, and we found no significant difference between experts and novices in terms of the cognitive load that these maps caused. It is important to mention that if the quality of the cartographic design fulfills its purpose of the design, it has a positive effect on users' experience.

Furthermore, the EEG power metrics used in the study and the procedures of extracting them have an influence on the results. To have more detailed insights on cognitive load and to detect the small changes in the EEG power spectrum, different procedures can be applied to the collected data. For instance, the seven-seconds-long study period can be segmented into sub-parts and the exact time points of the peak values of alpha and theta can be identified. These peak values can further be analyzed simply for the time periods of interest by visually inspecting the EEG time-frequency plots. Another interesting approach is to investigate the lower (8-10Hz) and upper (10-12Hz) alpha bands separately in order to indicate specific frequency effects that are not distinct when only looking at the broad alpha range as suggested by Morton et al. (Morton et al., 2019). There are a number of researches demonstrating different activity in upper and lower alpha bands in cognitive load conditions in a sense that upper alpha decreases when cognitive activity increases e.g. (Fink 2006; Klimesch et al., 1997; Sauseng et al., 2005. It is also possible to measure gamma oscillations, which are directly proportional to the cognitive activity e.g. (Fitzgibbon, 2004), and beta oscillations increase upon cognitive load e.g. (Güntekin et al., 2013).

EEG data might be overwhelming, and there are various other aspects to investigate further and a countless number of analysis to perform besides the ones mentioned above. However, the experimental design has a primary importance in a sense that deciding on where to pay attention to and what to expect from the collected data have to be well planned and tested, before conducting the main experiments. When integrated with other qualitative and quantitative user testing methods, EEG indeed suggests a valuable contribution to the understanding of the cognitive processes of individuals.

References

Al-Samarraie, H., Eldenfria, A., Price, M. L., Zaqout, F., & Fauzy, W. M. (2019). Effects of map design characteristics on users' search performance and cognitive load: An empirical study. The Electronic Library, 37(4), 667–679.

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using Electroencephalography to Measure Cognitive Load. Educational Psychology Review, 22(4), 425–438.

Behroozi, M., Lui, A., Moore, I., Ford, D., & Parnin, C. (2018). Dazed: Measuring the cognitive load of solving technical interview problems at the whiteboard. Proceedings of the 40th International Conference on Software Engineering New Ideas and Emerging Results - ICSE-NIER '18, 93–96.

Bombeke, K. (2017). Early Sensory Attention and Pupil Size in Cognitive Control: An EEG approach. Doctoral dissertation. Ghent University, Ghent.

Brouwer, A.-M., Hogervorst, M. A., Erp, J. B. F. van, Heffelaar, T., Zimmerman, P. H., & Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. Journal of Neural Engineering, 9(4), 045008.

Carmichael, A., Larson, A., Gire, E., Loschky, L., & Rebello, N. S. (2010). How Does Visual Attention Differ Between Experts and Novices on Physics Problems? AIP Conference Proceedings, 1289(1), 93–96. https://doi.org/10.1063/1.3515257

Castro-Meneses, L. J., Kruger, J.-L., & Doherty, S. (2020). Validating theta power as an objective measure of cognitive load in educational video. Educational Technology Research and Development, 68(1), 181–202.

Debue, N., & van de Leemput, C. (2014). What does germane load mean? An empirical contribution to the cognitive load theory. Frontiers in Psychology, 5. https://doi.org/10.3389/fpsyg.2014.01099

Delorme, A., & Makeig, S. (2012). EEGLAB tutorial outline chapter 1: Loading data in EEGLAB.

Di Stasi, L. L., Antolí, A., & Cañas, J. J. (2011). Main sequence: An index for detecting mental workload variation in complex tasks. Applied Ergonomics, 42(6), 807–813.

Dong, W., Zheng, L., Liu, B., & Meng, L. (2018). Using eye tracking to explore differences in map-based spatial ability between geographers and non-geographers. ISPRS International Journal of Geo-Information, 7(9), 337.

Fink, A., & Benedek, M. (2013). The creative brain: Brain correlates underlying the generation of original ideas. Neuroscience of Creativity, 207–232.

Fink, A., Grabner, R. H., Benedek, M., & Neubauer, A. C. (2006). Divergent thinking training is related to frontal electroencephalogram alpha synchronization. European Journal of Neuroscience, 23(8), 2241–2246.

Fitzgibbon, S. P., Pope, K. J., Mackenzie, L., Clark, C. R., & Willoughby, J. O. (2004). Cognitive tasks augment gamma EEG power. Clinical Neurophysiology, 115(8), 1802–1809.

Gedminas, L. (2011). Evaluating hurricane advisories using eye-tracking and biometric data. Master's Thesis, East Carolina University: North Carolina.

Güntekin, B., Emek-Savaş, D. D., Kurt, P., Yener, G. G., & Başar, E. (2013). Beta oscillatory responses in healthy subjects and subjects with mild cognitive impairment. NeuroImage: Clinical, 3, 39–46.

Haerem, T., & Rau, D. (2007). The influence of degree of expertise and objective task complexity on perceived task complexity and performance. Journal of Applied Psychology, 92(5), 1320.

Havelková, L., & Gołębiowska, I. M. (2020). What Went Wrong for Bad Solvers during Thematic Map Analysis? Lessons Learned from an Eye-Tracking Study. ISPRS International Journal of Geo-Information, 9(1), 9.

Herman, S. T., Abend, N. S., Bleck, T. P., Chapman, K. E., Drislane, F. W., Emerson, R. G., Gerard, E. E., Hahn, C. D., Husain, A. M., Kaplan, P. W., LaRoche, S. M., Nuwer, M. R., Quigg, M., Riviello, J. J., Schmitt, S. E., Simmons, L. A., Tsuchida, T. N., & Hirsch, L. J. (2015). Consensus Statement on Continuous EEG in Critically Ill Adults and Children, Part II: Personnel, Technical Specifications, and Clinical Practice. Journal of Clinical Neurophysiology, 32(2), 96–108.

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

Hunt, E. (1978). Mechanics of verbal ability. Psychological Review, 85(2), 109.

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. European Journal of Neuroscience, 15(8), 1395–1399.

Keil, A., Mussweiler, T., & Epstude, K. (2006). Alpha-band activity reflects reduction of mental effort in a comparison task: A source space analysis. Brain Research, 1121(1), 117–127.

Keskin, M., Ooms, K., Dogru, A., & De Maeyer, P. (2019). EEG & Eye Tracking User Experiments for Spatial Memory Task on Maps. ISPRS International Journal of Geo-Information, 8(12), 546.

Keskin, M., Ooms, K., Dogru, A. O., & De Maeyer, P. (2018). Digital sketch maps and eye tracking statistics as instruments to obtain insights into spatial cognition. JOURNAL OF EYE MOVEMENT RESEARCH, 11(3).

Klimesch, W, Doppelmayr, M., Pachinger, T., & Ripper, B. (1997). Brain oscillations and human memory: EEG correlates in the upper alpha and theta band. Neuroscience Letters, 238(1–2), 9–12.

Klimesch, W., Doppelmayr, M., Röhm, D., Pöllhuber, D., & Stadler, W. (2000). Simultaneous desynchronization and synchronization of different alpha responses in the human electroencephalograph: A neglected paradox? Neuroscience Letters, 284(1), 97–100.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. Brain Research Reviews, 29(2), 169–195.

Lanini-Maggi, S. (2017). Depicting movement data with animations for embodied and real-time decision-making: A user study with air traffic control displays and real-time movement data. Doctoral dissertation, University of Zurich, Zurich.

Li, W.-C., Chiu, F.-C., Kuo, Y., & Wu, K.-J. (2013). The Investigation of Visual Attention and Workload by Experts and Novices in the Cockpit. In D. Harris (Ed.), Engineering Psychology and Cognitive Ergonomics. Applications and Services (167–176). Springer.

Meghanathan, R. N., Nikolaev, A. R., & van Leeuwen, C. (2019). Refixation patterns reveal memory-encoding strategies in free viewing. Attention, Perception, & Psychophysics, 81(7), 2499–2516.

Missonnier, P., Deiber, M.-P., Gold, G., Millet, P., Pun, M. G.-F., Fazio-Costa, L., Giannakopoulos, P., & Ibáñez, V. (2006). Frontal theta event-related synchronization: Comparison of directed attention and working memory load effects. Journal of Neural Transmission, 113(10), 1477–1486.

Morton, J., Vanneste, P., Larmuseau, C., Van Acker, B., Raes, A., Bombeke, K., Cornillie, F., Saldien, J., & De Marez, L. (2019). Identifying predictive EEG features for cognitive overload detection in assembly workers in Industry 4.0. 3rd International Symposium on Human Mental Workload: Models and Applications (H-WORKLOAD 2019).

Mussel, P., Ulrich, N., Allen, J. J. B., Osinsky, R., & Hewig, J. (2016). Patterns of theta oscillation reflect the neural basis of individual differences in epistemic motivation. Scientific Reports, 6(1), 29245.

Ooms, K. (2012). Maps, how do users see them?: An in depth investigation of the map users' cognitive processes. Doctoral Dissertation, Ghent University, Ghent. Retrieved from https://biblio.ugent.be/publication/3103344

Ooms, K., De Maeyer, P., Fack, V., Assche, E. V., & Witlox, F. (2012). Interpreting maps through the eyes of expert and novice users. International Journal of Geographical Information Science, 26(10), 1773–1788.

Ooms, K., Maeyer, P. D., & Fack, V. (2014). Study of the attentional behavior of novice and expert map users using eye tracking. Cartography and Geographic Information Science, 41(1), 37–54.

Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. Clinical Neurophysiology, 110(11), 1842–1857.

Postigo, Y., & Pozo, J. I. (1998). The Learning of a Geographical Map by Experts and Novices. Educational Psychology, 18(1), 65–80.

Pratt, N., Willoughby, A., & Swick, D. (2011). Effects of Working Memory Load on Visual Selective Attention: Behavioral and Electrophysiological Evidence. Frontiers in Human Neuroscience, 5.

Sauseng, P., Klimesch, W., Schabus, M., & Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. International Journal of Psychophysiology, 57(2), 97–103.

Teplan, M. (2002). Fundamentals of EEG measurement. Measurement Science Review, 2(2), 1–11.

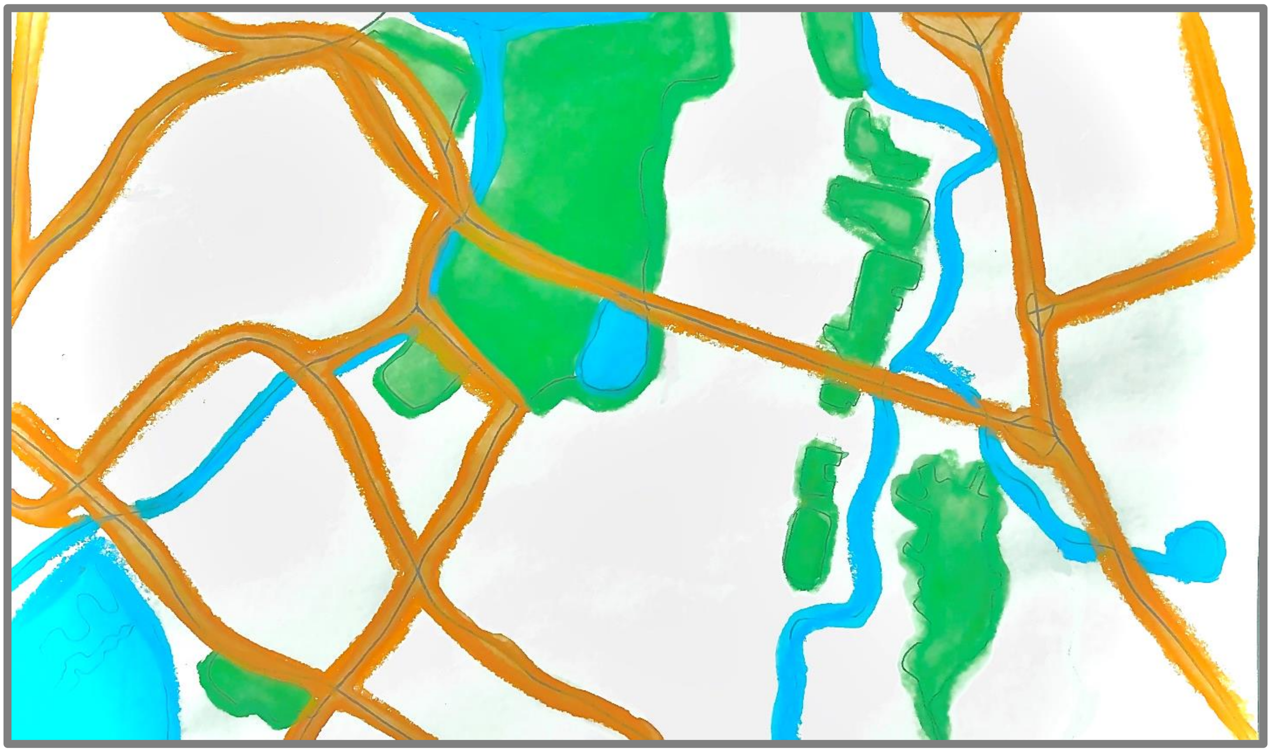
Thorndyke, P. W., & Stasz, C. (1980). Individual differences in procedures for knowledge acquisition from maps. Cognitive Psychology, 12(1), 137–175.

Townsend, J. T., Ashby, F. G., Castellan, N. J., & Restle, F. (1978). Cognitive theory.

Verhulst, N. (n.d.). A multi-method exploration of affective processes throughout the customer journey. 236.

Zagermann, J., Pfeil, U., & Reiterer, H. (2016). Measuring Cognitive Load using Eye Tracking Technology in Visual Computing. Proceedings of the Beyond Time and Errors on Novel Evaluation Methods for Visualization - BELIV '16, 78–85.

Chapter 6:
AoI-based Eye Tracking Analysis



"All perceiving is also thinking,
all reasoning is also intuition,
all observation is also invention."

Rudolf Arnheim, Art and Visual Perception, 2004

Abstract. *The contrast, geometry, color, location, shape and size of the map elements (i.e. visual variables), the distribution of the content within the map drawing area, and the task difficulty have an impact on the spatial cognition of linear and polygon map features. When measured and quantified based on the level of detail or the complexity of the map, all these characteristics can be associated with the cognition of the map element class of interest. Drawing AoIs around key elements of maps (i.e. green areas, water bodies, major rivers and roads, road junctions) is a more precise way to analyze the attention distribution of the participants. This chapter presents a preliminary AoI-based eye tracking analysis of Experiment 2 considering the participants' average fixation duration, time to first fixation and the number of objects they covered within AoIs.*

__Author Contributions:__ Conceptualization, Methodology: Merve Keskin, Ahmet Ozgur Dogru; Software: Merve Keskin; Validation: Ahmet Ozgur Dogru; Formal Analysis, Investigation: Merve Keskin; Resources: Philippe De Maeyer; Data Curation, Writing-Original Draft Preparation: Merve Keskin; Writing-Review & Editing: Merve Keskin, Ahmet Ozgur Dogru

6.1. Introduction

One of the important research questions is to elaborate on how the results of the conducted experiments can be interpreted in terms of the cartographic design principles, therefore, how they can contribute and drive map design. e.g., what kind of visualization/symbolization/generalization method can be recommended based on the distribution of the map objects? How to control advantaged and disadvantaged map objects equally/in the same way while visualizing, etc. The contrast, geometry, color, location, shape and size of the map elements (i.e. a subset of visual variables), the distribution of the content within the map drawing area, and the task difficulty have an impact on the spatial cognition of linear and polygon map features. When measured and quantified based on the level of detail or the complexity of the map, all these characteristics can be associated with the cognition of the map element class.

The results obtained in the conducted user experiments reveal that there are no big differences between the participants considering expertise and success rates. The reason for this might be that the maps used are compatible with the production purpose and the target users that is the general audience. Nevertheless, we intend to investigate further what results, whether known or unknown, could be drawn towards cartographic design from the existing findings. Due to the large size of the collected data in terms of the number of map stimuli and the main structuring elements included in them, we randomly selected 10 stimuli for which we conducted qualitative analysis with the focus/heat maps considering the map elements receiving the highest number of fixations. This provided a general overview of the participants' attentional behavior towards the map elements of interest and the similarities related to their map learning strategies (see Chapter 5.3.2). However, for measurable results, drawing AoIs around key elements of maps (i.e. green areas, water bodies, major rivers and roads, road junctions) is a more precise way to analyze the attention distribution of the participants.

6.2. Methodology

We chose one stimulus from Experiment 2 for detailed AoI-based eye tracking analysis considering the average fixation duration and time to first fixation metrics. The average fixation duration will reveal the amount of attention drawn to a certain AoI, and the shorter the time to first fixation, the earlier the attention is drawn to that region. AoIs were drawn on the selected map stimulus using SMI BeGaze, and similar to what was done in Chapter 5, related metrics were exported considering both classification of the recruited participants based on (i) expertise (i.e. expert vs novices), and (ii) overall success rates (i.e. good learners vs relatively poor learners). Not only the AoIs relevant with the task, but also the remaining AoIs within the map content were considered while calculating the eye tracking metrics. For instance, Block 2 is dedicated to the retrieval of main roads and hydrography; nevertheless, we investigated the attention distribution of the participants on green areas as well.

We present the average fixation duration and time to first fixation to the relevant AoIs for all seven experiment blocks that were grouped based on the task difficulty by emphasizing the differences and the similarities between participants. Additionally, the overall average fixation duration and the overall time to first fixation to the relevant AoIs were calculated by averaging these metrics for all blocks. To explore how many objects within an AoI group received the participants' attention can contribute to the understanding the attentional behavior of participants. For instance, how many roads are there, and how many of them received fixations and how many of them did not receive any? This could help us to understand which objects are the fundamental to focus, how much is enough in order to remember one information later or which objects are ignored and why? Therefore, using time to first fixation metrics, we counted the objects that received fixations separately for each block regarding three map object classes; roads, green areas, hydrography (rivers and water bodies). We finally discuss the attentional behavior of the participants towards the main structuring elements considering their geometry, size and distribution in the map drawing area.

6.2.1. Participants

We used the eye tracking data of 38 participants (*Mage*= 29.6, *SD*= 4.9) recruited for Experiment 2 and in this context; there were 21 novices and 17 experts. Next to level of expertise, we categorized the participants based on their overall success rates within the Experiment 2 for exploring the effect of their learning abilities on their cognitive behaviors. Therefore, 28 participants were assigned as good learners (GL) whereas 10 participants as relatively poor learners (RPL).

6.2.2. Stimulus & AoIs

AoIs depicted in Figure 6.1 correspond to the main structuring elements that were required to be remembered by participants within different difficulty levels. The total number of AoI is 53 including 25 main roads, 18 green areas and 10 hydrographic features (4 water bodies, 6 rivers). The roads cover the 27.0% of the AoIs, whereas green areas 53.6% and hydrography 19.4%.

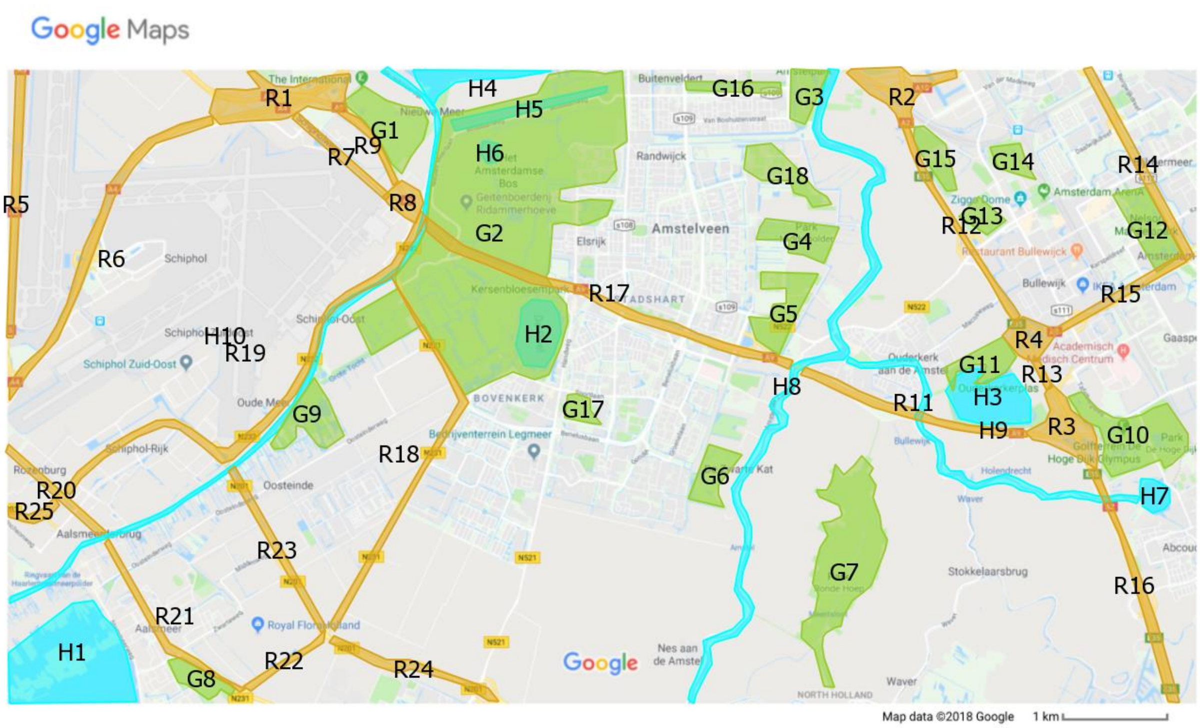


Figure 6.1. AoIs correspond to the main structuring elements of the selected map stimulus from Experiment 2.

6.3. Results

6.3.1. Average fixation duration

The average fixation duration for the tasks demanding the recall of the polygon features were the longest for all participants, i.e. Block 5 (green areas), Block 6 (hydography), whereas the participants exhibited the shortest average durations when the task required the retrieval of linear features, i.e. Block 7 (roads). On the one side, we observed similarities between participants in terms of the AoIs that did not attract attention at all (Table 6.1). R22, R24, R25 (except for Block 2 (roads and hydrography)) and G8 were the AoIs neglected by all participants for all blocks regardless of task demands. All of them were located within the lower left corner. Furthermore, R5, which is the road lies close to the upper left corner of the map drawing area, received fixations only for Block 7 when the task demanded the retrieval of the main roads (Figure 6.2).

Table 6.1. The full list of the common AoIs that did not receive any fixations

Blocks	Roads	Green Areas	Hydrography
Block 1 (all elements)	R25, R24, R23, R22, R20, R16, R5, R2	G17, G16, G15, G12, G8, G6, G1	H7, H6
Block 2 (Roads & Hydrography)	R24, R22, R21, R16, R15, R14, R13, R6, R5	G16, G14, G8, G6, G3	-
Block 3 (Roads & Green areas)	R25, R24, R23, R22, R18, R16, R15, R14, R9, R6, R5, R1	G8	H7
Block 4 (Green areas & Hydrography)	R25, R24, R23, R22, R20, R18, R16, R5	G14, G8, G3	H7
Block 5 (Green areas)	R25, R24, R23, R22, R21, R20, R16, R15, R13, R9, R8, R6, R5, R2	G16, G11, G8, G6, G3	H9, H7, H5, H1
Block 6 (Hydrography)	R25, R24, R22, R15, R9, R7, R6, R5, R2, R1	G16, G15, G14, G8, G7, G6	H7
Block 7 (Roads)	R25, R24, R22, R20, R16, R13, R11	G18, G16, G8, G7	H6, H5, H4, H1

On the other side, we focused on the differences between the participants within hard, moderate and easy tasks.

Experts vs. Novices

Hard tasks: Themost significant difference between novices and experts was observed at H5 for Block 1 (i.e. all map elements) as novices did not gaze at H5 at all, and at H1 for Block 2 (i.e. roads and hydrography) with longer fixation duration for novices. H5 is a small water canal inside G2, which is the biggest green area and the AoI of all. H1 is a water body located on the very bottom left corner (Figure 6.3).

Moderate tasks: The biggest difference between novices and experts occurred at H3 for Block 3 (i.e. roads and green areas) in a sense that novices exhibited longer fixation duration and at G1 for Block 4 (i.e. green areas and hydrography) with longer fixation duration for experts. H3 & G1 are in between two road junctions; (i) R3 and R4 that are located on the right side, close to the center, (ii) R1 and R8 lie on the upper left side, respectively (Figure 6.4).

Easy tasks: The biggest difference between novices and experts was remarked at R17 for Block 5 (i.e. green areas), at H2 for Block 6 (i.e. hydrography) both with longer fixation duration for experts, and at G11 for Block 7 (roads) with longer fixation duration for novices. R17 is the third longest road that is located in the middle of the map and partially contained within G2. H2 is a small water body also nested in G2, whereas G11 is a small area partially surrounding H3 that is the third largest water body (Figure 6.5).

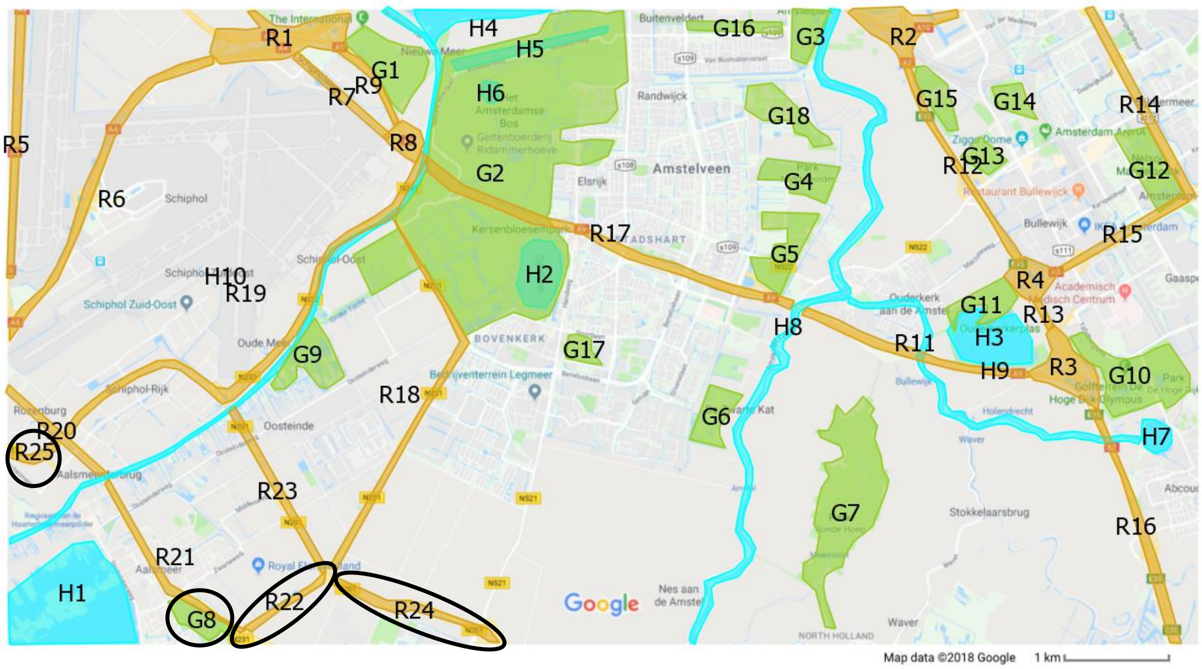


Figure 6.2. The common AoIs did not receive any fixations for none of the blocks.

Good learners vs. relatively poor learners

The differences in terms of average fixation duration was more visible between good learners (GL) and relatively poor learners (RPL) compared to those between experts and novices, especially for the hard tasks.

Hard tasks: The biggest difference between GL and RPL was observed at H3 for both Block 1 & Block 2 as RPL demonstrated longer fixation durations. H3 is in between two road junctions; R3 & R4 on the right side, close to the center (Figure 6.6).

Moderate tasks: The biggest difference between GL and RPL was observed at G13 for Block 3, and at R17 for Block 4 as in both occasions there occurred longer fixation durations for RPL. G13 is a small green area located approximately on the upper left corner of the map. R17 is the third longest road in the middle of the map, partially in contained within G2 that is the largest map element of all (Figure 6.7).

Easy tasks: On the one hand, the biggest difference between GL and RPL was observed at H2 for Block 5 as GL exhibited longer fixation duration. On the other hand, for Block 6 the biggest difference occurred at H4 and for Block 7 at G11, RPL demonstrated longer fixation durations in both cases. H2 is a small water body inside G2, H4 is a water body on the upper side of G2 and G11 is a small area partially surrounding H3 (Figure 6.8).

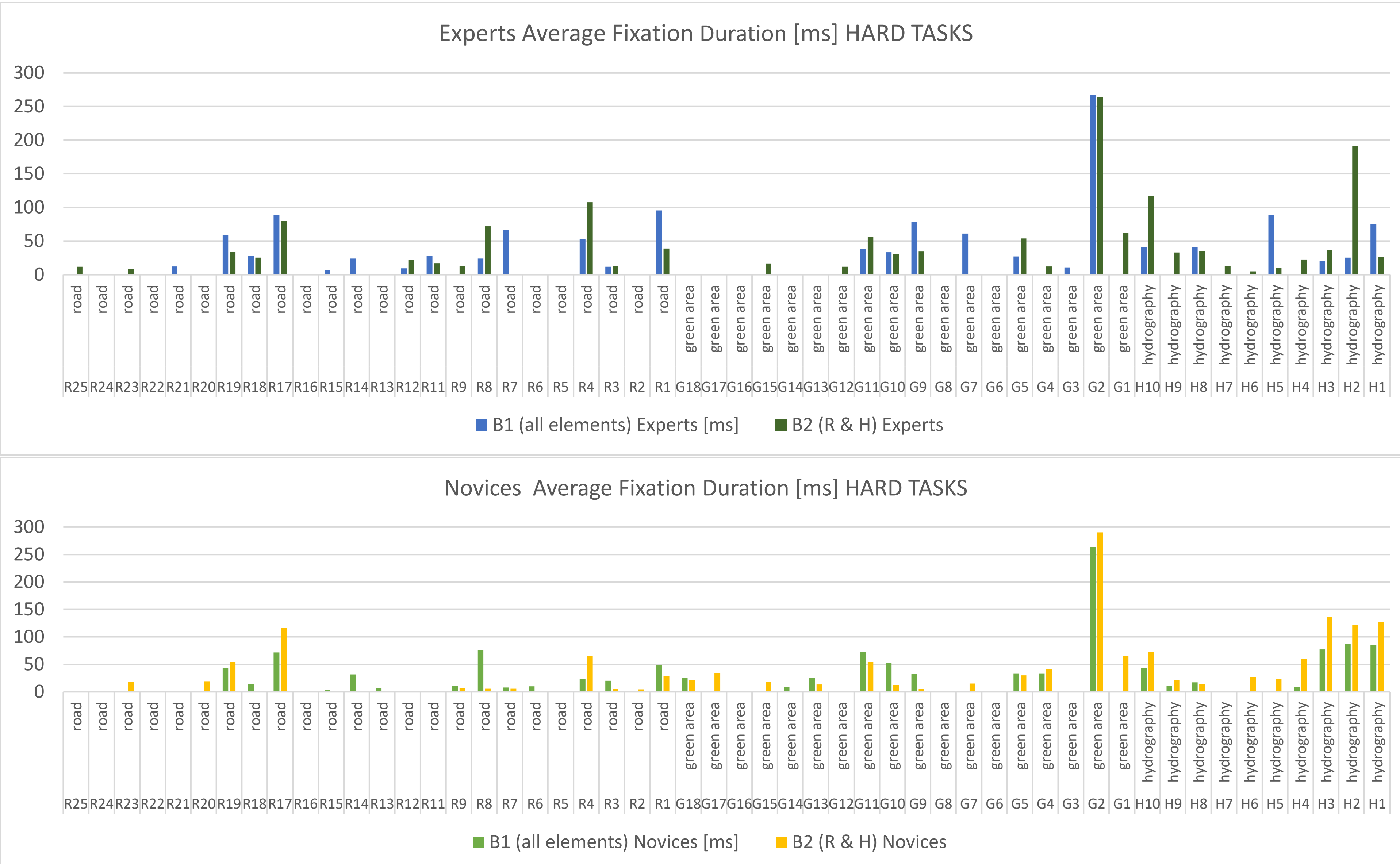


Figure 6.3. Average fixation duration for experts and novices at hard tasks (i.e. Block 1 & Block 2).

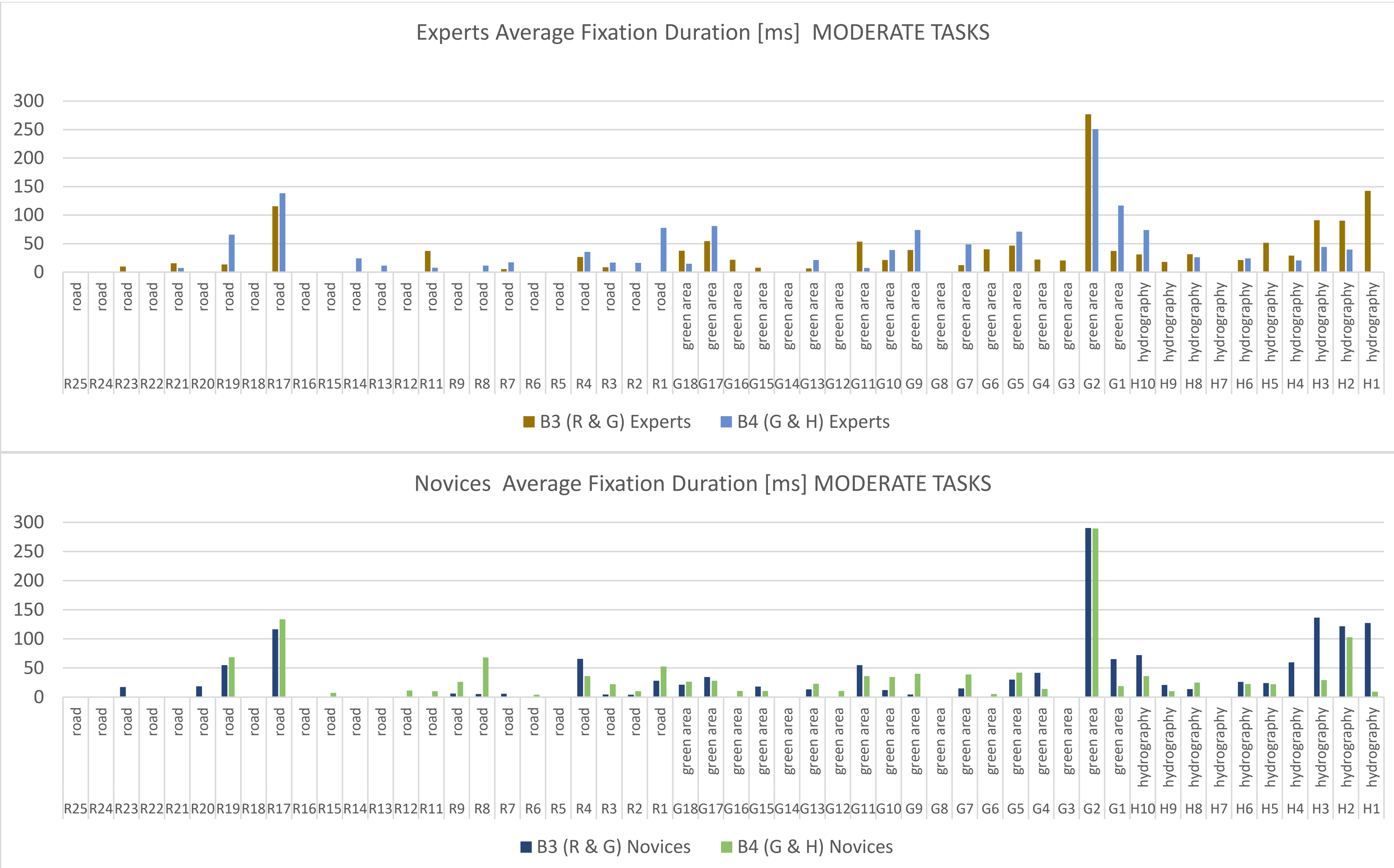


Figure 6.4. Average fixation duration for experts and novices at moderate tasks (i.e. Block 3 & Block 4).

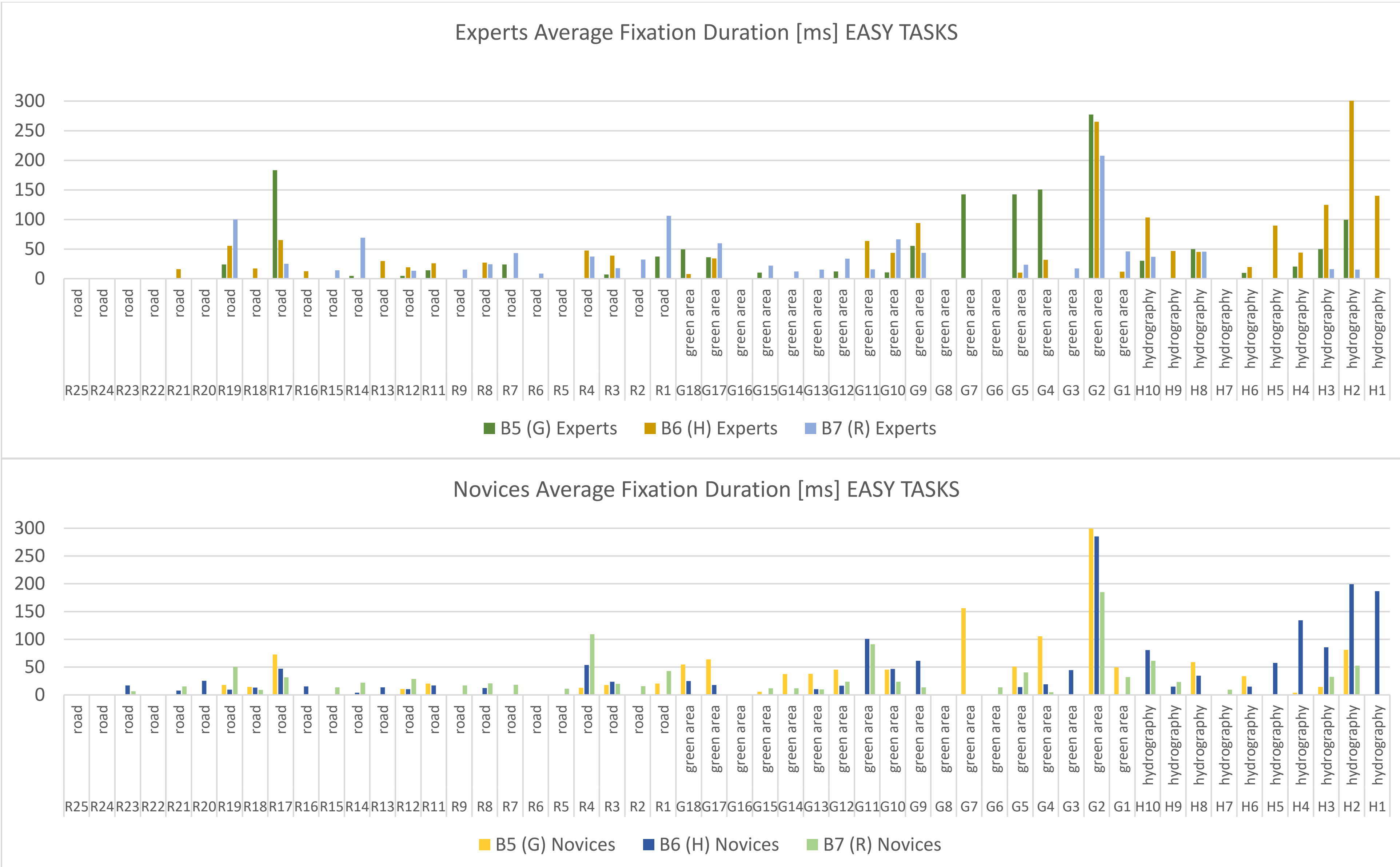


Figure 6.5. Average fixation duration for experts and novices at easy tasks (i.e. Block 5 & Block 6 & Block 7).

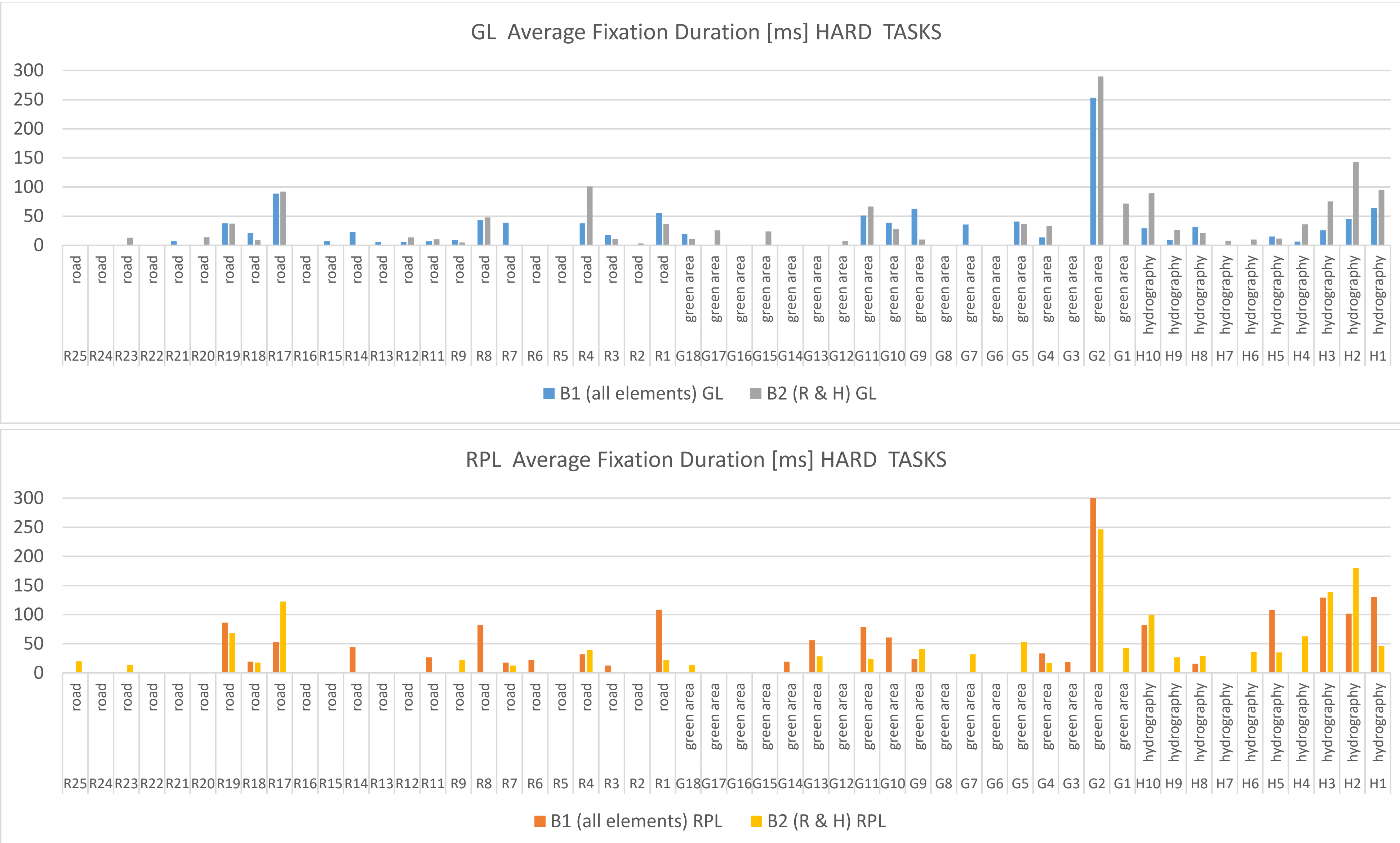
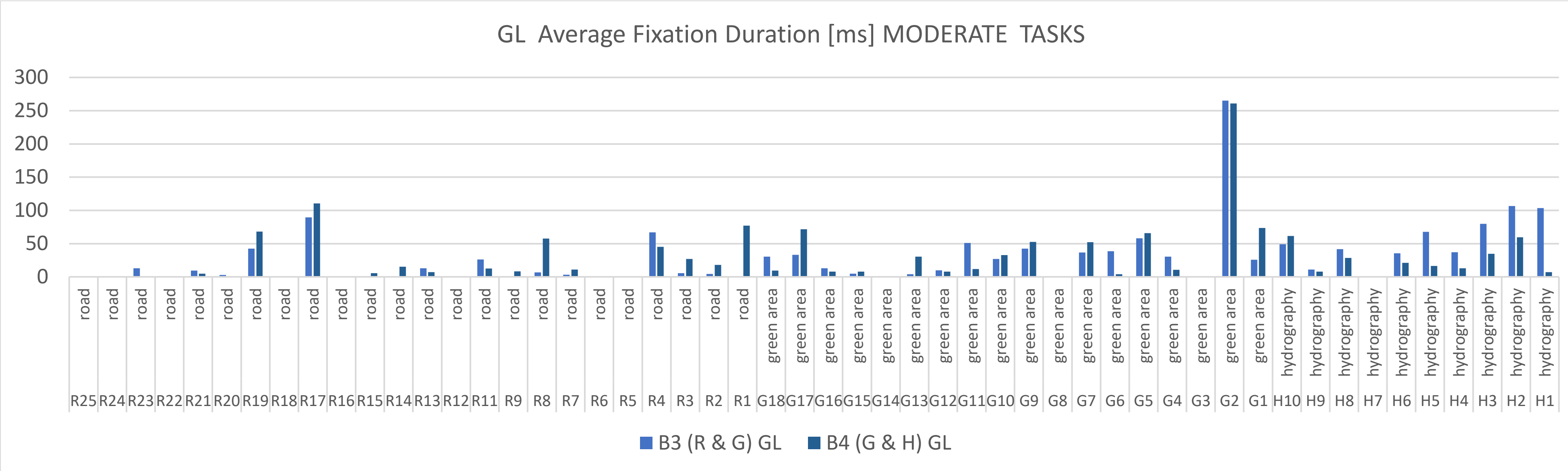


Figure 6.6. Average fixation duration for GLs and RPLs at hard tasks (i.e. Block 1 & Block 2).



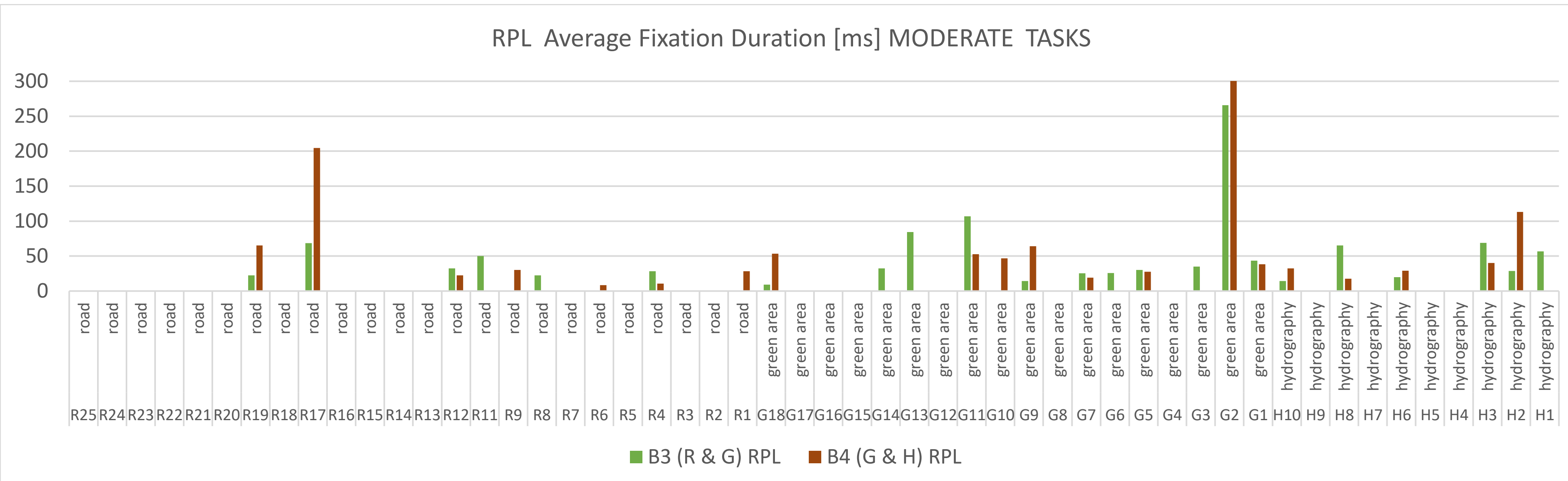
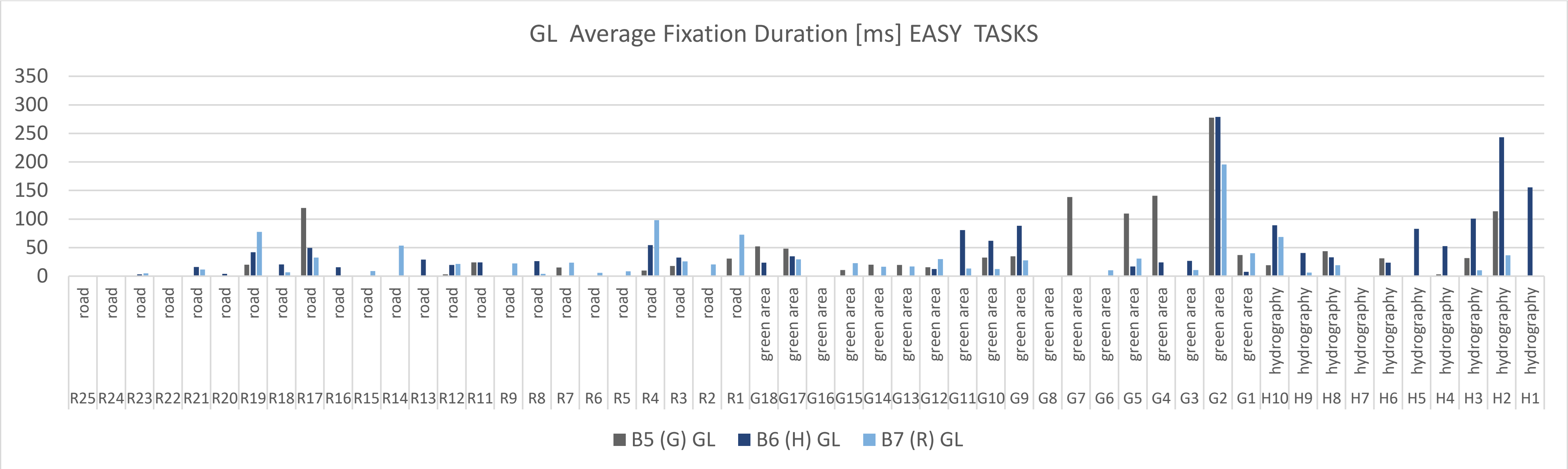


Figure 6.7. Average fixation duration for GL and RPL at moderate tasks (i.e. Block 3 & Block 4).



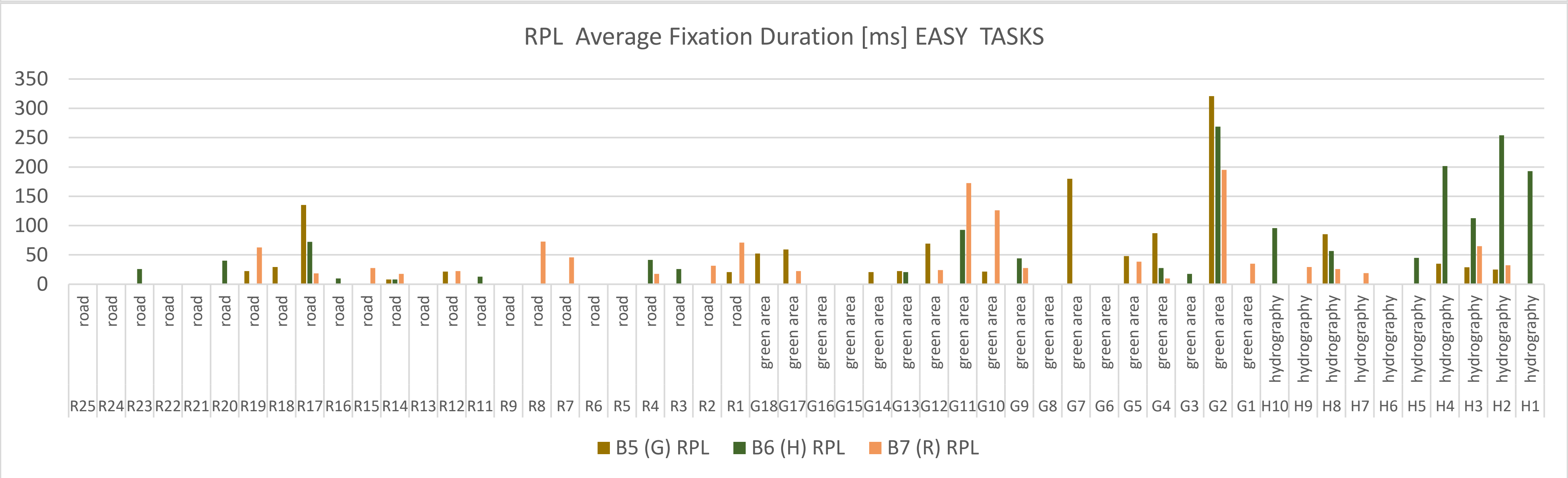


Figure 6.8. Average fixation duration for GL and RPL at easy tasks (i.e. Block 5 & Block 6 & Block 7)

We summarized the average fixation durations within all 53 AoIs based on seven experiment blocks for experts and novices (Figure 6.9a), for GL vs RPL (Figure 6.9b) and showed a comparison of their overall average fixation durations averaged for all blocks (Figure 6.9c). The average fixation durations within AoIs relevant with the task were additionally presented regarding both expert vs novice (Figure 6.9d), and GL vs RPL (Figure 6.9e) classifications. The overall average fixation durations of all groups averaged for each block are shown in Figure 6.9f. The overall average fixation durations within the AoIs relevant with the task were higher than those were within all 53 AoIs. We additionally observed that the differences between novices and experts are positive in sense that experts had longer fixation durations, whereas the differences between RPL and GL are negative in a way that RPL had longer fixation durations.

6.3.2. Time to first fixation

The participants took longer time to focus on AoIs associated with Block 1, 2 and 3 whereas they focused on the AoIs with Block 4 quicker. We do not have a specific pattern (e.g. decreasing/increasing with the task, decreasing/increasing based on the object geometry or the number of object classes within the block) for time to first fixation neither between experts and novices nor between GL and RPL. The biggest difference between experts and novices occurred for Block 3 (800ms) (Figure 6.10a) and between RPL and GL for Block 2 (618ms) (Figure 6.10b). Unlike experts, RPL spent the longest time to fixate on the relevant AoIs. This finding shows that the experts were the fastest of all to gaze at the map objects of interest (Figure 6.10c).

6.3.3. The number of objects covered within AoIs

The total number of objects covered within all seven blocks is 235 for GL, 219 for novices, 199 for experts and 160 for RPL. This result favors GL. We present the rest of the results as percentages in Figure 6.11. A similar trend is observed for all groups, namely experts, novices, GL and RPL in a way that all participants covered the objects within hydrography, the most; green areas, the second; and the roads, the last. When the task demanded the recall of all map elements (i.e. Block 1), all user groups paid attention to approximately (+/-10%) the half of the map objects of interest (i.e. roads, green areas, hydrography) as hydrography being the most.

The biggest difference between experts and novices in terms of the total number objects they covered was observed in Block 4 (i.e. green areas and hydrography) for green areas. The number of green areas and hydrography that novice participants focused on was 28% and 20% higher than that of expert users, respectively. (Figure 6.12a,c).

The differences in terms of the map elements covered within selected AoIs are much more visible between GL and RPL compared to those between experts and novices. The biggest difference between GL and RPL in terms of the total number of objects they covered was observed in Block 3 and 4 (i.e. green areas and hydrography) for hydrography. In both blocks, the number of hydrographic objects that GL focused on was 40% higher than that of RPL did. (Figure 6.12b,d).

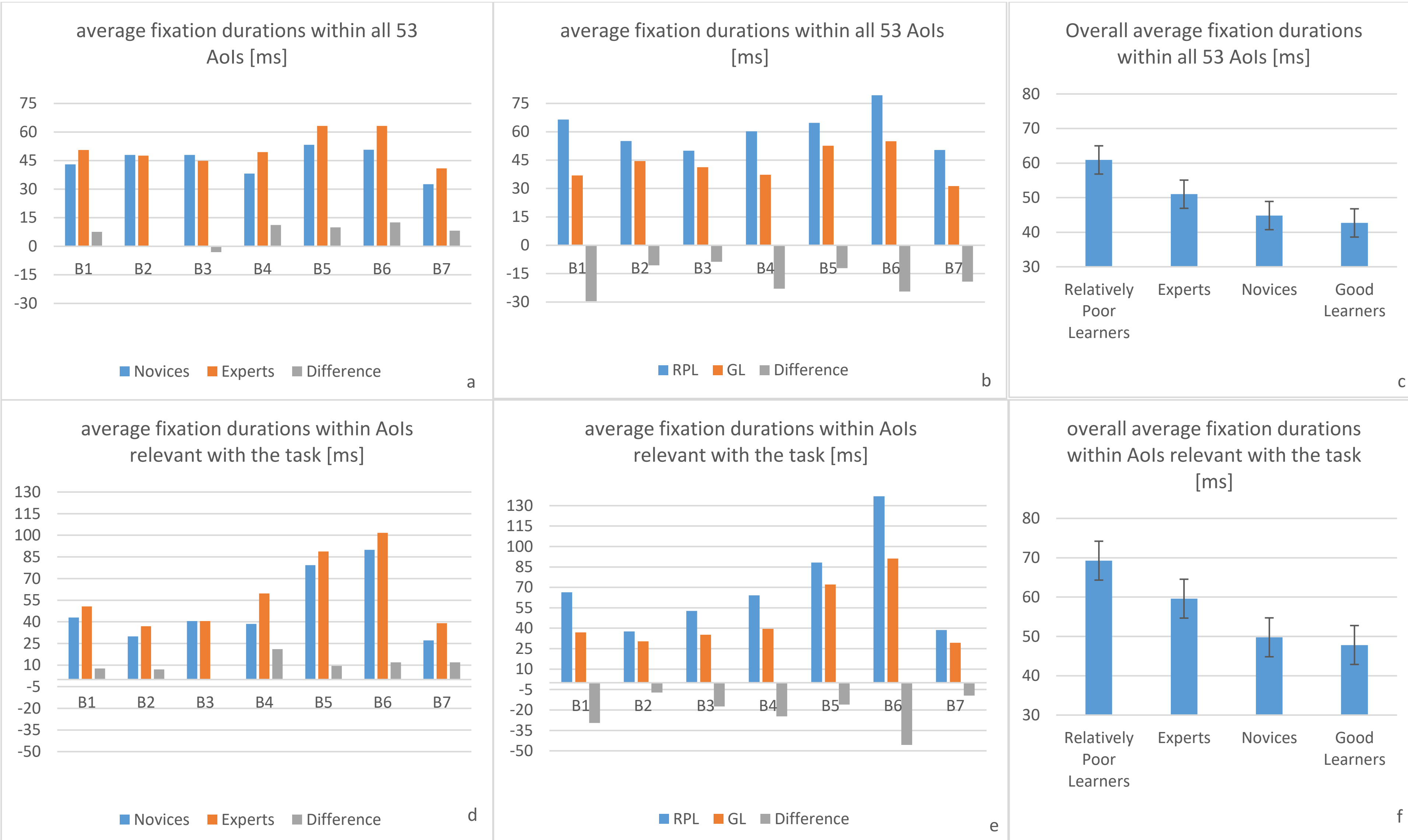


Figure 6.9. Average fixation durations within all 53 AoIs [ms] a. Experts vs. Novices; b. GL vs RPL; c. Overall average fixation durations within all 53 AoIs [ms] averaged for all blocks. Average fixation durations within AoIs relevant with the task [ms] d. Experts vs. Novices; e. GL vs RPL;. f. Overall average fixation durations within AoIs relevant with the task averaged for all blocks [ms]

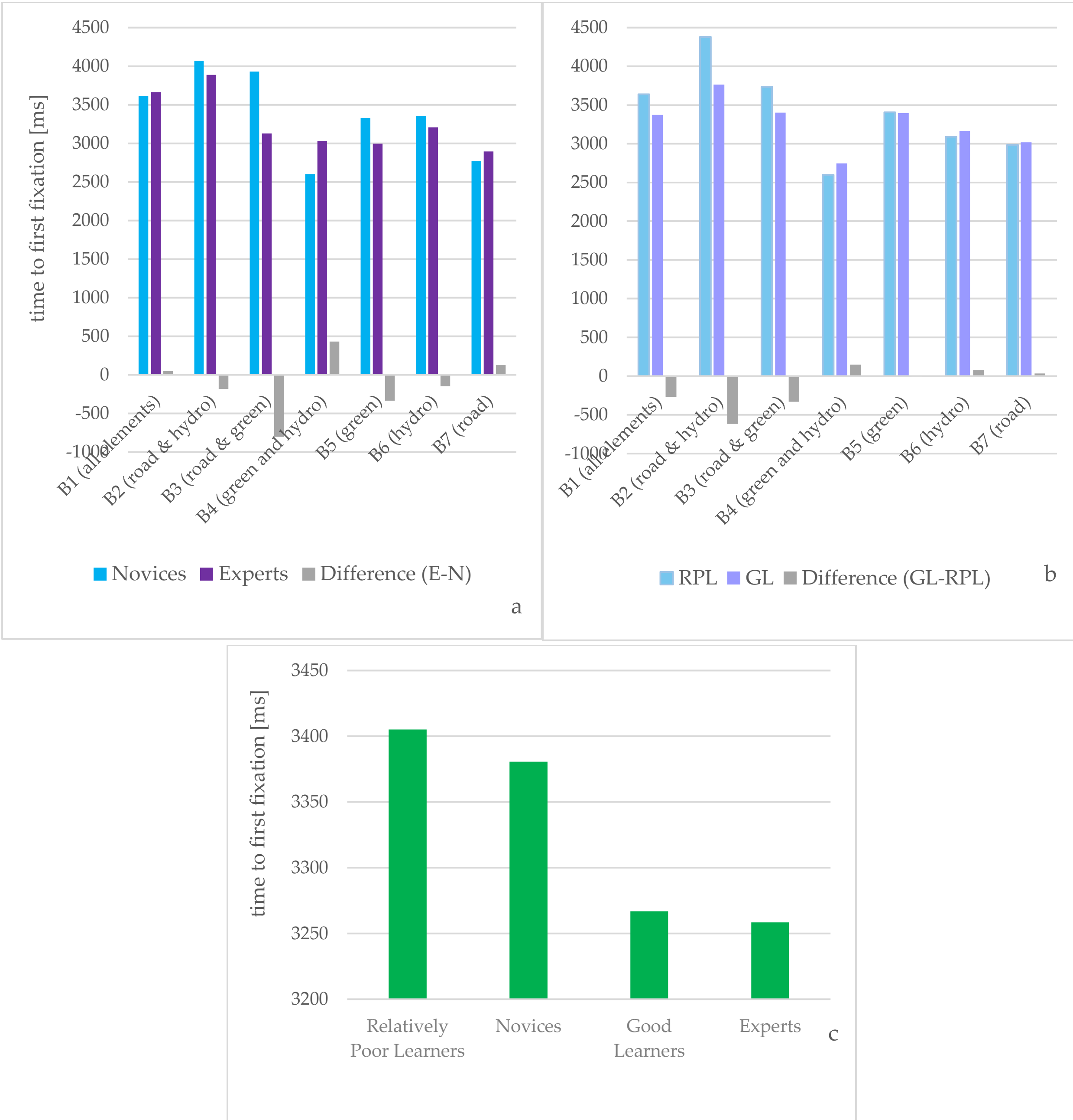


Figure 6.10. Time to first fixation to the AOIs relevant with the task [ms] a. Experts vs Novices; b. GL vs. RPL; c. All groups averaged for all blocks in comparison

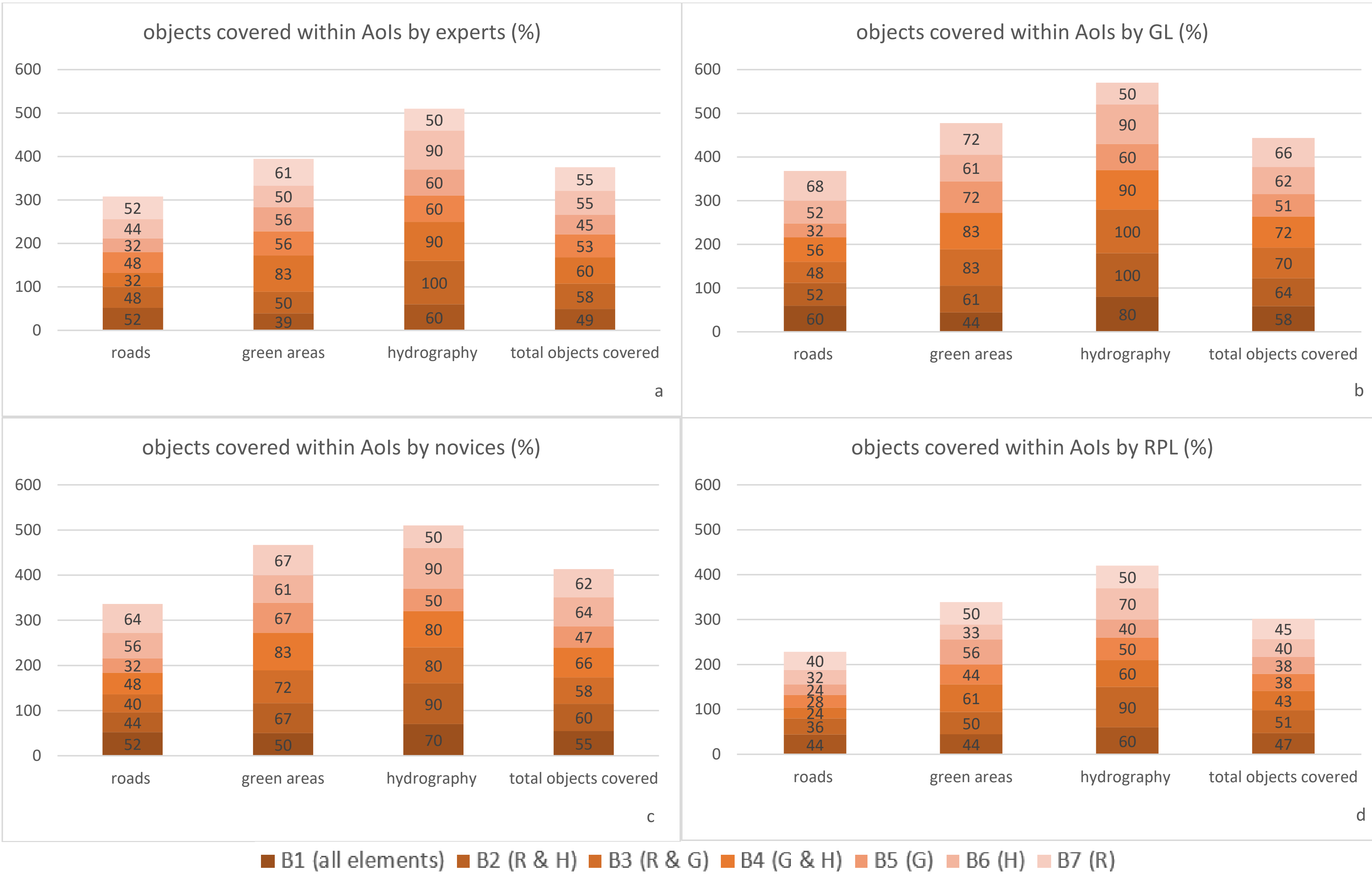


Figure 6.11. The rate of objects covered within AoIs (%) a. Experts, b. GL, c. Novices, d. RPL

Figure 6.12. The difference between groups in terms of the rate of the objects covered within AoIs (%) a. Experts vs Novice; b. GL vs RPL; The difference between groups in terms of the total number of objects covered within AoIs (%) c. Experts vs Novice; d. GL vs RPL

6.4. Discussion

We interpret the results of the average fixation duration, time to first fixation and the number of the objects covered within AoIs based on the objects' geometry, size, and the distribution in the map drawing area.

6.4.1. Average fixation duration

- The objects that did not receive any fixations from the participants are the ones located mostly in the corners, or close to the map frame. Their size or the area they covered on the map can be large, e.g. R5, R6, R14, R16 or small, e.g. G8, R25, G16, H7 (Table 6.1, Figure 6.2). This can be interpreted as the location of the map element is more influential on the participants' gaze behavior compared to its size.

- The biggest differences in average fixation durations either between experts and novices or between GL and RPL mainly occurred for the map features that are small and nested in large areas or those that are located close to the map frame or the corners of the map drawing area.
- The overall average fixation durations within the AoIs relevant with the task were higher than those were within all 53 AoIs. This can be interpreted as the task demands (e.g. the tasks in Block 7 require the retrieval of the main roads) has an impact on the attentional behavior of participants. The participants were able to neglect the map objects that were irrelevant with the task. Nevertheless, experts exhibited longer fixation durations (except for B3) compared to novices did, while GL always demonstrated lower fixation durations than the RPL did. This outcome shows that GL experienced lesser cognitive load than the rest did.
- There is no trend in terms of the differences occurred regarding to the blocks, therefore, the fixation durations within the (relevant) AoIs did not depend on the task difficulty.

6.4.2. Time to first fixation

The participants took longer time to focus on AoIs associated with Block 1, 2 and 3 whereas they focused on the AoIs with Block 4 quicker compared to other blocks. One common feature of the tasks in Block 1, Block 2 (i.e. hard tasks) and Block 3 (i.e. moderate task) is that they all required the retrieval of roads, which only contained linear objects. On the other hand, when the task demanded the retrieval of more polygon features, the participants were quicker to fixate on the relevant AoIs. For instance, Block 4 has more polygon features (i.e. green areas and hydrography) and the attentional behavior of the participants were similar towards these polygons. This might be an evidence that the participants treat linear and polygon objects differently in a sense that they tend to scan through linear objects and fixate/focus on the polygon objects. To perceive a linear object as a whole, one needs to know the start and the end of it, therefore, the eye will follow the linear object instead of focusing on parts of it unless distracted, whereas focusing at the center of a polygon object is sufficient to perceive its size and shape, especially considering the polygon AoIs in the example stimulus.

6.4.3. The number of objects covered within AoIs

- The number of objects covered within hydrography being the highest was expected due to the lowest number of objects belonging to that category (i.e. seven water bodies, three rivers) in the map drawing area. Experts and GL even covered 100% of the hydrographic objects for Block 2, which required the retrieval of road and hydrography (see Figure 6.11a, b). The second map object class in terms of the number of objects covered by participants was green areas whose areal coverage was the biggest in the map drawing area. Roads were the least covered objects compared to the other classes and this might be due to two reasons: (i) the road category having the highest number of AoIs, which makes it hard to pay attention to all of them in a short time, (ii) the distribution and the location of the roads in the map drawing area, e.g. roads located close to the map frame. However, when the task only demanded the retrieval of the roads (i.e. Block 7), the number of road objects covered was the highest above all other blocks for all groups (see roads for Block 1-7 in Figure 6.11a). GL covered the highest number of road objects in total and in Block 7. We do not observe a similar behavior when roads had to be remembered together with the other type of objects (i.e. Block 2: roads and hydrography; Block 3: roads and green areas). According to these findings, similar to what was observed within time to first fixation results; the eye scans through linear objects and fixates/focuses on the polygon objects.

- The difference between experts and novices was more visible when the task required the retrieval of more polygon features as in Block 4 (i.e. green areas and hydrography) for green areas (Figure 6.11a, c). However, it did not influence the success rates of the participants since the average success rate in Block 4 was 93% for experts and 91% for novices. Similarly, the biggest difference between GL and RPL was observed in Block 3 (i.e. roads and green areas) and Block 4 (i.e. green areas and hydrography) for hydrography (Figure 6.12b,d). This has an influence on the success rates in a sense that the average success rate in Block 4, 83.6% for RPL and 94.3% for GL. This outcome might be an evidence that 'good learners vs. relatively poor learners' classification was probably more realistic for usability testing of the general use maps for basic map learning tasks. For instance, Havelková and Gołębiowska (2020) claimed that novices/unsuccessful solvers are not one homogenous group; however, it is possible to categorize them into subgroups. Furthermore, expertise is not as influential as we think in these cases because '*the subject lacking domain-specific knowledge tends to construct a visual-spatial mental representation, as opposed to the semantic representation of the expert* (Postigo and Pozo [40] (p. 77-78), subsequently, the performances of non-expert subjects can be as good as the expert ones.

6.5. Conclusion

The obtained results are just preliminary and specific for the map stimulus used for the AoI analysis yet the anlaysis involved a large number of trials belonging to many participants. The results cannot be generalized due to uneven groups (28 GL, 10 RPL) and also because of the post-hoc analysis conducted with the same experiemental data; nonetheless, we list some of the hightlights of the AoI-based eye tracking analysis:

- The location of the map elements, as well as their distribution in the map drawing area, is more influential on the participants' gaze behavior compared to the size of them.
- The participants were usually able to neglect the map objects that were irrelevant with the task.
- Experts were the fastes to fixate on an AoI relevant with the task; however, their fixation durations were longer than of RPL.
- GL always demonstrated lower fixation durations and covered more map objects than the RPL did. GL experienced lesser cognitive load than the rest.
- We did not find a correlation between the average fixation duration and the task difficulty, hence, the fixation durations within the (relevant) AoIs did not depend on the task demands.
- The participants took longer time to focus on AoIs associated with the roads whereas were quicker to fixate on the relevant AoIs when the task demanded the retrieval of more polygon features.
- When the task only required the retrieval of the roads (i.e. Block 7), the number of road objects covered was the highest above all other blocks for all groups.
- The eye scans through linear objects and fixates/focuses on the polygon objects.
- Good learners vs. relatively poor learners' classification was probably more realistic for usability testing of the general use maps for basic map learning tasks.

According to the above results, GL's higher number of fixations towards the relevant AoIs and the objects covered within them and their shorter average fixation durations indicated that GL are good at formulating a learning strategy, therefore, spent less cognitive effort. Our results also showed that GL and expert participants had shorter time to first fixation to AoIs relevant with the task and that their selective attention was influenced by the task demands. These findings will further be discussed within the frame of the cartographic design recommendations made in Discussion (see Chapter 7, Research Objective 4). With the existing dataset, this study can be extended by increasing the number of map stimuli and the AoI metrics for more detailed insights on the attentional behaviors of map users. The influence of other visual variables can be included as well, for instance, the effect of the location of the objects within the map drawing area and relative to each other can be explored. Additionally, unlike what was applied in our preliminary analysis, the intertwining AoIs (e.g. polygon in polygon, line in polygon, line intersecting line) should be treated separately for more accurate conclusions.

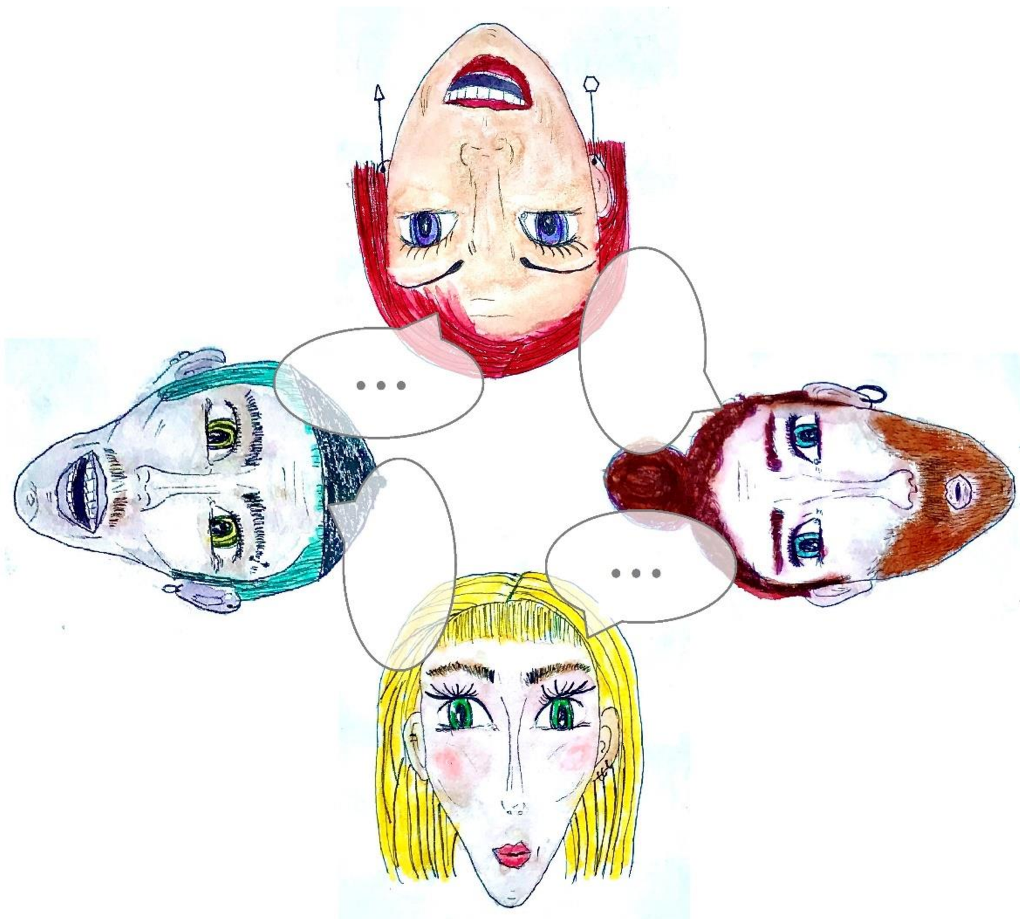
References

Havelková, L., & Gołębiowska, I. M. (2020). What Went Wrong for Bad Solvers during Thematic Map Analysis? Lessons Learned from an Eye-Tracking Study. ISPRS International Journal of Geo-Information, 9(1), 9.

Postigo, Y., & Pozo, J. I. (1998). The Learning of a Geographical Map by Experts and Novices. Educational Psychology, 18(1), 65–80.

Thorndyke, P. W., & Stasz, C. (1980). Individual Differences in Procedures for Knowledge Acquisition from Maps. Cognitive Psychology, 12(1), 137- 175.

Chapter 7: General Discussion



"Constant reminding ourselves that we not see with our eyes
but with our synergetic eye-brain system working as a whole
will produce constant astonishment as we notice,
more and more often,
how much of our perceptions emerge from our preconceptions."

David Eagleman, Incognito: The Secret Lives of the Brain, 2011

Abstract. *This chapter summarises and discusses the results obtained throughout the dissertation. Within each of the previous chapters, a detailed description of the achievements is presented. The aim of this chapter is to provide a link between these findings and place them in a broader context by discussing these findings in the light of the research objectives and questions that were identified in Chapter 1.2.1. Finally, this chapter concludes with the critical reflection on methodology, the challenges and venues for future work.*

In this discussion, we aim to reconsider the major findings of the user experiments mentioned in the previous chapters and provide concise overview of all the results. Table 7.1 demostrates the highlights of the two conducted user experiments. In the context of the study, certain research objectives, which were solidified though research questions were defined and described in Chapter 1. Each of the chapters in this dissertation focuses on different aspects of one or more research questions and, these different contributions are summarised and discussed in relation to each research question in the following sections.

Table 7.1. Overview of the dissertations' highlights with respect to expert vs. novice categorization

Experiment 1		Experiment 2	
Metrics used	**Highlights**	**Metrics used**	**Highlights**
reaction time (the study period)	Experts spent longer time to study the map stimulus.	**reaction time (the answering period)**	Experts spent longer time to choose the correct answer from the multiple choice graphical response screen.
score on drawn elements on sketch maps	Although not significantly, experts had higher scores and their sketch maps of experts were more detailed	**success rate**	Although not significant, experts had higher success rates, however, significant difference between GL and RPL in a sense that GL remembered more map elements.
the influence of the visual variables on the attentional behavior	Size, color, shape and location of a map object has an equally important role in recalling a map element	**the influence of the visual variables on the attentional behavior**	The location of the map elements, as well as their distribution in the map drawing area, is more influential on the participants' gaze behavior compared to the size of them.
drawing order	The majority of the participants first drew linear objects: experts: roads (red); novices: hydrography (blue) they might have used different strategies since hydrography contains both linear and polygon features.		
average fixation duration	Although not significantly, the average fixation duration of novices were longer.	**average fixation duration**	Although not significant, fixation durations of novices were longer, and the difference between experts and novices increased as the task difficulty increased
number of fixations per second	No significant difference between experts and novices	**number of fixations per second**	Although not significant, the number of fixations (per second) of experts was higher, and the difference increased as the task difficulty decreased.
		saccade velocity & amplitude	Novices had higher saccade velocity, which indicate a higher cognitive load and longer saccade amplitudes indicating that the search is less organized.

Experiment 1		Experiment 2	
qualitative analysis of focus maps	The focus map belonging to all participants highlighted the main structuring elements of the stimulus: main road construction, water bodies and large settlements.	**qualitative analysis of focus maps**	Both groups showed similar attentional behavior in terms of the map area covered and the map elements on which they focused
AoI-based eye tracking	Larger AoIs were gazed at earliest and the dwell times for such objects were much longer and the majority of participants drew these map elements on their sketch maps.	**AoI-based eye tracking**	The participants were usually able to neglect the map objects that were irrelevant with the task.
	The linear features were easier to learn and remember, although the viewer did not pay much attention.		The participants took longer time to focus on AoIs associated with the roads whereas were quicker to fixate on the relevant AoIs when the task demanded the retrieval of more polygon features.
			The participants tend to scan through the linear objects and fixate/focus on polygon objects
	The objects that were absent on the sketch map received the shortest fixation durations during the study phase		Experts were the fastest to fixate on an AoI relevant with the task
average alpha power	Experts spent considerably lesser cognitive load	**the event-related changes in power spectral density at theta frequency band**	Although not significantly, experts demonstrated higher theta power (except for easy tasks) which can be associated with a higher cognitive load
FAA (frontal alpha asymmetry)	All the participants exhibited negative FAA scores, which suggested a greater relative right activation, therefore, withdrawal-related motivation. No significant difference occurred between experts and novices.	**the event-related changes in power spectral density at alpha frequency band**	ERD in alpha was observed but alpha effect seem very small

7.1. Revisiting the Research Questions

Research Objective 1:

Contribute to the understanding of how different map users process the visual information on digital 2D static maps.

RQ 1: How do expert and novice map users "study and store" the visual information presented on digital 2D static maps?

We aimed to explore the attentional behavior of the participants during the map study/map-learning period. In this context, Experiment 1 included a simple map-learning task using one simplified digital 2D topographic map with no time constraints. Reaction times described in Chapter 2 show that experts allocated more time in studying than novices did. This result might contradict with the traditional expert-novice paradigm (e.g. Dong et al., 2018; Li et al., 2013; Ooms et al., 2012), however, it shows that experts take their time while comprehending a map stimuli they see for the first time. Allen and McGeorge (2011) studied the enumeration process of the air traffic controllers on stimuli containing the same number of objects whose distribution arranged differently (e.g. line, triangle, etc.) and found out that experts out-perform novices but their response times were significantly slower. This finding suggests that *"expertise is not rigid and automatic but, rather, is flexible and responsive to the specific situation, allowing experts to switch between strategies*". Experiment 2 was complex in a way that it contained large number of stimuli and tasks with varying difficulties and since all participants had a fixed amount of time to study the map stimulus, reaction time corresponded to the time when they decide which skeleton map to choose from the multiple choice graphical screen. Therefore, it is the retrieval of the main structuring elements and will be discussed in relation to the success rates within the RQ 2.

A visual stimulus can only be interpreted if the attention is subsequently concentrated on different sub parts of the stimulus and the attentional behavior of the users can be identified by registering the eye movements (Ooms, 2012). The visual analyses of the recorded eye movements were addressed in Chapter 2, 5 and 6. In this context, although not significantly, the average fixation duration of novices were longer in both experiments. Additionally, the difference between experts and novices increased as the task difficulty increased. Next to the fixation related metrics, analyzing saccades in the second experiement provided insights that the more traditional fixation analyses do not afford. The saccade-related analysis of Experiment 2 shows that novices had higher saccade velocity, which hints a higher cognitive load for novices and longer saccade amplitudes indicating that the search goes all across the image and is thus less organized. On the contrary, some of the previous studies (e.g. Gegenfurtner, Lehtinen & Säljö, 2011) claimed that the ability for holistic analysis should be reflected in longer saccade length; hence, experts should demonstrate longer saccade length. However, the reason of the shorter saccades of the experts in this experiment is that shorter saccades may increase the number of fixations per second and experts exhibited greater number of fixation per second than the novices did (Ooms, 2012).

Alternatively, in a post-hoc hypothesis, we examined high-performers' visual strategies against low-performers. In this context, the participants in Experiment 2 were grouped as good learners (GL) and relatively poor learners (RPL) based on their success rates. The saccadic eye movements showed that there may be differences between high and low performers in a way that GL appeared to have higher cognitive load which was also the case with experts in some metrics. Perhaps the experts were trying harder, because there is also a speed accuracy trade off and this outcome is in line with what we found in the first experiment. Although we followed a valid analytical approach, the issue with the comparison of GL and RPL is that it was based on the uneven sample sizes (i.e. 15 GL and 5 RPL). Therefore, we acknowledge that this analysis might lead statistical power issues and needs to be further studied with more participants to solidify the observations.

Additionally, all of the participants exhibited negative FAA scores, which suggested a greater relative right activation, therefore, withdrawal-related motivation and no significant difference occurred between experts and novices. The average EEG alpha power results of Experiment 1 indicated that experts spent significantly lesser cognitive load while studying the map stimulus. On the contrary, in Experiment 2, although not significantly, experts demonstrated higher theta power (except for easy tasks) than the novices did which can be associated with a higher cognitive load. The reason for experts experiencing low cognitive load in Experiment 1 and experts and novices showed no significant difference in Experiment 2 might be due to the time constraints in the second experiement.

RQ 2: *How do expert and novice map users "recall and use" the visual information presented on digital 2D static maps?*

We aimed to explore how the cognitive load varies between two groups during the retrieval stage and whether these two groups use different strategies while remembering the map-related information. The results of Experiment 1 shows that the majority of the participants first drew linear objects on their sketch maps in a sense that experts started with roads (depicted in red) and novices started with hydrography (depicted in blue). It is evident from the findings that expert and novice users might have used different retrieval approaches since hydrography contains both linear and polygon features. Furthermore, objects of a same category, for example lakes or rivers, were mostly drawn together as a group. As also found by Ooms (2012), this outcome agrees with the Gestalt Theory claiming that humans try to organise or group the objects in a visual image.

In Experiment 1, although not significantly, experts had higher scores on drawn elements on their sketch maps but the average drawing time for experts was observed to be greater than that for novices. The time spent on sketching the map might correspond to the richness of detail depicted in the sketch map, the difficulties encountered due to the lack of experience (e.g. unfamiliarity of the task and of the drawing tool), or recall issues. The fact that novices were faster may explain that novices were less involved with cartography, not aware of procedures involved in map production and did not exactly know what to remember, thus they might have paid less attention to having good results (see Chapter 2). Similar to the Experiment 1, experts generally had significantly longer time to choose the correct answer from the multiple-choice graphical response screen yet again higher success rates in Experiment 2. It seems the fact that experts being slower paid off with higher success rates.

The scores of the sketch maps in the first experiment and success rates in the second experiment implied that experts and novices showed no difference in map learning, unless the stimulus required specific map knowledge that only an expert possessed. We presumed that the general map knowledge, maplikeness of the stimulus (e.g. Thorndyke & Stasz, 1980), the simplicity of the map and the task (e.g. Kulhavy & Stock, 1996) had a great influence on their map learning (study and recall) process. In Experiment 1, the original map shown to participants was a simplified 1:10k Belgian topographic map which incuded no familiar places and both experts and novices observed for the first time. In Experiment 2, we used screenshots of Google's road maps (approximately 1:40k) that are produced for general audience and we paid attention that the regions covered in maps were not widely known. User characteristics of the recruited participants confirm that almost all participants use Google Maps everyday or once/twice a week (i.e. experts 94% and novices 90%) and they find it easy to use (i.e. experts 100%, novice 90%) (see Annex 6). Although our results present that experts and novices do not differ in terms of the amount of information they recall, the learning/recalling strategies of experts and novices may differ. The drawing order results could be evidence that they might use different approaches (Chapter 2). Longer reaction times are generally associated with the higher cognitive load but it is not accurate to interpret the slower responses of experts in both experiments as being purposeless guesses, less motivated performance or confusion, since they maintain a greater accuracy. As explained by Allen and McGeorge (2011), it takes time to mobilize and focus attention and experts obviously did something more attentionally demanding and spent additional time for a strategically beneficial purpose. Havelková and Gołębiowska (2020) also confirms this outcome by explaining as follows "*slow task solving is not a feature of inexperienced behavior or inefficient strategy, as it is attributable to their endeavor to solve the task correctly. It is characteristic to go back to the task formulation and the phase of solving the problem after already comparing the solution found with the possible solutions given, i.e., by verifying that the solution obtained is correct, even when a different problem-solving strategy was used*". In this context, experts might be a bit more ambitious and driven to accomplish the task compared to novices, and saved an extra time to review or verify their response before submitting it. (Chapter 5). When we look at the user characteristics of the recruited participants in Experiment 2, it is also clear that experts found the experiment more interesting compared to novices did, i.e. 65% of experts reported positive comments whereas only 24% of the novices gave positive feedback (see Annex 6).

When combining the discussion of RQ1 and RQ2, it is obvious that the main structuring elements act as a reference frame and the design of these elements essential in the whole communication process during both the interpretation and recall phase, in a way that they must be visible and organized to guide map users' attention. This is linked with which objects are stored in the long-term memory and whether they can be retrieved easily or in a structured way (Ooms, 2012).

The overall performances of experts and novices did not differ much; therefore, we assume that the influence of the expertise in a simple map-learning task is not substantial. This situation is not unexpected as also found by Gegenfurtner et al. (2011) that the performance differences are smaller when the visualization is static as in our experiments. Furthermore, the participants can be considered as sufficiently homogenous in terms of their spatial abilities or experience with maps for this specific map-learning task with simple map stimuli (e.g. Havelková & Gołębiowska, 2020). Comparing the performaces of GL and

RPL, we observed a significant difference between these groups in terms of the changes in EEG power spectral density. Accordingly, GL exhibited slightly higher overall cognitive load regardless of the difficulty of the tasks. On the contrary, AoI analysis indicated that GL remembered more map elements, had more fixations on the relevant information and demonstrated shorter average fixation durations than RPL did. Similar to what was found by Havelková and Gołębiowska (2020), unsuccessful participants differed in the general problem-solving approach, in a way that they tended to choose fast and less cautious strategies and they were also not as good as GL to distinguish relevant from irrelevant information. Nevertheless, all unsuccessful participants do not necessarily used inefficient strategies, since average successful participants needed more time to respond as explained above. As we explained previously, the distinction of GL and RPL is very much dependent on our specific sample in the experiment and since this sample size is quite small, we doubt the generalizability of the results, yet we find this disctinction quite useful and worth repeating with an even and larger sample. This classification gave us a hint that it might be better to group participants based on their spatial abilities next to their expertise, especially for further studies.

Research Objective 2:
Evaluate the potential of brain imaging techniques, integration of EEG with eye tracking for cartographic cognitive/usability research

RQ 3: What is the added-value of EEG in terms of cartographic usability research?

EEG and ET integration is not straightforward and time consuming especially when the experimenter lacks experience in EEG experimental design and analysis. Hence, we first intended to explain why the co-registration of eye movements and brain activity is important for cartographic usability research, specifically to explore the attentional behaviors of map users (Chapter 1). Later on, we strived to present a methodology for EEG and eye tracking experiment with a map-learning use case based on the technical availabilities and limitations in Chapter 3. We hope to contribute to the cartographic usability research by borrowing theories and methods from experimental and cognitive psychology and applying them in cartographic domain. In this context, the pros and cons of this methodology and possible analysis are clearly explainded based on the existing literature, knowledge of the domain experts and our hands-on experience.

Table 7.2 presents an overview of the time spent for the design, data collection and the analyses together with the data in numbers. When Experiment 1 and 2 are compared, two different approaches followed in the experimental design; either single trials but rather long EEG recordings per participant or multiple trials (e.g. at least 100 trials per condition) but short EEG recording per trial. It is clear that the preparation and the data management of Experiment 2 are much more complex, require expertise and might discourage the researchers who wants to conduct a similar research due to the requirement of great amount of time and labor. Many steps can be automatized such as converting, filtering, cleaning and analyzing the collected data with scripting (e.g. Annex 4 & 5). Nonetheless, some steps, such as excluding the bad EEG channels after filtering or removing bad epochs after segmentation, require visual inspection although it can be handled automatically. Furthermore, stimuli preparation might be automatized as well; however, it was a little tricky while identfying the main structuring elements on the map stimulus to compose the skeleton maps shown in the multiple-choice graphical answer panels as cartographic generalization was involved.

Table 7.2. Collected data in numbers and time spent for the user experiments

		Experiment 1	**Experiment 2**
Experimental Design		in collaboration with a cartographer who is an eye tracking expert (1 month)	in collaboration with an experimental psychologist who is an expert in ET and EEG (12 months)
Stimuli Preparation		one trial with one map stimulus that was existent before the study free-hand digitial sketch map drawings of participants	- 50 map stimuli - 50 trials * 7 blocks * 4 graphical options 1400 skeleton maps prepared by digitizing the map stimuli (4 months)
Pilot tests & EEG & ET recording		- 0.5 h (per participant) (5 participants per day) - 3 p. For pilots & 57 p. for the main experiment 0.5 * 60 = 30 h – 12 days	- 3 h (per participant) (1 or 2 participants per day) - 10 p. for pilots & 38 p. for the main experiment 3 * 48 = 144 h – 25 days
	Raw EEG data	*5.5 GB*	*32 GB*
	Raw ET data	*3 GB*	*10 GB*
Preprocessing the EEG data	Data export	10 days	17 days
	Filtering the EEG data	4 GB (15 days)	10 GB (15 days)
	Adding the events	-	8 GB (15 days)
	Segmentation of EEG data	-	~2.5 GB (21 days)
Re-referencing & Computations of EEG metrics		- Frontal Alpha Asymmetry - Average alpha power 4 participants per day (15 days)	- Power spectral density - Event-related changes 1 participants (7 blocks) per day (40 days)
TOTAL TIME SPENT		**~3 months**	**~21 months**

Chapter 4 and 5 provide the outcomes of the conducted EEG analysis to extract the cognitive load of the expert and novice participants. As discussed within RQ1, EEG results might reveal a different aspect into map users' cognitive strategies. For instance, the average EEG alpha power results of Experiment 1 indicated that experts spent significantly lesser cognitive load in the map learning task in which the trial duration is not restricted, although the eye tracking metrics and sketch map analysis claimed otherwise. The main takeaway from the co-registration of ET and EEG is that using EEG, one gets some nuance that performance data and/or eye movement analysis alone cannot capture, in other words, what FAA tells in Experiment 1 cannot be obtained with eye tracking. The negative alpha asymmetry scores corresponded to a greater relative right frontal activation which refers to withdrawal and avoidance whereas, positive ones indicate a greater relative left frontal activity which is associated with approach and motivation (e.g. Lanini-Maggi, 2017). In the first experiment, we observed that all of the participants exhibited negative FAA scores, which suggested a greater relative right activation, therefore, withdrawal-related motivation. Therefore, EEG is required to understand the motivation behind participants' behavior towards to the task at hand. On the contrary, in Experiment 2, although not significantly, experts demonstrated higher theta power (except for easy tasks) than the novices did which can be associated with a higher cognitive load whereas fixation-related eye tracking metrics indicated no significant difference between experts and novices. Hence, it is important to maintain a holistic approach by triangulating data coming from different sources while interpreting the cognitive load.

Research Objective 3:
Explore the influence of a subset of visual variables (i.e. location, size, shape, color) in spatial cognition and the use of this input to enhance the design and communication of cartographic products.

RQ 4: How does the participants' attentional behavior vary towards the map elements of interest?

The focus map belonging to all participants in Experiment 1 highlighted the main structuring elements of the stimulus: main road construction, water bodies, major rivers and large settlements. Both groups showed similar attentional behavior in terms of the map area covered and the map elements on which they focused as also in accordance with what was found by Ooms (2012). According to AoI analysis of the eye movements collected through Experiment 1, larger AoIs were found out to be gazed at earliest, the dwell times for such objects were much longer, and the majority of participants drew these map elements on their sketch maps. The linear features were easier to learn and remember, although the viewer did not pay much attention. The objects that were absent on the sketch map received the shortest fixation durations during the study phase as expected. It shows that the users first tried to store a general reference frame of map objects and later on more detailed objects were processed. This result is in correspondence with what was proposed by Kulhavy and Stock (1996) and Ooms (2012).

On the other hand, preliminary AoI analysis of Experiment 2 demonstrate that the participants took longer time to focus on AoIs associated with the roads whereas were quicker to fixate on the relevant AoIs when the task demanded the retrieval of more polygon features (e.g. Block 4 demands the retrieval of green areas and hydrography, which contains more polygon features). The average fixation duration for the tasks demanding the recall of the polygon features were the longest for all participants, i.e. Block 5 (green areas), Block 6 (hydography), whereas the participants exhibited the shortest average durations when the task required the retrieval of linear features, i.e. Block 7 (roads). Both experiments proved that the participants tend to scan through the linear objects and fixate/focus on polygon objects. Furthermore, the participants were usually able to neglect the map objects that were irrelevant with the task and experts were the fastest to fixate on an AoI relevant with the task. This finding is in line with what was found by Gegenfurtner, Lehtinen & Säljö (2011), explaining that "*experts devoted more fixations on task-relevant areas, and fewer fixations on task-redundant areas; owing to superiority in parafoveal processing and selective attention allocation*". This can be explained with information-reduction hypothesis of Haider and Frensch (1999) which explains how experts optimize the amount of information should be processed based on the task demands and strategically concentrating only on the relevant information in the visual span (i.e. selective attention).

RQ 5: How do we improve the maps based on the input collected through user experiments?

We interpreted the results of average fixation duration, time to first fixation and the number of the objects covered within AoIs based on the objects' geometry, size, and the distribution in the map drawing area for a single map stimulus. Since it was a general use map and the participants were encountered with a simple spatial memory task, we mostly found no significant difference between experts and novices, hence, we concentrated on the similarities on their attentional behaviors to make recommendations to improve the effectiveness of digital 2D static maps. Benefiting the results especially discussed within the Chapter 2 and 6, we kept the following questions in mind:

- What kind of visualization method(s) can be enabled based on the distribution, size and geometry of the map objects?
- How do we control advantaged and disadvantaged map objects equally/in the same way while visualizing?

By visualizing data in a structured way due to the attention having high importance to interpret the map content, the main structuring map elements should be recognizable at first glance by increasing their saliency or stressing the image or symbology in its general reference frame (Ooms, 2012). For instance, our results described that the participants can mostly define the main roads and their surroundings, however, the map elements that are located close to the map frame were easily neglected (Chapter 6). The simplest rule of good map design is that the primary information should be placed in or close to the center of the map drawing area; hence, important but neglectable information due to its size can be placed in the regions that are easily perceived by map users and the rest of the objects can be arranged accordingly. On the other hand, the participants do not pay attention to the map objects that are small and contained in larges areas. In this case, the perception and identification of these objects can be improved by providing contrast with other objects, therefore, by increasing the hue of the existing color used to depict that object class, e.g. using a darker blue for a water body inside a green area.

While continuous linear objects such as roads and rivers are easy to detect and recognize, the objects such as road intersections, small water bodies and green areas are more difficult to detect and recognize. To enhance the identification of the small objects, the outer frame of the polygon can be highlighted or emphasized using a darker tone of the current color, which corresponds to exaggeration and changing the contrast.

According to the preliminary AoI analysis for Experiment 2, it is clear that the eye scans through linear objects and fixates on the polygon objects. This finding shows the difference in perception between continuous objects (e.g. roads, rivers) and discontinuous objects (e.g. forests, lakes). In this case, exaggeration can be applied to small objects or the objects with similar characteristics can be aggregated into one feature to make them more recognizable. Accordingly, the multiple objects can be merged/amalgamated into one geometry to be able distinguish different object classes as they are clustered closer together on a map especially with changing zoom or scale levels. All of the suggested cartographic generalization operations help balancing the identification of the advantaged and disadvantaged map objects by map users.

Additionally, using grids might contribute to the interpretation and recalling the map-related information. Edler et al. (2014) confirmed that grids help reduce distortion and increase recall performance in topographic maps that are influenced by several factors adding complexity to a map. This complexity is not necessarily a bad thing for map users, on the contrary, a higher amount of visual objects guides spatial memory and the formation of detailed and accurate cognitive representations. For instance, a topographic base map showing a rural area provides less support than a topography consisting of more urban features. Moreover, recall performance is poorer on a map depicting only the main roads than the one containing additional point and polygon features. On the other hand, Thorndyke & Stasz (1980, p.171) explained the good learners' strategy of map learning as follows: "*they first segmented and focused systematically on subsets of information from the map. They demonstrated a variety of successful techniques for encoding both spatial relationships and verbal labels. Finally, they evaluated their learning progress consistently and accurately, using knowledge of their own uncertainties to determine their subsequent fixations and study behaviors.*" The "divide-and-conqure" strategy of good learners can be adapted into map design that stimulates the cognition of poor learners as well. In our case, the complexity of the maps can be balanced by using square grids, which will divide the attention into sub-parts and increase the object identification, therefore, improve the cognition of the map. Grids can especially be useful to enhance the retrieval of the map objects located in the corners or close to the map frame, however, the aesthetic aspects of map making should be considered while designing them. Keil et al. (2020) reported that holographic grids utilized with AR (Augmented Reality) improves distance estimation and location memory in 3D indoor environments when applied correctly.

7.2. Critical reflection on the methodology

The first experiment had an exploratory characteristic and we intended to explore the influence of gender next to expertise. Due to our lack of experience in psychological experimental design, the number of participants and their gender were not counterbalanced and we did not find a significant difference between males and females as opposed to many other existing research favoring males. However, we mentioned the effect of gender to argue on the traditional male-female paradigm. It was one of the lessons learned in the first experiment and this issue was fixed in the second experiment by counterbalancing the age, gender, number of the participants from expert and novice groups as much as possible to minimize the influence of these individual characteristics on spatial cognition. Hence, we no longer focused on gender. On the other side, the profile of the participants were not diverse enough to study the effect of the ethnicity and language, although they are not of least importance. Additionally, we introduced time constraints in the second experiment, which could have been done in the first experiment as well. Although it is the ecological valid way of studying the learning strategies of the participants, on the one hand, our decision to let the participants study the map as long as they like, had an important effect on the results, mainly because of the fact that time being a tool to manipulate cognitive load. The reason of experts experiencing low cognitive load in Experiment 1 and experts and novices showed no significant difference in Experiment 2 might be due to the time constraints in the second experiement. Restricting time adds a level of stress/pressure that might increase the cognitive load.

On the other hand, the role of expectations towards the experimenter might affected the task performance. For instance, it is possible that some participants felt that they had to do this as fast as possible which makes it more likely to forget things or some were more driven to provide accurate results instad of quick responses, so they devoted more time. Unstructuredly and informally, I discussed their strategy with the participants after the Experiment 1. One interesting comment was that they counted the number of objects belonging to the each object class considering their locations relevant to the line features on the map, in other words, a reference frame of the map. Their strategy was simply splitting the map drawing area (using roads or rivers) into sub-regions and quantifying the information within these zones. This feedback is also in line with our suggestion using grids to enhance map learning.

In the second experiment, due to the long experiment hours (approx. 2,5h) and voluntary participation, we did not want to take more time of the participants. However, in the post-test questionnaire, although not quite relevant with their map learning strategy, they were asked to put the experiment blocks in order from "easy to remember" to "hard to remember" according to their experience. However, their answers did not quite match with their performance (success rate and reaction time). Nonetheless, the experts were more consistent with how they performed and what they think about the difficulty of blocks. That is why subjective data cannot be trusted alone and needs to be verified with the quantitative data.

Accordingly, a future study that controls the trial time, i.e., giving everyone say 3 minutes to study the map, might show the effect of expertise somewhat more. It would be interesting to repeat the first experiment with more balance participant group, a particular trial time in the future, and a structured post-test questionnaire to collect information regarding the learning strategies of participants.

The first experiment had long EEG recordings but a single trial per participant and included no event marks to study the specific time points of the trial. Furthermore, I was inexperienced about the correct placement of EEG cap and the impedance check that had to be done for every single participant. These issues caused a much lower number of usable EEG data and the lessons learned within the Experiment 1 were incorporated in the design of the second experiment. While we were planning the second experiment which was a mixed design of between and within subjects, we collaborated with an experimental psychologist, who is specialized in EEG & eye tracking co-registration. He suggested recruiting 15 participants from each group, since approximately the half of them would be excluded due to the EEG signal noise, the muscle and blink artifacts or some other possible limitations such as not enough respondance, high attrition level, long experiment hours and recruitting volunteers. Accordingly, we were careful while selecting participants to minimize the influence of individual characteristics. In the second experiment, we recruitted 38 participants (17E, 21N) whose age and gender were counterbalanced. Since the participants we were able recruit, especially experts, were limited in number, we increased the number of trials. This strategy is also in line with the literature (e.g. Boudewyn, Luck, Farrens & Kappenman, 2017; Ito, Nikolaev & Van Leeuwen, 2005). Due to the signal/noise ratio in the EEG data, we had to exclude almost the half of the collected EEG data. There were 350 trials per participant (i.e. 100 hard, 100 moderate, 150 easy) and we paid attention to not to exclude more than 5-10% of the epoched data and the wrong answers not more than 10% of the all answers. In this context, we had at least 880 (i.e. 11*(100-20)) data points for each condition per group (expert and novice). The statistical power were calculated as 0.9 using using G*power (Wilcoxon-mann Whitney test: post-hoc, two tailed, effect size: 0.3) (Faul, Erdfelder, Lang, & Buchner, 2007). Nonetheless, while interpreting the results, it should be noted that due to the small group of participants, gender might be contaminated by expertise, although there were enough data points for statistical analysis.

7.3. Recommendations for further research

The experiments described in this dissertation create several possibilities for future research and although they are out of the scope of this dissertation, the variables used, the analysis conducted and the methodologies selected can be improved in order to obtain additional results. First, the data collected within the experiments is quite large and there are several other possible ways to extract information, especially, preliminary AoI-based eye tracking analysis can be extended using all the map stimuli involved in Experiment 2 to obtain more accurate and detailed information on the attentional behavior of this group of map users. Accordingly, a subset of visual variables (i.e. geometry, size, and the distribution in the map drawing area) and the eye tracking metrics considered within this analysis (i.e. average fixation duration, time to first fixation and the number of the objects covered within AoIs) can be increased to explore the influence of map design on users' cognition. On the other hand, additional analysis can be conducted with the collected EEG data in a way that the frequency bands can be decomposed into sub-parts to extract the cognitive load. As explained in Chapter 5, lower and upper alpha are differently associated to the cognitive activity (e.g. Fink 2006; Sauseng et al., 2005).

EEG and eye tracking usability research in cartography is not very common and one important methodological critique is that despite the dynamic reality of human cognition, the data collection methods used to understand cognitive procedures have been static and taken place in laboratory or an artificial environment that potentially restricts participants' behaviors. This traditional approach is advantegeous in terms of experimental control, however, lacks the real-world dimensionality and relevance of the findings. Nonetheless, emerging field of mobile cognition suggests significant added value to user experiments and the spatial memory abilities explained in this dissertation can be tested in real life conditions if technical and methodological requirements are met. The simultaneous recording and integration of mobile EEG and eye-tracking data can be utilized to obtain the actual timing of engagement with real-world stimuli (Ladouce, Donaldson, Dudchenko & Ietswaart, 2017). For example, a portable eye tracking and EEG headset can possibly be used to study the spatial memory abilities of map users in situ, e.g. in a car while using navigation maps or in an indoor way finding scenario using the landmarks and signs. Another example is using mobile EEG on the bike to study the attentional state of cyclists. Zink et al. (2016) collected data during an outdoor cycling condition in comparison to a fixed bike conditions and reported a decrease in P300 amplitude for the outdoor cycling, which suggests a reduced attention, hence can be attributed to an increased cognitive load due to being in a real-life scenario. This type of research can indeed benefit to understand the inner state of cyclists, how they perceive the danger and how their attention can be improved and so forth. However, there is no scientific publication reporting such integration in a real-world environment, yet in cartography where EEG is integrated with other user testing methods to explore cognitive processes of map users. The major reason for not having sufficient literature on this subject is the methodological and technical challenges of the co-registration of eye tracking and EEG systems in a mobile environment. The first issue is the optimization of the signal/noise ratio and more noise is simply inevitable in a mobile setting. Another critical concern is the definition of AoIs, which are essential to quantify the gaze dynamics across meaningful parts of the dynamic visual scene or environment, because AoIs would constantly be influenced by the displacement of the participants. The correct timing of events of interest requires much more complex algorithms compared to time-locking handled in laboratory settings. Moreover, mobile EEG systems use dry electrodes that decrease data quality and cause discomfort for the participants (Ladouce et al., 2017).

Our research was exploratory and limited to the resources in the neuro-lab, nonetheless, the spatial resolution of the EEG recording system can be enhanced to assure better data quality such as using 32, 64 or 128 electrodes EEG devices. Furthermore, the current analysis could be improved by separating the alpha band in lower and upper alpha to study the cognitive load or another interesting approach is to compute the frequency within the alpha band in which a participant shows the highest power (e.g. 10.5Hz) and center the alpha powerband around this number (e.g. plus/minus 2 Hz). The fact that the average power were calculated over an interval of 7 seconds might have been a missed opportunity to find effects we were looking for. To overcome this issue, one might consider calculating the alpha power every 2 seconds and relate these different intervals to specific map learning behavior.

On the other hand, the temporal accuracy of the EEG can improved in order to explore event-related potentials (ERP) or eye fixation related potentials EFRP as explained in Chapter 3. Focusing on the EEG activity that is time-locked to an event provides insights that are more precise and in this context, ERP-based study holds a great potential to further dig into the differences between novices and experts, and to study this with attentional paradigms making use of ERP components. For instance, how novice and expert map users differ in drawing roads (in color red) or water/lakes (in color blue) first? Which brain activity is originated from saccadic movements and which from the fixations -that are visual processing related-? We can sort EEG data based on saccadic and fixation related components, then match ERPs at desired fixations and explore if there are similar patterns for different participants. This would help to answer our main research question in more depth: ‘Do search strategies of map-related information differ for experts and novices?’ Consequently, it can be concluded that the EEG data is overwhelming and the possibilities are endless if experiment is designed well.

This research can also be followed up with an fMRI study specifically to find out which tasks use similar brain regions even though the tasks themselves seem to be very different or which tasks use different brain regions even though the tasks seem to be very similar (see Lobben et al, 2005). In this context, for instance, we can investigate:

- how (i) task difficulty, (ii) map complexity, (ii) number of objects, (iv) the use of visual variables to remember affect the spatial cognition, e.g. which regions engage during the retrieval of linear or areal features or features depicted with color red? so on.
- how spatial relationships are constructed during the map learning procedure on “digital 2D static topographic maps” or “screenshots of Google’s road maps” and compare it with what would happen while interacting e.g. “3D dynamic visualizations from Google earth”

Additionally, one can make use of hippocampal place cells that are grid cells in entorhinal cortex and they are efficient neural mechanism for encoding knowledge about the world, not only for spatial location but also for more abstract cognitive information. Human fMRI studies have observed grid-like signals that encode locations during mental imagery, features of abstract visual stimuli and eye position during 2D visual search (please read Kim & Maguire, 2019 for more detail). “*If grid cells are indeed involved in abstract cognitive mapping, the ‘space’ might not be limited to simple 2D physical space on which most grid cell research has to date been conducted, because cognitive tasks can involve more than two features or attributes. Grid cells should also be able to efficiently encode 3D and higher dimensional space (unless the high dimensional cognitive problem can be projected into low dimensional space, e.g. context-dependent encoding*” (Kim & Maguire, 2019, p1).

Within the scope of the thesis, we focused on digital 2D static maps. Although there can be quite some design challenges for new digital ways of representing geographical informations (e.g. 3D maps, GoogleStreetView, Augmented/Virtual Reality applications), there lies cartographic rules, mainly the symbolization and the manipulation of the visual variables which are used to construct the core information of maps, in the foundation of producing of either 2D or 3D cartographic representations. Moreover, the visual variables that were first listed by Bertin have been continuously amended by others with the advances in technology (e.g. perspective, transparency). In terms of spatial cognition and usability, 2D maps are still good and not outdated; hence, there is room to optimise their user experience. Previous studies evaluating the usability of 2D versus 3D representations are inconclusive (e.g. Lei et al., 2014; Popelka and Doležalová, 2015) and the advantages and disadvantages of both types of the representations have been reported depending on different modes of use and user contexts, therefore, task dependent. It may be possible to improve their usability by combining the advantages of each. For example, for 3D representations, cartographers can reduce the number of buildings to only important ones; for 2D maps, important landmarks should be included to help the users locate and orient themselves. The design and evaluation of representations that combine 2D and 3D features is a potentially interesting issue for further study (Liao et al., 2017).

References

Allen, R., & McGeorge, P. (2011). Enumeration: Experts take their time. Applied cognitive psychology, 25(4), 588-592.Dong, W., Zheng, L., Liu, B., & Meng, L. (2018). Using eye tracking to explore differences in map-based spatial ability between geographers and non-geographers. ISPRS International Journal of Geo-Information, 7(9), 337.

Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does it take to get a significant ERP effect? It depends. Psychophysiology, 55(6), e13049.

Edler, D., Bestgen, A. K., Kuchinke, L., & Dickmann, F. (2014). Grids in topographic maps reduce distortions in the recall of learned object locations. Plos one, 9(5).

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39, 175-191.

Fink, A., Grabner, R. H., Benedek, M., & Neubauer, A. C. (2006). Divergent thinking training is related to frontal electroencephalogram alpha synchronization. European Journal of Neuroscience, 23(8), 2241-2246.

Gedminas, L. (2011). Evaluating hurricane advisories using eye-tracking and biometric data. Master's Thesis, East Carolina University: North Carolina.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. Educational Psychology Review, 23(4), 523-552.

Haider, H., & Frensch, P. A. (1999). Information reduction during skill acquisition: The influence of task instruction. Journal of Experimental Psychology: Applied, 5(2), 129–151

Havelková, L., & Gołębiowska, I. M. (2020). What Went Wrong for Bad Solvers during Thematic Map Analysis? Lessons Learned from an Eye-Tracking Study. ISPRS International Journal of Geo-Information, 9(1), 9.

Ito, J., Nikolaev, A. R., & Van Leeuwen, C. (2005). Spatial and temporal structure of phase synchronization of spontaneous alpha EEG activity. Biological cybernetics, 92(1), 54-60.

Keil, J., Korte, A., Ratmer, A., Edler, D., & Dickmann, F. (2020). Augmented Reality (AR) and Spatial Cognition: Effects of Holographic Grids on Distance Estimation and Location Memory in a 3D Indoor Scenario. PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 1-8.

Kim, M., & Maguire, E. A. (2019). Can we study 3D grid codes non-invasively in the human brain? Methodological considerations and fMRI findings. NeuroImage, 186, 667-678.

Kulhavy, R. W., & Stock, W. A. (1996). How cognitive maps are learned and remembered. Annals of the Association of American Geographers, 86(1), 123-145.

Ladouce, S., Donaldson, D. I., Dudchenko, P. A., & Ietswaart, M. (2017). Understanding minds in real-world environments: toward a mobile cognition approach. Frontiers in human neuroscience, 10, 694.

Lanini-Maggi, S. (2017). Depicting movement data with animations for embodied and real-time decision-making: a user study with air traffic control displays and real-time movement data. Doctoral dissertation, University of Zurich, Zurich.

Lei, T. C., Wu, S. C., Chao, C. W., & Lee, S. H. (2016). Evaluating differences in spatial visual attention in wayfinding strategy when using 2D and 3D electronic maps. GeoJournal, 81(2), 153-167.

Li, X., Çöltekin, A., & Kraak, M.-J. (2010). Visual exploration of eye movement data using the space-time-cube. In: S. I., Fabrikant, et al.(Eds.) Lecture notes in computer science; Vol 6292(pp. 295-309). Berlin, Heidelberg: Springer

Liao, H., Dong, W., Peng, C., & Liu, H. (2017). Exploring differences of visual attention in pedestrian navigation when using 2D maps and 3D geo-browsers. Cartography and Geographic Information Science, 44(6), 474-490.

Lobben, A. K., Olson, J. M., & Huang, J. (2005, July). Using fMRI in cartographic research. In Proceedings of the 22nd International Cartographic Conference (p. 10).

Ooms, K. (2012). Maps, how do users see them?: An in depth investigation of the map users' cognitive processes. Doctoral Dissertation, Ghent University, Ghent. Retrieved from https://biblio.ugent.be/publication/3103344

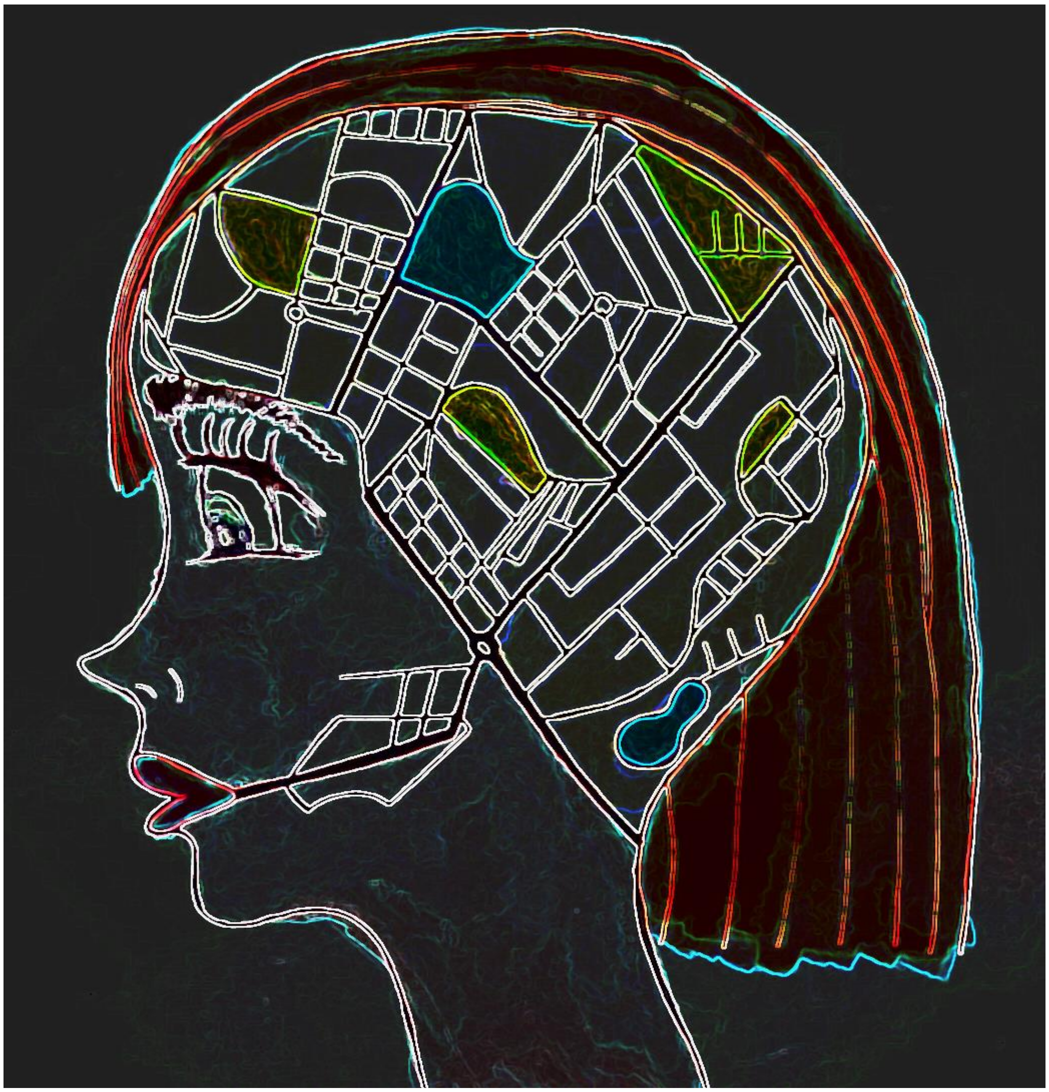
Popelka, S., & Doležalová, J. (2015). Non-photorealistic 3D visualization in city maps: an eye-tracking study. In Modern Trends in Cartography (pp. 357-367). Springer, Cham.

Sauseng, P., Klimesch, W., Schabus, M., & Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. International Journal of Psychophysiology, 57(2), 97–103.

Thorndyke, P. W., & Stasz, C. (1980). Individual Differences in Procedures for Knowledge Acquisition from Maps. Cognitive Psychology, 12(1), 137- 175.

Zink, R., Hunyadi, B., Van Huffel, S., & De Vos, M. (2016). Mobile EEG on the bike: disentangling attentional and physical contributions to auditory attention tasks. Journal of neural engineering, 13(4), 046017.

Chapter 8:
General Conclusion



"I don't know the future.
I didn't come here to tell you how this is going to end.
I came here to tell you how this is going to begin."

Neo, Matrix, 1999

Expertise plays a role equally important to maplikeness in map learning. To recall the locations and configurations of spatial objects from the memory usually requires experience with cartographic products in which topographic and topological information are represented by graphic symbols (Dickmann et al., 2016). Visual variables (i.e. position, size, shape, value, color hue, orientation, and texture), on the other hand, has a great impact on visual attention and perception depending on their property (i.e. selective, associative, ordered or quantitative) (Wolfe, 2000). In cartography, empirical studies focusing on map design and spatial cognition are increasing, however, only a number of them devoted to the exploration of cartographic elements (e.g. visual variables) which are the keys in cognitive map formation. Hence, we cannot yet formulate the cognitive map construction precisely and nevertheless, sketch maps, considering their complexity, can be utilized as one of the sources to assess the processes of cognitive map production.

While studying, a map-reader first perceptually divides the map into a number of spatial chunks. In this context, the structuring map elements, such as roads, hydrographic features or gridlines, initiates chunking process, thus, helps regionalizing the map and assists learning of map elements and their spatial relations. These structuring elements represent the spatial information of the map content in a hierarchically structured fashion and form fundamental units of cognitive maps, therefore, facilitate the perception and recognition of object locations (Edler et al., 2014).

The memory task in the first experiment required recalling the main structural elements of a screen map without any time constraints. This retrieval act involved WM-LTM transitions, such as retrieval of spatial information stored in WM through LTM or strategies for constructing hierarchy among map elements. Our findings (see Chapter 2) showed that the majority of participants drew hydrographic objects first. While experts drew roads first, novices focused more on hydrographic objects such as rivers and water bodies. The settlements and land-cover elements (i.e. forests) were drawn third and fourth, respectively, for both user groups. The fact that both experts and novices drew linear objects (hydrography and roads) first can be explained by the hierarchical structures of schemas in LTM in a way that the participants start drawing their sketch maps with the main linear structures and continue with other landmarks.

We regarded visual variables such as location, shape, size and color as though they were equally important for the drawing order which can be influenced by the use of visual variables Based on the scoring the size and shape of the objects depicted on sketch maps, the biggest difference between experts and novices was observed for settlements, favoring experts. A possible explanation could be that the depiction of settlements requires higher-level generalization knowledge. Since individual buildings come together to form a settlement or village, aggregation is needed to define a group of buildings as a settlement. Although not statistically significant, novices depicted the map elements slightly better using corresponding colors and the greatest difference between two groups was in hydrology (Figure 2.11). This result can be related to missing map elements on the sketch maps or to the fact that some experts did not prefer to use color. According to our findings, drawing a map element in the correct location was more difficult than describing its shape, size, and color.

To the best of our knowledge, in map design, important objects or the ones to emphasize are shown in red, and blue is a good color for backgrounds. This feature could be the reason why the experts drew the red linear objects (roads) first. On the other hand, having drawn the hydrographic elements first, novices might have found areal objects as important or interesting and thus as memorable as linear objects. We can infer that size is as important as color for the retrieval of an object. It is suggested that experts and novices use different strategies in spatial orientation, however, the common characteristic of the first drawn elements by all participants was that they both contained linear objects.

Based on the assessment of sketch maps considering the aggregated analysis of presence & location, shape, size, and color of drawn elements, we concluded that neither expertise, nor gender differences were influential on the retrieval of spatial information. Although our results present that experts and novices do not differ in terms of the amount of information they recall, the learning/recalling strategies of experts and novices may differ. The drawing order results could be evidence that they might use different approaches.

Eye tracking metrics provided valuable insight on how mental representations are formed. In this context, average fixation duration and the number of fixations per second revealed that there was no significant difference between the expert and novice groups. On the other hand, most participants paid less attention (late first fixation and less dwell time) to the relatively small linear (i.e. roads) and areal features (i.e. land cover) within the specified AoIs. The larger objects and the objects located in the upper middle of the screen caught a participant's attention earlier than the others did. However, when comparing the presence and accuracy scores of drawn elements, both groups mostly drew small roads on their sketch maps but not land-cover features. We could infer from this result that the linear features were easier to learn and remember, although the viewer did not pay much attention and as users tend to focus on polygon objects, they scan through linear objects (see Chapter 2, 6).

While eye tracking metrics and sketch maps analysis in the first experiment demonstrated that the differences between experts and novices were not significant, the EEG alpha power analysis contradicted those results. As the greater alpha power is associated with lower cognitive load, the results indicated that experts spent considerably less cognitive load on this memory task compared to novices. However, novices and experts showed no significant difference in FAA scores. Nevertheless, 70% of experts had negative scores on this metric, which reflects greater relative right activation, suggesting withdrawal-related motivation. Although the average FAA scores were negative for novices, 57% of them exhibited larger left-hemispheric activation, which is an indicator of approach-oriented motivation and positive affective states (see Chapter 2).

In the first experiment, we used simple and exploratory measurements for cognitive load extraction (e.g. average fixation duration, number of fixations per second, FAA, average alpha power). However, the map design and the level of complexity of maps might have an impact on cognitive load and even influence how difficult a particular task can be. Therefore, the second experiment was designed in a more complex way of addressing in-depth investigation of cognitive load of expert and novice participants within varying task difficulties regarding to the retrieval of the main structuring map elements in restricted trial periods

In this context, easy tasks involved Block 1 (all map elements) and Block 2 (roads & hydography); moderate tasks Block 3 (roads and green areas) and Block 4 (green areas and hydrography); easy tasks, Block 5 (green areas), Block 6 (hydrography) and Block 7 (roads) (see Chapter 4).

Our findings demonstrated that experts had longer reaction times (significantly longer for moderate and hard tasks), but higher success rates. This result is in line with the outcome of the first experiment confirming that experts allocated more time in studying the map stimulus. The reason for that might be that experts were a bit more ambitious and driven to accomplish the task compared to novices, and saved an extra time to review or verify their response before submitting it. Novices were observed to have longer fixation durations, mostly lower number of fixations per second and higher saccade velocity (except for hard tasks) which indicate a higher cognitive load for novices. Additionally, the saccade amplitudes of novices were longer. In longer saccades (i.e. larger amplitude), the search goes all across the image and is thus less organized. Experts exhibiting shorter saccades means a more targeted search from one focal point to the next, which are close to each other in the map, therefore, a less chaotic search pattern. The shorter fixations of experts also show that they needed less time to interpret what they saw. Although not significant, experts demonstrated higher theta power (except for easy tasks) which can be associated with a higher cognitive load. Qualitative analysis of the eye movements (i.e. through 10 randomly selected focus maps) indicated that both groups showed similar attentional behavior in terms of the map area covered and the map elements on which they focused. For instance, both experts and novices generally paid most attention to the green areas that are large and isolated (see Chapter 5). Based on the above findings, it is difficult to favor one user group in terms of their performance while retrieving the map-related information. The map-learning and recalling strategies of experts and novices and their approach to the task might not be similar, however, the overall performances of them did not differ much.

We alternatively grouped the participants in the second experiment as good learners (GL) and relatively poor learners (RPL) based on their success rates, and we observed that good learners exhibited significantly higher theta ERS considering their overall performance. Although the cognitive load of these groups did not differ based on the task difficulty within the frame of this research, classifying participants based on their spatial memory performances provided different insights on map user' cognitive processes. Similar to what was found by Havelková and Gołębiowska (2020), unsuccessful participants differed in the general problem-solving approach, in a way that they tended to choose fast, less cautious strategies and lacked motivation. Subsequently, we further analyzed the influence of the map design characteristics that play an important role in users' cognitive load and learning performance considering *expert vs. novices* and *GL vs. RPL* classifications. The qualitative analysis of the randomly selected 10 focus/heat maps considering the map elements receiving the highest number of fixations provided a general overview of the participants' attentional behavior towards the map elements of interest and the similarities related to their map learning strategies (see Chapter 5.3.2). However, for measurable results, drawing AoIs (area of interest) around key elements of maps (i.e. green areas, water bodies, major rivers and roads, road junctions) is a more precise way to analyze the attention distribution of the participants. Our preliminary findings showed that on the one side, there was no trend in terms of the differences occurred regarding to the blocks, therefore, the fixation durations within the (relevant) AoIs did not depend on the task difficulty. On the other side, the task demands (e.g. the tasks in Block 7 require the retrieval of the main roads) affect the attentional behavior of participants. Nevertheless, experts exhibited longer fixation durations (except for Block 3) compared to novices did, while GL always demonstrated lower fixation durations than the RPL did. This outcome shows that GL experienced lesser cognitive load than the rest.

The shorter the time to first fixation, the earlier the attention is drawn to that AoI and we did not observe a trend for time to first fixation neither between experts and novices nor between GL and RPL. While it took experts the shortest time to fixate on the relevant AoIs, RPLs spent the longest time to fixate on map elements of interest. This finding shows that experts were the fastest of all to gaze at the map objects of interest (see Figure 6.10c). According to AoI analysis of the selected map stimulus, we can infer that the eyes tend to focus on the polygon objects while they scan through linear objects. On the other hand, the differences in terms of the map elements covered within selected AoIs are much more visible between GL and RPL compared to those between experts and novices (Figure 6.12c, d).

Our results showed that EEG can be employed as a complementary technique to get a detailed insight about user actions and behaviors and reveal the information that we did not observe with eye tracking. Therefore, triangulating EEG and ET data seems useful to be able draw conclusions on user's behavior and also shows that the data require more investigation.

In our experiments, we used Google Maps, which is designed for everyone (i.e. regardless of the users' individual differences of spatial cognition) as stimuli in our experiments, and we found no significant difference between experts and novices in terms of the cognitive load that these maps caused. It is important to mention that if the quality of the cartographic design fulfills its purpose of the design, it has a positive effect on users' experience. Therefore, our findings confirm Google Maps were designed in a simple manner for general audience.

The more we know about the limitations and capabilities of visual perception and cognition of different map users, the higher the possibilities to design cartographic products in a more efficient, understandable and effective way. When integrated with other qualitative and quantitative user testing methods, EEG indeed suggests a valuable contribution to the understanding of the cognitive processes of individuals.

SUMMARY

Understanding how our brain copes with complex visual information is a challenge for both cognitive psychology and cartography. If we pursue to design better and usable maps, we require building a better knowledge on the cognitive processes of map users. This thesis aims to contribute to the understanding of the cognitive processes of a group of map users in learning, acquiring and remembering information presented via digital 2D static maps.

To be able to gain insight into the users' behaviors while they interact with maps, eye tracking (ET) and electroencephalogram (EEG) are enabled as synchronized data collection methods due to them being non-invasive and capturing direct responses of cognitive activities. Therefore, the preliminary goal of the research is to evaluate the use of ET and EEG for cartographic usability and spatial cognition research considering the technical and methodological aspects of this synchronization, also the limitations, possibilities and the contribution of EEG in the domain of cartography. The technical concerns refer to (i) the synchronization of ET and EEG recording systems, their accuracy and quality, and (ii) numerous processing steps (i.e. preprocessing, the alignment of the collected ET and EEG data, removal of non-cerebral activities from EEG data, segmentation and re-referencing). The methodological issues are situated in many aspects of the experimental design and its set-up, which includes identifying the research goals, participants, task and stimuli, psychological measures to use, evaluation methods and possible analyses of the collected data. These issues are pinpointed with respect to the existing literature, knowledge obtained from domain experts and hands-on experience in the neuro-lab.

The fundamental object of the thesis is to investigate on the traditional expert-novice paradigm as expertise being one of the individual characteristics that influences the users' performance on map-learning tasks. Since maps are widely used by both experts and novices, to study their differences in spatial cognition enables us to determine how to use this input to enhance the map design leveraging the map users' cognitive abilities. Therefore, our main research questions are: 'Do map learning strategies of experts and novices differ? How does the cognitive load vary between expert and novices?' In this context, we conducted two mixed-methods user experiments focusing on the cognitive strategies of a group of expert and novice map users and investigated their spatial memory capabilities through cognitive load measurements.

First experiment had a simple design and an exploratory characteristic, since we would initially assure that the eye tracking and EEG synchronization is of sufficient quality to explore users' cognitive behaviors towards map stimuli. Accordingly, it consisted of single trials and participants were instructed to study the main structuring elements of a map stimulus (i.e. roads, settlements, hydrography, and green areas) without any time constraints in order to draw a sketch map afterwards. On the one hand, the performance of the participants was assessed based on the order with which the objects were drawn on the digital sketch maps and the influence of a subset of visual variables (i.e. presence & location, size, shape, color). On the

other hand, trial durations and eye tracking statistics such as the average duration of fixations, and number of fixations per seconds were compared. Moreover, selected AoIs, which represent the main structuring elements of the map stimulus, were explored to gain a deeper insight on visual behavior of map users. Based on the evaluation of the drawing order, we observed that experts and males drew roads first whereas; novices and females focused more on hydrographic object. According to the assessment of drawn elements, no significant differences emerged between neither experts and novices, nor females and males for the retrieval of spatial information presented on 2D maps with a simple design and content. The differences in trial durations between novices and experts were not statistically significant while both studying and drawing. Similarly, no significant difference occurred between female and male participants for either studying or drawing. Eye tracking metrics also supported these findings. For average duration of fixation, there was found no significant difference between experts and novices, as well as between females and males. Similarly, no significant differences were found for the mean number of fixation. Furthermore, based on results of time to first fixation, dwell time, fixation count, the number of fixations per second, average fixation duration for selected AoIs, the larger AoIs were gazed at earliest and the dwell times for such objects were much longer compared to those for other AoIs. The linear features were easier to learn and remember, although the viewer did not pay much attention. Longer average fixation durations for a specific AoI indicated that the chances were higher to remember that object. The objects that were absent on the sketch map received the shortest fixation durations during the study phase. However, longer fixation durations may also indicate participants' difficulty to recognize the information in the map stimulus. Regarding to the EEG Frontal Alpha Asymmetry calculations, both user groups showed greater relative right frontal activation, which is in association with the less attentional, and focus performance. The difference between experts and novices was not significant, similar to the eye tracking results. On the contrary, alpha power averaged across all electrodes demonstrated that the novices exhibited significantly lower alpha power, indicating a higher cognitive load.

On the contrary, in Experiment 2, a complex and more structured approach was followed as a result of learning from the previous experiment and collaborating with the domain experts. This experiment contained a larger number of stimuli were used to study the effect of task difficulty (i.e. easy, moderate, hard) on the retrieval of map-related information. Next to the reaction time and success rate, we used fixation and saccade related eye tracking metrics (i.e., average fixation duration, the number of fixations per second, saccade amplitude and saccade velocity), and the event-related changes in EEG power spectral density (PSD) for alpha and theta frequency bands to identify the cognitive load. While fixation metrics and the qualitative analysis of the randomly selected focus/heat maps summarizing the participants' fixation behaviors indicated no statistically significant difference between experts and novices, saccade metrics proved the otherwise. EEG power spectrum analysis, on the other side, suggested an increase in theta power (i.e. event-related synchronization) and a decrease in alpha power (except moderate tasks) (i.e. event-related desynchronization) at all difficulty levels of the task for both experts and novices, which is an indicator of cognitive load. Although no significant difference emerged between two groups, we found a significant difference in their overall performances when the participants were classified as good and relatively bad learners. Triangulating EEG results with the recorded eye tracking data and the qualitative analysis of randomly selected focus maps indeed provided a detailed insight on the differences of the individuals' cognitive processes during this spatial memory task.

The qualitative analysis with the 10 randomly selected focus/heat maps provided a general overview of the participants' attentional behavior towards the map elements of interest and the similarities related to their map learning strategies. However, for measurable results, we selected one map stimulus and drew AoIs around key elements of maps (i.e. green areas, water bodies, major rivers and roads, road junctions) to analyze the attention distribution of the participants using average fixation duration, time to first fixation and the number of map objects covered within AoIs. Although the results are preliminary, we found out that the eye scans through linear objects and fixates/focuses on the polygon objects. The location of the map elements is more influential on the participants' gaze behavior compared to its size. The fixation durations within the (relevant) AoIs did not depend on the task difficulty. Additionally, our analysis showed that the GL experienced the least cognitive and this finding supports the evaluation of the participants by classifying them as "good learners and bad learners" during the usability tests of maps designed for general users with basic map learning tasks. In order to increase the understandability and usability of cartographic products, the results of this research can be used as guiding experiences in production processes where design methods that minimize the factors that negatively affect user perception (e.g. exaggeration, reduction of emphasis, utilizing the visualization elements to increase visual extraction such as grids).

SAMENVATTING

Begrijpen hoe onze hersenen omgaan met complexe visuele informatie is een uitdaging voor zowel de cognitieve psychologie als de cartografie. Als we ernaar streven om betere en bruikbare kaarten te ontwerpen, moeten we een betere kennis van de cognitieve processen van kaartgebruikers opbouwen. Deze thesis heeft als doel bij te dragen aan het begrijpen van de cognitieve processen van kaartgebruikers bij het leren, verwerven en onthouden van informatie die wordt gepresenteerd via digitale 2D statische kaarten.

Om inzicht te krijgen in het gedrag van de gebruikers tijdens de interactie met kaarten, worden eye tracking (ET) en electroencephalogram (EEG) ingeschakeld als gesynchroniseerde dataverzamelingsmethoden omdat ze niet-invasief zijn en directe responsen van cognitieve activiteiten vastleggen. Daarom is het voorlopige doel van het onderzoek om het gebruik van ET en EEG voor cartografisch bruikbaarheids- en ruimtelijk cognitieonderzoek te evalueren, rekening houdend met de technische en methodologische aspecten van deze synchronisatie, alsook met de beperkingen, mogelijkheden en de bijdrage van EEG in het domein van de cartografie. De technische aspecten hebben betrekking op (i) de synchronisatie van ET en EEG opnamesystemen, hun nauwkeurigheid en kwaliteit, en (ii) talrijke verwerkingsstappen (d.w.z. voorbewerking, de uitlijning van de verzamelde ET en EEG gegevens, het verwijderen van niet-cerebrale activiteiten uit EEG gegevens, segmentatie en re-referencing). De methodologische kwesties situeren zich in vele aspecten van het experimentele ontwerp en de opzet ervan, waaronder de identificatie van de onderzoeksdoelstellingen, de deelnemers, de taak en de stimuli, de te gebruiken psychologische maatregelen, de evaluatiemethodes en de mogelijke analyses van de verzamelde gegevens. Deze vraagstukken worden in kaart gebracht ten opzichte van de bestaande literatuur, de kennis van domeinexperts en praktijkervaring in het neuro-lab.

Het fundamentele doel van het proefschrift is het onderzoeken van het traditionele expert-noviceparadigma als expertise die een van de individuele kenmerken is die de prestaties van de gebruikers op het gebied van kaartleertaken beïnvloedt. Aangezien kaarten veel gebruikt worden door zowel experts als nieuwkomers, kunnen we, om hun verschillen in ruimtelijke cognitie te bestuderen, bepalen hoe we deze input kunnen gebruiken om het kaartontwerp te verbeteren door gebruik te maken van de cognitieve vaardigheden van de gebruikers van de kaart. Daarom zijn onze belangrijkste onderzoeksvragen: Verschillen kaartleerstrategieën van experts en beginnelingen? Hoe varieert de cognitieve belasting tussen expert en beginner? In deze context voerden we twee mixed-methods gebruikersexperimenten uit gericht op de cognitieve strategieën van deskundige en beginnende kaartgebruikers en onderzochten we hun ruimtelijke geheugencapaciteiten door middel van cognitieve belastingsmetingen.

Het eerste experiment had een eenvoudig ontwerp en een verkennend kenmerk, aangezien we in eerste instantie zouden verzekeren dat de eye tracking en EEG-synchronisatie van voldoende kwaliteit is om het cognitieve gedrag van de gebruikers te verkennen in de richting van kaartstimuli. Het bestond dan ook uit enkelvoudige proeven en de deelnemers kregen de opdracht om de belangrijkste structurerende elementen van een kaartstimulans (d.w.z. wegen, nederzettingen, hydrografie en groene gebieden) te bestuderen zonder enige tijdsdruk om daarna een schetskaart te tekenen. Enerzijds werd de prestatie van de deelnemers beoordeeld op basis van de volgorde waarin de objecten op de digitale schetskaarten zijn getekend en de invloed van visuele variabelen (bijv. aanwezigheid & locatie, grootte, vorm, kleur). Anderzijds werden de proefduur en de statistieken van de eyetracking, zoals de gemiddelde duur van de

fixaties, en het aantal fixaties per seconde, vergeleken. Bovendien werden geselecteerde AoI's (Area of Interests), die de belangrijkste structurerende elementen van de kaartstimulans vertegenwoordigen, onderzocht om een dieper inzicht te krijgen in het visuele gedrag van kaartgebruikers. Op basis van de evaluatie van de tekenvolgorde constateerden we dat experts en mannen eerst wegen trokken, terwijl novicen en vrouwen zich meer richtten op hydrografische objecten. Volgens de beoordeling van de getekende elementen zijn er geen significante verschillen tussen experts en novicen, noch tussen vrouwen en mannen voor het ophalen van ruimtelijke informatie die op 2D-kaarten met een eenvoudig ontwerp en inhoud wordt gepresenteerd. De verschillen in proefduur tussen beginnelingen en experts waren niet statistisch significant tijdens het bestuderen en tekenen. Evenzo was er geen significant verschil tussen vrouwelijke en mannelijke deelnemers voor zowel studeren als tekenen. Eye tracking metrieken ondersteunden ook deze bevindingen. Voor de gemiddelde duur van de fixatie werd er geen significant verschil gevonden tussen experts en novicen, evenals tussen vrouwen en mannen. Ook werden er geen significante verschillen gevonden voor het gemiddelde aantal fixaties. Bovendien werden op basis van de resultaten van de tijd tot de eerste fixatie, de verblijftijd, het aantal fixaties per seconde, de gemiddelde fixatieduur voor geselecteerde AoI's, de grotere AoI's op zijn vroegst bekeken en de verblijftijden voor dergelijke objecten waren veel langer vergeleken met die voor andere AoI's. De lineaire kenmerken waren gemakkelijker te leren en te onthouden, hoewel de kijker er niet veel aandacht aan besteedde. De langere gemiddelde fixatieduur van een specifiek AoI gaf aan dat de kans groter was dat het object zich zou herinneren. De objecten die afwezig waren op de schetskaart kregen de kortste fixatieduur tijdens de studiefase. Een langere fixatieduur kan echter ook wijzen op de moeilijkheid van de deelnemers om de informatie in de kaartprikkel te herkennen. Wat betreft de EEG Frontal Alpha Asymmetry berekeningen, toonden beide gebruikersgroepen een grotere relatieve rechter frontale activering, wat in verband staat met de minder aandachts- en focus prestatie. Het verschil tussen experts en beginners was niet significant, vergelijkbaar met de eyetracking resultaten. Integendeel, het gemiddelde alfavermogen over alle elektroden toonde aan dat de nieuwelingen significant minder alfavermogen vertoonden, wat duidt op een hogere cognitieve belasting.

Integendeel, in Experiment 2 werd een complexe en meer gestructureerde aanpak gevolgd als gevolg van het vorige experiment en de samenwerking met de domeinexperts. Dit experiment bevatte een groter aantal stimuli die werden gebruikt om het effect van taakmoeilijkheden (d.w.z. gemakkelijk, gematigd, moeilijk) op het terugvinden van kaartgerelateerde informatie te bestuderen. Naast de reactietijd en het succespercentage, gebruikten we fixatie en saccade gerelateerde eye tracking metrieken (dat wil zeggen, gemiddelde fixatieduur, het aantal fixaties per seconde, saccade amplitude en saccade snelheid), en de gebeurtenis-gerelateerde veranderingen in EEG-vermogen spectrale dichtheid (PSD) voor alfa en theta frequentiebanden om de cognitieve belasting te identificeren. Terwijl de fixatiemetingen en de kwalitatieve analyse van de willekeurig geselecteerde focus/warmtekaarten die het fixatiegedrag van de deelnemers samenvatten geen statistisch significant verschil tussen experts en beginners aangaven, bewees de saccade-metriek het tegendeel. EEG power spectrum analyse, aan de andere kant, suggereerde een toename in theta vermogen (d.w.z. gebeurtenis-gerelateerde synchronisatie) en een afname in alfa vermogen (met uitzondering van gematigde taken) (d.w.z. gebeurtenis-gerelateerde desynchronisatie) op alle moeilijkheidsgraden van de taak voor zowel deskundigen als nieuwelingen, wat een indicator is voor cognitieve belasting. Hoewel er geen significant verschil tussen twee groepen naar voren kwam, vonden we een significant verschil in hun algemene prestaties toen de deelnemers werden geclassificeerd als goede en relatief slechte leerlingen. De EEG-resultaten kruisend met de geregistreerde eyetrackinggegevens en de kwalitatieve analyse van willekeurig gekozen focuskaarten gaven inderdaad een gedetailleerd inzicht in de verschillen in de cognitieve processen van de individuen tijdens deze ruimtelijke geheugentaak.

De kwalitatieve analyse met de 10 willekeurig gekozen focus/warmtekaarten gaf een algemeen overzicht van het aandachtige gedrag van de deelnemers ten aanzien van de kaartelementen die van belang zijn en de overeenkomsten met hun kaartleerstrategieën. Voor meetbare resultaten selecteerden we echter één kaartstimulans en tekenden we AoI's rond de belangrijke kaartelementen (d.w.z. groene gebieden, waterlichamen, grote rivieren en wegen, wegenknooppunten) om de aandachtsverdeling van de deelnemers te analyseren met behulp van de gemiddelde fixatietijd, de tijd tot de eerste fixatie en het aantal kaartobjecten dat binnen AoI's aan bod komt. Hoewel de resultaten voorlopig zijn, kwamen we erachter dat het oog de lineaire objecten scant en fixaties/focussen op de veelhoekige objecten scant. De locatie van de kaartelementen heeft meer invloed op het kijkgedrag van de deelnemers in vergelijking met hun grootte. De fixatieduur binnen de (relevante) AoI's was niet afhankelijk van de moeilijkheidsgraad van de taak. Bovendien toonde onze analyse aan dat de GL de minst cognitieve ervaring had en deze bevinding ondersteunt de evaluatie van de deelnemers door hen te classificeren als "goede en slechte leerlingen" tijdens de bruikbaarheidstests van kaarten ontworpen voor algemene gebruikers met basiskaartleertaken. Om de begrijpelijkheid en bruikbaarheid van cartografische producten te vergroten, kunnen de resultaten van dit onderzoek worden gebruikt als leidraad voor ervaringen in productieprocessen waarbij ontwerpmethoden worden gebruikt die de factoren die de gebruikersperceptie negatief beïnvloeden, minimaliseren (bijv. overdrijving, accentvermindering, gebruik van visualisatie-elementen om de visuele extractie te vergroten, zoals rasters).

ÖZET

Beynimizin karmaşık görsel bilgilerle nasıl başa çıktığını anlamak hem bilişsel psikoloji, hem de bilişsel kartografya için bir zorluktur. Daha iyi ve kullanılabilir haritalar tasarlamayı hedefliyorsak, harita kullanıcılarının bilişsel süreçleri hakkında daha iyi bilgi sahibi olmamız gerekir. Bu tez, harita kullanıcılarının iki boyutlu (2B) sayısal statik haritalar aracılığıyla sunulan içeriği öğrenme, bilgi çıkarımı yapma ve hatırlamadaki bilişsel süreçlerinin anlaşılmasına katkıda bulunmayı amaçlamaktadır.

Kullanıcıların haritalarla etkileşim halindeyken sergilediği davranışlara ilişkin içgörü elde etmek amacıyla, invazif olmayan ve bilişsel aktiviteyi doğrudan ölçmeyi sağlayan göz izleme (eye tracking) ve elektroensefalogram (EEG), senkronize veri toplama yöntemleri olarak kullanılmıştır. Bu nedenle, araştırmanın öncül amacı, bu senkronizasyonun teknik ve metodolojik unsurlarına ilişkin kısıt ve imkanlarını dikkate alarak, kartografik kullanılabilirlik (usability) ve mekansal biliş araştırmaları özelinde sunduğu olanakları değerlendirmek ve ayrıca EEG'nin geomatik ve kartografya disiplinlerine potansiyel katkısını ortaya koymaktır. Teknik unsurlar, (i) göz izleme ve EEG kayıt sistemlerinin senkronizasyonu, doğruluğu ve veri kalitesi ile (ii) çok sayıda veri işleme adımlarından oluşur (örn. ön işleme, toplanan göz izleme ve EEG verilerinin çakıştırılması (offline synchronization), serebral olmayan EEG aktivitelerin ayıklanması, segmentasyon, yeniden referanslama). Metodolojik unsurlar ise araştırma hedefleri, katılımcılar, görev ve uyaranlar, analizlerde kullanılacak psikolojik ölçme metrikleri ve toplanan verilerle gerçekleştirilecek olası analizlerin saptanması gibi konuları içermek üzere, deneysel tasarım ve kurulum ilkeleri etrafında toplanmıştır. Tez kapsamında bu konular, mevcut literatür, alanında uzman kişilerden edinilen bilgiler ve deneylerin gerçekleştirildiği nöro-laboratuvardaki deneyimler göz önünde bulundurularak ele alınmıştır.

Doktora çalışmasının temel amacı, harita kullanıcılarının harita öğrenme (map learning) performanslarını etkileyen bireysel özelliklerden biri olan uzmanlığın etkisini geleneksel uzman-acemi paradigması çerçevesinde araştırmaktır. Haritalar, günümüzde hem uzmanlar, hem de uzman olmayanlar tarafından yaygın olarak kullanıldığından, harita kullanıcılarının mekansal bilişteki (spatial cognition) farklılıklarını incelemek, bu girdinin onların bilişsel yeteneklerinden yararlanarak harita tasarımını geliştirmek için nasıl kullanılacağını belirlenmesine öncü olur. Bu nedenle, temel araştırma soruları: “Uzman ve uzman olmayan kullanıcıların harita öğrenme stratejileri farklı mıdır? Bilişsel yük (cognitive load), bu kullanıcılar arasında nasıl farklılık gösterir?” Bu bağlamda, uzman ve uzman olmayan harita kullanıcılarının bilişsel stratejilerini tespit etmeye yönelik, iki tane karma yöntemli kullanıcı deneyi (mixed-methods user experiment) gerçekleştirilmiş ve bilişsel yük ölçümleri aracılığıyla harita kullanıcılarının mekansal bellek (spatial memory) yetenekleri araştırılmıştır.

Başlangıçta göz izleme ve EEG senkronizasyonunun, kullanıcıların harita uyaranlarına yönelik bilişsel davranışlarını saptamak için yeterli kalitede olduğu şartını sağlamak amacıyla, ilk deney basit bir tasarıma fakat keşifsel (exploratory) bir özelliğe sahiptir. Buna göre, ilk deney tek bir harita öğrenme görevinden (single trials) oluşmakta ve bu kapsamda katılımcılara daha sonradan dijital ortamda taslak (eskiz) haritasını (sketch map) çizmek üzere, herhangi bir zaman kısıtlaması olmadan basitleştirilmiş bir topografik harita uyaranının temel yapı elemanlarını (yollar, yerleşim bölgeleri, hidrografik objeler ve yeşil alanlar) ezberlemeleri talimatı verilmiştir. Bir yandan, katılımcıların performansı harita nesnelerinin taslak haritadaki çizim sırasına ve görsel değişkenlerin (örn. konum, boyut, şekil, renk) ifade edilişine göre değerlendirilirken, diğer yandan katılımcıların deney tamamlama süreleri ile ortalama sabitleme süresi

(average fixation duration) ve saniyedeki sabitleme sayısı (number of fixations per second) gibi göz izleme istatistikleri karşılaştırılmıştır. Ayrıca, harita kullanıcılarının görsel davranışları hakkında daha detaylı bir fikir edinmek amacıyla, haritanın temel yapı elemanlarını temsilen seçilen AoI'ler (ilgi alanları) incelenmiştir. Çizim sırasının değerlendirmesine göre, uzmanların ve erkeklerin önce yol nesnelerini; uzman olmayanların ve kadınların ise önce hidrografik nesneleri çizdikleri görülmüştür. Taslak haritalar üzerine çizilmiş nesnelerin değerlendirmesine göre, basit bir tasarım ve içeriğe sahip 2B haritalarda sunulan mekansal bilgilerin hatırlanması görevinde, ne uzman ve uzman olmayanlar, ne de kadın ve erkek katılımcılar arasında performans açısından anlamlı bir fark ortaya çıkmıştır. Aynı durum, harita içeriğini ezberleme ve taslak haritaların çizimini kapsayan deney sürelerindeki farklılıklar için de geçerlidir. Göz izleme metrikleri de bu bulguları desteklemekte; ne ortalama sabitleme süresi, ne de ortalama sabitleme sayısı için uzman ve uzman olmayanlar ile kadın ve erkek katılımcılar arasında anlamlı bir fark bulunmaktadır. Ayrıca, ilk sabitleme süresi, bakma süresi (dwell time), sabitleme sayısı, saniye başına sabitleme sayısı, seçilen AoI'ler için ortalama sabitleme süresi sonuçlarına dayanarak, katılımcılar öncelikle alansal olarak büyük AoI'lere dikkatlerini yöneltmişlerdir ve bu nesneler için bakma sürelerinin çok daha uzun olduğu saptanmıştır. Katılımcıların fazla dikkat harcamadan çizgisel detayları öğrenmesi ve hatırlaması daha kolay olmuştur. Belirli bir AoI'ye olan ortalama sabitleme süresinin uzaması, o nesneyi hatırlama şansının da artması şeklinde yorumlanabilir. Taslak haritalarda bulunmayan nesneler, harita içeriğini ezberleme aşamasında da en kısa sabitleme sürelerini almıştır. Ancak, daha uzun sabitleme süreleri, katılımcıların harita uyaranın içerdiği bilgileri tanıma güçlüğü ile de ilişkilendirilebilir. EEG frontal alfa asimetrisi (FAA) hesaplamaları, hem uzman, hem de uzman olmayan harita kullanıcılarının daha büyük göreceli sağ frontal aktivasyon sergilediklerine işaret etmektedir; sağ frontal aktivasyon düşük dikkat ve odak performansı ile ilişkilendirilmektedir. Göz izleme sonuçlarına benzer şekilde, FAA sonuçları da bu iki grup arasındaki farkın anlamlı olmadığını göstermesine karşın; tüm elektrotlar üzerinden hesaplanan ortalama alfa gücünün, uzman olmayanlarda önemli ölçüde daha düşük olması, harita öğrenme görevinini uzman olmayanlarda daha yüksek bir bilişsel yüke yol açtığı şeklinde yorumlanmıştır.

İkinci deneyde, önceki deneyden edinilen deneyimler ve deneysel psikoloji uzmanlarıyla işbirliği sonucunda, ilk deneye göre daha karmaşık ve yapılandırılmış bir yaklaşım izlenmiştir. Bu deney kapsamında görev zorluğunun (e.g. kolay, orta, zor) harita içeriğinin hatırlanması üzerindeki etkisini incelemek amacıyla çok sayıda uyaran kullanmıştır. Bilişsel yükü tanımlamak için reaksiyon süresi ve başarı oranlarının yanı sıra, sabitleme ve sekme (saccade) ile ilişkili göz izleme metrikleri (i.e. ortalama sabitleme süresi, saniyedeki sabitleme sayısı, sekme genliği ve sekme hızı) ve EEG'de görülen olaya bağlı değişiminin alfa ve teta bantları için hesaplanan güç spektral yoğunluğu (PSD) kullanılmıştır. Sabitleme metrikleri ve katılımcıların sabitleme göz hareketlerini özetleyen ve rastgele seçilen odak/ısı haritalarının (focus/heat maps) nitel analiz sonuçları, uzman ve olmayan katılımcılar arasında istatistiksel olarak anlamlı bir fark göstermemekle birlikte, sekme metrikleri aksini göstermektedir. Öte yandan, EEG güç spektrumu analizleri, her iki grup için de tüm görev zorluk seviyelerinde, tetada artışa (yani olayla ilgili senkronizasyon) ve alfada ise (orta düzey görevler hariç) azalmaya (yani olayla ilgili senkronizasyonun azaltılması) işaret etmektedir. Uzman ve olmayan grup arasında anlamlı bir fark görülmezken, katılımcılar iyi öğrenenler (good learners) ve görece kötü öğrenenler (relatively poor learners) olarak sınıflandırıldığında bu iki grubun genel performansları arasında anlamlı bir farka rastlanmıştır. EEG sonuçlarını, göz izleme metrikleri ve odak/ısı haritalarının analiz sonuçları ile birleştirilerek aslında bu mekansal bellek görevi sırasında bireylerin bilişsel süreçlerinin farklılıkları hakkında ayrıntılı bir içgörü elde edilmiştir.

Rastgele seçilen 10 adet odak/ısı haritası ile yapılan nitel analiz, katılımcıların ilgilenilen harita öğelerine karşı dikkate yönelik davranışlarına ve harita öğrenme stratejileriyle ilgili benzerliklere sadece genel bir bakış sağlamıştır. Ölçülebilir sonuçlar için, seçilen bir harita uyaranına ait temel yapı elemanları (i.e. yeşil alanlar, su kütleleri, büyük nehirler ve yollar ve yol kavşakları) etrafına AoI'ler çizilmiş ve bu AoI'lere ilişkin ortalama sabitleme süresi, ilk sabitleme süresi ve AoI'lerde kapsanan harita nesnesi sayısı metriklerinden yararlanılarak katılımcıların dikkat dağılımını analiz edilmiştir. Elde edilen birincil sonuçlara göre, gözün çizgisel objeleri taradığı ve alansal objelere sabitlendiğii/odaklandığı, harita objelerinin konumunun, katılımcıların algısı üzerindeki etkisinin, objelerin boyutuna oranla daha fazla olduğu saptanmıştır. Görevle ilgili AoI'lere olan sabitleme sürelerinin görev zorluğuna bağlı olmadığı gözlenmiştir. Ayrıca yapılan analizlerde iyi öğrenenler grubunun en az bilişsel yükü sergilediği görülmüştür; bu durum, temel harita öğrenme görevleriyle genel kullanıcılar için tasarlanmış haritaların kullanılabilirlik testleri sırasında katılımcıların "iyi öğrenenler ve kötü öğrenenler" şeklinde sınıflandırılarak değerlendirilmesini desteklemektedir. Bu araştırmanın sonuçları; kartografik ürünlerin anlaşılabilirliğini ve kullanılabilirliğini artırmak için kullanıcı algısını olumsuz etkileyen faktörlerin en aza indirgenmesini sağlayan tasarım yöntemlerinin (abartma, vurguyu azaltma, grid/kareler ağı gibi görsel çıkarımını artırmaya yönelik görselleştirme öğelerinden yararlanılması gibi) uygulandığı üretim süreçlerinde yol gösterici deneyimler olarak kullanılabilir.

ANNEX 1:

Orientation Script for Experiment 1 & 2

INTRODUCTION

First, thank you for participating in this study.

My name is Merve Keskin. I will be here throughout the study. If you need clarification regarding to any steps of the experiment, you can ask me.

The purpose of this study is to verify which regions in the brain are activated when end users perform different types of tasks on digital 2D static maps. In this test, you are considered as an end user.

During the experiment, your brain activity and eye movements will be registered.

Your brain activity will be measured using an EEG cap, which will be placed on your head. To do so:

i. Your skull will be measured to ensure a proper placement of the EEG cap.
ii. After you wear the cap, an electro-gel will be syringed to each electrodes on the cap to ensure proper signal acquisition.

Your eye movements will be captured using an eye tracker, which is mounted underneath your screen. The eye tracker will be calibrated in the beginning and in the middle of the test, to take into account the specific characteristics of each user.

I would like to emphasize that it is not your performance that is evaluated, but your eye movements and brain activity while performing tasks. You do not need to rush or try to complete the tasks as fast possible.

After the test, I will ask you to fill out a questionnaire, which helps me to understand your background. In addition, your data will be processed anonymously.

As a final remark, I ask each user to read this document and sign it. This document is an authorization record your data - in this case, the eye movements and brain signals - and use it for this study.

Please press <u>space bar</u> to continue.

ANNEX 2:

Instructions for Experiment 1 & 2

Experiment 1 - INSTRUCTIONS

Memory Task consists of two parts.

PART I
You will see a map on the screen like the one below. You are asked to remember the general structure of this map, so you that can draw in the second part. You are not asked to memorize and remember everything to the smallest detail, but the structural elements such as forests, rivers, roads, villages, railways, etc.

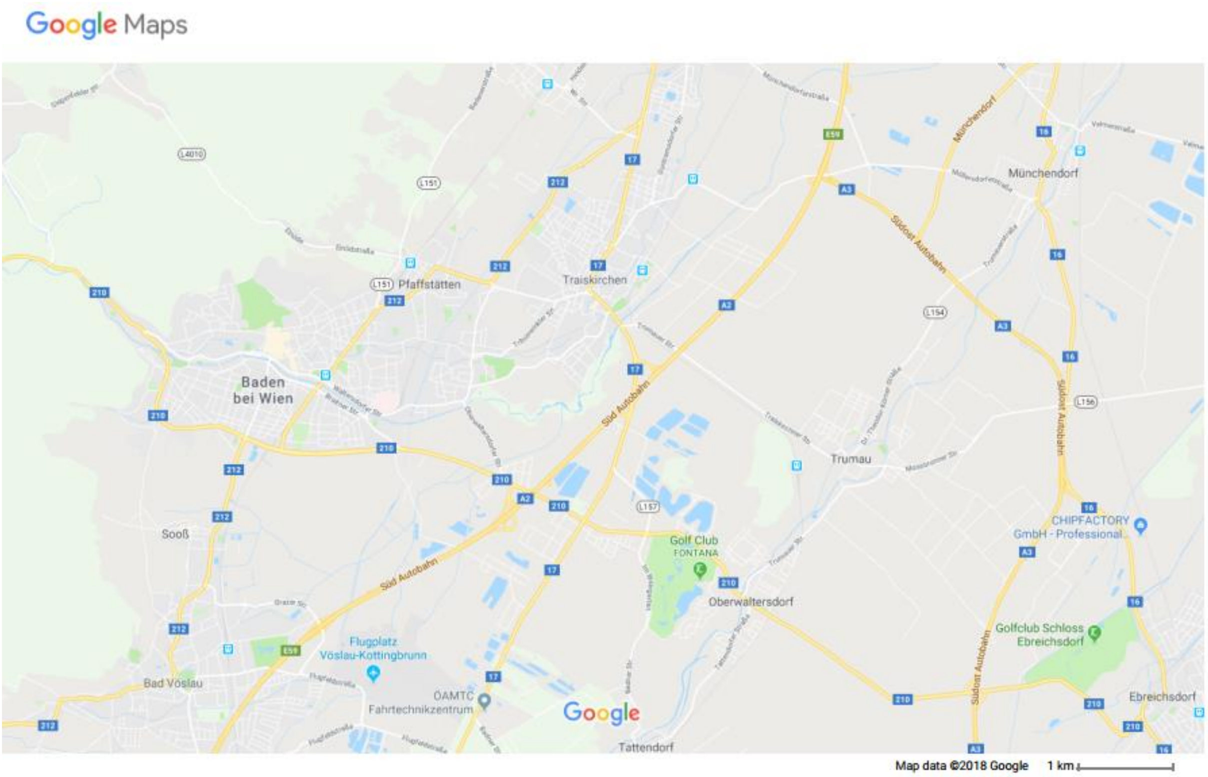
Once you think that you have studied the map enough, you may hit the space bar. The map will disappear from the screen and then you can start the second part of the test.

PART II
After the map disappears, a blank drawing screen (in MS Paint) will appear. There you will be able to draw what you remember related to the structural elements on the map presented to you previously.

Once you think you are finished with MS Paint, please <u>do not close the MS Paint window and press F11 to continue.</u>

Experiment 2 - INTRODUCTION

The experiment consists of 7 question blocks; each including 50 trials. In the beginning of each block, you will perform a training task where you will first see a map similar to the one below and have 10 seconds to study it.



Second, a question related to some main structuring map elements *(i.e. roads, green areas and hydrography)* will appear on the screen for 3 seconds.

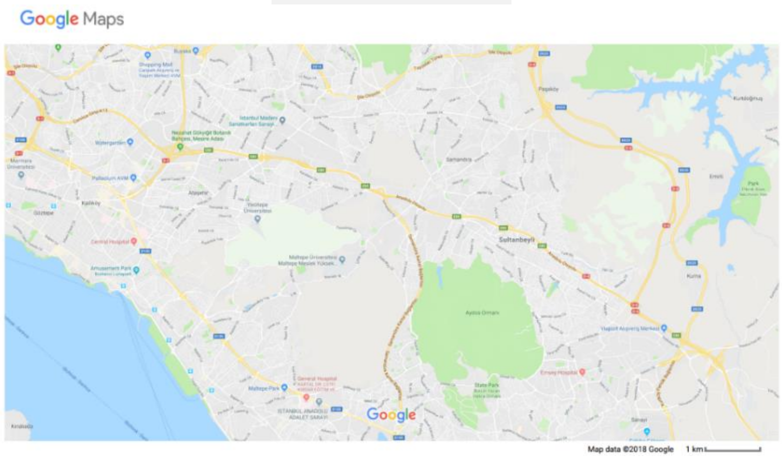
Third, you will be directed to a response screen with 4 graphical options (*named as* a, b, c, d) featuring the relevant main structuring map elements. One of the options will correspond to the map you just studied. **You will have to find the correct option and remember the corresponding letter (*i.e.* a, b, c, d) it. Once you find the correct option, you have to press space bar immediately.**

Pressing space bar will direct you to a second response screen where you will see only letters. Here, **you should click on the letter which you kept in your memory to complete the task**.

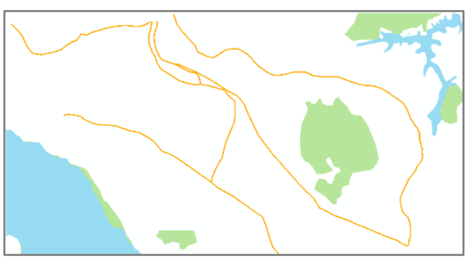
After completing the training task, you will continue to the main task where you will answer the same question for the following 50 map stimuli in the block.

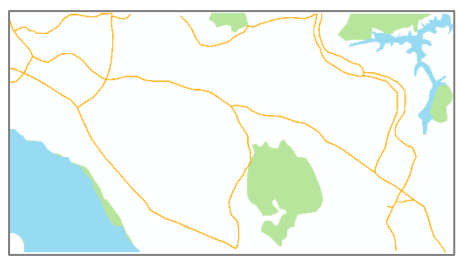
Please press <u>space bar</u> to continue with the training task.

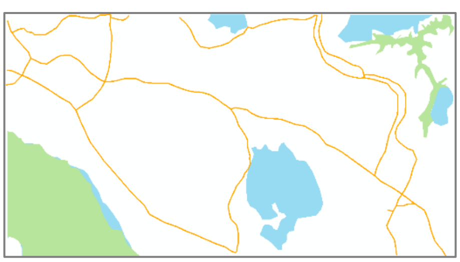
ANNEX 3: Task Structure of Experiment 2

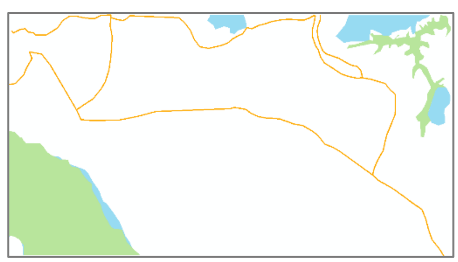


Which of the following sketches is similar
to the map you studied?

a.


b.


c.


d.


Which of the following sketches is similar to the map you studied?

- ☐ a
- ☐ b
- ☐ c
- ☐ d

ANNEX 4:
Python Script for organizing EEG and ET data

```
import os

'''
Created on 1-jun.-2017

@author: Kristien
'''

#--------------VARIABLES----------------------
dir_org="data_org"
dir_filter="data_filter"
dir_res="data_res"
dir_rawET="_rawET"
dir_eegEvents="_eegEvents"
list_events = ['train_vi_easy', 'train_vi_mod_1', 'train_vi_mod_2', 'train_vi_mod_3', 'train_vi_hard_1', 'train_vi_hard_2', 'train_vi_hard_3', 'cross', '1_ist6 Page 1','2_ist2 Page 1','3_istanbul3 Page 1','4_la3 Page 1', '5_la4 Page 1', '6_la5 Page 1', '7_la2 Page 1', '8_am2 Page 1', '9_cr1 Page 1', '10_li1 Page 1', '11_lo1 Page 1', '12_nc1 Page 1', '13_pa1 Page 1', '14_am1_v2 Page 1', '15_lv1 Page 1', '16_sd1 Page 1', '17_fr1 Page 1', '18_mi1 Page 1', '19_mic1 Page 1', '20_ad Page 1', '21_an Page 1', '22_bu Page 1', '23_bue Page 1', '24_chi2_2 Page 1', '25_egy Page 1', '26_hn Page 1','27_in Page 1','28_in2 Page 1','29_it Page 1','30_ka Page 1','31_ko Page 1','32_mb Page 1','33_my Page 1','34_pal Page 1','35_pi Page 1','36_pl Page 1','37_pl2 Page 1','38_sa Page 1','39_ser Page 1','40_tas1 Page 1','41_tw Page 1','ve Page 1','43_bi Page 1','44_da Page 1','45_mu Page 1','46_per Page 1','47_ph Page 1','48_pin Page 1','49_vin Page 1','50_xi Page 1']


#--------------FUNCTIONS----------------------
### perform conversions
### convert rawET data
def processRawET(f_et):
    print("FT: convert rawET")
    print (f_et)
    notHeader= False
    message=''
    syncEventNr = 100
    lastStimulus=''
    stimulusNr=0

    f_et_data= open(dir_filter+dir_rawET+'/'+f_et, 'r')

    f_w_et_data = open(dir_res+dir_rawET+'/'+f_et, 'w')
```

```
    for f_et_l in f_et_data:
        f_et_l_split=f_et_l.split("\t")
        if notHeader:
            #process data
            #if it is a message
            if(f_et_l_split[1]=="MSG"):
                for e in list_events:
                    if e in f_et_l_split[3]:
                        syncEventNr +=1
                        message = ("# Message: SYNC %s" %(syncEventNr))
                        print(message)
                        print(f_et_l_split[3])
                        f_w_et_data.write("%s\t%s\t%s\t%s\n"
%(f_et_l_split[0],f_et_l_split[1],f_et_l_split[2],message))

            #if it is data
            elif(f_et_l_split[1]=="SMP"):
                if(syncEventNr>100): #only record data after first stimulus has started
                    #print first columns - no change
                    f_w_et_data.write("%s\t%s\t%s\t" %(f_et_l_split[0],f_et_l_split[1],f_et_l_split[2]))

                    #convert each new stimulus to number
                    if(lastStimulus != f_et_l_split[3]):
                        stimulusNr+=1
                        lastStimulus=f_et_l_split[3]
                    f_w_et_data.write("%s\t" %(stimulusNr))


                    #write next columns, no change
                    for i in range(4, 18):
                        f_w_et_data.write("%s\t" %(f_et_l_split[i]))

                    #convert AOI data - L and R - to number - 0,0
                    f_w_et_data.write("0\t0\t")


                    #write next columns, no change
                    for i in range(20, len(f_et_l_split)-3):
                        f_w_et_data.write("%s\t" %(f_et_l_split[i]))

                    etEvent=f_et_l_split[len(f_et_l_split)-1].strip('\n').strip('\r')
                    #alter last column -  ET events to number
                    if(etEvent== "Blink"):
                        f_w_et_data.write("0\t0")
                    elif(etEvent== "Saccade"):
                        f_w_et_data.write("1\t1")
                    elif(etEvent== "Fixation"):
```

```
                f_w_et_data.write("2\t2")
            else:
                f_w_et_data.write("3\t3")


            f_w_et_data.write("\n" )

        #error handling in case no SMP or MSG messages are found
        else:
            print("ERROR: no SMP or MSG type: "+ f_et_l_split[1])

    #comments before actual data
    elif (f_et_l[0]=='#'):
        print("COMMENTS BEFORE DATA "+ f_et_l)
    #first line - header
    else:
        #process header
        f_w_et_data.write(f_et_l)
        notHeader=True

  f_et_data.close()
  f_w_et_data.flush()
  f_w_et_data.close()

### convert eegEvents
def processEegEvents(f_eeg):
  #print("FT: convert eegEvents")
  #print (f_eeg)
  notHeader= 0
  syncEventNr_eeg = 100
  timeSec=0

  #in case of txt data
  f_eeg_data= open(dir_org+dir_eegEvents+'/'+f_eeg, 'r')

  f_w_eeg_data = open(dir_res+dir_eegEvents+'/'+f_eeg, 'w')

  #write new header
  f_w_eeg_data.write("latency\ttype\tposition\n")

  for f_eeg_l in f_eeg_data:
    print(f_eeg_l)
    f_eeg_l=f_eeg_l.replace("  ","\t")
    f_eeg_l_split=f_eeg_l.split("\t")

    #process data
    if notHeader>1:
        #print(f_eeg_l_split[3])
```

```
            for e in list_events:
                if e in f_eeg_l_split[3]:
                  syncEventNr_eeg +=1
                  if ("ns" in f_eeg_l_split[0]):
                    timeSec=f_eeg_l_split[0].strip(" ns")
                  elif ("sec" in f_eeg_l_split[0]):
                    timeSec=f_eeg_l_split[0].strip(" sec")
                  elif ("min" in f_eeg_l_split[0]):
                    timeMin=f_eeg_l_split[0].strip(" min")
                    timeSec=float(timeMin)*60

                  line="%s\t%s\t%s\n" %(timeSec,syncEventNr_eeg,syncEventNr_eeg)
                  #/print(line)
                  f_w_eeg_data.write(line)

        #skip header (2 lines)
        else:
          notHeader+=1

    f_eeg_data.close()

#--------------MAIN CODE----------------------

print("START SCRIPT")
#read all files in folder containing original raw et data
for f_et in os.listdir(dir_filter+dir_rawET):
  processRawET(f_et)


#read all files in folder containing original eegEvent files
for f_eeg in os.listdir(dir_org+dir_eegEvents):
  processEegEvents(f_eeg)

print("END SCRIPT")
print("-------------------------------")
```

ANNEX 5:

Python Script for filtering fixations and saccades

```
import os

'''
Created on 17 Nov 2017

@author: Kristien Ooms

conditions:
The eye data that has to be filtered out is:

- saccade size <20 deg (and also remove the corresponding following fixation)
- fixation duration: 50-1000 ms.  (keeping  the fixation duration only between 50-1000ms and remove other values, and also the corresponding preceding saccade.

distance: 70 cm

'''
from numpy import integer
from cmath import sqrt, atan
from math import *


#--------------VARIABLES----------------------
dir_org="data_org"
dir_filter="data_filter"
dir_rawET="_rawET"

screenSizePixY=1050
screenSizeCmY=30
viewingDistCm=70
minFixDurMs=50
maxFixDurMS=1000
minSacSizeDeg=20
minSacSizePix=viewingDistCm*tan(radians(minSacSizeDeg))*(screenSizePixY/screenSizeCmY)

#--------------FUNCTIONS----------------------
### perform conversions
### convert rawET data
def filterRawET(f_et):
    #print("FT: filter rawET")
    #print (f_et)
```

```
etEvent_prev=''
et_R_Raw_X=''
et_R_Raw_Y=''
firstColumn=''
et_R_Raw_X_first=''
et_R_Raw_Y_first=''
firstColumn_first=''
et_data_currentEvent=[]
et_data_list5events=[]
et_list5_flags=[]
flagCurrentEvent=True
flagNextEvent=True


f_et_data= open(dir_org+dir_rawET+'/'+f_et, 'r')

f_w_et_data = open(dir_filter+dir_rawET+'/'+f_et, 'w')

for f_et_l in f_et_data:
  f_et_l_split=f_et_l.strip('\n').strip('\r').split("\t")
  # print(f_et_l_split)

  if f_et_l[0]=='#' :
    print("COMMENTS BEFORE DATA: "+ f_et_l)

  elif f_et_l_split[0]=='Time':
    # print("HEADER: "+ f_et_l)
    f_w_et_data.write(f_et_l)

  else:
    etEvent=f_et_l_split[len(f_et_l_split)-1]


    if(etEvent==etEvent_prev):
      # print("same event")
      et_data_currentEvent.append(f_et_l_split)


    else:
      #print("new event")

      if(etEvent_prev=="Fixation"):
        #print("event: %s - prev event: %s" %(etEvent,etEvent_prev))
        #print("starttime: %s - endtime: %s" %(firstColumn_first,et_t_prev))
        #print("CHECK: fixation duration: 50-1000 ms.  (keeping  the fixation duration only between
50-1000ms and remove other values, and also the corresponding preceding saccade." )
```

```
            fixTimeStart=et_data_currentEvent[0][0] #time first element current event
            fixTimeEnd=et_data_currentEvent[-1][0] # time last element current event
            fixDuration= (float(fixTimeEnd) - float(fixTimeStart))/1000
            #print("CHECK: %s - %s = %s" %( fixTimeEnd,fixTimeStart,fixDuration))


            if(fixDuration < 50 or fixDuration>1000):
              flagCurrentEvent=False
              #check event type before fixation
              if(len(et_data_list5events)>1): #there is data available in the list
                prev2Event=et_data_list5events[-1][0][-1] #last event, first row, last coloumn
                #print(prev2Event)
                if(prev2Event=='MSG'):
                  print("MSG - do nothing") #if the event before the fixation is a message: do nothing: start
of new stimulus
                elif(prev2Event=='Saccade'):
                  et_list5_flags[-1]=False #if the event before this fixation is a saccade: not to be printed
                elif(len(et_data_list5events)>2): #Blink or -
                  prev3Event=et_data_list5events[-2][0][-1] #last event, first row, last coloumn
                  if(prev3Event=='MSG'):
                    print("MSG - do nothing") #if the event before the fixation is a message: do nothing:
start of new stimulus
                  elif(prev3Event=='Saccade'):
                    et_list5_flags[-2]=False #if the event before this fixation is a saccade: not to be printed
                  elif(len(et_data_list5events)>3): #Blink or -
                    prev4Event=et_data_list5events[-3][0][-1] #last event, first row, last coloumn
                    if(prev4Event=='MSG'):
                      print("MSG - do nothing") #if the event before the fixation is a message: do nothing:
start of new stimulus
                    elif(prev4Event=='Saccade'):
                      et_list5_flags[-3]=False #if the event before this fixation is a saccade: not to be printed
                    else:
                      print("ERROR - NO PREVIOUS SACCADE prev5Event???")



          elif (etEvent_prev=="Saccade"):
            #print("event: %s - prev event: %s" %(etEvent,etEvent_prev))
            #print("start xy: (%s,%s) - end xy: (%s,%s)" %( et_R_Raw_X_first,
et_R_Raw_Y_first,et_x_prev,et_y_prev))
            #print("CHECK: saccade size >20 deg (or %s pix) (and also remove the corresponding following
fixation)" %(minSacSizePix) )

            sacStartX=float(et_data_currentEvent[0][4]) #X position of start saccade, first line in list (pixels)
            sacStartY=float(et_data_currentEvent[0][5]) #Y position of start saccade, first line in list (pixels)
            sacEndX=float(et_data_currentEvent[-1][4]) #X position of end saccade, last line in list (pixels)
            sacEndY=float(et_data_currentEvent[-1][5]) #Y position of end saccade, last line in list (pixels)
```

```
        sacDistPix=sqrt((sacStartX-sacEndX)*(sacStartX-sacEndX)+(sacStartY-sacEndY)*(sacStartY-
sacEndY))
        #print("CHECK: (%s,%s) - (%s,%s)) = %s (pix)" %(
sacStartX,sacStartY,sacEndX,sacEndY,sacDistPix))
        if(sacDistPix>minSacSizePix):
          flagCurrentEvent=False
          flagNextEvent=False
      else:
        print("event: %s - prev event: %s" %(etEvent,etEvent_prev))


      etEvent_prev=etEvent

      #add data current event to list with events - and whether is should be printed (flag)
      et_data_list5events.append(et_data_currentEvent)
      et_list5_flags.append(flagCurrentEvent)

      #empty list current event
      et_data_currentEvent=[]
      if(etEvent=='Fixation'):
        flagCurrentEvent=flagNextEvent
        flagNextEvent=True
      elif(etEvent=='MSG'):
        flagCurrentEvent=True
        flagNextEvent=True
      else:
        flagCurrentEvent=True

      #write first line of data of new current event
      et_data_currentEvent.append(f_et_l_split)


      #if there are more than 5 events in the list, write and remove last if flag is true
      if(len(et_data_list5events)>5):
        printData=et_data_list5events.pop(0)
        flagPrintData=et_list5_flags.pop(0)
        #print('PRINT DATA--')
        #print(flagPrintData)
        if(flagPrintData):
          for l in printData:
            #print(l)
            for element in l:
              f_w_et_data.write("%s\t" %(str(element)))
            f_w_et_data.write("\n" )
```

```
    for i in range(len(et_data_list5events)):
        printData=et_data_list5events.pop(0)
        flagPrintData=et_list5_flags.pop(0)
        #print('PRINT DATA')
        if(flagPrintData):
            for l in printData:
                #print(l)
                for element in l:
                    f_w_et_data.write("%s\t" %(str(element)))
                    f_w_et_data.write("\n" )

    f_w_et_data.flush()
    f_w_et_data.close()
#--------------MAIN CODE----------------------

#print("START SCRIPT")
#read all files in folder containing original raw et data
for f_et in os.listdir(dir_org+dir_rawET):
    filterRawET(f_et)




print("END SCRIPT")
print("-------------------------------")
```

ANNEX 6:

User characteristics of the recruited participants in Experiment 2

%	Q1: Please choose the highest level of education you have completed	N	Q2: How often do you use Google maps?	N	Q3: On a scale of 1-5, with 5 being "strongly agree" and & being "strongly disagree" please answer: Do you think Google maps is easy to use?	N	Q4: What do you think about the experiment?	N
Experts (N=17)	PhD	1	everyday	10	5	13	Positive	11
	MSc	16	once/twice a week	6	4	4	Neutral	2
			once a month	1	3<=	0	Negative	4
Novices (N=21)	MSc	11	everyday	8	5	8	Positive	5
	BSc	8	once/twice a week	11	4	11	Neutral	10
	High School	2	once a month	2	3	1	Negative	6
					2	1		

CURRICULUM VITAE

Merve Keskin is specialized in cognitive cartography, cartographic user experiment design, HCI/UX/UI principles, neuroscientific user testing methods (i.e. eye tracking, EEG) and experienced in GIS (i.e. visualization, map design, data management, geostatistical analysis, classification). With strong written & oral communication and training expertise, she is a lifelong learner, self-motivated and resourceful as well as able to motivate others with interpersonal skills. As a part of her PhD studies, she had an opportunity to work in a multicultural and interdisciplinary environment (e.g. experimental psychology, marketing, information and communication technologies).

She obtained her BSc in Geomatics Engineering at Istanbul Technical University (ITU) in 2011. In her BSc thesis, which was funded under EU 7th Framework Programme, she studied the GIS-based spatial/geostatistical interpolation methods (i.e. IDW, NN, Kriging) for meteorological data (i.e. temperature, precipitation, wind speed) and published a book chapter and journal articles related to that topic in the domains of (renewable) energy and environmental sciences.

Between 2011 and 2017, she worked as a research assistant at the geomatics engineering department of ITU with the responsibility of assisting cartography, GIS, computer aided graphic & map design courses, and fieldwork applications. She finished her masters in Geomatics Engineering at the same university in 2013 with a thesis topic of Investigating the Potential of Satellite Images in Topographic Map Production. In this context, she created requirement matrices for satellite image mapping by exploring topographic maps of small, medium, large scales and linking them with the satellite images of different resolutions. In 2013, she started her - PhD at the same department, and between 2014 and 2016; she collaborated as a remote-sensing scholar with urban planners and architects in an interdisciplinary project that aimed to propose a methodology to identify landscape identity indices of rural settlements.

For her PhD, she has focused on the usability of the cartographic products and since 2016, she has been working on the design and the user issues of maps, cognitive abilities and limitations of map users through neurophysiological techniques such as eye tracking and EEG at the geography department of Ghent University (UGent). Her joint-PhD between UGent and ITU is an interdisciplinary research incorporating cognitive cartography and experimental psychology, focusing on the spatial memory abilities of different map users (e.g. varying in gender and expertise) towards map learning tasks at different difficulty levels. In this context, she has gained an experience in psychological experiment design, data collection, processing and qualitative & quantitative analysis as well as programming.

Since March 2019, she has assited several GIS projects in the water management unit of Antea Group with the responsibility of research, management, visualization and presentation of the relevant geospatial data mainly with the framework of the United Nations Development Program in South Africa and South-Central America. Accordingly, she has been involved in several projects such as flood hazard vulnerability assessment for climate change, ecosystem valuation and mapping, land-cover classification and took a role in different teams for varied task.