

Eye-tracking analysis of the influence of relief shading on finding labels on tourist maps

STANISLAV POPELKA

Palacký University Olomouc, Faculty of Science, Department of Geoinformatics, Czechia, e-mail: stanislav.popelka@upol.cz

ABSTRACT This paper evaluates the influence of relief shading in finding labels on tourist maps. Two types of maps were compared, one in which terrain was depicted with contour lines and spot heights, and the other enhanced with relief shading. The task was to find specific hills and villages. Two aspects were investigated – whether shading helped users find hills, and whether shading made it more difficult to find villages. The eye-tracking method was used for this study. The results indicate that respondents prefer shaded maps from an aesthetic point of view. Pair-wise comparison of individual stimuli pairs and groups of stimuli was performed with the use of five eye-tracking metrics. Most of the eye-tracking metrics were significantly different for most of the stimuli. The results of the experiment show that shaded maps are less suitable for finding hills and villages. The least effective result was observed in finding villages on a shaded map.

KEY WORDS cartography – maps – eye-tracking – relief shading – visual perception – map reading

POPELKA, S. (2018): Eye-tracking analysis of the influence of relief shading on finding labels on tourist maps. *Geografie*, 123, 3, 353–378.

Received November 2017, accepted June 2018.

1. State of the art

Contouring is the most frequently used technique to provide map readers with elevation information (Dušek, Miřijovský 2009). Contours are isarithms – lines connecting points of equal elevation. Most casual map readers cannot imagine landforms indicated by contour lines. As Castner and Wheate (1979, p. 78) stated: “Relief depiction with contours is not usually immediately interpretable or imaginable, especially for inexperienced users.” Instead, most people recognize shapes primarily by the interplay of light and dark. This method of portraying the land-surface form is called shading (Robinson et al. 1995).

Shaded reliefs are already found in early manuscript maps from the seventeenth century. With the invention of lithography (1798), it became possible to print half-tones. Cartographers began to combine relief shading with other means of terrain display. After World War II, many reliefs were drawn by airbrush. Since the end of the twentieth century, relief shading has mostly been generated from digital elevation models (Jenny, Råber 2015).

The angle of lighting must be defined when using shading in maps. Conventional lighting comes from the upper left corner of the map – from the northwest in maps of the northern hemisphere (Bernabé-Poveda, Sánchez-Ortega, Çöltekin 2011). A recent publication (Biland, Çöltekin 2016) suggested that NNW illumination is better than NW. Imhof (2007) states that this may be due to people writing from left to right, light being on the left hand side when holding a pen while the right being in shadow. Even though this situation cannot occur in reality (in the northern hemisphere), this light is psychologically the most effective in perceiving terrain plasticity (Imhof 2007). Most users are accustomed to the light from the northwest, therefore a map illuminated from the south will be perceived negatively (Imhof 2007). This was demonstrated with the use of eye-tracking in the author’s dissertation (Popelka 2015).

The main aim of relief shading in maps is to provide information about height, but it also has an aesthetic function. Ortag (2009) published a chapter focusing on the variables of aesthetics in maps. In this chapter, the results of focused interviews with more than 150 participants were summarized. Relief or 3D impression was mentioned as a the third most important reason for describing a map as beautiful (after color and readability).

Usability evaluation of visual representations have gained much attention in recent cartographic and visual analytics research (Coltekin et al. 2009, Fabrikant et al. 2008). “Usability evaluation allows us to obtain data, often quantitative, about aspects of a system or user performance with that system which can be used to identify aspects that are problematic for users and highlight potential fixes. These methods can also be used for comparative purposes, for example, against established benchmarks or alternative designs or products in order to identify

which is easier to use or identify their relative advantages and disadvantages.” (Fuhrmann et al. 2005, p. 559)

The evaluation of shading on tourist maps is important because it is not known whether shading helps map readers or not. Both map types are available on the market – shaded and non-shaded – but the study of usability of these types of relief visualization has not been previously performed. The motivation for this case study was to learn whether shading affects search performance when looking for a specific object on a tourist map. The research questions that guided this study focus on the basic search tasks that can be done with tourist maps. The first question investigates the search for relief-related objects (hills), the second, the search for non-relief-related objects (villages).

The research questions are:

1. Does shading help participants find hills, as participants need only scan the darkest areas (representing hills), not the whole map?
2. Is searching for villages slower on shaded maps, as these maps are darker by design than maps without shading and may negatively impact its legibility?

Maps that include a comprehensive representation of terrain together with a landform relief (topographic or tourist maps) have not been studied from a user perspective as deeply as other types of geovisualizations (i.e., city maps or urban plans; Burian, Štávková 2009), although their importance remains high in many common tasks, necessitating a high-level understanding of terrain (Putto et al. 2014).

In the 1970s, experiments with terrain visualization methods concentrated on legibility studies of different methods of representing topographic information (Chang, Antes, Lenzen 1985). Phillips, Lucia, and Skelton (1975) performed a questionnaire study testing four different types of relief maps (contours, contours with hill shading, layer tints, and spot height maps). In most of the questions, statistically significant differences were found, but no single map type was the best for all 13 map reading questions. Visualization using contour lines with hill shading was an advantage only in questions requiring visualization of the landscape (e.g., visibility, finding the steepest slope). This result corroborates the study of DeLucia (1972), who found a significant increase in time necessary when extracting required information from a map with hill shading. Similar visualization techniques were investigated in the study of Potash, Farrell, and Jeffrey (1978), who analyzed contour maps and contour maps supplemented by layer tints and shading. The results of the study showed that layer tints increased reading speed, whereas shading did not and caused a decrease in accuracy. In the study of Castner and Wheate (1979), contour maps and shaded relief were analyzed. The results showed that in tasks where a search target was associated with a topographic situation, shaded relief was an advantage.

In a more recent study, Petrovič and Mašera (2004) created a questionnaire to find out how different 3D cartographic representations of terrain could fulfil a user's needs. Savage, Wiebe, and Devine (2004) compared performance between 2D and 3D topographic representations in solving different tasks. The authors learned that "there was no apparent advantage in 3D maps for tasks requiring elevation information, nor was there a disadvantage for integrated tasks which did not require elevation information". Schobesberger and Patterson (2007) conducted a study comparing conventional (2D) and perspective (3D) trail maps of an outdoor area of the Zion National Park in Utah. Respondents generally agreed that 3D maps depict reality better than conventional maps. Wilkening and Fabrikant (2011) investigated how varying time constraints and different map types influence people's visuospatial decision making.

According to the scheme introduced by Rohrer (2014), evaluation methods can be divided into behavioral (objective) and attitudinal (subjective) methods. In many previous studies, methods such as questionnaires or interviews have been used. These methods are attitudinal methods because they show "what people say". By contrast, eye-tracking can be considered a behavioral (objective) method, because it shows "what people do". Without an objective evaluation method, it is not possible to reveal the true efficiency of shading in tourist maps. The advantage of an eye-tracking study over a questionnaire is that not only can a respondent's answers in solving a task be analyzed, but also their strategy.

Eye-tracking has been used in the field of cognitive cartography since the 1970s, but work with devices of the time was time consuming and expensive. With cheaper technology, eye-tracking has become more widely used in cartographic studies (Brychtová et al. 2013; Coltekin et al. 2014; Fuhrmann, Komogortsev, Tamir 2009; Kubíček et al. 2017a; Kubíček et al. 2017b; Popelka, Brychtová 2013).

One of the first eye-tracking studies dealing with terrain visualization was performed by Chang, Antes, Lenzen (1985), who analyzed the effect of experience on reading topographic relief maps.

Putto et al. (2014) performed an eye-tracking study evaluating three different terrain visualizations (contour lines, relief shading, and oblique view) in solving three types of spatial tasks (visual search, area selection, and route planning). The results showed that performance on contour line visualization and shaded relief were comparable (oblique view was the slowest). Popelka and Brychtová (2013) investigated the differences in reading 2D maps with contour lines and perspective 3D views by cartographers and non-cartographers. Statistically significant differences were observed for Scanpath Length. Longer scanpaths were recorded in the cartographer group.

The tasks in the case study focus on searching for a specific point on a map – a text label for a particular hill or village. Incoul, Ooms, and De Maeyer (2015) conducted an eye-tracking study comparing paper and digital maps. The task was

also to locate (three) labels on a map. The results of the study contained a statistical evaluation of Search Times, Fixation Count, and Fixation Duration. The results of the study did not prove whether lower screen resolution for digitally presented maps had any influence on the user's behavior. Searching for a specific label on a map was a task in the study of Ooms et al. (2012), who analyzed different label placement in maps. Participants had to locate five names in the displayed stimuli. The aim of the study was to analyze the effectiveness of an improved label placement algorithm. Response times, Fixation Count, and Fixation Duration were analyzed. These measurements are the most common in cartographic research (Popelka 2015). A qualitative analysis of the Scanpath was also performed. No significant influences for a new label placement algorithm were found.

2. Methods

2.1. Design of the experiment

Conventional lighting from the upper left corner of a map is the most commonly used and is also used in tourist maps on the Czech map portal *Mapy.cz*. Respondents were asked to find a certain hill or village on the map as quickly as possible. The tasks were devised as a label finding task. Whether shading had a positive influence on finding hills or whether it distracted respondents in finding villages was examined. The following hypotheses were tested:

- H1: In terms of aesthetics, respondents prefer a shaded map in the questionnaire
- H2: In terms of suitability, respondents prefer a non-shaded map in the questionnaire
- H3: Finding a hill is more effective on a shaded variant of the map
- H4: Finding a village is less effective on a shaded variant of the map

The reason for hypothesis H1 was to corroborate the findings of Ortog (2009) that relief is one of the main reasons for describing a map as beautiful. H2 was based on that most tourist maps are designed without shading, so participants are more familiar with them. Hypotheses H3 and H4 are consistent with the research questions stated above. Shading could help participants find hills because they need only focus on darker areas. However, shaded maps are darker overall, which could negatively impact their legibility.

The remote eye-tracking device SMI RED 250, developed by SensoMotoric Instruments, was used for the study. This device operates at a frequency of 120 Hz.

The experiment was performed with nine-point calibration and the SMI Experiment Center software. Calibration with less than 1° deviation of the visual

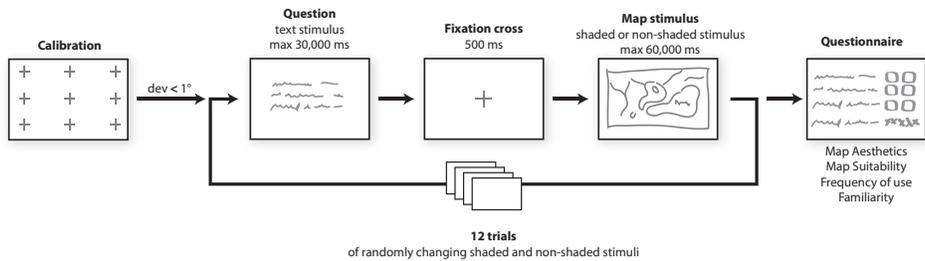


Fig. 1 – Design of the experiment

angle was considered successful. Most of the participants had a much smaller deviation. Average deviation along the X-axis was 0.41° , on the Y-axis 0.43° . For data analysis, perfect spatial accuracy was not necessary as cells in an AOI grid have 190×190 px dimensions. After calibration, respondents had 30 seconds to read and remember the name of the target feature (hill or village). Target names were selected for the same level of difficulty in order to prevent problems for participants (difficult names, too long, too similar to another name on the map, etc.). Before each stimulus, a fixation cross was displayed for 500 ms. The fixation cross was centered on the screen as a neutral location for the eye at the start of the experiment. Stimuli were then displayed on a shaded or non-shaded map. Respondents had 60 seconds to find the target and mark it with a mouse click (see Fig. 1). A time limit (60 s) was chosen based on pilot testing, which was sufficient for most of the stimuli. Respondents only had problems finding the target within the time limit on maps with the sixth pair.

For fixation identification, the I-DT algorithm was used. IDT takes into account the close spatial proximity of eye position points during an eye movement trace. “The algorithm defines a temporal window which moves one point at a time, the spatial dispersion created by the points within this window being compared against the threshold. If such dispersion is below the threshold, the points within the temporal window are classified as a part of a fixation. Otherwise, the window is moved by one sample, and the first sample of the previous window is classified as a saccade.” (Komogortsev et al. 2010) The threshold value for minimum fixation duration was 80 ms, and the dispersion threshold was 50 px. More information about this setting is available in the author’s study (Popelka 2014).

2.2. Stimuli

The experiment contained six pairs of stimuli on shaded and unshaded tourist maps of the Northern part of Czechia found on map portal Mapy.cz (Seznam.cz 2015). The ratio of tasks to find hills or villages was also balanced (6 vs. 6). All

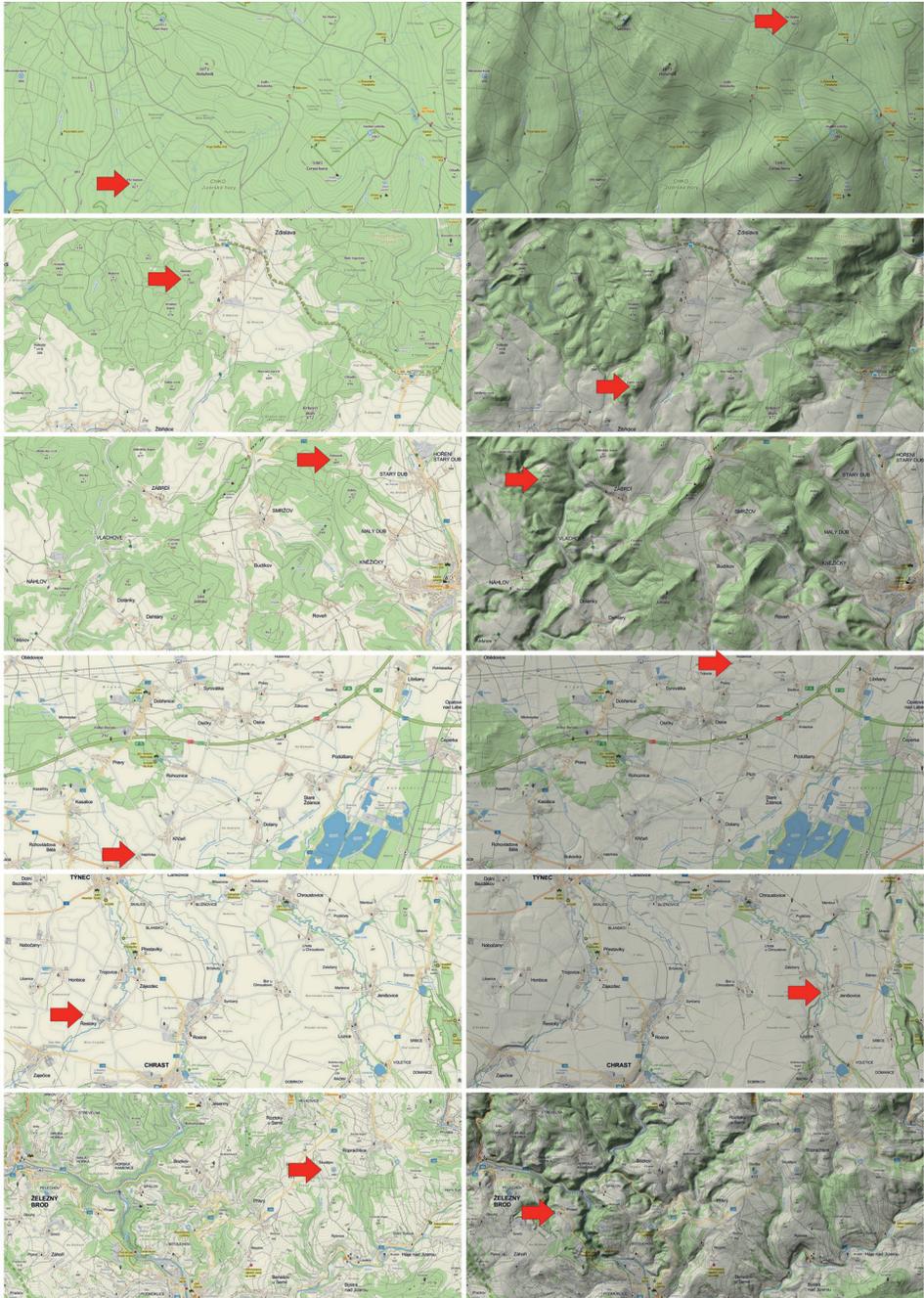


Fig. 2 – Pairs of maps from the eye-tracking experiment. Each figure is a separate stimulus displaying a map of the same area. On the left is a non-shaded variant, on the right a shaded map. Red arrows mark correct answers. Original image source: Mapy.cz © Seznam.cz, a. s.

stimuli pairs (Fig. 2) showed the same territory and differed only in the method of visualization (shaded, non-shaded). The depicted territory lies in the Northern part of Czechia and was relatively unknown to the respondents. Insignificant peaks and villages were also selected for the tasks. Most of the participants came from different parts of Czechia and all were asked whether they knew or were familiar with the depicted area. None confirmed this, so the results were not influenced by familiarity with the region. All stimuli from the experiment are accessible in higher resolution at Eye-tracking Group (2012–2015).

The experiment had a within-subject design, so all respondents saw all the stimuli. To avoid any learning effect, the tasks had to be modified. Therefore, respondents were required to look for different locations on the maps. When the stimuli and tasks were selected, special care was taken to choose targets of similar distance from the center of stimulus across different stimuli. Targets alternating location on the left and right-hand sides of the stimulus were selected, and all stimuli were presented in random order. Before the start of the experiment, respondents were instructed how the task would look, what targets they would search for, and how to mark the correct answer. Practical training was not needed.

2.3. Participants

A total of 40 respondents (24 women and 16 men) were involved in the experiment. Most of the participants were students of geoinformatics with an average age of 20–25 years. The rest were university students of different specializations. The selection of participants was not strictly dependent on their field of study (i.e., cartographers, non-cartographers). All participants had normal vision and were not paid for testing. At the beginning of the experiment, they completed a questionnaire about their knowledge of the Mapy.cz portal. The portal is very popular in Czechia, so even respondents without cartographic knowledge may use

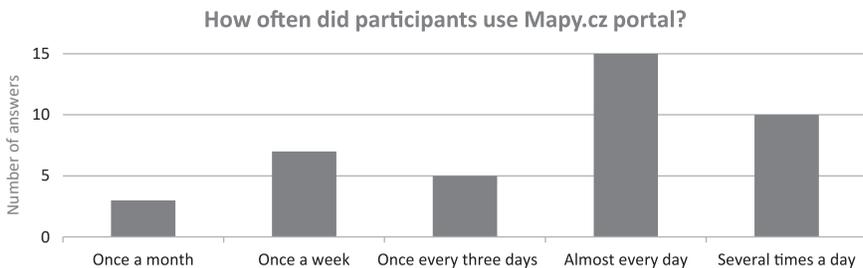


Fig. 3 – Summary of the results of the questionnaire focusing on the frequency of use of the Mapy.cz portal by respondents.

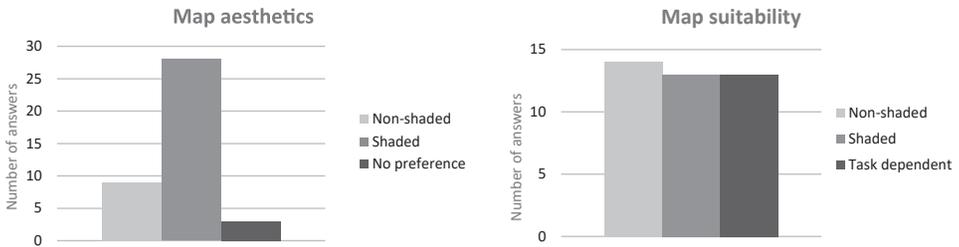


Fig. 4 – Evaluation of the suitability and aesthetics of stimuli

it daily (see Fig. 3). At the end of the experiment, respondents were asked whether they were familiar with the displayed region. None of them had deep knowledge of or familiarity with this area of Czechia.

3. Results

At the end of eye-tracking data recording, respondents completed a short questionnaire with two questions. The first asked about the subjective opinion of specific visualizations – whether the respondents preferred the shaded or unshaded map. The second question concerned the suitability of the map to solve the task – finding the hill or village.

An analysis of answers on the first question concerning map aesthetics (H1) showed that most of the respondents (27) liked the shaded map more than the non-shaded map. Only nine respondents preferred the non-shaded map. The remaining three respondents had no preference (Fig. 4; left). The results of the questionnaire assessment confirmed that users preferred more realistic visualizations, regardless whether they were less efficient for a given task (Hegarty et al. 2009).

For map suitability, the answers were almost equal. Fourteen respondents chose the non-shaded map as more suitable. Thirteen respondents chose the shaded map as more suitable. The remaining respondents (13) answered that it depended on the task – whether they had to find a hill or a village (Fig. 4; right).

3.1. Eye-movement data visualization

To visualize the recorded eye-tracking data, the FlowMap and Gridded AOI methods were used.

3.1.1. FlowMap visualization

The FlowMap method, introduced by Andrienko et al. (2012), is available in the V-Analytics software and provides discrete spatial and spatiotemporal aggregation of gaze trajectories. A visualization was created with arrows representing the movement of gazes of all the participants between Voronoi polygons in the stimulus. These polygons are generated according to the distribution of fixations within the stimulus. Although the same settings for generating Voronoi polygons was applied in both cases, the resulting Voronoi polygons differed because of the different distribution of fixations in the source data. The thickness of the arrows is derived from the number of gaze movements between the Voronoi polygons.

Figure 5 displays the FlowMap visualization of the second pair of maps. The figure shows that the strategy of all respondents was similar in both cases – they looked at the same places. The amount and thickness of arrows (representing the number of moves) is higher in the shaded variant of the stimuli.

In the non-shaded variant, respondents performed a total of 618 moves between the generated Voronoi polygons. For the purpose of visualization, only arrows representing more than three moves were filtered. The upper part of Figure 5 contains 263 moves (43%). In the shaded variant, a greater number of moves between Voronoi polygons was registered (820). After filtering, the lower part of Figure 5 contained 364 arrows (44%).

3.1.2. Gridded AOI visualization

The Gridded AOI method was first introduced by Brodersen, Andersen, and Weber (2002). A regular square grid is overlaid on the analyzed stimulus. Eye-tracking metrics are related to the cells of this grid.

In this eye-tracking experiment, a grid with 5x10 cells was created over each stimulus. The number of rows and columns in the grid was set to create square cells. The software COTOS was used to create the grid. COTOS was programmed as a bachelor thesis output (Kučera 2014). The software exported the coordinates of the cell corners, which were then imported into the OGAMA open source application (Voßkühler et al. 2008). In OGAMA, the number of recorded fixations was calculated for each cell of the grid. Data were exported as a text document and inserted into MS Excel, where the conditional formatting function was used to color the grid according to a normalized number of recorded fixations. The numbers in cells represent the proportion of fixations recorded in the stimuli. The blue table on the left displays data for non-shaded variants (Fig. 6). The red table in the middle shows the proportion of fixations for shaded stimuli. The advantage of this type of visualization is the ability to further compute the numbers in cells. In this experiment, the proportion of fixations in the non-shaded stimuli

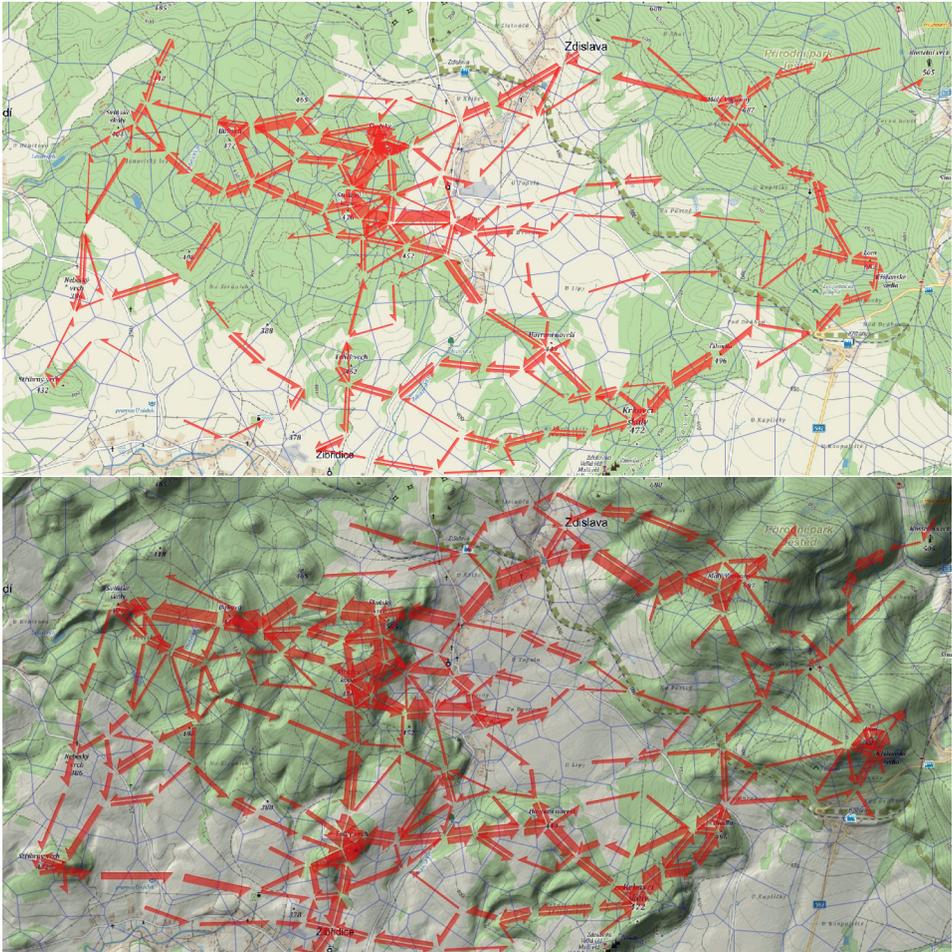


Fig. 5 – FlowMap visualization of aggregated gaze movement of all participants. FlowMaps of both images were generated with the same settings (0; 0; 0; 25; $r = 50$). Data are filtered, so only more than three moves are displayed. Original image source: Mapy.cz © Seznam.cz, a. s.

were divided by the proportion of fixations in the shaded stimuli (and multiplied by 100). The resulting grid was again colored using conditional formatting. The values higher than 100 (shades of blue) represent the dominance of fixations on the non-shaded stimuli. When the value in the cell was lower than 100, relatively more fixations in this cell were recorded in a shaded stimulus.

A visual analysis of Figure 6 shows the similar approach of participants on both stimuli. Even when looking for different targets, the respondents clearly focused on (performed a higher number of fixations) the same parts of the stimuli. The clearest example is shown in the first pair of stimuli, in which a higher proportion

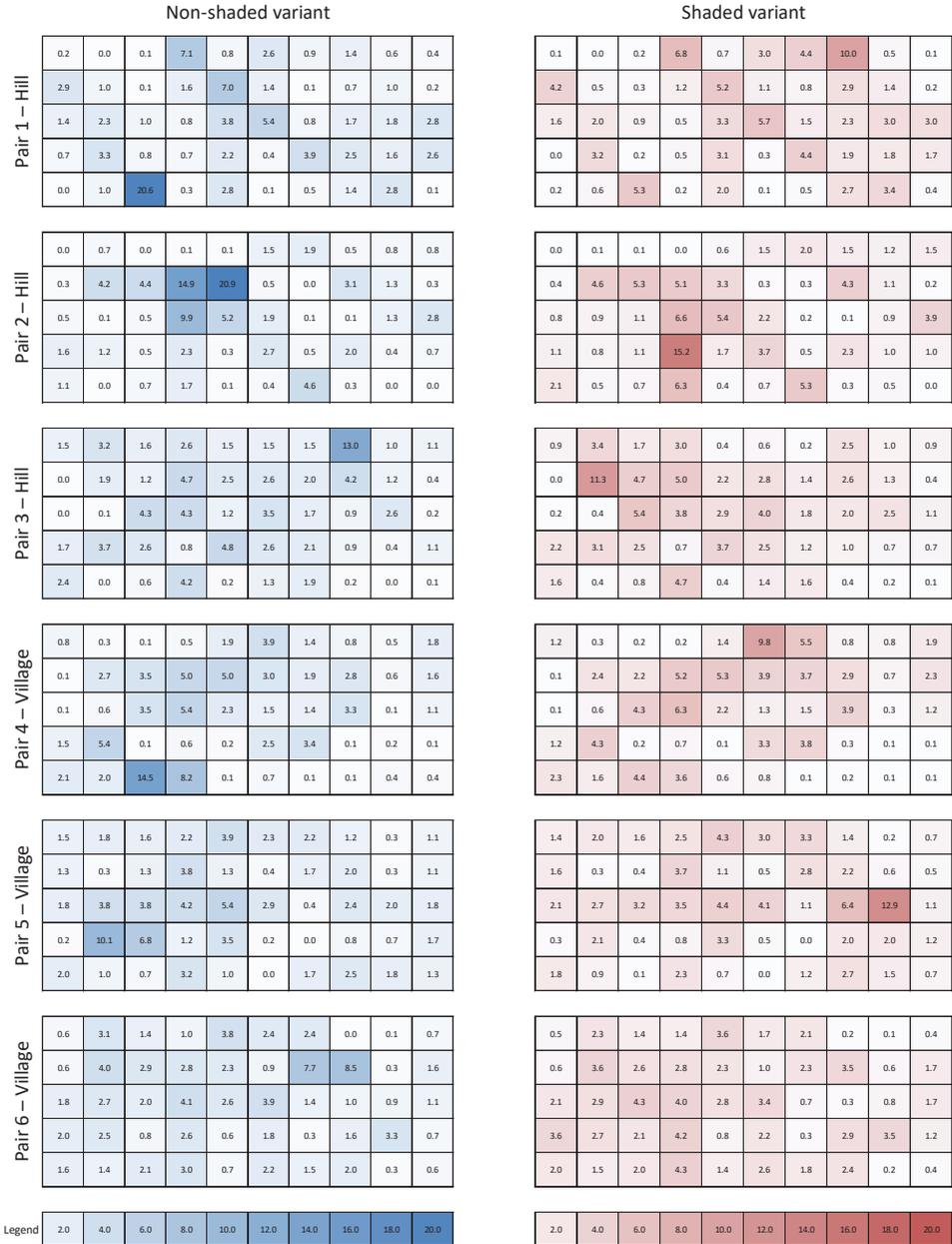


Fig. 6 – Percentage of fixations recorded in individual cells of stimuli displayed using the Gridded AOI visualization method. The tables on the left contain the distribution of fixations in the non-shaded stimuli. The tables in the middle contain data from the shaded stimuli. The tables on the right display the difference between the previous tables (generated as a subtraction of the non-shaded grid values from the shaded grid values).

Difference Shaded – Non-shaded

110.6	100.0	94.8	104.4	104.0	90.9	34.8	22.3	102.1	128.5
74.6	127.0	83.6	117.4	130.7	116.9	61.6	43.1	81.7	103.2
95.5	109.4	101.4	114.5	111.3	95.5	71.3	83.8	70.0	95.7
167.5	101.9	143.5	108.3	78.1	105.6	90.4	122.1	94.4	132.6
81.0	120.9	343.0	104.5	125.3	101.7	100.9	63.0	85.6	78.9

100.0	152.2	91.1	113.4	71.5	100.3	97.3	62.3	83.0	73.2
91.2	92.2	86.4	261.0	506.8	118.8	77.3	77.0	112.8	106.1
86.2	60.3	74.0	142.8	97.7	91.3	94.9	103.3	124.5	77.7
125.7	123.8	74.0	20.3	47.6	78.1	103.2	90.0	70.9	84.5
67.9	67.2	99.2	37.8	81.5	83.2	88.5	98.1	67.2	100.0

130.8	95.0	97.5	90.9	175.8	156.7	209.0	405.9	102.7	111.4
100.0	23.5	39.6	94.3	108.8	95.1	126.4	141.8	93.4	94.8
85.1	84.8	82.2	110.9	57.4	89.2	96.6	64.5	104.7	59.3
86.0	115.7	104.7	105.8	123.9	102.6	139.6	99.0	84.4	122.9
131.9	74.0	88.4	92.4	84.7	96.1	112.1	90.2	85.1	98.6

85.8	101.6	94.3	123.6	120.4	45.2	37.4	100.5	81.3	94.1
96.9	109.7	139.4	96.1	96.4	82.8	62.3	95.5	90.3	79.4
96.9	96.1	85.1	87.4	101.9	110.0	95.8	88.6	84.8	93.4
112.6	121.5	92.5	93.1	109.7	81.7	92.0	84.8	115.2	96.9
93.7	115.8	289.6	201.6	70.0	95.5	103.3	98.6	128.9	122.5

107.0	93.6	98.4	89.9	92.3	82.4	74.5	90.9	104.2	119.1
88.7	98.6	161.4	101.7	108.8	93.6	70.0	93.6	81.8	145.3
89.5	131.9	114.8	116.3	118.2	76.0	69.2	46.2	21.4	133.2
88.9	355.5	554.4	121.6	105.5	80.7	100.0	59.3	57.2	119.3
107.3	103.7	149.6	126.2	116.3	100.0	122.2	95.3	111.4	130.0

105.8	122.2	103.5	83.8	103.9	123.6	107.8	84.1	99.5	125.3
97.2	108.2	108.4	99.4	101.3	94.0	261.5	212.5	77.8	97.5
89.6	95.8	56.3	102.1	96.3	109.5	141.5	146.7	109.1	77.7
66.1	93.6	56.8	68.5	90.9	88.1	99.5	65.4	93.9	76.4
87.1	93.6	103.6	75.5	70.4	89.3	88.4	90.1	112.9	111.7

20.0	40.0	60.0	80.0	100.0	120.0	140.0	160.0	180.0	200.0
------	------	------	------	-------	-------	-------	-------	-------	-------

of fixations was recorded in the cells on the diagonal (from the fourth cell to the bottom-right).

The right-hand part of Figure 6 shows mainly red-shaded cells. It means that the proportion of fixations was higher in the majority of cells for shaded stimuli.

3.2. Statistical analysis of Eye-movement data

3.2.1. Statistical analysis of eye-tracking metrics for each pair of stimuli

The first step in statistical analysis of the recorded eye-tracking data was to compare the Trial Duration metric. Trial Duration describes how much time the respondents need to find the correct answer. The boxplot in Figure 7 shows the median values of this metric. In most cases (except the third pair of stimuli), a higher value was observed in the shaded variant of the map.

In the next step, eye-tracking metrics Scanpath Length (length of the gaze trajectory) and Fixation Frequency (number of fixations per second) were analyzed. More overall fixations during the same time interval and a longer scanpath indicate a less efficient search (Goldberg, Kotval 1999; Goldberg et al. 2002).

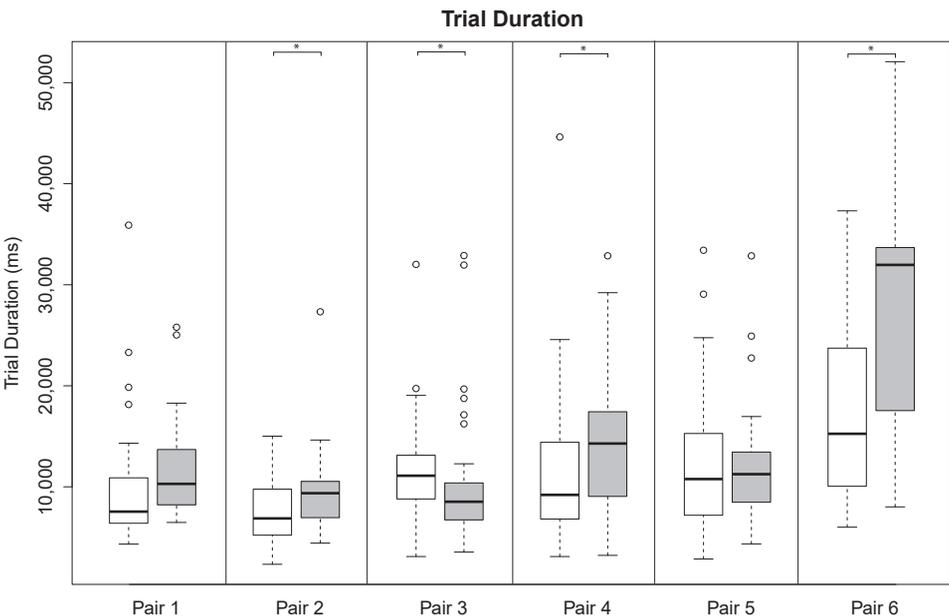


Fig. 7 – Trial Duration values for all six pairs of stimuli in the experiment. Grey marks a shaded variant of the stimuli. An asterisk marks a statistically significant difference.

In almost all pairs of stimuli, a longer scanpath was observed for the shaded variant of the stimuli. The only exception again is the third pair of stimuli (Fig. 8).

The number of fixations a user can have per second is closely related to the duration of fixations. If the duration of a fixation is very long, the number of fixations per second will decrease (Ooms et al. 2012). The fixation frequency was higher with the shaded stimulus in all cases except pairs 3 and 5. The only statistically significant value was observed for the fourth pair (Fig. 9). According to Jacob and Karn (2003), a higher number of single fixations, or clusters of fixations, are often an index of greater uncertainty in recognizing a target item.

The data were analyzed using the Wilcoxon rank sum test as the measured data did not have a normal distribution, which was proven by applying the Shapiro-Wilk test. In the first part of statistical testing, the differences between non-shaded and shaded maps were tested separately for each pair of stimuli. Statistically significant differences are marked with an asterisk in the boxplot figures. In almost all cases, the higher values of eye-tracking metrics were observed in the shaded variant.

A summary of the Wilcoxon test results is shown in Table 1. Statistically significant differences were found in almost all pairs of stimuli for Trial Duration, Fixation Count, and Scanpath Length metrics. In the case of Fixation Duration and

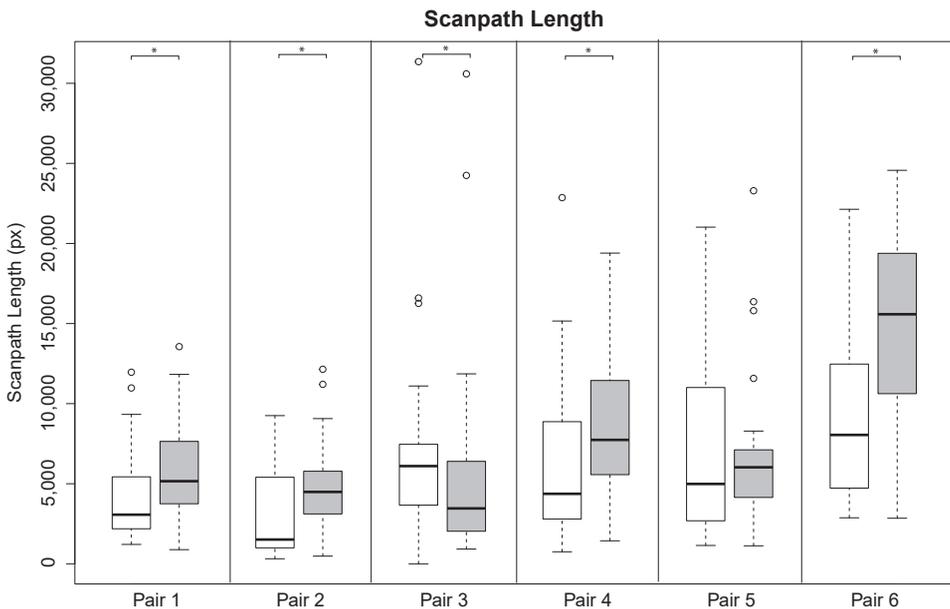


Fig. 8 – Scanpath Length values for all six pairs of stimuli in the experiment. Grey marks a shaded variant of the stimuli. An asterisk marks a statistically significant difference.

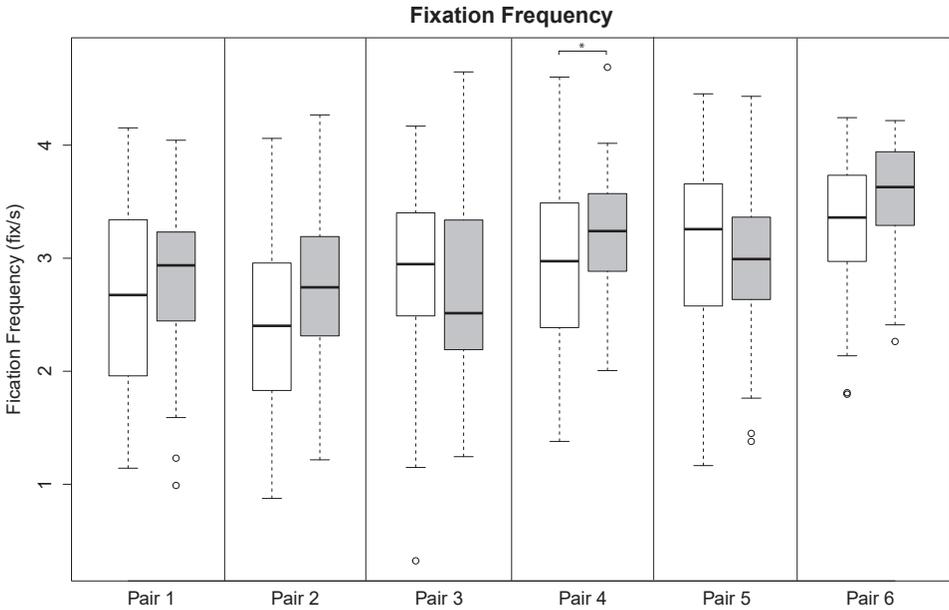


Fig. 9 – Fixation Frequency values for all six pairs of stimuli in the experiment. Grey marks a shaded variant of the stimuli. An asterisk marks a statistically significant difference.

Fixation Frequency, statistically significant differences were observed only in the case of the second, respectively third, pair of stimuli.

The previous analyses investigated the differences between shaded and non-shaded stimuli separately for each task. In the next step of the experiment, the differences were tested for shaded and non-shaded stimuli together. In the case of Trial Duration, Fixation Count, and Scanpath Length, differences were statistically significant at significance level $\alpha = 0.05$ (Fig. 10).

Table 1 – Results of statistical testing of the differences between the non-shaded and shaded variant of each pair of stimuli for five analyzed eye-tracking metrics. Results with a p-value less than 0.05 are marked in italic.

	Trial Duration	Fixation Count	Fixation Duration	Fixation Frequency	Scanpath Length
Pair 1	0.069	0.058	0.650	0.485	0.014
Pair 2	<i>0.039</i>	<i>0.031</i>	<i>0.004</i>	0.106	<i>0.019</i>
Pair 3	<i>0.022</i>	<i>0.028</i>	0.080	0.232	<i>0.019</i>
Pair 4	0.029	0.037	0.357	<i>0.050</i>	0.024
Pair 5	0.879	0.922	0.330	0.493	0.837
Pair 6	< 0.001	< 0.001	0.078	0.087	< 0.001

3.2.2. The dependence of eye-tracking metrics on the type of task

The next step of statistical analysis was testing the difference between the values of measured eye-tracking metrics and the type of task (searching for a village or a hill). A comparison of the stimuli focusing on searching for hills (pairs 1–3) and villages (pairs 4–6) showed statistically significant differences for all eye-tracking metrics. Most of the eye-tracking metrics had higher values in the case of searching for a village (Fig. 11). Searching for a village was more time consuming, respondents performing a higher number of shorter fixations and their gaze trajectory being longer. Finding the village took longer, regardless of the map used for visualization.

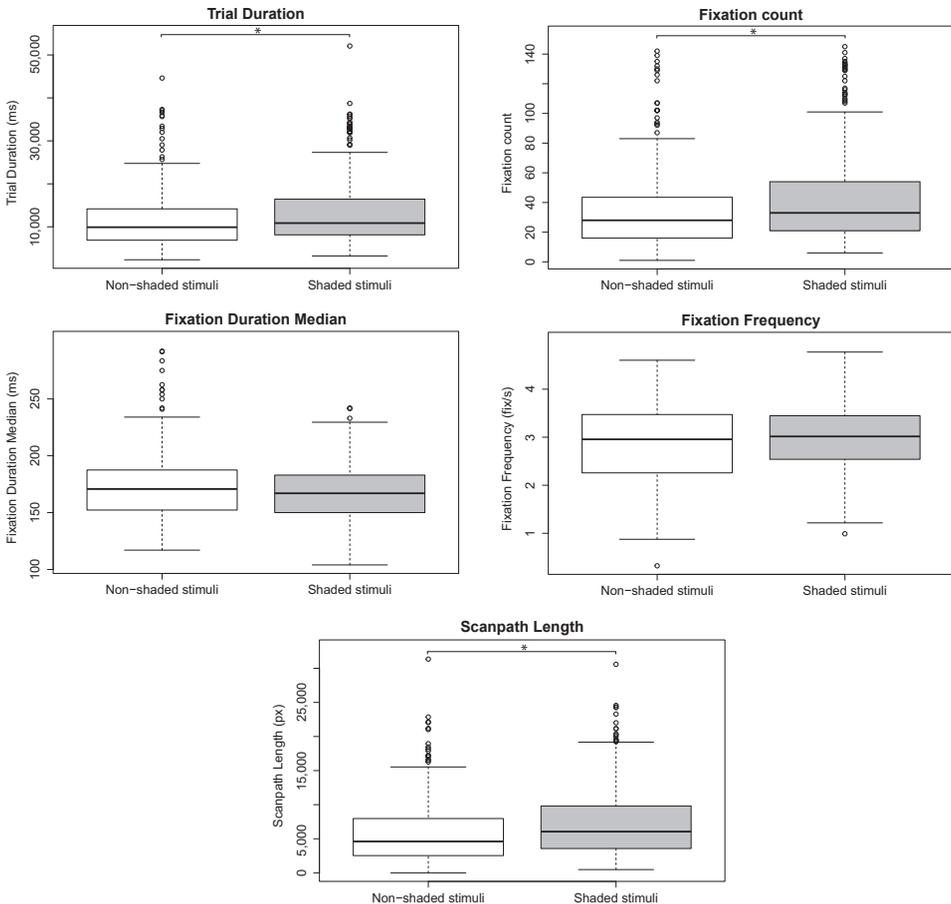


Fig. 10 – Values of five eye-tracking metrics for all six pairs of stimuli in the experiment in total. Grey marks a shaded variant of the stimuli. An asterisk marks a statistically significant difference.

3.2.3. Trial Duration dependence on the type of task and type of visualization

Dependence on the type of task (village vs. hill) and the type of visualization (shaded vs. non-shaded) was evaluated using a two-way analysis of variance (ANOVA) and multiple paired comparison (TukeyHSD). Because the residues of the ANOVA did not have a normal distribution, the correctness of results was confirmed with bootstrapping. Bootstrapping was made as a modification to Manly's approach - Unrestricted Permutation of Observations (Manly 2006). An analysis of variance and multiple paired comparisons were calculated for all eye-tracking metrics. For the eye-tracking metrics Trial Duration, Fixation Count, and Scanpath

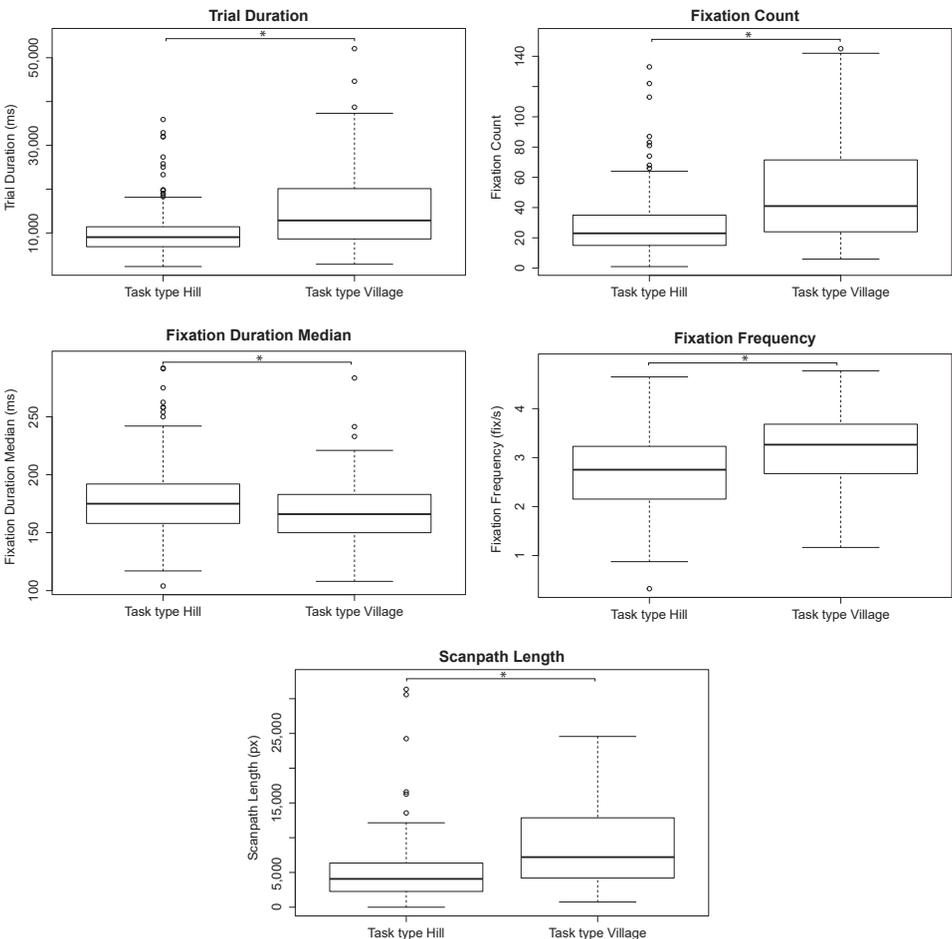


Fig. 11 – Comparison of five eye-tracking metrics for tasks aimed at finding hills (pairs 1–3) and villages (pairs 4–6). An asterisk marks a statistically significant difference.

Table 2 – Comparison of effect of the type of task (village vs. hill) and the type of visualization (shaded vs. non-shaded) for all investigated eye-tracking metrics (ANOVA + TukeyHSD)

Pair	Trial Duration	Fixation Count	Fixation Duration	Fixation Frequency	Scanpath Length
Shaded hill Non-shaded hill	0.905	0.908	0.536	0.691	0.710
Non-shaded village Non-shaded hill	< 0.001	< 0.001	0.008	< 0.001	< 0.001
Shaded village Non-shaded hill	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Non-shaded village Shaded hill	0.002	< 0.001	0.253	< 0.001	0.001
Shaded village Shaded hill	< 0.001	< 0.001	0.013	< 0.001	< 0.001
Shaded village Non-shaded village	< 0.001	< 0.001	0.622	0.567	0.001

Length, statistically significant differences were observed on all possible combinations except on the pair of tasks aimed at hill finding (shaded and non-shaded), in which a statistically significant difference was not found. For the eye-tracking metrics Fixation Duration Median and Fixation Frequency, statistically significant differences were not found in other combinations (see Table 2).

Figure 12 shows the output of one test, but it was addressed from two different perspectives, the first (Fig. 12; left) focusing on the mean Trial Duration value depending on the visualization method, and the second (Fig. 12; right) focusing on the Trial Duration value depending on the question type. The difference between

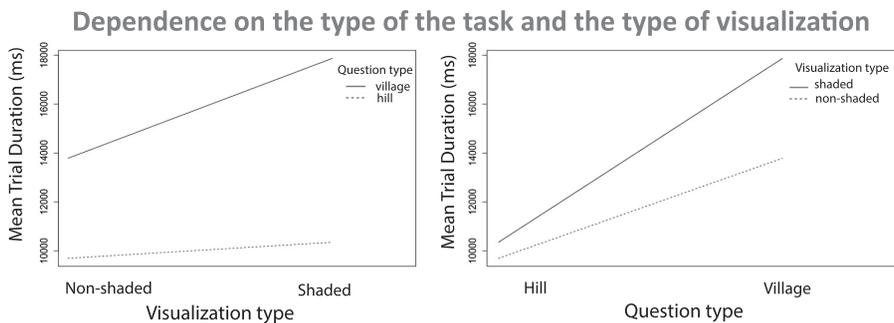


Fig. 12 – Dependence on the type of task (village-hill) and the type of visualization (shaded-non-shaded) on tourist maps in the experiment. Eye-tracking metric Trial Duration was analyzed by using a two-way ANOVA and a TukeyHSD.

hill finding and village finding in shaded vs non-shaded maps is statistically significant, as is village finding on both variants of the map. No statistically significant difference was observed for hill finding in the shaded and non-shaded variants of the map. Hills were found more quickly than villages. This was probably due to the maps containing fewer hills than villages.

It can be stated that the most difficult task for respondents was finding a village on the shaded map. The easiest (fastest) task was finding a hill on the non-shaded map without. The difference between the time needed to find the hill on the shaded map and the non-shaded map was not statistically significant (Table 2).

4. Discussion

As Lloyd (2005) states, visual attention is directed in two basic ways – a top-down process and a bottom-up process. In the first case, attention can be directed consciously by the map reader using information already stored as prior knowledge in memory. In the second, visual attention is directed by information found on the map. The results show that a respondent's attention was directed by information on the map – they focused on labels and did not give priority to places where hills should be (according to shading). This finding contrasts with the results of the study by Castner and Wheate (1979), in which shaded relief provided an advantage in tasks associated with topography (as hill-finding is).

By contrast, the results of the study agree with other previous studies. Savage, Wiebe, Devine (2004) conducted a survey comparing 2D and 3D topographic representation for different tasks. Orthogonal maps and perspective views with color hypsometry were used in this study. The results showed that 3D maps provided no apparent advantage for tasks requiring elevation information. In tasks where elevation information was not necessary, the use of 3D maps was disadvantageous.

Wilkening and Fabrikant (2011) analyzed user decision making with four types of maps (contour lines, hill-shading, dark hill-shading, and colored slope classes). Participants were required to find a place where a helicopter could safely land (not steeper than 14%). Compared to the other maps, the accuracy of answers was significantly higher with the slope map. In maps with relief shading, accuracy was worse compared to the other map types.

The only study using a similar type of stimuli (shaded and non-shaded maps) and eye-movement recording is the study of Putto et al. (2014). In this study, three types of stimuli were used – contour lines, shaded relief, and oblique view enhanced by a triangular grid. The results of the study show that the oblique view method was the slowest. Performance on shaded relief and contour line visualization were comparable.

The different results that Castner and Wheate (1979) obtained may have had several reasons, such as different contrast in both map types or different user experience with maps. Those will be looked at in this discussion. Several methods for eye-movement data analyses exist. This part of the discussion describes the reasons for selecting the methods used as well as possible problems.

Higher values of analyzed eye-tracking metrics and a more problematic search on shaded maps might be caused by the lower contrast between the map background and labels. An empirical study focusing on the readability of map labels placed on a background with different distances between colors was performed by (Brychtová, Coltekin 2014), who discovered that smaller distances between colors (lower contrast) between labels and the background slow participants down. Näsänen, Ojanpää, and Kojo (2001) showed that contrast strongly affects the speed of visual searches. The task in their eye-tracking study was to find a letter in an array of numerals with five different contrasts. Their results showed that the threshold search time and average number of fixations decreased with increasing contrast. Contrast was expressed as Michelson contrast with the use of a Minolta Luminance Meter. This measurement is not possible for complex stimuli such as maps, so the saliency was calculated. The saliency of an object is the state or quality by which it stands out relative to its neighbors. Visual saliency typically arises from contrast between items and their neighborhood. Saliency for the third pair of maps was calculated in OGAMA software using an iLab toolkit (Itti, Koch, Niebur 1998) with a standard, predefined combination of channels (Fig. 13). Warmer colors represent higher values of the orientation channel. Labels are more distinct on the non-shaded variant of the map (Fig. 13; left). For this reason, Staněk et al. (2010) and Kubíček et al. (2011) recommend suppressing the visual appearance of a background for certain special map types (i.e., for emergency management).

User experience with the map portal (Mapy.cz) was assessed only by questionnaire, in which respondents estimated how often they used the portal (Fig. 3). This approach does not prove that people can successfully interpret a map. It would

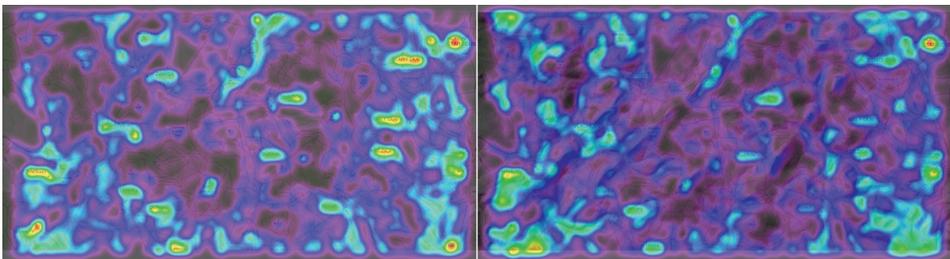


Fig. 13 – Saliency comparison of the third pair of maps. Labels are more distinct on the non-shaded variant of the map (left) than on the shaded variant (right).

probably be better to perform an additional test to understand how familiar people are with topographic maps and shaded relief representations.

Several methods were used for eye-tracking data analysis. For statistical analysis, five frequently used eye-tracking metrics were selected (Trial Duration, Fixation Count, Fixation Duration, Fixation Frequency, and Scanpath Length) based on the author's review of eye-movement studies in cartography (Popelka 2015). For data visualization, the Gridded AOI method was used. The advantage of using this method, for example, instead of an Attention Map is that a synthesis output can be derived from resultant grids to visualize the differences between maps.

Eye-tracking metric Fixation Count correlates with Trial Duration, so the number of recorded fixations can reflect Trial Duration. The time and number of fixations required to solve the task were the main indicators of its complexity. In most cases, though, the results of these two metrics are similar, and both are included in the results to show the results for all five most frequently used eye-tracking metrics.

The results show that shaded maps are less suitable than their non-shaded variant for tasks involving finding villages. In the case of finding hills, the results of statistical testing of eye-tracking metrics show only a small difference between shaded and non-shaded maps. The hypothesis that shading helps find hills has not been proven, although the use of shading is not a disadvantage in a hill-finding task. In the case of finding villages, a darker shaded map negatively impacts the efficiency of the search.

However, more testing with different map types and other tasks will be necessary to generalize this finding. Further testing will be conducted with more complex tasks in which respondents will have to imagine the terrain – for example, analyses of visibility or finding the steepest path.

5. Conclusion

This paper describes a user eye-tracking study focusing on the differences between reading a shaded and a non-shaded tourist map. The task given to respondents was to find and mark a specific hill or village on a map. The aim of the study was to analyze whether shading affected performance when searching for these targets. At the beginning, four hypotheses were defined.

Hypothesis H1, that respondents would prefer the shaded variant of the map from an aesthetic point of view, was confirmed. The results are consistent with the results of the interview performed by Ortag (2009) in which shading was marked as an important factor in map aesthetics. On the question of suitability of the map to solve the task (finding a hill or village), the respondents' answers were

balanced (H2 not confirmed). Eye-tracking data measured during the experiment were visualized with the Gridded AOI and FlowMap methods.

An analysis of the eye-tracking data shows that statistically significant differences between shaded and non-shaded maps is observable in three of the five evaluated eye-tracking metrics (Trial Duration, Fixation Count, Fixation Duration, Fixation Frequency, and Scanpath Length). Higher values were recorded for shaded variants of the map. The influence of the type of task on the values of eye-tracking metrics was also analyzed. For most of the eye-tracking metrics, statistically significant differences were found for all possible combinations of task and visualization methods. The only exception was in finding hills on a shaded or non-shaded map, in which no statistically significant differences were found for any analyzed eye-tracking metric. Hypothesis H3, that respondents would find a hill faster on the shaded variant of the map, was not confirmed. Finding a village on the shaded map was significantly slower than on a non-shaded map (H4 confirmed).

The results of the experiment show that a shaded map is less suitable than a non-shaded variant of the map when finding a village, and that only a small difference was observed between a shaded and non-shaded map when finding a hill.

References

- ANDRIENKO, G., ANDRIENKO, N., BURCH, M., WEISKOPF, D. (2012): Visual Analytics Methodology for Eye Movement Studies. *IEEE Transactions on Visualization and Computer Graphics*, 18, 12, 2889–2898.
- BERNABÉ-POVEDA, M. A., SÁNCHEZ-ORTEGA, I., ÇÖLTEKIN, A. (2011): Techniques for Highlighting Relief on Orthoimaginery. *Procedia-Social and Behavioral Sciences*, 21, 346–352.
- BILAND, J., ÇÖLTEKIN, A. (2016): An empirical assessment of the impact of the light direction on the relief inversion effect in shaded relief maps: NNW is better than NW. *Cartography and Geographic Information Science*, 1–15.
- BRODERSEN, L., ANDERSEN, H.H., WEBER, S. (2002): Applying eye-movement tracking for the study of map perception and map design. Copenhagen: National Survey and Cadastre.
- BRYCHTOVÁ, A., COLTEKIN, A. (2014): An Empirical User Study for Measuring the Influence of Colour Distance and Font Size in Map Reading Using Eye Tracking. *The Cartographic Journal*, 1–11.
- BRYCHTOVÁ, A., PÁSZTO, V., MAREK, L., PÁNEK, J. (2013): Web-design evaluation of the crisis map of the Czech Republic using eye-tracking. Paper presented at the International Multidisciplinary Scientific GeoConference: SGEM: Surveying Geology & mining Ecology Management, Albena, Bulgaria.
- BURIAN, J., ŠTÁVOVÁ, Z. (2009): Kartografické a geoinformatické chyby v územních plánech. *Geografie*, 114, 3, 179–191.
- CASTNER, H.W., WHEATE, R. (1979): Re-assessing the role played by shaded relief in topographic scale maps. *The Cartographic Journal*, 16, 2, 77–85.

- COLTEKIN, A., DEMSAR, U., BRYCHTOVÁ, A., VANDROL, J. (2014): Eye-hand coordination during visual search on geographic displays. Paper presented at the 2nd International Workshop on Eye Tracking for Spatial Research (ET4S 2014), Vienna, Austria.
- COLTEKIN, 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.
- DELUCIA, A. (1972): The effect of shaded relief on map information accessibility. *The Cartographic Journal*, 9, 1, 14-18.
- DUŠEK, R., MIŘIJOVSKÝ, J. (2009): Visualization of geospatial data: chaos in the dimensions. *Geografie*, 114, 3, 169-178.
- Eye-tracking Group at Department of Geoinformatics Palacký University in Olomouc, Czech Republic. Olomouc 2012-2015, <http://eyetracking.upol.cz/3d-shading-mapy-cz/> (15. 9. 2018).
- 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. *Cartographic Journal*, 45, 3, 201-215.
- FUHRMANN, S., AHONEN-RAINIO, P., EDSALL, R.M., FABRIKANT, S.I., KOUA, E.L., TOBÓN, C., WILSON, S. (2005): Making useful and useable geovisualization: Design and evaluation issues. *Exploring Geovisualization*, 553-566.
- FUHRMANN, S., KOMOGORTSEV, O., TAMIR, D. (2009): Investigating Hologram-Based Route Planning. *Transactions in GIS*, 13, s1, 177-196.
- GOLDBERG, J.H., KOTVAL, X.P. (1999): Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics*, 24, 6, 631-645.
- GOLDBERG, J.H., STIMSON, M.J., LEWENSTEIN, M., SCOTT, N., WICHANSKY, A.M. (2002): Eye tracking in web search tasks: design implications. Paper presented at the Proceedings of the 2002 symposium on Eye tracking research & applications.
- HEGARTY, M., SMALLMAN, H.S., STULL, A.T., CANHAM, M.S. (2009): Naïve cartography: How intuitions about display configuration can hurt performance. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 44, 3, 171-186.
- CHANG, K.-T., ANTES, J., LENZEN, T. (1985): The effect of experience on reading topographic relief information: Analyses of performance and eye movements. *The Cartographic Journal*, 22, 2, 88-94.
- IMHOF, E. (2007): *Cartographic relief presentation*: ESRI, Inc.
- INCOUL, A., OOMS, K., DE MAEYER, P. (2015). Comparing paper and digital topographic maps using eye tracking *Modern Trends in Cartography*, 339-356.
- ITTI, L., KOCH, C., NIEBUR, E. (1998): A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 11, 1254-1259.
- JACOB, R.J., KARN, K.S. (2003): Eye tracking in human-computer interaction and usability research: Ready to deliver the promises *The Mind's Eye - Cognitive and Applied Aspects of Eye Movement Research*, Elsevier, 573-605.
- JENNY, B., RÄBER, S. (2015): Relief Shading, <http://www.reliefshading.com/history/> (16. 6. 2018).
- KOMOGORTSEV, O.V., JAYARATHNA, S., KOH, D.H., GOWDA, S.M. (2010): Qualitative and quantitative scoring and evaluation of the eye movement classification algorithms. Paper presented at the Symposium on Eye-Tracking Research & Applications (ETRA 2010), Austin, Texas, USA.
- KUBÍČEK, P., MULIČKOVÁ, E., KONEČNÝ, M., KUČEROVÁ, J. (2011): Flood management and geoinformation support within the emergency cycle (EU example) *Environmental Software Systems. Frameworks of eEnvironment* (pp. 77-86): Springer.

- KUBÍČEK, P., ŠAŠINKA, Č., STACHOŇ, Z., HERMAN, L., JUŘÍK, V., URBÁNEK, T., CHMELÍK, J. (2017): Identification of altitude profiles in 3D geovisualizations: the role of interaction and spatial abilities. *International Journal of Digital Earth*, 1–17.
- KUBÍČEK, P., ŠAŠINKA, Č., STACHOŇ, Z., ŠTĚRBA, Z., APELTAUER, J., URBÁNEK, T. (2017): Cartographic Design and Usability of Visual Variables for Linear Features. *The Cartographic Journal*, 54, 1, 91–102, doi:10.1080/00087041.2016.1168141.
- KUČERA, M. (2014): Využití open-source nástrojů pro přípravu. průběh a vyhodnocení eye-tracking experimentů. Thesis, Univerzita Palackého, Olomouc.
- LLOYD, R.E. (2005): Attention on maps. *Cartographic Perspectives*, 52, 28–57.
- MANLY, B.F. (2006): Randomization, bootstrap and Monte Carlo methods in biology, 70, CRC Press.
- NÄSÄNEN, R., OJANPÄÄ, H., KOJO, I. (2001): Effect of stimulus contrast on performance and eye movements in visual search. *Vision Research*, 41, 14, 1817–1824.
- OOMS, K., DE MAEYER, P., FACK, V., VAN ASSCHE, E., WITLOX, F. (2012): Interpreting maps through the eyes of expert and novice users. *International Journal of Geographical Information Science*, 26, 10, 1773–1788.
- OOMS, K., MAEYER, P.D., FACK, V., ASSCHE, E.V., WITLOX, F. (2012b): Investigating the Effectiveness of an Efficient Label Placement Method Using Eye Movement Data. *Cartographic Journal*, The, 49, 3, 234–246.
- ORTAG, F. (2009): Variables of Aesthetics in Maps Cartography and Art, Springer Berlin Heidelberg, Berlin, Heidelberg, 1–9.
- PETROVIČ, D., MAŠERA, P. (2004): Analysis of user's response on 3D cartographic presentations. Paper presented at the 7th meeting of the ICA Commission on Mountain Cartography.
- PHILLIPS, R.J., LUCIA, A., SKELTON, N. (1975): Some objective tests of the legibility of relief maps. *The Cartographic Journal*, 12, 1, 39–46.
- POPELKA, S. (2014): Optimal eye fixation detection settings for cartographic purposes. *International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management, SGEM*, 1, 2, 705–712.
- POPELKA, S. (2015): Hodnocení 3D vizualizací v GIS s využitím sledování pohybu očí. Univerzita Palackého v Olomouci, Olomouc.
- POPELKA, S., BRYCHTOVÁ, A. (2013): Eye-tracking study on different perception of 2D and 3D terrain visualisation. *The Cartographic Journal*, 50, 3, 240–246.
- POTASH, L., FARRELL, J., JEFFREY, T. (1978): A technique for assessing map relief legibility. *The Cartographic Journal*, 15, 1, 28–35.
- PUTTO, K., KETTUNEN, P., TORNIAINEN, J., KRAUSE, C.M., TIINA SARJAKOSKI, L. (2014): Effects of cartographic elevation visualizations and map-reading tasks on eye movements. *The Cartographic Journal*, 225–236.
- ROBINSON, A.H., MORRISON, J., MUEHRCKE, P., KIMERLING, A., GUPTILL, S. (1995): *Elements of Cartography*. New York: John Wiley&Sons.
- ROHRER, C. (2014). When to Use Which User-Experience Research Methods, <https://www.nngroup.com/articles/which-ux-research-methods/> (16. 6. 2018).
- SAVAGE, D.M., WIEBE, E.N., DEVINE, H.A. (2004): Performance of 2d versus 3d topographic representations for different task types. Paper presented at the Human Factors and Ergonomics Society Annual Meeting.
- Seznam.cz (2015): Mapy.cz, <http://www.mapy.cz/> (16. 6. 2018).
- SCHOBESBERGER, D., PATTERSON, T. (2007): Evaluating the Effectiveness of 2D vs. 3D Trailhead Maps. Paper presented at the 6th ICA Mountain Cartography Workshop, Lenk, Switzerland.

- STANĚK, K., FRIEDMANNOVÁ, L., KUBÍČEK, P., KONEČNÝ, M. (2010): Selected issues of cartographic communication optimization for emergency centers. *International Journal of Digital Earth*, 3, 4, 316–339.
- VOŽKÜHLER, A., NORDMEIER, V., KUCHINKE, L., JACOBS, A.M. (2008): OGAMA (Open Gaze and Mouse Analyzer): open-source software designed to analyze eye and mouse movements in slideshow study designs. *Behavior research methods*, 40, 4, 1150–1162.
- WILKENING, J., FABRIKANT, S.I. (2011): How do decision time and realism affect map-based decision making? *Spatial Information Theory*, Springer, 1–19.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Culture of the Czech Republic (NAKI II research programme; project No. DG18P02OVV017), Ministry of Education, Youth and Sports (project no. CZ.02.3.68/0.0/0.0/16_011/0000669) and The Czech Science Foundation (project No. 18-05432S).