

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

K. Ooms<sup>a\*</sup>, P. De Maeyer<sup>a</sup> and V. Fack<sup>b</sup>

<sup>a</sup>Department of Geography, Ghent University, Krijgslaan 281, S8, B-9000 Ghent, Belgium; <sup>b</sup>Department of Applied Mathematics and Computer Science, Ghent University, Krijgslaan 281, S9, B-9000 Ghent, Belgium

(Received 1 June 2012; accepted 6 October 2013)

The aim of this paper is to gain 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, the participants studied four screen maps which had been manipulated to introduce deviations. The eye movements of 24 expert and novice participants were tracked, recorded, and analyzed (both visually and statistically) based on a grid of Areas of Interest. These visual analyses are essential for studying the spatial dimension of maps to identify problems in design. In this research, we used visualization of eye movement metrics (fixation count and duration) in a 2D and 3D grid and a statistical comparison of the grid cells. The results show that the users' eye movements clearly reflect the main elements on the map. The users' attentive behavior is influenced by deviating colors, as their attention is drawn to it. This could also influence the users' interpretation process. Both user groups encountered difficulties when trying to interpret and store map objects that were mirrored. Insights into how different types of map users read and interpret map content are essential in this fast-evolving era of digital cartographic products.

**Keywords:** user study; eye movement; cognitive cartography

### Introduction

Cartography has undergone a tremendous technological evolution during the last two decades. Already at the beginning of the twenty-first century, it was estimated that the number of maps distributed through the Internet daily 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 is 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, cartographic communication involves 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 information to users using a special code, the cartographic syntax. They are the channels that allow (and should facilitate) transmission of information. Consequently, if a map's design is not optimal, it could introduce noise in the communication process (Montello 2002). A communication model for maps presented by Kolácný (1969) was most influential on cartographic research and was considered to be an essential aid in studies on how can maps be interpreted easily (MacEachren 1995).

New technologies have profoundly not only impacted the display of cartographic products but also the perception and interpretation of information. Screen maps create new possibilities, such as animations and user interactions, but are also inherently linked to certain critical limitations in terms of resolution, size, color use, etc. (Peterson 2003). Recently, some concerns have arisen regarding this evolution in cartography (and GIScience). 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 on cognitive issues in cartography and geographic information visualization in general (e.g., Fabrikant and Lobben 2009; Harrower 2007; Montello 2002, 2009; Slocum et al. 2001). In the following sections, the cognitive structures and processes necessary to interpret visual information (such as maps) are described.

### How can we process and interpret visual stimuli?

Montello (2002) stated that cognitive cartography includes the study of knowledge structures involved in map reading, such as perception, learning, and memory. According to the cartographic communication model, the first step in map use is the interpretation of visual information encoded in maps. The interpretation process consists of a number of subsequent steps or levels which are linked to

\*Corresponding author. Email: Kristien.Ooms@UGent.be

the structure of human memory. Atkinson and Shiffrin (1968) identified three components of memory: *sensory memory*, *short-term memory*, and *long-term memory* (LTM). This model had a major impact on the early studies in cognitive psychology and is often called the *modal model*.

Sensory memory records input from each of our senses, including vision, but such input is quickly forgotten (less than 2 seconds). Some of the information in sensory memory is transferred to short-term memory which also has a limited capacity. Short-term memory is also referred to as *working memory* (WM), and this term will be used in the remainder of this paper. In order to transfer the information from WM to the LTM, it has to be rehearsed and, thus, learned (Cowan 2001; Matlin 2002; Miller 1956). The capacity of LTM is considered to be virtually infinite.

In order to explain the processes that take place in the WM, Baddeley (1999) proposed a WM structure (see Figure 1) which 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). The WM is thus much more than a database that can store a certain number of information chunks. WM comprises information obtained through sensory memory but it can also hold "old" information retrieved from LTM. The central executive makes it possible to manipulate the three different sources of information (Matlin 2002). Atkinson and Shiffrin's model is essential to understanding how humans process visual stimuli and subsequently interpret them.

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

To understand how users perceptually organize visual scenes, another well-established 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. It also assumes that 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 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 recognizes 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 is 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, described as *other knowledge about* or *verbal level*. These levels are subsequently processed while interpreting a certain object. Olson (1976) also identified three levels of processing taking place during the interpretation (or recognition) of symbols on maps. These are: compare symbol pairs, recognize groups of symbols, and use the symbols to retrieve information from the map.

The interpretation process is based on a combination of bottom-up and top-down processing of information. Top-down processing is closely linked to information already stored in the users' memory, i.e., experiences, familiarity, etc. Bottom-up processing relies solely on visual input; i.e., without using knowledge (MacEachren 1995). Hegarty, Canham, and Fabrikant (2010), for example, examined the influence of salience (bottom-up) and domain knowledge (top-down) on map comprehension (in case 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.

The interpretation process "consumes" part of the limited capacity of WM. Another part of the capacity of the WM is used to transfer the processed information to the LTM during the learning process. Bunch and Lloyd (2006) and Harrower (2007) give an excellent description of the "consumption" of WM's limited capacity while processing geographic and cartographic information. Their explanations are based on the *cognitive load theory*. According to this theory, WM is limited in its processing capabilities, i.e., the amount of cognitive load it can take. Different types of cognitive load can be identified, all contributing

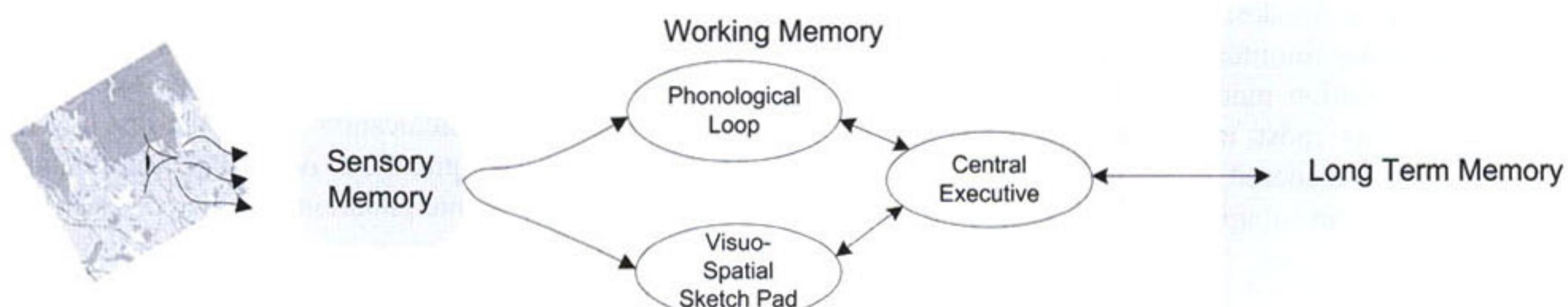


Figure 1. Memory structure: an integrated model according to Atkinson and Shiffrin (1968) and Baddeley (1999).

to the total amount of cognitive load. Maps containing overly complex information (causing a high intrinsic cognitive load) represented in a chaotic way (causing a high extraneous cognitive load) will thus be very difficult to interpret by the user. They cannot be interpreted (or learnt from) in an efficient way because there is no room left for the germane cognitive load. Germane cognitive load is addressed during the learning process (Bunch and Lloyd 2006; Harrower 2007).

Ooms et al. (2012a) studied the reaction time measurements and eye movements of expert and novice map users performing visual search on screen maps with a very basic design. Significant differences were identified between the two user groups: experts were able to interpret the contents of a map more efficiently than novice users. The authors relied on the cognitive load theory to explain this finding. Alvarez and Cavanagh (2004) investigated the *visual information load* which relates 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 and no complex objects; only points with labels and three polygons visualized in pastel color. The results obtained with these maps cannot therefore be generalized. In this research, however, we extend Ooms et al.'s (2012a) study with the incorporation of more complex (topographic) maps to gain deeper insights into how different types of users process visual information. Such research is essential, given recent technological developments in cartography in general (e.g., Fabrikant and Lobben 2009; Harrower 2007; Montello 2002, 2009; Slocum et al. 2001).

### **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 into their attentive behavior were obtained without any user interference. Reflection on one’s own thoughts results in subjective and unreliable results, often because participants do not know how their thoughts had been formed; it is an automatic process (Nielsen 1993; Rubin and Chisnell 2008).

With eye tracking, the position where a user is looking (point of regard (POR)) is recorded at a certain sampling rate. This provides insights into the user’s attentive behavior, i.e., where the focus of attention is at a given moment in time. As mentioned before, attention is an essential step in object recognition and thus in the interpretation of the map content. Besides finding the POR, other usable metrics (such as fixation duration) can be derived from eye movement data which provide insight into the user’s cognitive processes during the interpretation of visual

content. Based on previous research regarding the eye tracking method, it can safely be assumed that a close link exists between cognitive processes and eye movement metrics (Duchowski 2007; Jacob and Karn 2003; Poole and Ball 2006).

The use of eye tracking is not new. It was used to study the movements of pilots’ eyes as early as the 1950s (Fits, Jones, and Milton 1950). In the 1970s (e.g., Dobson 1977; Jenks 1973) and the 1980s (e.g., Castner and 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 they found that it could not be used to derive new knowledge. The use of the eye tracking method in cartography almost disappeared after 1985. However, recently, renewed interest in the method in cartographic research has been noticed (e.g., Brodersen, Andersen, and Weber 2001; Çöltekin et al. 2009; Fabrikant et al. 2008).

This “rediscovery” of the eye tracking method can be explained by the technical evolution of the eye tracking systems themselves. They have become smaller, less intrusive, more accurate, and less expensive. Not only the POR, but also the length and duration of fixation and saccades can be derived from these more accurate measurements. A fixation is a time interval (of at least 80 milliseconds) 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 into 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., Duchowski 2007; Holmqvist et al. 2011). Saccades are rapid eye movements between two fixations, during which no information is processed. 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 (e.g., Duchowski 2007; Goldberg et al. 2002; Holmqvist et al. 2011; Jacob and Karn 2003; Poole and Ball 2006; Rayner 1998).

This renewed interest is 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. Such knowledge is key to linking the visualization of future maps to the cognitive structures of the map users and, as a result, creating more effective maps (Cartwright 2012; Fabrikant and Lobben 2009; Montello 2009). If we can understand how map users read, process, interpret, and store (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 to process the information. This visualization helps the user to process the information.

Eye movement data can be analyzed in a number of different ways which can broadly be grouped in two main methodological categories: quantitative and qualitative methods. Quantitative methods often use statistics (e.g.,

ANOVA) to identify significant differences between two categories: these can be differences in map design tested with a homogeneous group of participants (within-user design) (e.g., Ooms et al. 2012b) or differences in user characteristics tested with a homogeneous map design (between-user design) (e.g., Nielsen 1993; Ooms et al. 2012a; Rubin and Chisnell 2008). These types of analyses have a higher level of objectivity because they are carried out on the actual numbers, applying a set of standard tests to compare them and resulting in a level of significance (*P*-value). The *P*-value is linked with a sample size and its power. All these elements make it possible to (objectively) interpret the data and compare similar analyses. Such comparisons are not available with qualitative analyses which are more “exploratory” in nature.

However, quantitative analyses often do not allow the values to be studied in the context of 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, not the results, give an indication of “where” the measures are taken or “where” the differences are. This issue is discussed in more detail in Ooms et al. (2012), who present a visual analytic approach for studying eye movement data. This can handle the spatial dimension (the where-question), 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. This is because the interpretations are done “at sight”, which does not give any information about whether the perceived difference is based on coincidence or not.

The combination of these different techniques sheds light on other aspects of the cognitive processes taking place during map reading. This means that in order to obtain the most accurate picture of the cognitive processes during the interpretation of maps, it is good practice to combine different techniques.

## **Study design**

### **Participants**

Two groups of participants were selected to take part in the study. Each group comprised 12 persons, equally divided into males and females. The first group consisted of experts in map use and cartography. All participants in this group had at least a master’s degree in geography or geomatics and received cartographic training during their studies. At the time of the study, members of group one 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 a mean of 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.

### **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.

Simple instructions were provided. The participants were told that a map would be shown on a screen and that they had to remember the structure of this map. They did not have to remember all details on the map (such as individual houses), but certainly the main features such as roads, rivers, and forests. The instructions were in Dutch, the native language of all participants. A translated version is presented below:

First you may take place at the chair in front of the eye tracker. A map will be shown to you on the screen. Your aim should be to remember the map’s general structure, so you can draw it during the second part of the study. You don’t have to remember every detail; make sure you can account for such main features as the location of forests, rivers, roads, villages, railways, etc.

Once 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 bias due to (time) pressure, the participants could study the map at their own pace. When they had studied the map long enough, they were free to remove the map display by pushing a button. Participants were informed that the map would be displayed for a maximum of 10 minutes. This limit was imposed to keep the study manageable. A pilot study was carried out before the actual study in order to determine the different parameters in the study, including the time limit. In the pilot study, we found that participants needed on average 5 minutes to complete the task of committing the map to memory. We decided to double the map exposure time in order to avoid putting pressure on participants. (It should be noted that only a few participants needed the full 10 minutes for this task). During the map memorization task, the participants’ eye movements were recorded.

In order to force the participants to interpret the map contents, they had to execute a second task. Prior to the study, the participants were instructed that they would have to draw the map they had just seen, using a paper and a pencil. No time limit was set on the drawing task;

the participants just had to indicate when they were ready. Hence, they would need to use (retrieve) the stored information again. In order to be able to use this information later on, it has to be stored in (working and long-term) memory in the form of chunks of information (or schemata) which are linked to other information stored in the LTM. This requires that the objects are read, recognized, and interpreted (given meaning) (see above).

The process of “remembering the map – drawing the map” was repeated four times. After the completion of the fourth trial, participants were asked to fill out a questionnaire. This post-study questionnaire was used to obtain personal characteristics (expertise, age, gender, etc.) to verify their familiarity with the presented regions, and to receive feedback.

In this paper, we did not test users’ memory performance; this was done in Ooms et al. (forthcoming). We were interested in finding out where (and how) they looked at stimuli (or maps), namely how information is retrieved, how is it structured, and how much is retrieved. Thinking aloud, sketch maps and a questionnaire were used to study the information retrieval process. These findings confirm that the participants would have had interpreted the information on the maps.

### **Stimuli**

Four maps from the Belgian 1:10,000 topographic map series were displayed on screen during the user study. The selected maps were not crowded with information but some obvious structures were visible (roads, rivers, forests, etc.), and the region is not well known. The percentage of the map covered with large uniform areas such as forests and meadows was an important criterion. All selected maps 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, respectively, in map 1, map 4, map 2, and map 3).

Familiarity with a certain area influences the interpretation process and should be avoided. The participants all live in the northern part of Belgium (Flanders), but the selected maps cover regions located in the southern part. Therefore, it is unlikely that the participants would know the depicted regions by heart. The post-study questionnaire confirmed this.

The four maps were displayed in the same order to each participant. This fixed order was necessary to ensure that certain stimuli (map 1 and map 4) would not be depicted right after each other. Figure 2 shows five maps, even though only four were presented to the participants. This is due to a variation introduced with 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 others 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.

As can be seen in Figure 2, map 4 is a mirrored version of map 1; this time over the horizontal central axis. 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 into how familiarity, due to the mirrored map image, influences the map interpretation process. Mirroring of map images (e.g., map 1 vs. map 4 and the two versions of map 3) is done at random. Both map 1 and map 4 were shown to all users so that they may see the influence of (controlled) familiarity which is linked to both bottom-up and top-down processing. The content of map 3 was only depicted once (mainly bottom-up processing), ruling out the familiarity element. In both cases however, we can compare the users’ eye movements (e.g., scanpaths).

Finally, the second topographic map (map 2) 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 familiar with. It is thus important to know how users react during the interpretation process to such deviations. We chose to adapt the color for the village background and water bodies because these elements are present on all displayed maps. The map with the deviating color was used in the second trial, that is, after all participants (novices and experts) had already seen a map with a “normal” color scheme.

For those participants who may have been familiar with the color scheme of the 1:10,000 topographic map used in Belgium, deviations from the familiar color scheme could distract or confuse users and thus influence the interpretation process. It is for this reason that participants were asked in the post-study questionnaire to indicate their level of familiarity with Belgian topographic maps drawn at 1:10,000. Its results confirmed that most experts used such maps on a regular basis, whereas the novices did not, which could influence their reaction to deviations in the map design (color use).

### **Apparatus and recordings**

The participants’ eye movements were recorded using an EyeLink1000 eye tracking device from SR Research (Mississauga, Ontario, Canada) installed at the eye tracking laboratory of the Department of Experimental Psychology at 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.

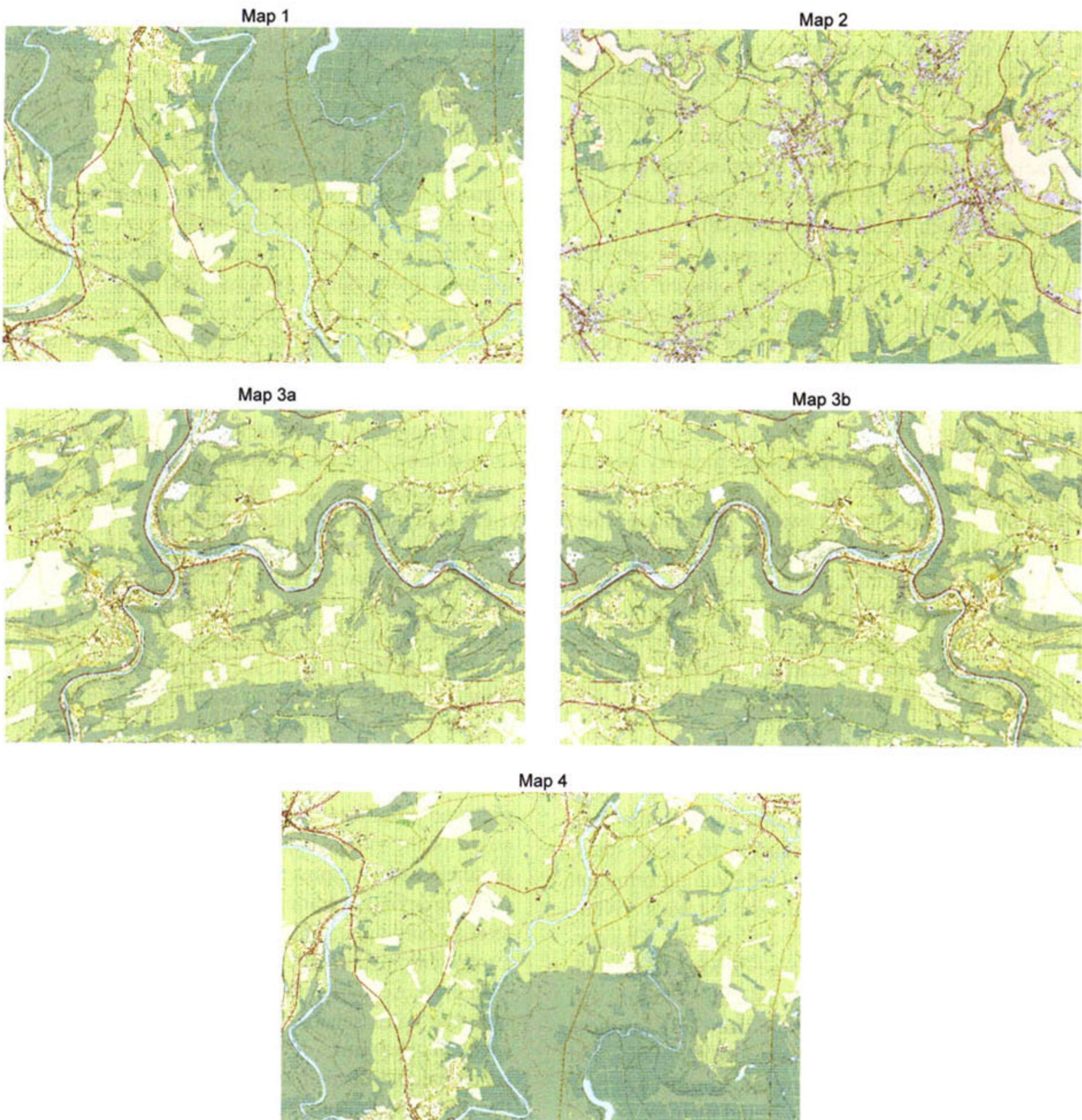


Figure 2. Stimuli depicted during the trials of the user study.

DataViewer software from SR Research was used to aggregate raw data into meaningful measurements, such as fixations and saccades. Fixations correspond to time periods when the POR is relatively stable. Because this is the time when the user is interpreting the visual information presented, these metrics are of utmost importance. The DataViewer has tools for reporting the number of fixations and the average duration of these fixations for each trial. Detailed information regarding fixation metrics can also 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 fixations and their total

duration within its boundaries are listed separately for each trial and participant.

### Methodology and results

#### *Statistical comparison: experts versus novices*

DataViewer can also be used to export a trial report. This report aggregates eye movement measurements per trial. The metrics of particular interest were the average duration of fixations and the number of fixations per second. The former can reveal difficulty with which the visual stimulus is processed (e.g., Duchowski 2007; Goldberg

et al. 2002; Holmqvist et al. 2011; Jacob and Karn 2003; Poole and Ball 2006; Rayner 1998). Complex or chaotic stimuli (which may be difficult to process due to a rise in the cognitive load) typically result in longer fixation durations. If a user finds a part of the visual stimulus particularly interesting, the duration of the fixations usually increases when the observer finds a particularly interesting visual stimulus. The number of fixations a user can have per second is closely linked to the average fixation duration for that user. Longer fixation durations result in fewer fixations per second. A study of both metrics can be useful in explaining the results.

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, for both expert and novice study participants. The third column gives the results of a one-way ANOVA for the two user groups. The last column provides more information regarding the effect size of the ANOVA test: Cohen's  $d$ . "Medium and large" effects indicate that a sufficient large sample size has been tested. The ANOVA tests carried out in this study show that the experts had significantly shorter fixations than the novice participants. Furthermore, experts can have significantly more fixations per second. These findings are in line with the results described by Ooms et al. (2012a) who analyzed eye movement metrics resulting from a visual search on a very basic map design. The results obtained in this study confirm 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 map on the screen. The last row in Table 1 indicates that experts chose to study the map for a

longer period of time than novice users did (335.7 seconds or 5.6 minutes vs. 205.7 seconds or 3.4 minutes).

### 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 comprised of three colors is used: green (areas with lower fixation intensities), yellow (areas with higher fixation intensities), and 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.

Heatmaps not only provide a good initial overview of eye movement data, but they also have a number of serious drawbacks. First, it is very difficult to compare heatmaps objectively; nowadays this is often done just at sight (qualitative analysis). Second, in most software it is not possible to adapt the classification system. Depending on the topic under investigation, the focus of the visualization might be on the general pattern of fixation intensities or on extreme values. Different classification schemes are thus highly desirable but, since adaptations of the standard classification scheme are often not possible, this continues to be a problem. Third, heatmaps are not suitable for detecting extreme values between, for example,

Table 1. 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  |

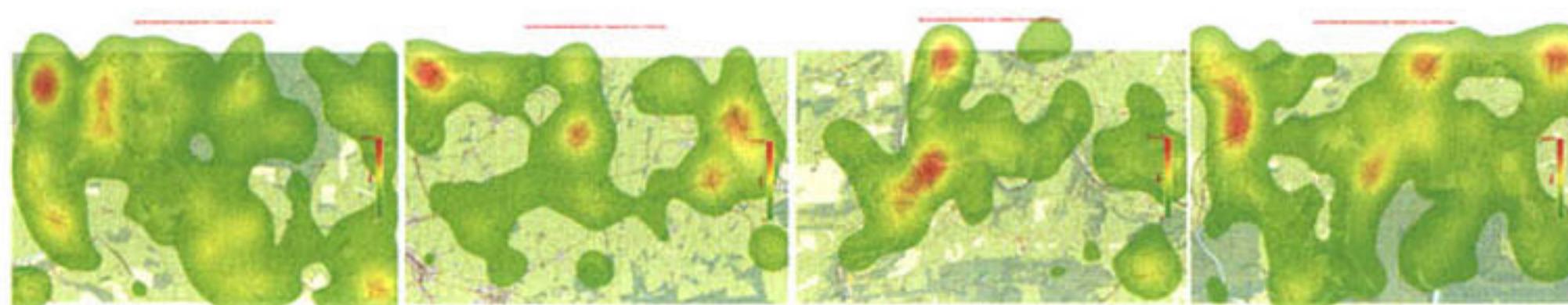


Figure 3. Heatmaps from the same participant for the four trials.

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 6.211, 11.444, 15.445, and 13.048 seconds. Although the maximum value of the first heatmap is about half of those of the other heatmaps, the same color scheme is applied (see Figure 3). As a result, it is impossible to determine which fixation intensity is lower or higher. In our case, looking at the heatmaps in Figure 3 does not tell us that the fixation intensity on map 1 is much lower. Because of these drawbacks, an alternative to the heatmap visualization is proposed in the next section – the gridded visualization.

## *Gridded visualization: methodology*

A similar approach to visually analyze eye movement data was described in Brodersen, Andersen, and Weber (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 such that detailed information could be obtained (small enough), taking into account the accuracy of the eye tracker. A maximum acceptable deviation of  $0.5^\circ$  on the calibration results in a deviation of 0.6 cm on the screen (at a viewing distance of 70 cm). Therefore, the cell sizes should preferably be no less than 1.2 cm (or about 34.9 pixels). Based on the

size of the map image ( $1280 \times 800$  pixels), it was decided to use square AOIs measuring  $40 \times 40$  pixels.

This means that a grid of  $20 \times 32$  cells is placed over the map image, resulting in 640 AOIs. The DataViewer software by 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 AOI report's structure is 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 and the four trials they participated in.

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 metrics to be able to detect deviations in the participant's behavior, which could be linked to the scanpaths (e.g., Duchowski 2007; Holmqvist et al. 2011).

However, as mentioned before, a significant difference in the duration of the trials was observed between the expert and novice users. As a consequence, these absolute values are not comparable between the 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 the fixations. In order to be able to compare these measurements objectively, normalized values linked to a uniform trial duration were used. The mean trial duration of the experts was 335.7 seconds

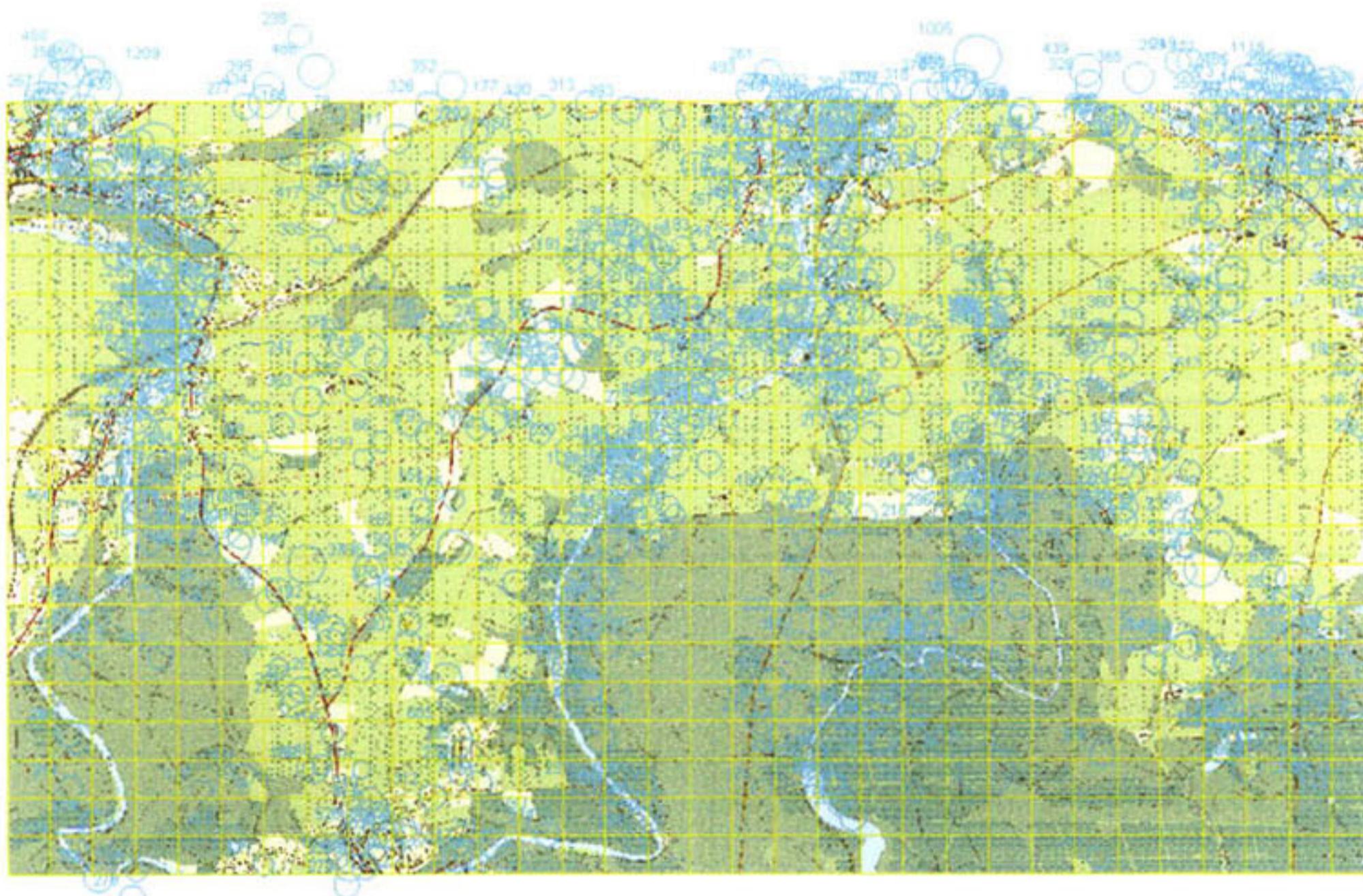


Figure 4. Grid of AOIs (yellow) placed over each map image to visually analyze the users' fixations (cyan).

compared to that of 205.7 seconds for the novice users. Therefore, it was decided to use a uniform trial duration of 300 seconds (or 5 minutes). All data – total fixation count and total dwell time – were recalculated based on the initial and the 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 the adapted AOI-reports and restructure the data to obtain a grid of 32 by 20 cells. Based on the (x, y)-position of the corresponding fixations, all values were, placed on their correct (spatial) position in the original grid for each map. The 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 were aggregated for each stimulus, separately for each user group (experts vs. novices), and for each dependent variable (fixation count and dwell time). To get an idea 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 32 grids: four maps, two user groups (experts vs. novices), two variables (fixation count and dwell time), and two aggregation types (average and maximum).

Finally, the aggregated values were grouped into eight different classes. A grayscale color was assigned to each class, based on ColorBrewer, an online tool of usable color schemes for maps (Brewer 2012). The addition of this visual component facilitated the interpretation of the grids. The classification of the values was chosen as such that patterns in the fixations' distributions could be detected, including extreme values. The classification and color scheme applied to fixation count and dwell time are presented in Table 2.

The *FixDur* in this table corresponds to the fixation durations summed over the whole trial, recalculated to the uniform trial duration of 300 seconds. This total 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 could be recalculated based on this average. Linking the classification of the fixation durations to that of the fixation counts makes it possible to detect regions where users are staring. In this latter case, the classification of the dwell time is higher than the classification of the fixation count in the corresponding cell. The resulting grids are discussed in detail in the next sections.

#### **Gridded visualization: total fixation count**

The aggregated gridded visualizations are depicted in Figure 5 for the average values and in Figure 6 for the maximum values. A similar pattern in the fixation counts (both for the average and the maximum values) was noticed between both user groups. This fixation pattern reflects the structure of each stimulus. In the grid that corresponds to map 1, two vertical, linear clusters with a higher fixation count can be identified. They correspond to the two leftmost rivers on this map. Both the expert and the novice map users focused on these linear structures. The cluster of fixations resulted in a higher fixation count, indicating that participants tried to remember a reference frame in which other map elements could be placed. The focus on these linear elements is stronger for the experts than it is for the novices. Another element of interest (mostly to the expert group) was a village flanked by a major road that is in the lower right corner of the map.

Map 4 in the experiment is a mirrored version of map 1. The structure in the fixation pattern for the fourth map could thus be the mirrored equivalent of the structure for map 1. In map 4, a similar, mirrored pattern of the two rivers on the left is indeed visible, as is that of the village in the (now) upper right corner. In map 4 (Figure 6), the grid for the novice group has a darker background than that for the experts. This is due to a high maximum fixation count for this group over the entire map, while the experts seemed to focus more on particular items. However, this pattern cannot be derived from a grid with the average values (Figure 5). A statistical comparison of

Table 2. Classification and color scheme for fixation visualization.

| Variable                         | Classification and color schemes |               |               |               |               |               |               |             |
|----------------------------------|----------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|-------------|
| <i>FixCount</i>                  | [0–1]                            | [1–2]         | [2–4]         | [4–6]         | [6–8]         | [8–10]        | [10–20]       | [20–...]    |
| <i>FixDur</i><br>(dwell<br>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–...] |
| <i>Color</i><br>(RGB)            | 255                              | 247           | 217           | 189           | 150           | 99            | 37            | 0           |

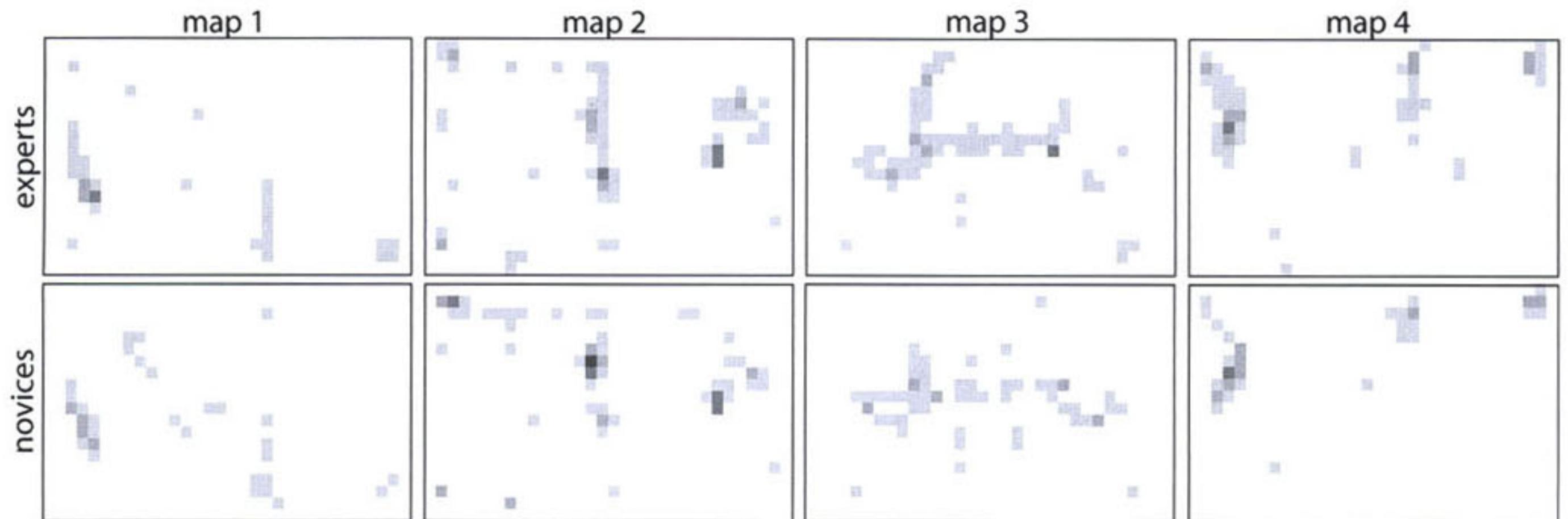


Figure 5. Grid of AOIs depicting the average fixation count for expert (top row) and novice (bottom row) users for each stimuli (left to right).

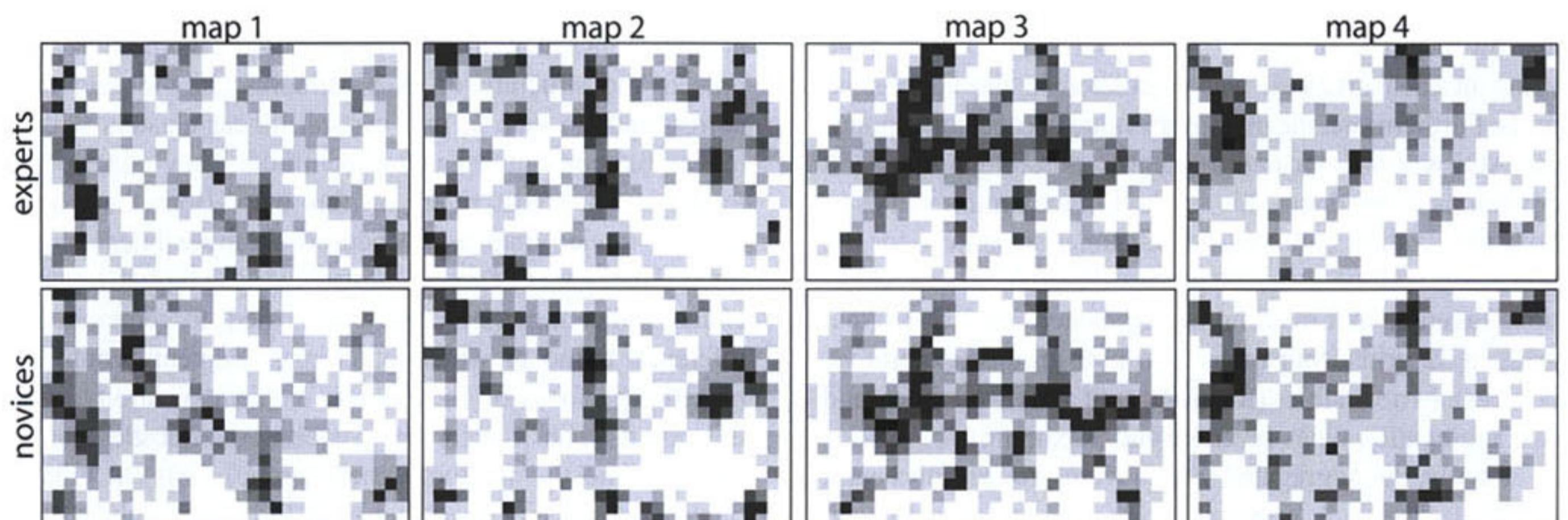


Figure 6. Grid of AOIs depicting the maximum fixation count for expert (top row) and novice (bottom row) users for each stimuli (left to right).

fixation counts for map 1 and map 4 is discussed in a following section.

In the grids related to map 2, there is a vertical linear structure in the middle of the grid. This structure corresponds to the location of a major road on the map. The higher fixation count for the road means that study participants focused on this road the most during the trial, perhaps because they wanted to remember especially this road as a reference frame. This linear structure is more obvious in the grids for the expert group than in those for 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 map readers' attention, resulting in a higher fixation count. The deviating color of 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 while the other half saw a mirrored version (over its central vertical axis). The measurements of these two stimuli were aggregated (average or maximum) and presented in one grid. A fixation pattern results that is similar (but mirrored) on the left and right side of the map. The fixation count of the expert group clearly reflects the linear structure of the vertical and horizontal road/river combination on the map. This structure is also present in the grids of the novices, but it is less pronounced. Both the experts and novices seem to have a higher fixation count on the left side of the map, although only half of the participants of each group saw the mirrored version. This pattern with a greater focus on the right is again more pronounced in the expert group, particularly with regard to the maximum

fixation count. This could indicate that (all) map users tend to fixate more on the left side of the map.

When averaging the fixation count for each of the four map quadrants (upper left, upper right, lower left, and lower right) instead of for each AOI, the upper part of each map always has the highest fixation count. Furthermore, the number of fixations on the left side of each map is higher than on the right side. More detailed information on how users looked at the map over time can be gleaned by studying the evolution of their scanpaths. This is discussed below, under scanpath visualization.

### **Gridded visualization: total dwell time**

Besides the number of fixations at a certain location, the total dwell time might provide important insights into the users' cognitive processes taking place during map interpretation. Longer fixation duration might indicate that the user finds a certain region of the visual stimulus particularly interesting. But longer fixation duration might also indicate difficulty with interpreting the content. When the user has difficulty recognizing an object, cognitive load increases, which in turn is linked to longer fixation duration (e.g., Duchowski 2007; Goldberg et al. 2002; Holmqvist et al. 2011; Jacob and Karn 2003; Poole and Ball 2006; Rayner 1998). Two such grids constructed for 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. A comparison of map 2 in Figure 5 (fixation count) and Figure 7 (dwell time) confirms that the classification distribution of fixation counts corresponds to the classification of dwell time. This indicates that there is a strong relationship between 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 two sections.

### **3D gridded visualization: dwell time**

With 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 whose height corresponds to the value in that cell. In this case, these values correspond to the total dwell time in that cell. The 3D gridded visualizations of average dwell time are depicted in Figure 8 for maps 1 and 4 and in Figure 9 for maps 2 and 3. Because the data is not classified, this visualization makes it possible to compare the dwell times (summed fixation durations) between experts and novices in more detail. The downside of this approach is that the spatial distribution of the values on the grid is not all that clear. The perspective view and the fact that the higher bars are in front of the image conceal the lower bars. However, 3D graphs are useful when

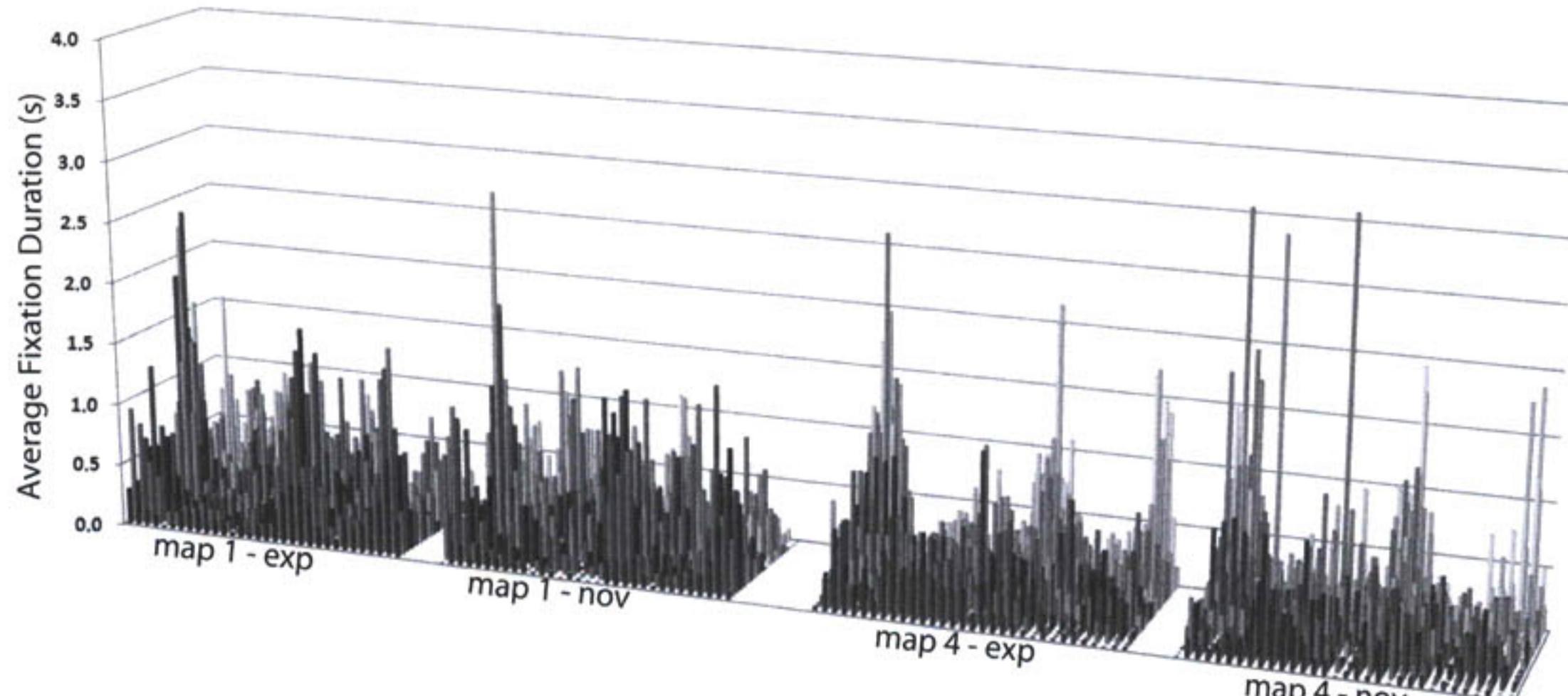


Figure 8. 3D representation of the average dwell time for map 1 and map 4.

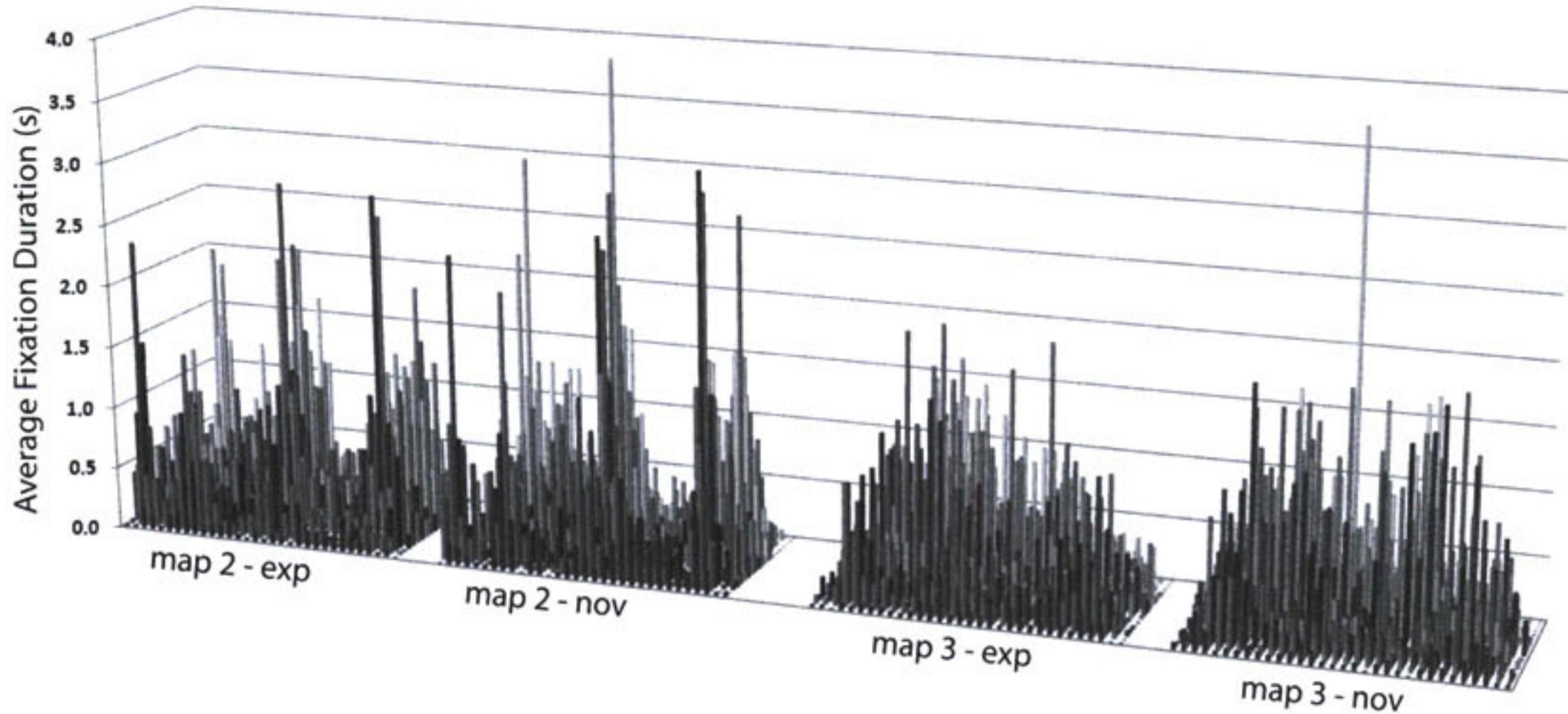


Figure 9. 3D representation of the average dwell time for map 2 and map 3.

studying differences in the main pattern of values and extreme values between two user groups, without considering their precise spatial location. The (normal) gridded visualization is less suitable for obtaining detailed insight into 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 represents a 3D gridded visualization of the average dwell time for maps 1 and 4. The graphs for the two user groups are placed together to allow better comparison. Similarly, map 1 and map 4 are in the same figure to facilitate a comparison of their corresponding (mirrored) values. The results of the remaining stimuli (map 2 and map 3) are depicted in Figure 9. Both figures show that the dwell times in the depicted AOIs are very similar between the two user groups, which is consistent with the results obtained by 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, which 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 the middle of map 2 represent 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 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 a higher total dwell time. The attention of all participants was attracted by these "strange objects", but the novice group had higher values. Surprisingly, the deviating background color of the villages (light purple instead of light yellow) did 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 the 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 attracted by something in that region, or it can also indicate regions that are difficult to interpret by the user, resulting in longer fixations but not necessarily a higher

count. To distinguish between these two options, an additional eye movement metric was studied: the average fixation duration of a single fixation, which can then be compared to the map's content at that location using the (3D) gridded visualization.

### **3D gridded visualization: fixation duration**

The average fixation duration (of a single fixation) was already statistically analyzed. As mentioned before, these statistical analyses miss the spatial dimension that is inherently linked to maps and their design. The difficulty with which a user interprets the visual content 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., Duchowski 2007; Goldberg et al. 2002; Holmqvist et al. 2011; 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 content at that location.

Similar to the average dwell time, a 3D gridded visualization was created in which each bar height corresponds to the average fixation duration (of a single fixation) at that location. These graphs are depicted in Figures 10 and 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 content of a map (for all maps), causing a higher cognitive load (e.g., Bunch and Lloyd 2006; Harrower 2007; Holmqvist et al. 2011; Jacob and Karn 2003; Poole and Ball 2006; Rayner 1998).

Although map 4 is the mirrored version of map 1, the difference between experts and novices was much more pronounced in map 4 (see Figure 10). The expert group

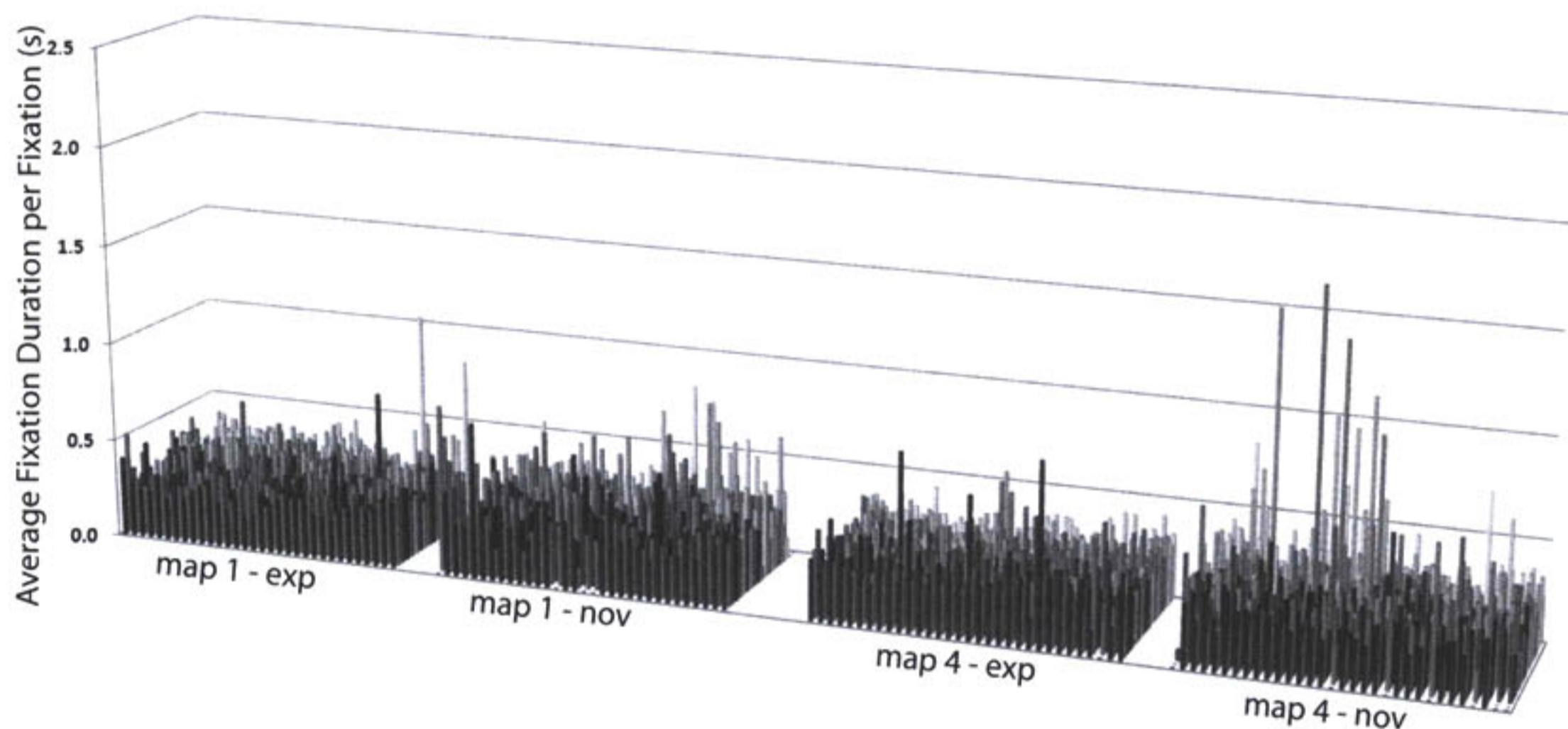


Figure 10. 3D representation of the average fixation duration per fixation for the expert users.

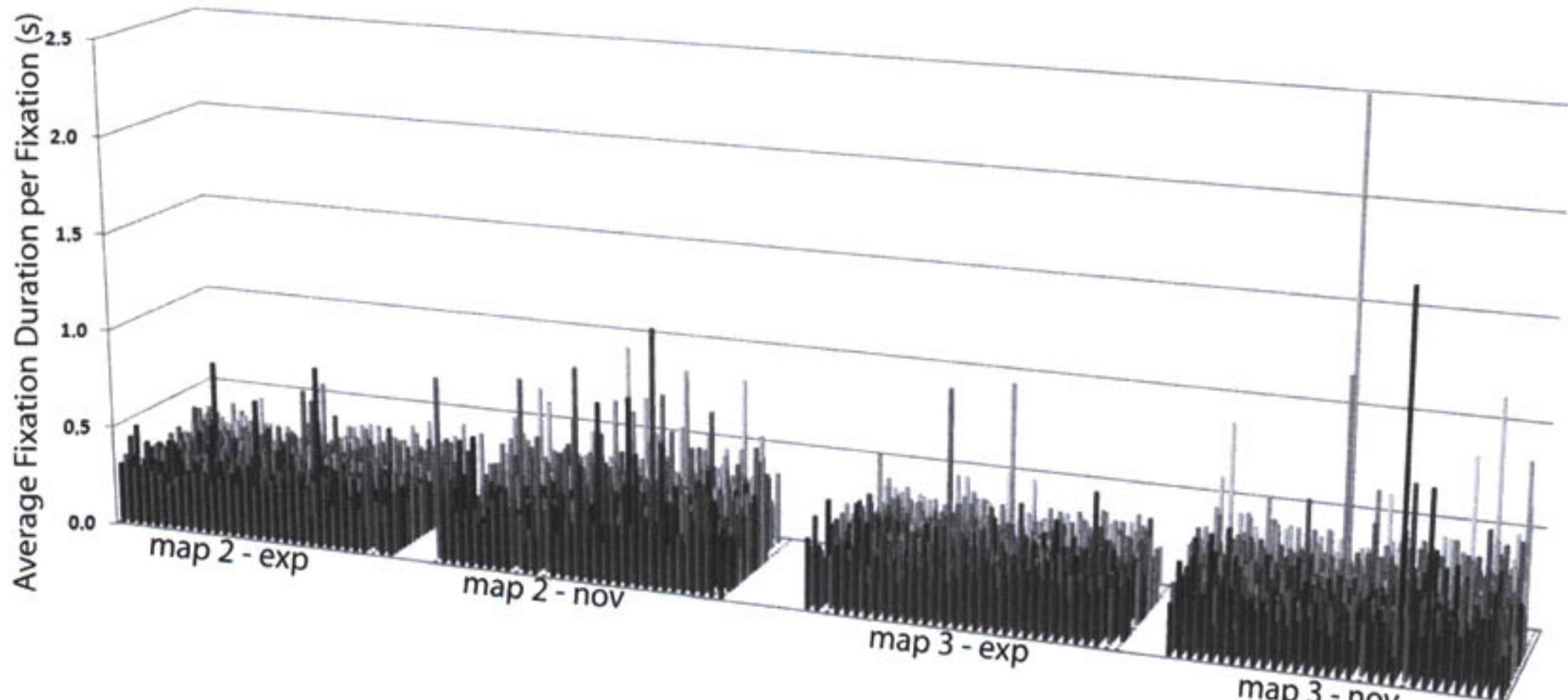


Figure 11. 3D representation of the average fixation duration per fixation for the novice users.

had a number of fixation durations which were longer than normal, but the novices had 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 were looking when the map was displayed. It can thus be concluded that the novice users had longer fixation durations when the map was first displayed, which indicates confusion. This confusion might be explained by the recognition of map 1 that is, however, 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 in the novices’ graph, they are distributed over the entire map image. No particular deviations were noticed at the location of water bodies (depicted in a deviating color).

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

#### **Statistical grids: mirrored maps**

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

Previously, we concluded that the distribution of the total fixation count reflected the general structure of the

map. Important and engaging items drew attention more often than other items. The 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 other’s mirrored equivalents. 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).

The statistical grid 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 many significant and highly significant differences in the fixation counts of the corresponding cells in the grid. This amount of significant differences 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 versions overlap, resulting in a lighter line in the statistical grids: not significantly or nearly significantly different. The  $P$ -values show a clearly similar mirrored pattern on both sides of the axis (above vs. below for map 1 and map 4; left vs. right for map 3).

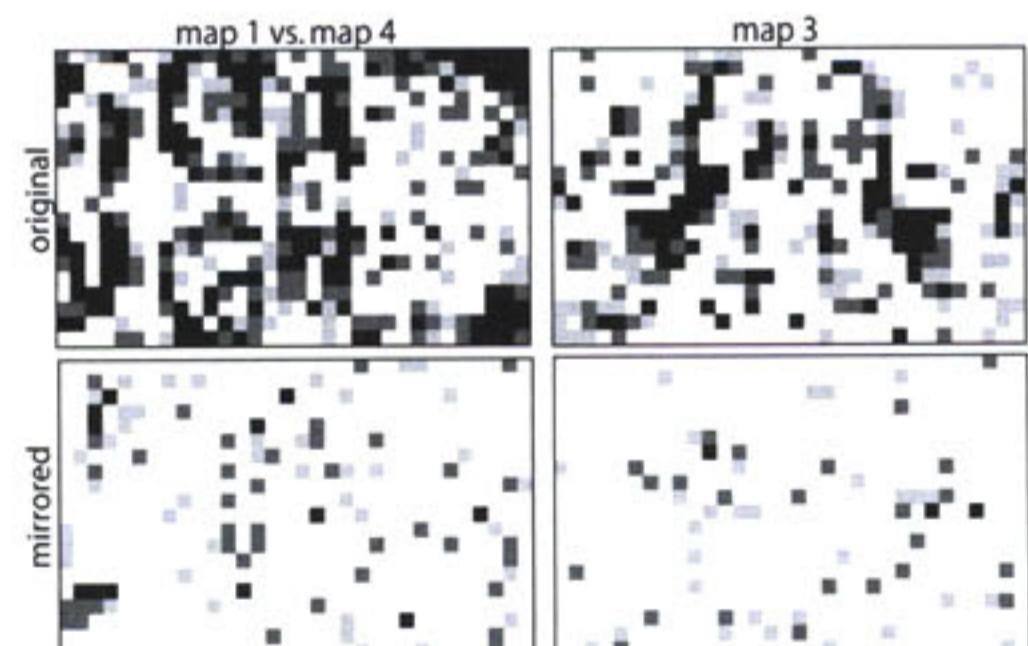


Figure 12. Statistical comparison of the fixation count between two mirrored maps.

Table 3. Classification and color scheme for the ANOVA results.

| Variable           | Classification and color schemes |                  |             |                    |
|--------------------|----------------------------------|------------------|-------------|--------------------|
| <i>Sign. (P)</i>   | > 0.1                            | [0.1–0.05]       | [0.05–0.01] | <0.01              |
| <i>Description</i> | not significant                  | near significant | significant | highly significant |
| <i>Color (RGB)</i> | 255                              | 217              | 150         | 37                 |

The lower left grid in Figure 12 shows the statistical comparison between map 1 and the mirrored version of the grids for map 4. When the grids for map 4 are mirrored over their horizontal axis, their pattern would be expected to reflect the structure of map 1. The statistical grid in Figure 12 confirms this. 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.

A comparison of 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) shows that there are no highly significant differences. The better result obtained for 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. Only half of the participants saw the original version of map 3, the others saw the mirrored version. This avoids influences on the cognitive processes due to recognition or confusion. The statistical grids at the bottom of Figure 12 indicate that the patterns of the users' fixations are guided primarily by the main (linear) structures on the map.

### 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 of the trial) were depicted – 0 to 10 seconds; 0 to 30 seconds; 30 to 60 seconds.

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 to these main (often linear) elements. During the second half of the first minute, the participants also fixate other map objects, but the structuring elements still receive a lot of attention.

When comparing the scanpaths during the first 10 seconds for maps 1 and 4, a striking difference is observed. In map 1, the participants direct their attention to the main structuring elements in the map. This holds true especially for the participants in the expert group. However, the scanpaths associated with map 4 are rather chaotic during the first 10 seconds. The structure of the two leftmost vertical lines is not present in map 1. There is a high number of horizontal scanpath lines zigzagging across the image during the first 30 seconds of the map's display, especially for the novice map users. These chaotic scanpaths likely indicate that the users were confused by the mirrored map

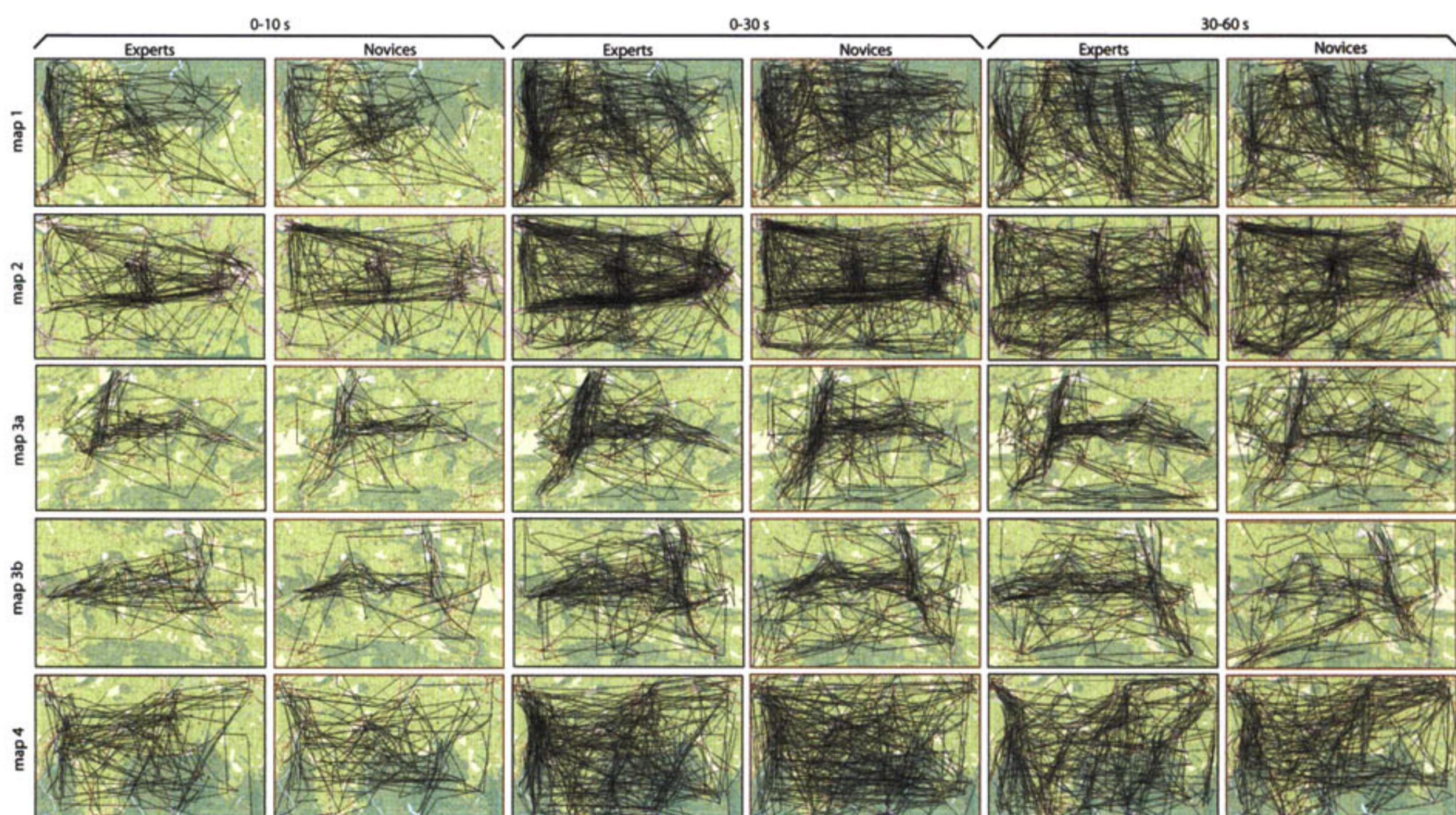


Figure 13. Evolution of the participants' scanpaths for the different maps, and user groups.

image of map 4 (Çöltekin, Fabrikant, and Lacayo 2010; Fisher, Monty, and Senders 1981; Stark and Ellis 1981).

The experts' scanpaths on map 2 show a cluster, both horizontally and vertically, on the major road on the map during the first 10 seconds. The novices seem to be more distracted by the water bodies depicted with the deviating color. In the first 30 seconds, the expert users focus more on these water bodies, but they do so 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 used to depict the background of the villages does not seem to influence the participants' attentive behavior.

The scanpaths for maps 3a and 3b are dissimilar during the first 10 seconds, although the same structures are found in the map image. This is especially noticeable for 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 zeroes in 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 visible in a longer time interval and corresponds to the findings of the fixation counts and fixation durations presented in this paper. The experts, in particular, tend to be more attracted to the left side of the map.

## Discussion and conclusion

The study described in this paper is an extension of the work done by Ooms et al. (2012a) and aims to verify whether their results could be generalized to wider array, and thus more complex, map types. The main aim of the experiment was to gain a better understanding of the way expert and novice map users process and interpret the (complex) visual information on maps. Deviations in the stimuli were introduced in order to study their effects on the movement of map users' eyes, 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 those by novices. This indicates that their interpretation process, including the different stages of object recognition, is much faster than that of the novice users. 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 may explain this phenomenon. Shorter fixations leads to more fixations per second, and since this is true for the experts, they interpret a larger part of the map in the same amount of time than the novices. It can thus be concluded that expert map users can interpret maps (be they simple or complex) more efficiently.

In order to spatially analyze the results, different approaches are presented, all of which complement each other: gridded visualizations, 3D gridded visualizations,

statistical grids, and scanpath analyses. These visual and (mainly) qualitative methods shed light on different aspects of the users' interpretation process and made it possible to study differences in the attentive behavior of expert and novice users. A number of eye movement metrics related to the users' fixations were analyzed and compared, including average fixation count in one trial, average dwell time in one trial, and average fixation duration (of a single fixation). These analyses took the spatial distribution of the fixations across the map image into account, and were 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. The main structure of the map is thus 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 the general structuring elements, just to remember the map as well as possible. The users' very intent focus on these elements is an interesting finding, especially when comparing differences in attentive behavior between experts and novices. The focus on these linear structures is more pronounced in the grids for the expert group. The visualization of the users' scanpaths during the first 10 and 30 seconds also shows that user's attention is immediately directed toward the structuring elements. Nevertheless, the novices' eye movement measurements show more extreme values than those for the experts in respect of the number and duration of the fixations.

The deviating color use for the water bodies in map 2 influences the users' attentive behavior. Both user groups were attracted by these objects, as the eye movements show. 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 than in the expert group. Experts focus more on the central horizontal and vertical reference frame (major roads) on the map image. The deviating color use of the villages' background did not seem to have any influence on the users' attentive behavior or their interpretation process. In the gridded visualizations, the cell that covered the villages did not contain more or longer fixations in comparison to the other maps 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) were found on the left side of the superimposed result. This indicates that the

users tended to fixate more on the left side of the map, regardless of the contents of the map. A quadrant analysis for all maps indicates that this holds true for all other stimuli presented during the study. This observation of higher fixations on the left side is more distinct in the measurements of the expert group. Nevertheless, based on the recorded 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 to those of the original map and, therefore, 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 for map 4 corresponds to those for map 1. Nevertheless, the statistical grid showed more significant differences in the corresponding cells of map 1 and map 4 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 were detected. However, the 3D gridded visualization of the dwell times shows a number of extreme values. These values were explained in the 3D graphs representing the average fixation durations, where a number of peaks are observed. This confusion can also be seen in the structures of the scanpaths. During the first 30 seconds, eye movements were not immediately directed toward the structuring elements (as was shown in the case for map 1). The scanpaths show a more chaotic distribution over the map image.

It can thus be concluded that the eye movements of both user groups show a similar pattern; they reflect the general structuring elements on the map. This was confirmed by comparing the eye movement patterns for the original map with those of its mirrored version. However, a tendency to have more fixations on the left side of the map was noticed. The attraction of the users' eye movements or attention to the major structuring elements on the map can be influenced by striking deviations in the map image. Nevertheless, a number of substantial differences were noticed between the two user groups, indicating that the expert users can process 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. Their eye movement measurements have less extreme values, which indicates 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., and P. Cavanagh. 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., and R. M. Shiffrin. 1968. "Human Memory: A Proposed System and Its Control Processes." In *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 2, edited by K. W. Spence and J. T. Spence, 89–105. New York: Academic Press.
- Baddeley, A. D. 1999. *Essentials of Human Memory*. Hove: Psychology Press.
- Bertin, J. 1967. *La Sémiologie Graphique*. Paris: Mouton-Gauthier-Villars.
- Brewer, C. A. 2012. *Online Tool*. Accessed February 15. <http://ColorBrewer.org>
- Brodersen, L., J. H. K. Andersen, and S. Weber. 2001. "Applying the Eye-Movement Tracking for the Study of Map Perception and Map Design." In *Kort og Matrikelstyrelsen*, 98. Copenhagen: National Survey and Cadastre.
- Bunch, R. L., and R. E. Lloyd. 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., and J. R. Eastman. 1984. "Eye-Movement Parameters and Perceived Map Complexity I." *American Cartographer* 11 (2): 107–117.
- Castner, H. W., and J. R. Eastman. 1985. "Eye-Movement Parameters and Perceived Map Complexity II." *American Cartographer* 12 (1): 29–40.
- Çöltekin, A., S. I. Fabrikant, and M. Lacayo. 2010. "Exploring the Efficiency of Users' Visual Analytics Strategies Based on Sequence Analysis of Eye Movement Recordings." *International Journal of Geographical Information Science* 24 (10): 1559–1575.
- Çöltekin, A., B. Heil, S. Garlandini, and S. I. Fabrikant. 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.
- 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*. London: Springer.
- Fabrikant, S. I., and A. Lobben. 2009. "Introduction: Cognitive Issues in Geographic Information Visualization." *Cartographica* 44 (3): 139–143.

- Fabrikant, S. I., S. Rebich-Hespanha, N. Andrienko, G. Andrienko, and D. R. Montello. 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., and Senders, J. W., eds. 1981. *Eye Movements: Cognition and Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fits, P. M., R. E. Jones, and J. L. Milton. 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., M. J. Stimson, M. Lewenstein, N. Scott, and A. M. Wichansky. 2002. "Eye Tracking in Web Search Tasks: Design Implications." *Proceedings of the Eye Tracking Research and Applications Symposium*, New York, March 25–27, 2002.
- Harrower, M. 2007. "The Cognitive Limits of Animated Maps." *Cartographica* 42 (4): 349–357.
- Hegarty, M., M. Canham, and S. I. Fabrikant. 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., M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van De Weijer. 2011. *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford: Oxford University Press.
- Jacob, R., and K. Karn. 2003. "Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises." In *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, edited by R. Radach, J. Hyona, and H. Deubel, 573–605. Amsterdam: Elsevier.
- 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*. New York: Guilford Press.
- Matlin, M. W. 2002. *Cognition*. 5th ed. Hoboken, NJ: Wiley & Sons.
- 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*. San Francisco, CA: Morgan Kaufmann.
- Olson, J. M. 1976. "A Coordinated Approach to Map Communication Improvement." *American Cartographer* 3 (2): 151–159.
- Ooms, K., G. Andrienko, N. Andrienko, P. De Maeyer, and V. Fack. 2012. "Analysing the Spatial Dimension of Eye Movement Data Using a Visual Analytic Approach." *Expert Systems with Applications* 39 (11): 1324–1332.
- Ooms, K., P. De Maeyer, V. Fack, E. Van Assche, and F. Witlox. 2012a. "Interpreting Maps Through the Eye of Expert and Novice Users." *International Journal of Geographical Information Science* 26 (10): 1773–1788.
- Ooms, K., P. De Maeyer, V. Fack, E. Van Assche, and F. Witlox. 2012b. "Investigation the Effectiveness of an Efficient Label Placement Method Using Eye Movement Data." *The Cartographic Journal* 49 (3): 234–246.
- Ooms, K., P. De Maeyer, V. Fack, E. Van Assche, and F. Witlox. Forthcoming 2013. "Listen to the Map User: Cognition, Memory, and Expertise." *The Cartographic Journal*.
- Peterson, M. P. 2003. "Maps and the Internet: An Introduction." In *Maps and the Internet*, edited by M. P. Peterson, 1–16. Oxford: Elsevier Science.
- Poole, A., and L. J. Ball. 2006. "Eye Tracking in Human Computer Interaction and Usability Research: Current Status and Future Prospects". In *Encyclopedia of Human Computer Interaction*, edited by C. Ghaoui, 211–219. Hershey, PA: Idea Group.
- Rayner, K. 1998. "Eye Movement in Reading and Information Processing: 20 Years of Research." *Psychological Bulletin* 124 (3): 372–422.
- Rubin, J., and D. Chisnell. 2008. *Handbook of Usability Testing. How to Plan, Design and Conduct Effective Tests*. 2nd ed. Indianapolis, IN: Wiley.
- Slocum, T. A., C. Blok, B. Jiang, A. Koussoulakou, D. R. Montello, S. Fuhrman, and N. R. Hedley. 2001. "Cognitive and Usability Issues in Geovisualisation." *Cartography and Geographic Information Science* 28 (1): 61–75.
- Stark, L. W., and S. R. Ellis. 1981. "Scanpaths Revisited: Cognitive Models Direct Active Looking." In *Eye Movements: Cognition and Visual Perception*, edited by C. Fisher, R. Monty, and J. Senders, 193–226. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Steinke, T. R. 1987. "Eye Movement Studies in Cartography and Related Fields." *Cartographica* 24 (2): 40–73.