Published as: Ooms, K., De Maeyer, P., Fack, V., 2014, Study of the attentive behavior of novice and expert map users using eye tracking. *Cartography and Geographic Information Science*, 41 (1), p. 37- 54 (doi:10.1080/15230406.2013.860255).

Studying the attentive behavior of novice and expert map users using eye tracking

K. Ooms[1*], P. De Maeyer[1], V. Fack[2]

[1]Ghent University, Department of Geography, Krijgslaan 281, S8, B-9000, Ghent, Belgium

Email: {Kristien.Ooms; Philippe.DeMaeyer}@UGent.be

[2]Ghent University, Department of Applied Mathematics and Computer Science, Krijgslaan 281, S9, B-9000, Ghent, Belgium

Email: Veerle.Fack@UGent.be

*Corresponding Author. Tel: +32 (0)9 264 4636 - Fax: +32 (0)9 264 4985

Abstract

The aim of this paper is to gain a better understanding of the way map users read and interpret the visual stimuli presented to them and how this can be influenced. In particular the difference between expert and novice map users is considered. In a user study, with a between user design, the participants had to study four different screen maps, which had also been manipulated for introducing deviations. The eye movements of expert and novice participants, 24 in total, were studied during these trials. Based on a grid of Areas of Interest, these eye movement recordings were analyzed both visually and statistically. These two dimensional visual analyses are essential for studying the spatial dimension of maps to locate problems in its design. The proposed analyses include the visualization of eye movement metrics (fixation count and duration) in a 2D and 3D grid, including a statistical comparison of these grid cells. The results show that the users' eye movements clearly reflect the main structuring elements on the map. The interpretation process of both user groups is influenced by deviating color use on the map as their attention is drawn to it. Furthermore, both user groups encounter difficulties when trying to interpret and store familiar map objects, which are not exactly the same as the ones they stored previously (caused by a mirror operation on the objects). These influences are found to be more pronounced with novices than with experts. Insights in how different types of map users read and interpret the map content are essential in this fast evolving era of digital cartographic products.

Keywords: user study, eye movement, cognitive cartography

1. Introduction

Due to the permanent advancements in technology, cartography has undergone a tremendous evolution during the last few decades. Already at the beginning of the new century, is was estimated that the number of maps distributed through the Internet on a daily basis exceeded the number of paper maps printed each day (Peterson, 2003). Moreover, the Internet has made maps and cartography a lot more accessible to the general public. However, the main goal of these 'modern' cartographic products remains the same: communication.

Communication can be seen as a process in which different steps are involved and as a consequence, cartography includes more than mapmaking. Several models of cartographic communication have been proposed since 1960. In its simplest form, the cartographic communication process can be seen as a transmission of source information (the world around us) to a recipient (the map reader). Maps visually represent the (spatial) information surrounding us and communicate this to the users by applying some kind of code, the cartographic syntax. They are the channels that allow (and should facilitate) this transmission. Consequently, if a map's design is not optimal, it could introduce noise in the communication process (Montello, 2002). A more elaborate model presented by Kolácný (1969) was most influential on cartographic research and was considered to be an essential aid in the studies regarding the way maps can be interpreted easily (MacEachren, 1995).

New technologies have a profound impact on the display of cartographic products to the user, and thus on how the information is perceived and interpreted. Screen maps create new possibilities, such as animations and user interactions, but are also inherently linked to some critical limitations: resolution, size, color use, etc. (Peterson, 2003). Recently, some concerns regarding this evolution in cartography (and GIScience) have risen. How effective are these new map displays? What effect do they have on the users' cognitive processes? What are the limits of the map reader's visual and cognitive processing abilities? Several authors expressed these concerns and concluded that more research is necessary regarding the cognitive issues in cartography and geographic information visualization in general (e.g. Fabrikant & Lobben, 2009; Harrower, 2007; Montello, 2002, 2009; Slocum, et al., 2001). In the next sections the theoretical background regarding the cognitive structures and processes, necessary to interpret visual information (such as maps), is described.

1.1. How can we process and interpret visual stimuli?

Montello (2002) stated that cognitive cartography includes the study of knowledge structures and processes of map reading, such as perception, learning, and memory, to understand mapping. Taking into account the cartographic communication model, the first step in map use is the interpretation of the visual encoded information on these maps. This interpretation process consists of a number of subsequent steps or levels, which are linked to the structure of human memory. Atkinson and Shiffrin (1968) divided the memory in three structural components: the *sensory memory*, the *short term memory*, and the *long term memory*. This model had a major impact on the early studies in cognitive psychology, and is often called the *modal model*.

The sensory memory records input from each of our senses, including vision, but this is quickly forgotten (less than two seconds). Some of the information in the sensory memory will be transferred to the short term memory, which also has a limited capacity. Nowadays, the short term memory is called the *working memory* (WM), and this term will also be used in the remainder of the paper. In order to transfer the information from the WM to the long term memory (LTM), it has to be rehearsed and thus learned (Cowan, 2001; Matlin, 2002; Miller, 1956). The capacity of the LTM is considered virtually infinite.

2 paragraphs removed

In order to be able to explain the processes that take place in the WM, Baddely (1999) proposed a WM structure that consists of three separate components: the *phonological loop* (which stores sounds), the *visuo-spatial sketch pad* (which stores visual and spatial information), and the *central executive* (which processes the information). This structure is depicted in Figure 1. Consequently, the WM is much more than a database that can store a certain number of chunks containing information. The information contained in the WM can constitute information obtained through the sensory memory, but 'old' information retrieved from the LTM can also be held in the WM. The central executive enables the manipulation of these different information sources (Matlin, 2002). This memory structure is essential to understanding how humans process visual stimuli, and thus subsequently interpret them.

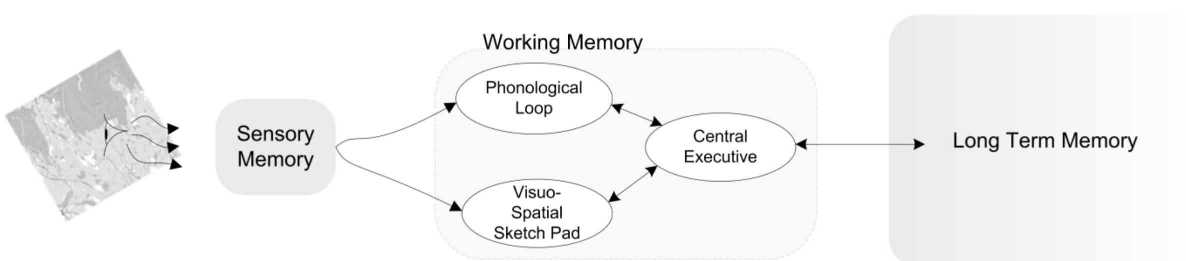

Figure 1: Memory structure: an integrated model according to

Atkinson and Shiffrin (1968) and Baddely (1999)

Title removed

In order to interpret maps, and the visual information depicted on them, the map reader will use the previous knowledge to process the stimuli registered by his (visual) senses. Two important cognitive processes form the basis of the interpretation process: *attention* and *object recognition* (Matlin, 2002). Attention can be defined as a concentration of mental activity. To interpret a certain object, users have to focus their attention on it, which normally happens automatically. The next phase in the interpretation process is object recognition.

Paragraph removed

To understand how users perceptually organize visual scenes, another influential approach has to be explained at this point: the *Gestalt approach*. The basic idea of this approach is that the whole is greater than the sum of the parts. Furthermore, humans will (unconsciously) try to organize what they see. Consequently, in order to interpret a visual scene, it is not sufficient to independently recognize the separate parts that are present in this scene (MacEachren, 1995; Matlin, 2002).

Several authors have also described different levels at which object recognition can take place. For example, Gerber (1981) identified three levels of successful object recognition. At the first level, the *perception-recipe* or *pictorial level*, the user can recognize the shape of an object that he has seen previously. At a higher level, the user also knows the name of the object; hence this called the *label* or *pictorial-verbal level*. Finally, if the user possesses other knowledge regarding the object, object recognition takes place at an even higher level: *other knowledge about* or *verbal level.* These levels are subsequently processed during the interpretation process of a certain object. When considering the interpretation or recognition of symbols on maps, Olson (1976) also identified three levels of processing: (1) compare symbol pairs, (2) recognize groups of symbols, (3) use the symbols to retrieve information from the map.

The interpretation process is based on a combination of bottom-up or top-down processing of information. Top-down processing of is closely linked with information that was already stored in the users' memory: experiences, familiarity, etc. Bottom-up processing is only based on the visual input, without using knowledge (MacEachren, 1995). Hegarty, et al. (2010), for example, examined the influence of salience (bottom-up) and domain knowledge (top down) on map comprehension (in cased of weather map displays). They discovered that eye movements are mainly guided by top-down factors and that a good design facilitated the processing of task relevant visual features.

This interpretation process ‘consumes’ part of the limited capacity of the WM. Another part of the WM’s capacity will be used to transfer the processed information to the LTM: the learning process. Bunch and Lloyd (2006) and Harrower (2007) give an excellent description of the ‘consumption’ of the limited capacity of the WM to process geographic and cartographic information. Their explanations are based on the *cognitive load theory*. This theory states that the WM is also limited in its processing capabilities; the amount of cognitive load it can take. Different types of cognitive load can be identified, all contributing to the total amount of cognitive load. Consequently, maps containing overly complex information, represented in a chaotic way, will be very difficult to interpret by the user (high intrinsic and extraneous cognitive load). These types of maps cannot be interpreted or learned from in an efficient way, because there is no room left to address the germane cognitive load that facilitates the learning process (Bunch & Lloyd, 2006; Harrower, 2007).

Previous paragraph: shorter

Ooms, et al. (2012a) studied the reaction time measurements and eye movements of expert and novice map users who performed a visual search on screen maps with a very basic design. Significant differences were identified between both user groups: experts seem to be able to interpret the map contents more efficiently than novice users. The authors relied on the cognitive load theory to explain these findings. Alvarez and Cavenagh (2004) investigated the *visual information load,* which is related to the amount of detail of the perceived objects. A higher level of detail results in a slower processing rate for each object. The maps presented during the user study conducted by Ooms, et al. (2012a) had a very simple map design, containing no complex objects: only points with labels and three polygons visualized in a pastel color. Consequently, these results cannot be generalized to a wide range of maps. The aim of this paper is to extend the study of Ooms, et al. (2012a) with the incorporation of more complex (topographic) maps to gain even deeper insights in how different types of users process this visual information. As mentioned before, insights in these cognitive processes are considered essential in the light of the recent technological developments in cartography in general (e.g. Fabrikant & Lobben, 2009; Harrower, 2007; Montello, 2002, 2009; Slocum, et al., 2001).

1.2. How to study 'map interpretation'?

Eye tracking is a 'direct' method to study users' cognitive processes. The participants in the user study do not have to reflect on their thoughts, using introspection or retrospection. Insights in their attentive behavior are thus obtained without any user interference. Participants reflecting on their own thoughts results in subjective and unreliable results, often because they do not exactly know how their thoughts had been formed: it is an automatic process (Nielsen, 1993; Rubin & Chisnell, 2008).

Using eye tracking, the position where a user is looking (Point of Regard, POR) is registered at a certain sampling rate. This provides insights in the user's attentive behavior: the location on which the focus of attention is directed at that moment in time. As mentioned before, attention is an essential step in object recognition and thus to interpret the whole map contents. Besides the position of the POR, other usable metrics (such as fixation duration) can be derived from the eye movement data which provide insight in the user's cognitive processes during the interpretation of the visual contents. Based on previous research regarding the eye tracking method, it can safely be assumed that a close link exists between these cognitive processes and the eye movement metrics (Duchowski, 2007; Jacob & Karn, 2003; Poole & Ball, 2006).

New paragraph below (older eye tracking studies)

The use of eye tracking is not new. In the 1950s it was already used to study the movements of pilots' eyes (Fits, et al., 1950). In the 1970s (e.g. Dobson, 1977; Jenks, 1973) and the 1980s (e.g. Castner & Eastman, 1984, 1985; Steinke, 1987) eye tracking was also applied in cartographic research (e.g. dot maps). These authors confirmed the method's applicability, but the found that it could not be used to derive new knowledge. The use of the method in cartography almost disappeared after 1985. However, recently a renewed interest in the use of eye tracking in cartographic research is noticed (e.g. Brodersen, et al., 2001; Çöltekin, et al., 2009; Fabrikant, et al., 2008).

New paragraph below (new eye tracking studies)

This 'rediscovery' of the eye tracking method can on the one hand be explained by the technical evolutions of the eye tracking systems themselves. They have become smaller, less intrusive for the participants, more accurate and less expensive. Not only where a user is looking (POR) is nowadays registered, but the length and duration of the fixations and saccades can also be derived from these more accurate measurements. A fixation is a time interval (of at least 80ms) during which the POR is relatively stable and the user is interpreting the information. Studying, for example, the duration of the fixation can give insights in how difficult it is to interpret the information: e.g. longer fixations can indicate that the user finds it difficult to process the information (e.g. Duchowsky, 2007; Holmqvist, et al., 2011). These different eye movement metrics, their meaning and their link to the users' cognitive processes are discussed in detail in a number of books and journal articles: Duchowsky (2007); Holmqvist, et al. (2011); Goldberg, et al. (2002); Jacob and Karn (2003); Poole and Ball (2006); Rayner (1998); among others.

New paragraph below (eye tracking metrics)This renewed interest is, on the other hand, also closely linked to the recent need to gain better understanding of the cognitive processes (and limits) of map users while working with highly dynamic, interactive, animated screen maps. These insights are essential to be able to key the visualization of future maps to the cognitive structures of the map users and to create more effective maps as a result (Cartwright, 2012; Fabrikant and Lobben, 2009; Montello, 2009). If we can understand how map users read, process, interpret, store, etc. (their cognitive structures) the information on the maps (and what influences this), we can design the maps in such a way that it is easier for the them to process the information. This visualization (symbolization) will be an aid the user to process the information.

Eye movement data can be analyzed in a number of different ways, in which two main categories can be identified: quantitative and qualitative methods. Quantitative methods often use statistics (e.g. ANOVA) to identify significant differences between two categories: differences in map design (within user design) (e.g. Ooms, et al., 2012b) or differences in user characteristics (between user design) (e.g. Ooms, et al., 2012a). These types of analyses have a higher level of objectivity because these are carried out on the actual numbers, applying a set of standard tests to compare them, resulting in a level of significance (P-value). Furthermore, this P-value is linked with a sample size and its power. All this elements make it possible to (rather objectively) interpret the data and compare between similar analyses. This is not available when considering qualitative analyses, which are more 'exploratory' in nature.

However, the quantitative analyses often do not allow to study the values in the context their spatial relationships, which should not be ignored when studying maps and their design. Statistical tests give lists of data which can be compared to each other. Mostly, these data, nor the results, give any indication of 'where' the measures are taken, or 'where' the differences are. This issue is discussed in more detail in Ooms et al. (2012). A visual analytic approach for studying eye movement data can handle this spatial dimension, as the distribution of the eye movements is considered. The visualization of the participants' scanpaths, for example, allows detecting patterns (Ooms, et al., 2012). However, caution is necessary when interpreting the visual data to avoid subjective conclusions. These interpretations are done 'at sight', but this cannot be justified whether the perceived difference is based on coincidence or not.

The combination of these different techniques shed light on other aspects of the eye movement data and thus on the related cognitive processes. In order to obtain the most elaborate overview of the cognitive processes taking place during the interpretation process of maps, it is good practice to combine different techniques.

2. Study design

2.1. Participants

Two groups of participants were selected to take part in the study, containing 12 persons each, with an equal share of males and females. The first group consisted of experts in map use and cartography. All participants in this group had at least a Master degree in Geography or Geomatics and received cartographic training during their studies. Furthermore, at the time of the study they were employed at the Department of Geography at Ghent University. The second group consisted of participants who did not receive any previous cartographic training and did not work with maps on a professional level. The average age of the participants was 23.8 years, with 25.9 years for the expert group and 21.4 years for the novice group. All participants took part in the study on a voluntary basis.

2.2. Tasks

The instructions were read out loud to each participant in order to avoid differences in task interpretation due to a different use of wording. Furthermore, at the start of the test, the participant could read through the instructions again on the screen. At this point, the participant could ask any questions if the instructions were not clear.

The instructions for the task were simple and clear. The participants were told that a map would be shown on a screen and that they had to remember the general structure of this map. They did not have to remember all details on the map (such as individual houses), but certainly the general structure such as roads, rivers, forests, etc. The user could read these instructions from a printed paper and they were read out loud by the moderator of the test. These instructions were in Dutch, the native language of all participants. A translated versions is presented below:

"First you may take place at the eye tracker. A map will be shown to you on the screen. Your aim should be to remember its general structure, so you can draw it during the second part of the study. You don't have to remember everything to the smallest detail, but still the general structure: location of forests, rivers, roads, villages, railways, etc.

Once you think of yourself that you have studied the map long enough, you can press one of the buttons on the joystick. The map will disappear from the screen and you can start the second part of the test."

In order to avoid any biases due to (time) pressure, the participants could execute this task at their own pace. If they found that they had studied this map long enough, they could remove the map display themselves by pushing a button. The participants were informed that the map was displayed for a maximum time of 10 minutes. This limit was introduced to keep the study manageable. Before the actual study was carried out, a pilot study was done to determine the different parameters in the study, such as this time limit. In this pilot study, it was found that the participants needed on average 5 minutes to complete this task. Therefore, it was decide to double this average time interval to avoid any time pressure on the participants as it should be more than enough to perform the task. It should be noted that this limit was reached only a few times. During this task, the participants' eye movements were registered.

In order to force the participants to interpret the map contents, they had to execute a second task during which they would need to use (retrieve) the information again. In order to be able to use this information later on, it has to be stored in (working and long term) memory. This would mean that the information is stored in the form of chunks of information (or schemata) which are linked to other information stored in the long term memory. In order to be able to do this, the objects have to be read, recognized and interpreted (give a meaning to it) (see Section 1.1). Before the start of the study, the participants were also instructed that they would have to draw the map they had just seen, using paper and pencil. Likewise, no time limit was set on the drawing task; the participants just had to indicate when they was ready.

In this paper we did not test or analyze the users' memory performance, only where (and how) they looked at the stimuli. The findings of the memory test are described in Ooms et al., 2013. In this paper the authors present how the information is retrieved, how this is structured and how much is retrieved. A combination of thinking aloud, sketch maps and a questionnaire was used to studied the 'information retrieval process'. These findings confirm that the participants would have to had interpreted the information.

This process of ‘remembering the map – drawing the map’ was repeated four times. After the completion of the fourth trial, the participant was asked to fill out a questionnaire. This post study questionnaire is used to obtain personal characteristics (expertise, age, gender, etc.), to verify potential familiarity of the participant with the presented regions, and to receive feedback.

2.3. Stimuli

During the user study, four different maps were presented on a screen. These stimuli were selected out of the Belgian topographic map series on 1 : 10 000. These maps were chosen so that the map image would not be crowded with information; some obvious structures were visible (roads, rivers, forests, etc.) and the region was not well known. The selection of the map was, among others, based on the percentage of the map image covered with large uniform areas such as forests and meadows. All selected stimuli had a coverage of more than 75% for these two types of land use (48.4% and 36.7%, 73.6% and 14.0%, 50.7% and 25.6%; meadow and forest coverage in map 1/map 4, map 2 and map 3 respectively).

Familiarity with a certain area influences the user’s interpretation process caused by previous knowledge and should therefore be avoided. Previous knowledge (regarding the depicted region) is essential in top down processing. The participants of the user study all live in the northern part (Flanders) and the selected topographic maps cover regions located in the southern part. Therefore, it is rather unlikely that the participants know the depicted regions by heart. Furthermore, in a post-study questionnaire it was verified if the participants knew the depicted regions. None of the participants indicated that they were familiar with the regions. The four stimuli were depicted in the same order to all participants. This fixed order was necessary to ensure that certain stimuli (map 1 and map 4) would not be depicted right after each other. The maps used during the study are depicted in Figure 2. This figure shows five different maps, although only four of them were presented to each user. This is due to a variation that had been introduced in the third stimulus, which was only shown to half of the participants. Six experts and six novices saw map 3 in its normal orientation; the other half of the test persons saw the map mirrored over its vertical central axis. This allows detecting whether the users' scanpaths – which result from the interpretation process for this map – would also be mirrored.

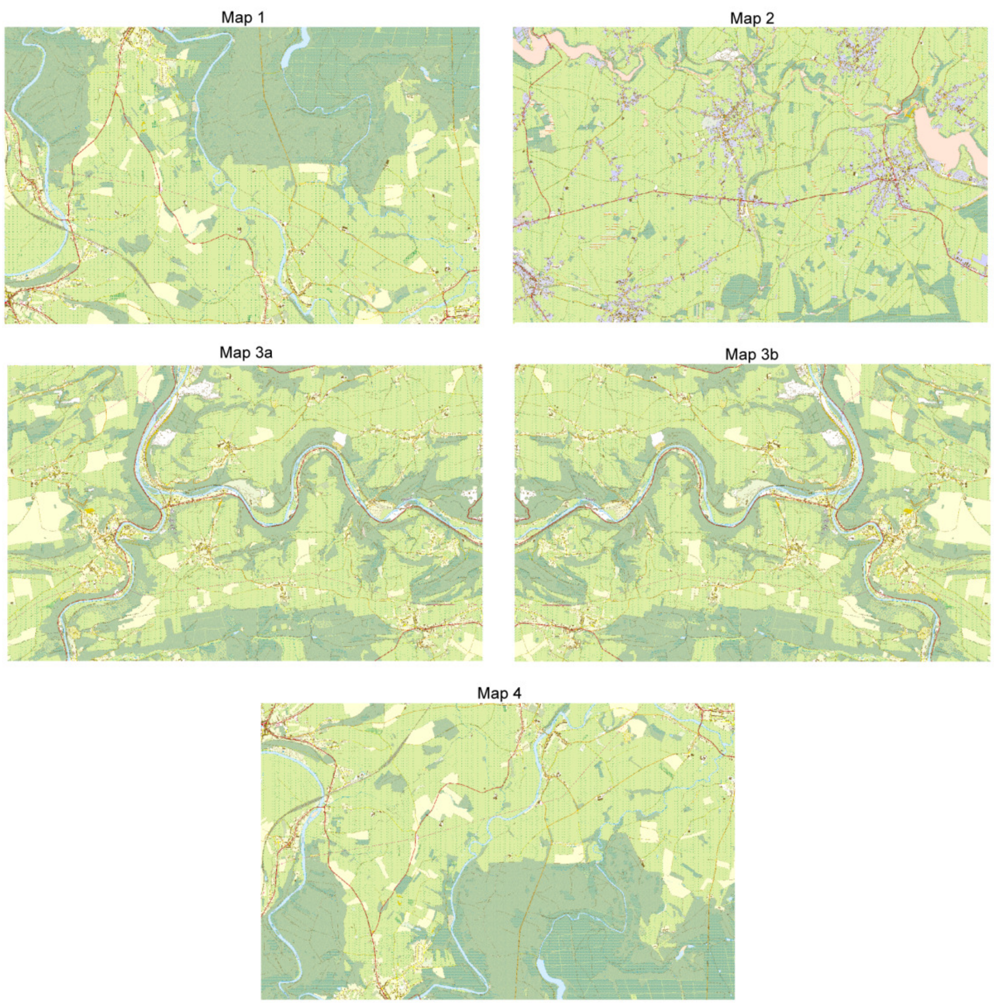

Figure 2: Stimuli depicted during the trials of the user study

As can be seen in Figure 2 map 4 is a mirrored version of map 1; this time over the horizontal central axis. This means that each participant saw both the original map and the mirrored version, separated by two other stimuli (map 2 and map 3a or map 3b). This would provide insights in how the familiarity of the map image, which is also significantly different, influences the interpretation process of the map users. Mirroring the map images (map 1 vs. map 4 and the two versions of map 3) is done at random. With map 1 and map 4, all users saw both maps to see the influence of (controlled) familiarity: bottom up and top down processing. The contents of map 3 was only depicted once (mainly bottom up processing), ruling out the familiarity element. In both case however, we can compare the influence on the users' eye movements (e.g. scanpaths).

Finally, the second topographic map is characterized by a deviating use of colors to depict water bodies and village backgrounds. The hue of both original colors (cyan and light yellow) has been changed over 180° into a light orange and purple respectively. When a cartographer wants to improve the design (symbology) of a map, he has to alter something in it (e.g. the color scheme). To map users, this is a deviation to what they are used to (or familiar with). Therefore, it is also important to know how users react (during the interpretation process) to such deviations. We selected the village background color and the color of the water bodies, because these are elements that are (rather prominently) present on all depicted stimuli. The map with the deviating color, was depicted in the second trial. As a consequence, all participants (novices and experts) already saw a map with a 'normal' color scheme.

Likewise to the familiarity of a region, the users might be familiar with the color use of the Belgian topographic map (at a scale of 1 : 10 000). Deviations in this color use could distract or confuse the users and thus influence the interpretation process. In the post-study questionnaire, the participants also had to indicate their level of familiarity with the Belgian topographic maps (on 1 : 10 000). Most experts indicated that they used these maps on a regular basis, whereas the novices did not. Therefore, this could influence in how both user groups react to deviations in the maps design (color use).

2.4. Apparatus and recordings

The participants' eye movements were registered with an EyeLink1000 eye tracking device from SR Research (Mississauga, Ontario, Canada). This device is installed in the eye tracking laboratory of the Department of Experimental Psychology (Ghent University). This desk mounted device with a chin rest can sample a user's POR at a rate of 1000 Hz. The maps were presented on a 21 inch monitor.

The software DataViewer (SR Research) was used to aggregate the raw data into meaningful measurements, such as fixations and saccades. Fixations correspond to time intervals at which the POR is relatively stable. At this moment the user is interpreting the visual information, and as a consequence, the metrics related to this measurement are of utmost importance. The DataViewer has tools to create reports listing the number of fixations and the average duration of these fixations for each trial. Furthermore, detailed information regarding these fixation metrics can be obtained separately for indicated Areas of Interest (AOIs). These AOIs are regions (squares in this case) that are subsequently compared with the eye movement data. For each AOI the number of fixation and the total duration of the fixations within its boundaries are listed (separately for each trial and participant).

3. Methodology and results

3.1. Statistical comparison: experts vs. novices

In the DataViewer software, a trial report can be exported. This report aggregates the available data (including eye movement measurements) per trial. The eye movement metrics of interest are the average duration of the participants' fixations and the number of fixations per second. The average duration of the fixations can indicate the difficulty with which the visual stimulus is processed (e.g. Duchowsky, 2007; Holmqvist, et al., 2011, Goldberg, et al., 2002, Jacob and Karn, 2003, Poole and Ball, 2006; Rayner, 1998). Complex or chaotic stimuli, which are difficult to process by the user due to a rise in the cognitive load, typically result in longer fixation durations. If a user would find a part of the visual stimulus particularly interesting, the duration of the fixations would also increase. The number of fixations a user can have per second is closely linked to the average fixation duration of that user. Longer fixation durations result in fewer fixations per second. However, studying both metrics can be useful to explain the results.

Table1: statistical comparison between expert and novice map users

(average fixation duration; number of fixations per second; duration of the trial)

	Experts (N=48)		***Novices (N=48)***		***ANOVA***		***Cohen's d***	
	M	***SD***	***M***	***SD***	***F***	***P***	***d***	***effect***
AvgFixDur(s)	.308	.046	.339	.062	7.578	**.007**	0,568	*medium*
Fix/s	2.804	.467	2.566	.469	6.235	**.014**	0,509	*medium*
TrialDur (s)	335.7	128.0	205.7	128.5	10.516	**.002**	1,014	*large*

Table 1 lists the mean values (*M*) and standard deviations (*SD*) for the average fixation durations, the number of fixations per second and the duration of the trial, both for the expert and novice users. The third main column in this table shows the results of a one way ANOVA between the two user groups. The last column provides more information regarding the effect size of the ANOVA test: Cohen's d. These latter results (medium and large effect) that a sufficient large sample size has been tested. From the ANOVA tests it can be derived that the expert users have significantly shorter fixations than the novice users. Furthermore, experts can have significantly more fixations per second. These findings are in correspondence with the results described by Ooms, et al. (2012a), who analyzed the eye movement metrics resulting from a visual search on a very basic map design. The results obtained in the present study indicate that the findings of Ooms, et al. (2012a) can be generalized to a wider range of maps and applications.

As mentioned in the description of the tasks, the participants could decide for themselves how long they wanted to study the screen map. The last row in Table 1 indicates that experts chose to study the map for a longer period of time than novice users (335.7 s or 5.6 minutes vs. 205.7 s or 3.4 minutes).

3.2. Heatmaps – density maps

Eye movement data are regularly visualized by what is often called *heatmaps* in eye tracking software. These are actually density maps, but we will continue to use the term 'heatmap' as this is most commonly used in eye tracking research. In Figure 3, four of such heatmaps from the same participant are depicted (each associated with a different stimulus). Almost all software accompanying eye tracking systems contain tools to create heatmaps. These 'maps' visualize the intensity levels where the participant was looking at the stimulus. Typically, a color scale ranging over three colors is used: green (areas with lower fixation intensities) over yellow (areas with higher fixation intensities) to red (areas with very high fixation intensities). It must be noted here that in most software packages it is possible to change this color range to a more cartographically acceptable representation – using one color (hue) (cfr. Bertin, 1967) – but this option is rarely used.

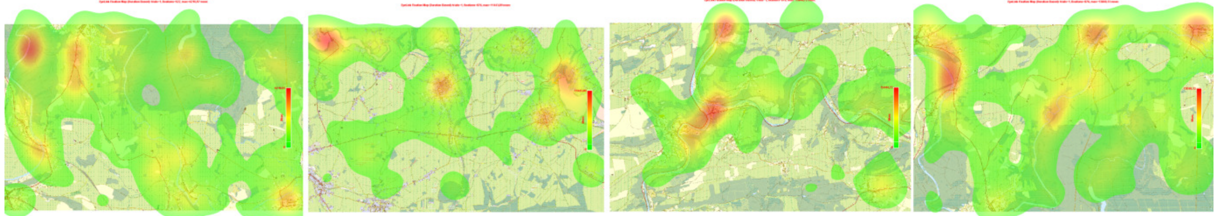

Figure 3: Heatmaps from the same participant for the four trials

Heatmaps are a good aid to get an initial overview of the eye movement data, but they imply a number of serious drawbacks. First, it is very difficult to compare the heatmaps objectively; nowadays this is often done just at sight (qualitative analysis). Second, most software do not allow operators to adapt the classification system. Depending on the topic under investigation, the focus of the visualization might be on the general pattern of the fixation intensities or on the extreme values. As a consequence, different classification schemes are desirable. However, adaptations in the standard classification scheme are often not possible. Third, it is not possible to detect extreme values between, for example, different user groups. This is again a consequence of the application of the standard classification scheme. The software determines the maximum value (total dwell time in this case) and applies the same color scheme on all maps based on this criterion. For example, the maximum values for the heatmaps in Figure 3 are respectively 6.211 s, 11.444 s, 15.445 s, and 13.048 s. Although the maximum value of the first heatmap is about half these of the other heatmaps, the same color scheme is applied. Consequently, from these heatmaps in Figure 3 it cannot be derived that the fixation intensity is much lower on map 1. Because of these drawbacks, an alternative to the heatmap visualization is proposed in the next section: the gridded visualization.

3.3. Gridded visualization: methodology

A similar approach to visually analyze eye movement data was described in Brodersen, et al. (2001). A grid of AOIs was placed over each map image to obtain detailed information on the participants' fixation in each of the grid cells. In Figure 4, this grid of AOIs is depicted in yellow; the cyan circles represent the participant's fixations. The size of the cells was chosen as such that detailed information could be obtained (small enough), considering the accuracy of the eye tracker. A maximum (acceptable) deviation of 0.5° on the calibration results in a deviation of 0.6 cm on the screen (at a viewing distance of 70cm). Therefore, the cell sizes should preferable be no less than 1.2cm (or about 34.9 pixels). Based on the size of the map image (1280x800 pixels), it was decided to use square sized AOI's with the round number of 40 by 40 pixels.

This means that a grid of 20 by 32 cells is placed over the map image, resulting in 640 AOIs. The DataViewer software (SR Research) can create so called AOI reports that list, among others, the fixation count and the total dwell time in each of the AOIs related to one trial. The AOIs report's structure is as such that all data related to a specific AOI are represented on a single row and all AOIs are listed underneath each other. One report was created for each participant, containing the four trials.

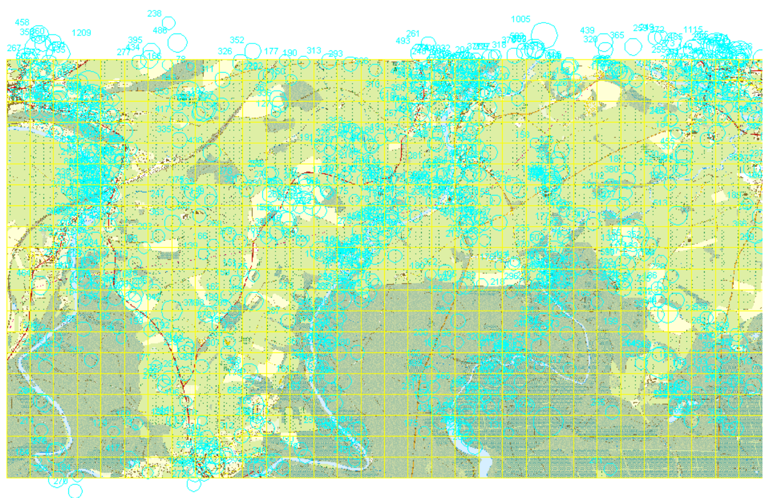

Figure 4: Grid of AOIs (yellow) placed over each map image

to visually analyze the users' fixations (cyan)

Two columns in this report are of particular interest: the total count of fixations and the total dwell time per trial, within each AOI respectively. As was mentioned in section 1.2, longer fixations can indicate that the user finds it difficult to process the information. The number of fixations is closely linked to their duration: longer fixations should result in less fixations and vice versa. However, it is good practice to study both to be able to detect deviations in the participants behavior, which could be linked to the scanpaths (e.g. Duchowsky, 2007; Holmqvist, et al., 2011).

However, as mentioned before (section 3.1), a significant difference in the duration of the trials was observed between the expert and novice users. As a consequence, these absolute values as such are not comparable between both user groups. The longer trials of the experts can have a significant influence on the number of fixations counted during each trial and thus on the total duration of these fixations. In order to be able to compare these measurements objectively, normalized values linked to a uniform trial duration, are used. The mean trial duration of the experts was 335.7 seconds, whereas this was 205.7 seconds for the novice users. Therefore, it was decided to use a uniform trial duration of 300 s (or 5 minutes). All data – total fixation count and total dwell time – were recalculated based on the initial and this uniform trial duration: original value/trial duration*uniform trial duration.

In order to be able to present the results visually and spatially, a program was written (in JAVA) that could read these (adapted) AOI-reports and restructure the data to obtain the grid of 32 by 20 cells. All values were, , based on the (x, y)-position of the corresponding fixations, placed on its correct (spatial) position in the original grid for each map. These grids were constructed for the adapted total fixation count and dwell time, resulting in a total of 192 grids (24 participants, 4 maps, and 2 dependent variables). Next, the values in all corresponding AOIs are aggregated over for each stimulus, separately for each user group (experts vs. novices) and each dependent variable (fixation count and dwell time). To obtain a general overview of these data, the average value in each grid cell was calculated in the aggregated grids. Furthermore, the maximum value in each corresponding AOI was also located to identify possible outliers and deviations in the data. This resulted in maximum values of both the fixation count and dwell time, separately for the four maps and the two user groups: 16 grids. These operations (average and maximum) aggregated the 192 original grids in in 32 grids: 4 maps, 2 user groups (experts vs. novices), 2 variables (fixation count and dwell time), and 2 aggregation types (average and maximum).

Finally, the obtained (aggregated) values were categorized into eight different classes. A grayscale color was assigned to each class, based on the ColorBrewer, an online tool to create usable color scheme for maps (Brewer, 2012). The addition of this visual component facilitates the interpretation of the grids. The classification of the values was chosen as such that a general overview could be obtained regarding the location of fixation patterns and extreme values. The classification and color scheme applied to the data (fixation count and dwell time) is presented in Table 2.

Table 2: Classification and color scheme for fixation visualization

Variable	***Classification and color schemes***							
FixCount	[0-1]	[1-2]	[2-4]	[4-6]	[6-8]	[8-10]	[10-20]	[20-...]
FixDur (dwell time)	[0.000-0.325]	[0.325-0.650]	[0.650-1.300]	[1.300-1.950]	[1.950-2.600]	[2.600-3.250]	[3.250-6.500]	[6.500-……...]
Color (RGB)	255	247	217	189	150	99	37	0

The *FixDur* in this table corresponds to the total (summed) fixation durations for the whole trial, recalculated to the uniform trial duration of 300 seconds. This total (summed) fixation duration is also called ‘dwell time’. The color scale used to visually enhance the (spatial) presentation of these dwell time distributions can be keyed to the color scale of the fixation counts. The average fixation duration (for a single fixation) for this assignment was about 0.325 seconds. Next, the boundaries of the fixation count classification could be recalculated based on this average. Linking the classification of the fixation durations to these of the fixation counts allows detecting regions where users are staring: the classification of the (summed) fixation duration is higher than the classification of the fixation count in the corresponding cell. The resulting grids are discussed in detail in the next sections.

3.4. Gridded visualization: total fixation count

The resulting aggregated gridded visualizations are depicted in Figure 5 for the average values and in Figure 6 for the maximum values. In general, a similar pattern in the fixation counts (both for the average and the maximum values) is noticed between both user groups. This fixation pattern reflects the general structure of each of the stimuli. In the grid that corresponds with map 1, two vertical linear clusters with a higher fixation count can be identified. These correspond to the two leftmost rivers on this map. The users from both user groups focus on these linear structures, resulting in a cluster of fixations and thus a higher fixation count, to make sure that they remember a reference frame in which other map elements can be placed. The focus (or the users' attention) on these linear elements is somewhat stronger with the experts than with the novices. Another interesting element (mostly considered by the expert group) seems to be located in the lower right corner of the map. This is a village that is flanked by a major road.

Map 4 in the experiment is the mirrored version of map 1. The structure in the fixation pattern for this fourth map could thus be the mirrored equivalent of this for map 1. In map 4, a similar (mirrored) pattern regarding the two leftmost rivers can indeed be identified, along with the village in the (now) upper right corner. Regarding the maximum fixation count of map 4 (Figure 6), it can be noticed that the grid of the novice group has a darker background than this of the expert group. This means that the maximum values regarding the fixation count are higher in the whole map image, whereas the experts seem to direct their focus more on particular items. This pattern cannot be derived from the grid with the average values (Figure 5). A more detailed (statistical) comparison of the fixation counts between map 1 and map 4 is discussed in a separate section (section 3.8).

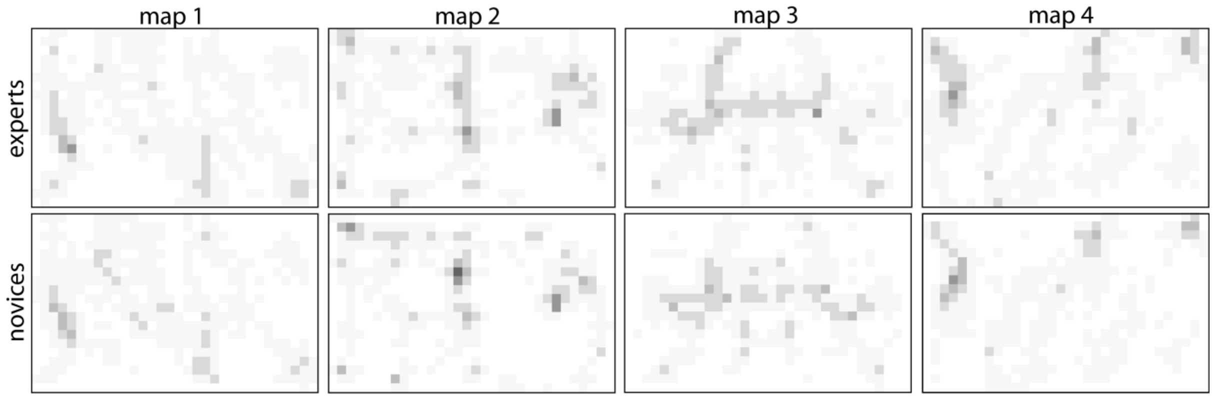

Figure 5: Grid of AOIs depicting the average fixation count

for experts (top row) and novices (bottom row) users for each stimuli (left to right)

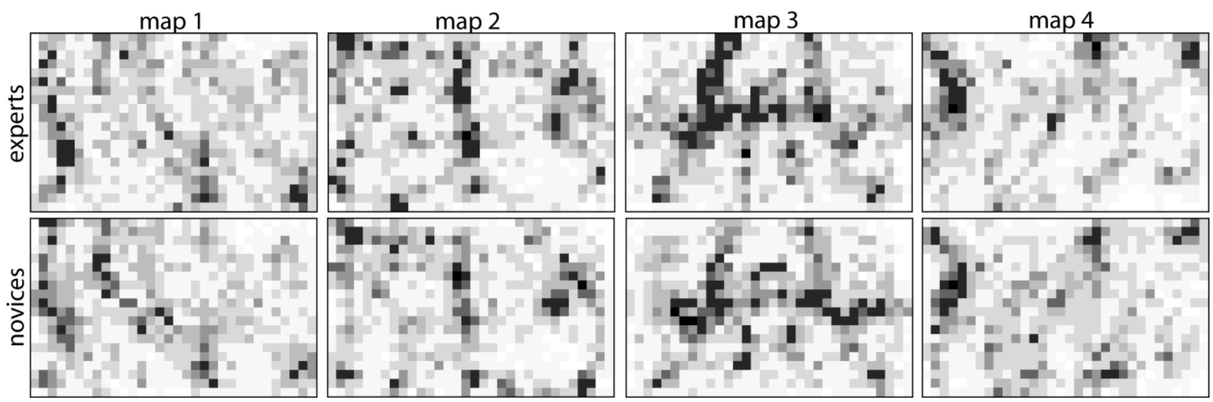

Figure 6: Grid of AOIs depicting the maximum fixation count

for experts (top row) and novices (bottom row) users for each stimuli (left to right)

In the grids related to map 2, a vertical linear structure can be identified in the middle of the grid. This corresponds to the location of a major road on the map. The higher fixation count related to this road indicates that this road is fixated the most during the trial, which suggests that the users especially wanted to remember this road as a reference frame. This linear structure is more obvious in the grids of the expert group than in these of the novice group. This indicates that the experts tend to focus their attention more on this reference frame than the novices. Other points of focus can be found in the top left corner and the upper right side of the map. These fixation clusters correspond to the location of the water bodies on the map, which were visualized in a deviating color. This deviation in color use seems to attract the users' attention, resulting in a higher fixation count. The deviating color use regarding the village background does not seem to have an influence on the attentive behavior of both user groups.

During the study, two types of stimuli were used in the third trial. Half of the participants saw the original map, the other half saw a mirrored version (over its central vertical axis). The measurements of these two stimuli are aggregated (average or maximum) and presented in one grid. This results in a fixation pattern that is similar (but mirrored) on the left and right side of the map. Especially the fixation count of the expert group reflects the linear structure of the vertical and horizontal road/river combination on the map. This is also present in the grids of the novices, but less pronounced. Both the experts and novices seem to have a higher fixation count on the left side of the map, although half of the participants of each group saw the mirrored version. This pattern is again more pronounced in the expert group, particularly when considering the maximum fixation count. This could indicate that the map users tend to fixate more on the left side of the map.

When averaging the fixation count for each of the four quadrants (upper left, upper right, lower left, and lower right) instead of for each AOI, the fixation count is always the highest in the upper part of each map as opposed to the lower part. Furthermore, a higher number of fixations is counted in the left part of each map opposite to the right part. More detailed information on how users looked at the map over time can be identified by studying the evolution of their scanpaths. This is discussed in section 3.9.

3.5. Gridded visualization: total dwell time

Besides the number of fixations at a certain location, the total dwell time might provide important insights in the users' cognitive processes while trying to interpret the map's contents. Longer fixation durations might, on the one hand, indicate that the user finds a certain region of the visual stimulus particularly interesting. On the other hand, these longer fixation durations might also indicate that the user finds it difficult to interpret the contents. This means that the user has difficulty to recognizing the object, resulting in a higher cognitive load, which is also linked to longer fixation durations (e.g. Duchowsky, 2007; Holmqvist, et al., 2011, Goldberg, et al., 2002, Jacob and Karn, 2003, Poole and Ball, 2006; Rayner, 1998). Two of such grids, related to map 2 for experts and novices, are depicted in Figure 7.

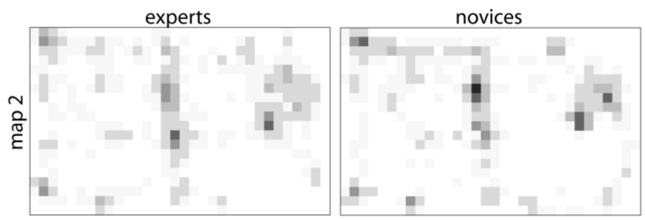

Figure 7: Grid of AOIs depicting the average dwell time (map 2)

This visualization provides a good overview regarding the dwell time, linked to the counted number of fixations (see Figure 5). Similar to the fixation counts, the general structure of the map is reflected in the grid: the main linear structures are linked to longer dwell time, both for the experts and novices. When comparing map 2 in Figure 5 (fixation count) and Figure 7, it can be confirmed that the classification distribution of the fixation counts is in correspondence with the classification of the dwell time. This indicates that there is a strong relationship between the fixation count and fixation duration: longer fixations result in fewer fixations. This, in turn, indicates that there are no deviations in the duration of the saccades between the fixations. This observation holds true for both user groups and is similar for the other stimuli. In order to investigate the differences between both user groups on a more detailed level, other visualization methods are indispensable. Two of such methods are discussed in the next sections: 3D gridded visualization of the dwell time (Section 3.6) and calculation of the average fixation duration (per fixation) (Section 3.7).

3.6. 3D gridded visualization: dwell time

With the 3D gridded visualizations, an extra dimension is added to the original (aggregated) grids. In each cell of the grid (or AOI), a bar is constructed which height corresponds to the value in that cell. The 3D gridded visualizations of the average dwell time are depicted in Figure 8 for map 1 and map 4; and in

Figure 9 for map 2 and map 3. This visualization allows comparing the values for the dwell times (summed fixation durations) between experts and novices in more detail, as the data is not classified. The downside of this approach is that the spatial distribution of the values on the grid is not that clear. The perspective view, in combination with the higher bars in front of the image, conceals the lower bars behind them. However, these 3D graphs are a good aid to study differences in the general and extreme values between two user groups, without considering their spatial location in much detail. The (normal) gridded visualization is less suitable for obtaining detailed insight in the differences of the actual values between the two user groups, but their spatial distribution is clearly visible. As a consequence, both approaches complement each other in the type of information that can be obtained.

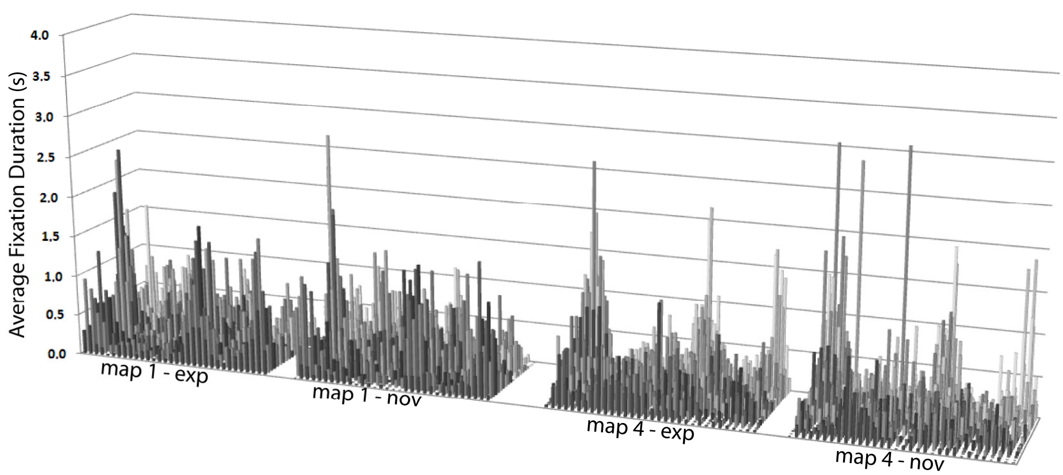

Figure 8: 3D representation of the average dwell time for map 1 and map 4

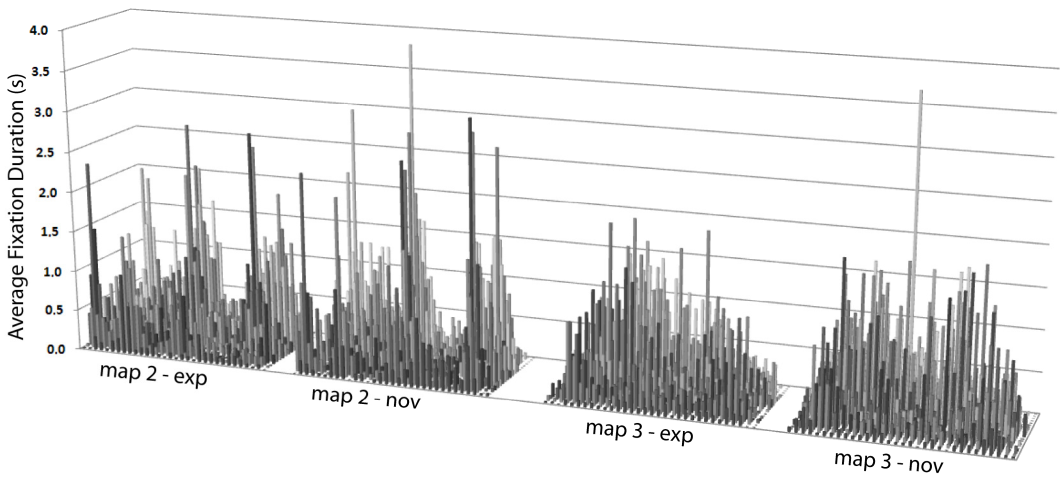

Figure 9: 3D representation of the average dwell time for map 2 and map 3

Figure 8 represents the 3D gridded visualization of the average dwell time related to map 1 and map 4. The resulting graphs for the two user groups are placed next to each other to allow a better comparison. Furthermore, map 1 and map 4 are placed in the same figure to be able to compare their corresponding (mirrored) values. The results of the remaining stimuli (map 2 and map 3) are depicted in Figure 9. These figures indicate that the resulting dwell times in the depicted AOIs are very similar between the two user groups, which is consistent with the gridded visualizations of the fixation counts. However, the novice group seems to have more extreme values – rather long dwell times or higher bars – in each of the maps. Hardly any of the measurements related to the expert group are higher than 2.5 seconds; a threshold which is more often crossed by the novice group.

The extreme values related to map 4 might be explained by the fact that this map was the mirrored version of map 1. The users recognized the structure of the map, but this was displayed upside down. This latter element could cause confusion. This is a typical case of *proactive interference* (Matlin, 2002). The users find it difficult to interpret and learn new material (map 4), because of previously learned material (map 1) that keeps interfering with the current interpretation and learning process. This negative influence on the users' cognitive processes (a higher cognitive load) results in longer dwell times.

The peaks observed in map 2 are, on the one hand, situated in the middle of the map, which corresponds to the location of the main vertical road on the map. The higher dwell times on the top left and upper right side of the map, on the other hand, correspond to the locations of the two water bodies which were depicted in a deviating color. From the fixation counts in the gridded visualization, it could be derived that these regions were clustered with fixations, which sum up to this higher total dwell time. The users' attention is attracted by these 'strange objects', but still higher values are observed for the novice group. However, the deviating background color of the villages (light purple instead of light yellow) does not seem to influence the attentive behavior of the users.

The 3D graph depicting the average dwell times of the expert group looking at map 3 shows a more homogeneous distribution. This could be explained by the fact that half of the participants saw the mirrored version of the map. However, the bars are higher on the left side of the map, which is in correspondence with the gridded visualization of the fixation counts. The expert users spend more time fixating the left side of the map than the right side, despite the mirrored map image. This observation does not hold true for the novice users; the height of the bars is nearly equal on the left and right side of the map. An extreme value is noticed in the middle of the map, which cannot be linked to an extreme measurement in the fixation counts. This indicates that the users were staring at this location on the map. However, no special or deviation color use or objects are located at that position on the map. The cause of this outlier can thus not be brought back to anything on the map itself.

Extreme values in the dwell times would imply that the user fixates a certain region during an abnormal amount of time. This would indicate that the user is, on the one hand, attracted by something in that region. It can, on the other hand, also indicate regions that are difficult to interpret by the user, resulting in longer fixations but not necessary a higher count. To distinguish these two options, an additional eye movement metric is studied in more detail: the average fixation duration of a single fixation, which can then be compared to the map's contents at that location using the (3D) gridded visualization).

3.7. 3D gridded visualization: fixation duration

The average fixation duration (of a single fixation) was already statistically analyzed in Section 3.3. As mentioned before, these statistical analyses miss the spatial dimension inherently linked to maps and their design. The difficulty with which a user interprets the visual contents at a certain location on a map can identify problems in the map's design. This difficulty is typically reflected in longer fixation durations, due to a higher cognitive load. However, longer fixation durations might also indicate that a certain object is more engaging in some way (e.g. Duchowsky, 2007; Holmqvist, et al., 2011; Goldberg, et al., 2002; Jacob and Karn, 2003; Poole and Ball, 2006; Rayner, 1998). The difference between both interpretations can be made by linking the (deviating) results to the actual map contents at that location.

Similar to the average dwell time, a 3D gridded visualization is created in which each bar height corresponds to the average fixation duration (of a single fixation) at that location. These graphs are depicted in Figure 10 and Figure 11. Comparing the results for all maps between the expert and the novice users reveals that novices tend to have more deviating (longer) fixation durations. This general trend shows that the novices find it more difficult to interpret (and thus learn) the contents of the map (for all maps), causing a higher cognitive load (e.g. Bunch & Lloyd, 2006; Harrower, 2007; Holmqvist, et al., 2011; Jacob and Karn, 2003; Poole and Ball, 2006; Rayner, 1998).

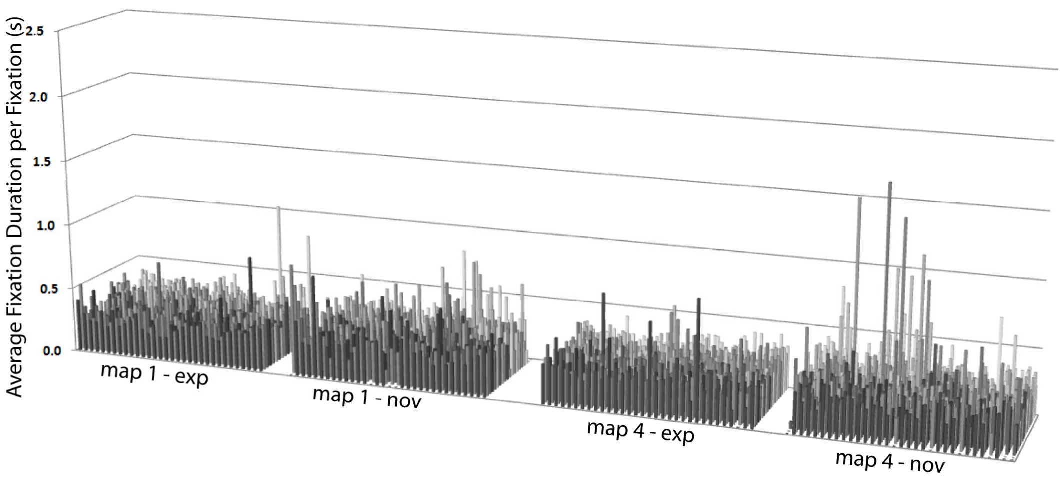

Figure 10: 3D representation of the average fixation duration per fixation for the expert users

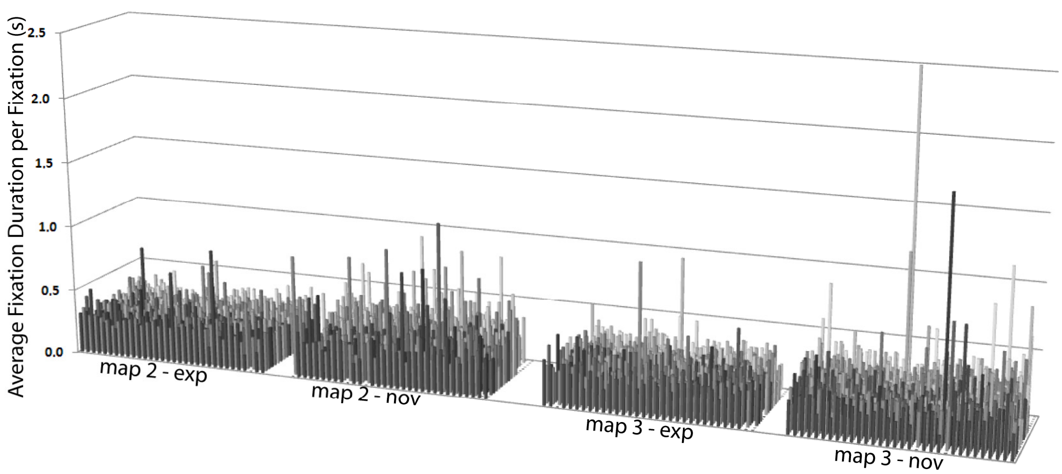

Figure 11: 3D representation of the average fixation duration per fixation for the novice users

Although map 4 is the mirrored version of map 1, the difference between experts and novices is much more pronounced on map 4 (see Figure 10). The expert group has a number of fixation durations which are longer than normal, but the novices show a cluster of very high fixation durations near the middle of the map. The position of these grid cells corresponds to the location of the calibration target that was displayed between each map to check the validity of the calibration. Consequently, this was also the region where the users are looking at the moment that the map was displayed. It can thus be derived that the novice users had longer fixation durations when the map was first displayed. This indicates confusion, which might be explained by the recognition of map 1; however this is displayed ‘upside down’.

The difference between expert and novice users is also clearly visible in the 3D graphs related to map 2. The higher bars in the experts’ graph are centered on the main vertical and horizontal road, whereas these of the novices’ graph are distributed over the entire map image. No particular deviations are noticed at the location of water bodies (depicted in a deviating color).

In map 3, the novice group is characterized by the same outlier that was also present in the 3D graphs of the total fixation durations. When studying the original (not aggregated) data, it could be concluded that this high fixation duration was caused by a single participant who stared at this AOI for an extreme long period of time (14.2 s). This measurement distorted the average over all novice participants for this AOI. This type of ‘staring’ could be explained by the cognitive process of rehearsal in order to remember (or learn) the map image: transferring information from the WM to the LTM. The expert group also shows a number of longer fixation durations on this map, but clearly more are found in the novice group.

3.8. Statistical grids: mirrored maps

The statistical grids add a statistical component to the gridded visualization. The values of all corresponding cells (AOIs) are compared statistically, using a one way ANOVA. For each grid comparison, 640 significance (P) values are thus obtained. These results are again placed in the gridded structure. To visually represent these ANOVA results, a classification scheme with four classes (and thus colors) is applied on the grid. This classification scheme is listed in Table 3.

Table 3: Classification and color scheme for the ANOVA results

Variable	***Classification and color schemes***			
Sign. (P)	> 0.1	[0.1-0.05]	[0.05-0.01]	<0.01

Description	not significant	near significant	significant	highly significant
Color (RGB)	255	217	150	37

From Section 3.4 it could be concluded that the distribution of the total fixation count reflected the general structure of the map. Important and engaging items are fixated more often than other items. These important items corresponded to the main linear structures on the map, which might be used as a reference frame. What is more, map 1 and map 4 are each other's mirrored equivalents. As a consequence, it could be expected that the patterns found in the related grids are also each others' mirrored equivalent. The same can be expected from the fixation distributions related to map 3. Half of the participants saw the original map; the other half saw the mirrored version (this time over the vertical axis).

Statistical grids method is used to test whether this hypothesis holds true. The first column in Figure 12 depicts two of such statistical grids, related to map 1 and map 4; the second column contains the tests related to map 3. The top images depict the comparison between the original grids. The comparison between map 1 and map 4 shows a lot of significant and highly significant differences in the fixation counts of the corresponding cells in the grid. This amount is less in the comparisons related to the map 3, but the map's pattern is still clearly visible. This can be explained by users' focus on the main structuring elements. In both statistical grids, the horizontal and vertical axis of the mirror operation can also be distinguished. This is the location where the original and mirrored version overlap, resulting in a lighter line in the statistical grids: not significantly or nearly significantly different. The *P*-values show a clear similar (mirrored) pattern on both sides of this axis (above vs. below for map 1 and map 4; left vs. right for map 3).

The lower left grid in Figure 12 shows the statistical comparison between map 1 and the mirrored version of the grids related to map 4. By mirroring the grids related to map 4 over their horizontal axis, it could be expected that their pattern reflects the structure of map 1. This is confirmed by the depicted statistical grid. Very few (highly) significant differences are found in the grid. The majority of the grid is populated with not significant or near significant values ($P<0.1$). A cluster of significant differences is found in the lower left corner of the grid, which corresponds to the location of a village and a crossroads.

An even more similar result is obtained when comparing the grids of the original version of map 3 with the mirrored grids of the adapted version of map 3 (lower right corner in Figure 12). In this grid, almost no highly significant differences are found. The better result regarding map 3 might be explained by the fact that the participants saw both map 1 and map 4, which caused confusion (particularly among the novice users). The two versions of map 3 were looked upon by two separate user groups, avoiding influences on the cognitive processes due to recognition or confusion. The statistical grids at the bottom of Figure 12 thus indicate that the patterns of the users' fixations are guided primarily by the main (linear) structures on the map.

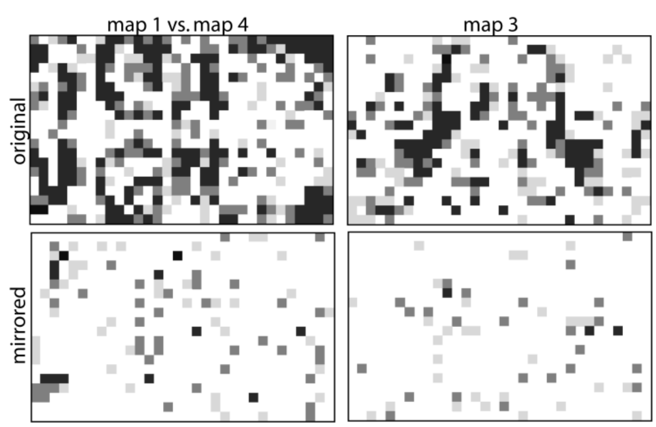

Figure 12: Statistical comparison of the fixation count between two mirrored maps

3.9. Scanpath visualization

Another way to explore and analyze the spatial dimension of the eye movements visually is to study the scanpaths of the participants. These scanpaths are sequences of subsequent fixations and saccades. The Visual Analytics Toolkit was used to visualize the participants' scanpaths on top of the actual stimuli. Filter operations based on attributes (stimuli and user group) and on time intervals facilitate the visual analyses of the eye movements (Ooms, et al., 2012). Figure 13 illustrates these scanpaths, separately for each map and user group. What is more, to study the evolution of these scanpaths, different time intervals (during the first minute) are also depicted: 0 to 10 s; 0 to 30 s; 30 to 60 s.

The location of the participants' scanpaths seems to be clustered on the main structuring elements of the map (major roads and rivers), with a very similar pattern between the experts and novices. This pattern remains visible during the entire first minute: the participants keep directing their attention on these main (often linear) elements. During the second half of the first minute, the participants also fixate other objects, but the structuring elements still receive a lot of attention.

When comparing the scanpaths during the first ten seconds between map 1 and map 4, a striking difference is observed. In map 1 (similarly to the other maps) the participants already direct their attention on the main structuring elements in the map. This holds especially true for the participants in the expert group. However, the scanpaths associated with map 4 are rather chaotic during the first ten seconds. The structure of the two leftmost vertical lines is not present, which was the case for map 1. Especially the novice map users show a high number of horizontal scanpath lines, zigzagging across the image during the first thirty seconds of the map's display. These chaotic scanpaths likely indicate that the users are confused by this mirrored map image (Çöltekin, et al., 2010; Fisher, et al.,1981; Stark & Ellis, 1981)

The experts' scanpaths on map 2 show a cluster on the horizontal and vertical major road on the map image during the first ten seconds. The novices seem to be more distracted by the water bodies (with the deviating color use). In the first thirty seconds, the expert users also fixate more on these water bodies, but again less during the second half of the first minute. The villages at the bottom (left) also receive more attention during this latter interval. The deviating color use to depict the background of the villages does not seem to influence the participants' attentive behavior.

The scanpaths on map 3a and map 3b are very dislike during the first ten seconds, although the same structures are found in the map image, which is especially striking for the expert users. On map 3a the experts focus on the vertical river/road on the left side of the image and less on the horizontal road/river. On map 3b the focus is more situated on the horizontal main linear structure and not so much on the vertical road/river on the right side of the map. This pattern is still visible in a longer time interval and corresponds to the findings of the fixation counts and fixation durations. The users, particularly the experts, tend to be more attracted to the left side of the map.

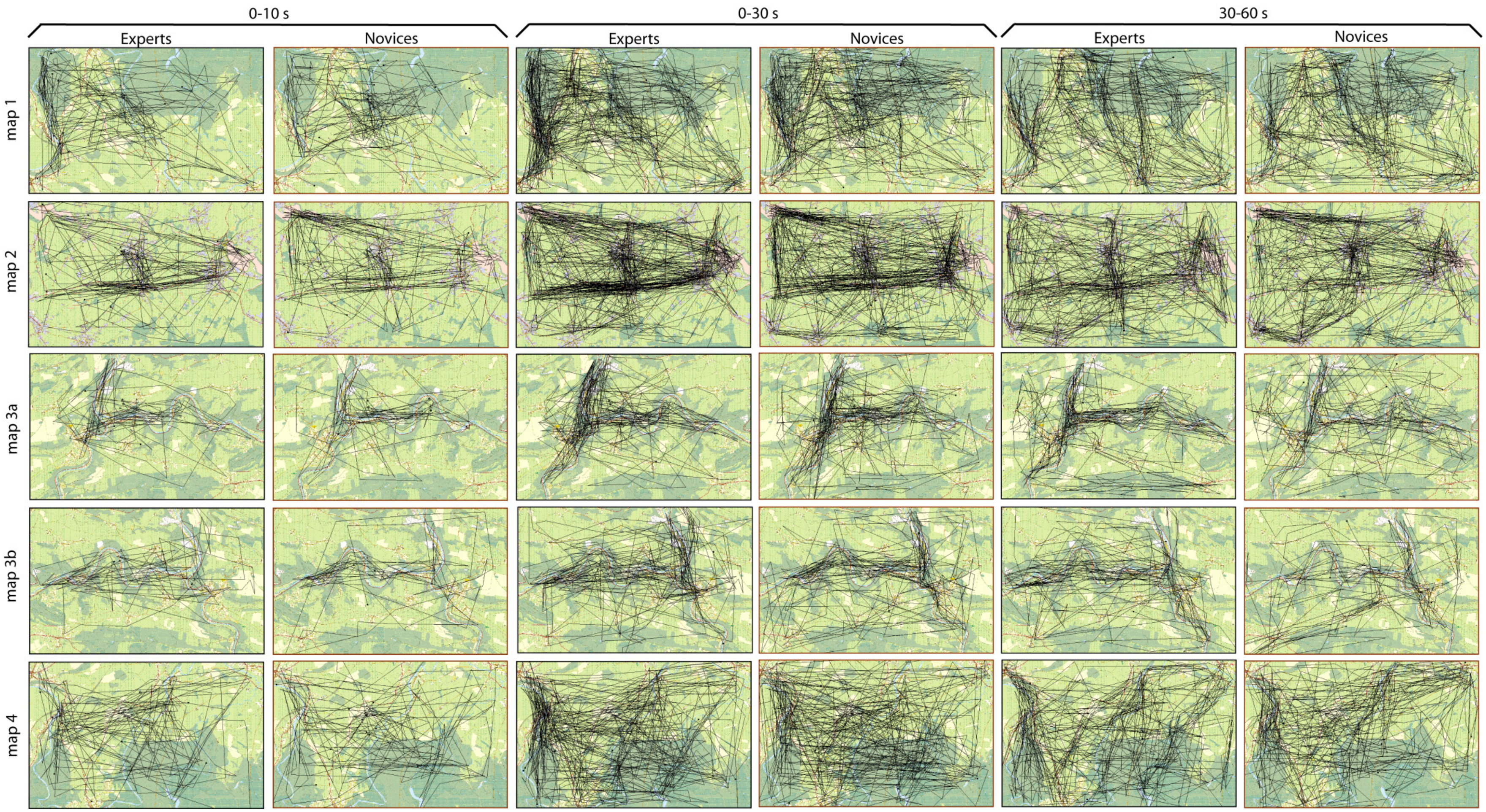

Figure 13: Evolution of the participants' scanpaths for the different maps, and user groups

4. Discussion and conclusion

The study described in this paper is an extension of the work done by Ooms, et al. (2012a) in order to verify whether their results could be generalized to wider array, and thus more complex, map types. The main aim of the experiment is to gain a better understanding of the way expert and novice map users process and interpret the (complex) visual information on maps. Furthermore, deviations in the stimuli are introduced in order to study their effects on the map users' eye movements, and thus on their attentive behavior.

The statistical analyses confirm the findings of Ooms, et al. (2012a). The experts' fixation durations are significantly shorter than these from the novices. This indicates that their interpretation process, including the different stages of object recognition, is much faster than this of the novice users. This could be explained by the level of experience, and thus the amount of background knowledge stored in the LTM, that the expert users have in comparison to the novices. Due to these shorter fixations, the experts can have more fixations per second, and can therefore interpret a larger part of the map in the same amount of time. As a consequence, it can be concluded that expert map users can interpret the maps more efficiently, both when simple and complex maps are considered.

In order to spatially analyze these results, different approaches are presented, which all complement each other: gridded visualizations, 3D gridded visualizations, statistical grids, and scanpaths analyses. These visual and (mainly) qualitative methods shed light on different aspects of the user's interpretation process and allow studying differences in the attentive behavior of expert and novice users. Furthermore, a number of eye movement metrics related to the users' fixations are analyzed and compared: average fixation count in one trial, average dwell time in one trial, and average fixation duration (of a single fixation). These analyses take the spatial distribution of the fixations across the map image into account, and are based on a grid of square AOIs.

From these analyses it could be concluded that both user groups focus their attention on a reference frame in the map image, resulting in a higher number of fixations and thus longer dwell times. This reference frame mostly consists of major linear structures, such as roads and rivers. As a consequence, the main structure of the map is reflected in the gridded visualization of the total fixation counts and durations. It is important to note that, the assignment did not instruct the participants to look at these general structuring elements, just to remember the map as good as possible. The very detailed focus of the users attention on these elements in an interesting finding, especially when comparing differences in attentive behavior between the two user groups. The focus on these linear structures is more pronounced in the grids of the expert group than with the novices. The visualization of the users' scanpaths during the first ten and thirty seconds also shows that the user's attention is immediately directed towards these structuring elements. Nevertheless, the novices' eye movement measurements show more extreme values than the experts regarding the number and duration of the fixations.

The deviating color use of the water bodies in map 2 also has an influence on the users' attentive behavior. Both user groups are attracted by these objects, which is reflected in the location of the eye movements. A higher number of fixations are found in the top left and upper right region on map 2. However, this attraction seems to be stronger in the novice group than in the expert group. This latter user group focuses more on the central horizontal and vertical reference frame (major roads) on the map image. The deviating color use of the villages' background does not seem to have any influence on the users' attentive behavior or interpretation process. In the gridded visualizations, the cell that covered the villages did not contain more or longer fixations in comparison to the other stimuli which showed the villages with their 'normal' background color.

Two types of maps were used for the third stimulus: half of the participants saw the original map; the other half saw the mirrored version (over its central vertical axis). The superimposed results show a mirrored pattern in the total fixation count and dwell times. However, more fixations (and thus longer dwell times) are found on the left side of the superimposed result. This indicates that the users tend to fixate more on the left side of the map, regardless the contents of the map. A quadrant analyses on all maps also indicates that this also holds true for all other stimuli presented during the study. This observation is more distinct in the measurements of the expert group. Nevertheless, based on the findings of the registered eye movements it could also be concluded that the users' POR are mainly guided by the main structures on the map. When mirroring the results of the adapted map 3 again over its central vertical axis, the fixation clusters corresponded with these of the original map and therefore to the general structure of the original map.

This latter observation also holds true for the results related to map 1 and map 4. The mirrored version of the results related to map 4 corresponds to these of map 1. Nevertheless, the statistical grid showed more significant differences in the corresponding cells than in map 3. This could be explained by the fact that the participants saw both the original map and the mirrored map during the study. Other results related to map 4 also show evidence of proactive interference (Matlin, 2002). The map users are confused by this map because it is very similar to information they have processed before. However, the visual information is also significantly different because it is upside down. These confusions and difficulties in the interpretation process translate into longer fixation durations, which is especially visible in the novices' eye movement recordings.

In the gridded visualization of the total fixation count related to map 4, no obvious deviation values are detected. However, the 3D gridded visualization of the dwell times shows a number of extreme values. These values are explained in the 3D graphs representing the average fixation durations (for a single fixation), where also a number of peaks are observed. This confusion can also be derived from the scanpath visualizations. During the first thirty seconds, eye movements are not immediately directed towards the structuring elements (as was the case for map 1). The scanpaths show a more chaotic distribution over the map image.

As a result, it can be concluded that the eye movements of both user groups show a similar pattern: they reflect the general structuring elements on the map. This is also confirmed by comparing the eye movement patterns of the original map to those of its mirrored version. However, a tendency to have more fixations on the left part of the map is noticed. This attraction to the major structuring elements on the map can be influenced by striking deviations in the map image. Nevertheless, a number of substantial differences are noticed between both user groups which indicate that the expert users can process the spatial and visual information on the map more efficiently. They are more focused on the structuring elements on the map, and thus less distracted by other elements and deviations. What is more, their eye movement measurements have less extreme values, which indicate that the experts experience fewer difficulties during the interpretation process. This results in a lower cognitive load, which in turn facilitates the learning process.

References

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

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A Proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89-105). Acadeic Press, New York.

Baddeley, A. D. (1999). *Essentials of human memory*. Psychology Press, Hove, England.

Bertin, J. (1967). *La sémiologie graphique*, Paris: Mouton-Gauthier-Villars.

Brewer, C. A. (2012) http://ColorBrewer.org, 15/02/2012

Brodersen, L., Andersen, J. H. K., & Weber, S. (2001). Applying the eye-movement tracking for the study of map perception and map design. In *Kort and Matrikelstyrelsen* (pp. 98). National Survey and Cadastre, Copenhagen, Denmark.

Bunch, R. L., & Lloyd, R. E. (2006). The Cognitive Load of Geographic Information. *The professional geographer, 58* (2), 209-220.

Cartwright, W. (2012). Neocartography: Opportunities, issues, and prospects. *South African Journal of Geomatics*, *1*(1), 14-31.

Castner, H. W., & Eastman, J. R. (1984). Eye-movement parameters and perceived map complexity I. *American Cartographer, 11*(2), 107-117.

Castner, H. W., & Eastman, J. R. (1985). Eye-movement parameters and perceived map complexity II. *American Cartographer, 12*(1), 29-40.

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

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.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24* (1), 87-114.

Dobson, M. W. (1977). Eye movement parameters and map reading. *Cartography and Geographic Information Science, 4*(1), 39-58.

Duchowski, A. T. (2007). *Eye tracking methodology - Theory and practice*. Springer, London.

Fabrikant, S. I., & Lobben, A. (2009). Introduction: cognitive issues in geographic information visualization. *Cartographica, 44* (3), 139-143.

Fabrikant, S. I., Rebich-Hespanha, S., Andrienko, N., Andrienko, G., & Montello, D. R. (2008). Novel method to measure inference affordance in static small-multiple map displays representing dynamic processes. *The Cartographic Journal, 45* (3), 201-215.

Fisher, D. F., Monty, R. A., & Senders, J. W. .eds., *Eye movements: cognition and visual perception*, Hillsdale. NJ: Lawrence Erlbaum Associates.

Fits, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review, 9*(2), 24-29.

Gerber, R. V. (1981). Competence and performance in cartographic language. *The Cartographic Journal, 18* (8), 104-111.

Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002). Eye tracking in web search tasks: Design implications. *Proceedings of the Eye Tracking Research and Applications Symposium*, New York.

Harrower, M. (2007). The cognitive limits of animated maps. *Cartographica, 42* (4), 349-357.

Hegarty, M., Canham, M., & Fabrikant, S. I. (2010). Thinking about the weather: How display salience and knowledge affect performance in a graphic inference task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 37-53.

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

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* (pp. 573-605). Elsevier, Amsterdan.

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

Kolácný, A. (1969). Cartographic information - A fundamental concept and term in modern cartography. *The Cartographic Journal, 6* (1), 47-49.

MacEachren, A. M. (1995). *How maps work : representation, visualization, and design*. Guilford Press, New York.

Matlin, M. W. (2002). *Cognition* (5th ed.). Wiley & Sons, Hoboken, NJ, USA.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *The Psychological Review, 63* (2), 81-97.

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

Montello, D. R. (2009). Cognitive research in GIScience: recent achievements and future prospects. *Geography Compass, 3* (5), 1824-1840.

Nielsen, J. (1993). *Usability Engineering*. Morgan Kaufmann, San Francisco.

Olson, J. M. (1976). A coordinated approach to map communication improvement. *American Cartographer, 3* (2), 151-159.

Ooms, K., Andrienko, G., Andrienko, N., De Maeyer, P., & Fack, V. (2012). Analysing the Spatial Dimension of Eye Movement Data Using a Visual Analytic Approach. *Expert Systems with Applications, 39* (11), 1324-1332.

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

Ooms, K., De Maeyer, P., Fack, V., Van Assche, E., & Witlox, F. (2012b). Investigation the effictiveness of an efficient label placement method using eye movement data. *The Cartographic Journal, 49*(3), 234-246.

Ooms, K., De Maeyer, P., Fack, V., Van Assche, E., & Witlox, F. (2013). Listen to the map user: Cognition, memory, and expertise. *The Cartographic Journal, in press*.

Peterson, M. P. (2003). Maps and the Internet: an introduction. In M. P. Peterson (Ed.), *Maps and the Internet* (pp. 1-16). Elsevier Science, Oxford.

Poole, A., & Ball, L. J. (2006). Eye tracking in human computer interaction and usability research: current status and future prospects. In C. Ghaoui (Ed.), *Encyclopedia of Human Computer Interaction* (pp. 211-219). Idea Group.

Rayner, K. (1998). Eye movement in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.

Rubin, J., & Chisnell, D. (2008). *Handbook of Usability Testing. How to Plan, Design and Conduct Effective Tests* (second ed.). Wiley Publishing, Indianapolis.

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

Steinke, T. R. (1987). Eye movement studies in cartography and related fields. *Cartographica, 24*(2), 40-73.