Where do people look at during multi-scale map tasks?

Laura Wenclik[1] and Guillaume Touya [1]

[1]LASTIG, Univ Gustave Eiffel, IGN-ENSG, F-77420 Champs-sur-Marne, France

Correspondence: Laura Wenclik (laura.wenclik@ign.fr)

Abstract. In order to design better pan-scalar maps, i.e. interactive, zoomable, multi-scale maps, we need to understand how they are perceived, understood, processed, manipulated by the users. This paper reports an experiment that uses an eye-tracker to analyse the gaze behaviour of users zooming and panning into a pan-scalar map. The gaze data from the experiment shows how people look at landmarks to locate the new map view after a zoom. We also identified different types of behaviours during a zoom when people stare at the mouse cursor, or during a pan where the gaze follows a landmark while the map translates.

Keywords. cartography, pan-scalar maps, eye-tracking, zooming, landmarks

1 Introduction

Pan-scalar maps, i.e. interactive, zoomable multi-scale maps (Gruget et al., 2023), accessible through web map services such as Google Maps (Skopeliti and Stamou, 2019), are now clearly the most used maps around the world. But contrary to traditional static maps, their design is not guided by our cognition of these maps, because these objects are too recent to be as studied as static paper maps have been in the past century (Dumont et al., 2020). How do people really understand, explore, interact with these pan-scalar maps? What changes compared to our well-documented use of static maps? Researchers have to address these questions to design better pan-scalar maps, where disorientation occurs less often (Touya et al., 2023). In order to have a better understanding of the behaviour of pan-scalar map users, which explains why some feel disoriented during their use, this paper explores the visual behaviour of these users.

Using eye-tracking during user surveys is one of the most promising methods to address these questions. Since the first experiments in the 1960s (Yarbus, 1967), eye-trackers are frequently used to analyze the gaze behaviour of map readers. Several literature reviews show the diversity of such research projects (Kiefer et al., 2017; Krassanakis

and Cybulski, 2021; Keskin and Kettunen, 2023). Eye-trackers provide different kinds of information about the gaze, but researchers mainly use the fixation points and the scan paths between these fixation points (Çöltekin et al., 2009; Cybulski and Horbiński, 2020). The fixation and saccades can be massive and complex to analyze (Ooms et al., 2017), and areas of interest can also be derived from the fixation point to obtain a more generalised view of the gaze (Çöltekin et al., 2009; Cybulski and Horbiński, 2020; Keskin et al., 2023). Blinks can also be used, for instance to measure the cognitive load during a visual task (Zagermann et al., 2016).

In this paper we have two main hypotheses. The first hypothesis is that pan-scalar map readers rely on landmarks to guide their exploration through multiple map view. This paper presents an experiment designed to explore this hypothesis, or more generally to understand the behaviour of gaze during the zooming interactions.
The second hypothesis is based on the behaviour of the user according to the map background. It is supposed that the user will use the map differently depending on the representation of the landmarks and whether or not certain landmarks are highlighted.

The paper is structured as follows. Section 2 describes the reported eye-tracking experiment. Section 3 briefly describes how the eye-tracking data were post-processed, and Section 4 presents a selection of our results.

2 Materials and Methods

2.1 Description of the tasks

The aim of the experiment is to understand what a person looks at to find their way around a pan-scalar map and to understand the different strategies used. This experiment is exploratory and qualitative, we do not expect quantitative results to validate hypothesis. Our goal is to reproduce a real use of the pan-scalar map, i.e. the users can use pan and zoom at will to fulfill their tasks. Four types of tasks

are proposed in order to cover more uses of the pan-scalar map.

The first task simulates the use of web map service, with a user searching for a specific address. The map application will then zoom in very strongly on the address entered in the text panel. The user has little spatial context and it often takes some time to find their way around. To simulate this classical use, a point is placed on the map and the display is very zoomed (zoom level 18). The user is then asked to interact with the map (zooming and panning) until they feel they know where they are. Four different locations are shown iteratively during this task, the same four locations for all participants.

The second task is to find a place in the map from an aerial image oriented north/south. The aerial image of a specific area is displayed and the map is zoomed out to the city where the location is. The user then tries to find the location in the image by zooming in and out to explore the map. This task is repeated in two different cities.

The third task also consists in finding a precise location in the map, using this time textual indications such as "the stadium which is located north of the main river". This task is repeated in two different cities.

The last task builds on tasks 2 and 3. The user is asked to find a location using an aerial image of the location (the image has little geographical context) and textual information is given to help the user find the area where the location is located. This task is repeated for two different cities.

The background maps are alternated between two stages and two candidates (according to the following table). We select two different pan-scalar maps, OpenStreetMap (OSM) and Google Maps. (Table 1).

Table 1. Distribution of the map backgrounds during the experiment stages, for the participants that were alternatively assigned to experiment 1 and 2.

stage	experiment 1	experiment 2
task 1, stage 1	Google Maps	OSM
task 1, stage 2	OSM	Google Maps
task 1, stage 3	Google Maps	OSM
task 1, stage 4	OSM	Google Maps
task 2, stage 1	Google Maps	OSM
task 2 stage 2	OSM	Google Maps
task 3, stage 1	Google Maps	OSM
task 2, stage 2	OSM	Google Maps
task 4, stage 1	Google Maps	OSM
task 4, stage 2	OSM	Google Maps

2.2 Apparatus and participants

To understand the behaviour of a person in this experiment, we will use an eye-tracker. The eye-tracker used is a Pupil Core device from Pupil Labs. The basic eye tracker configuration is used for this experiment, i.e. a fix-

ation time of 80 ms to 200 ms. Decreasing the minimum time does not bring any significant change while increasing it makes most of the fixation points disappear. The eye-tracker is calibrated at the beginning of the survey with targets clips on the corners of the screen used for the survey (the same 23.8 inches screen is used for all participants, with a resolution of 1920x1080 pixels).

The participants were mainly students from the University and also a few administrative staff members. There were a total of 20 candidates, 3 of whom were not usable (device malfunctioning during the experiment or the screen was not fully visible during the experiment because the participant moved too much on their chair). There were no particular conditions for participating in the survey.

The survey follows a similar protocol for each participant. The participant begins by reading a consent form on data collection. The eye-tracker is calibrated and in parallel the different tasks and instructions are explained (5 to 10 minutes total for this preliminary part). The participant begins the survey and the instructions are recalled for each task and textual indications are given for tasks 3 and 4. At the end of the experiment, the data is recorded and anonymised.

The participants used a web application that we developed for the survey, using Open Layers when OSM maps are rendered, and the Google Maps API, when Google Maps are rendered.

2.3 Data and Software Availability

The code of the web application, as well as the scripts used to post-process the data are available on Github[1]

The data that we collected during the survey, including gaze data and videos, is made available on Zenodo [2].

3 Post-processing

3.1 Geolocation of fixation points

The eye-tracker used in the survey makes it possible to identify the points of fixation of the gaze in a 3D scene. It is therefore necessary to pass it through the 2D reference frame of the screen in order to locate these points on the map displayed on the screen. Having geographical coordinates for the fixation points is particularly useful because it enables multi-scale analyses and also because it enables GIS analyses using the vector objects in the map. We are not interested in the fixation points outside the map. The aim is thus to transform a fixation point in a 3D scene into a geolocated point. The first stage is to know the coordinates of the fixation points in the coordinates system that originates at the bottom left part of the map in the screen,

---

[1]https://github.com/LostInZoom/eye_tracking
[2]https://zenodo.org/record/7584092

with pixel units. These coordinates $(x_{fixation}, y_{fixation})$ are computed using Equation 1, where the envelope of the map in the screen is given by the two points $(x_{min}^{map}, y_{min}^{map})$, and $(x_{max}^{map}, y_{max}^{map})$. With the help of physical markers it is possible to identify a surface in the 3D scene which allows us to know the position of the point on the screen (x_{screen}, y_{screen}). The map view displayed on the screen during the experiment is of fixed size and static on the screen.

$$\begin{cases} x_{fixation} &= \dfrac{x_{screen} - x_{min}^{map}}{x_{max}^{map} - x_{min}^{map}} \\ y_{fixation} &= \dfrac{y_{screen} - y_{min}^{map}}{y_{max}^{map} - y_{min}^{map}} \end{cases} \quad (1)$$

Then, in a second stage , we record the record the coordinates of the extent of the map at each time of the survey. It is thus possible to geo-reference each fixation point (x_{final}, y_{final}) using Equation 2, where the extent of the map at the time of the fixation is given by the points (x_{min}, y_{min}), and (x_{max}, y_{max}).

$$\begin{cases} x_{final} &= x_{fixation} \times (x_{max} - x_{min}) + x_{min} \\ y_{final} &= y_{fixation} \times (y_{max} - y_{min}) + y_{min} \end{cases} \quad (2)$$

3.2 Visualization of the data

A large part of the analysis is the visual analysis of the data. The fixation points were visualised in QGIS after the georeferencing. To better visualise the data in relation to the different scales, a buffer was applied to the points in relation to the margin of error of the eye-tracker in pixels and the zoom (pixel/m), i.e. larger circles correspond to fixations at smaller zoom levels.

4 Results

4.1 Visual analysis

After a visual analysis of the results of task 1, certain behaviours were noticed for several participants. The first one is the use of landmarks to find one's way. Users will look at the landmarks around the point (Figure 1). They change the scale very little to find their way around or at a fairly large zoom. People who use landmarks to locate themselves have a relatively detailed mental map of Paris. It was possible to identify many landmarks such as the name of a metro station or the Montparnasse train station, especially in stage 2 of task 1. The participants used these landmarks at several scales (several fixation points at different scales) to locate themselves.

Conversely, people who have a poor mental map of Paris will not look for landmarks and will zoom out to get a
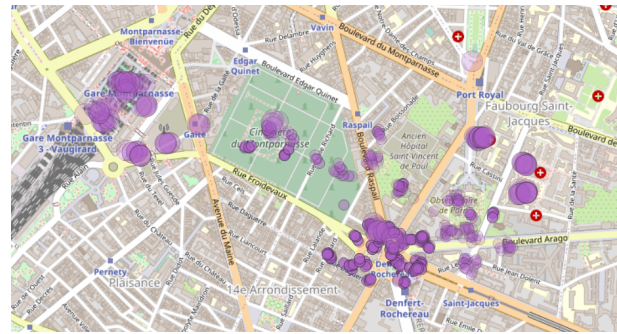

Figure 1. Point of fixation of a person looking at place names, stations during task 1 on one of the places

global view of the city. They will look directly for a global view to find their way around (Figure 2).
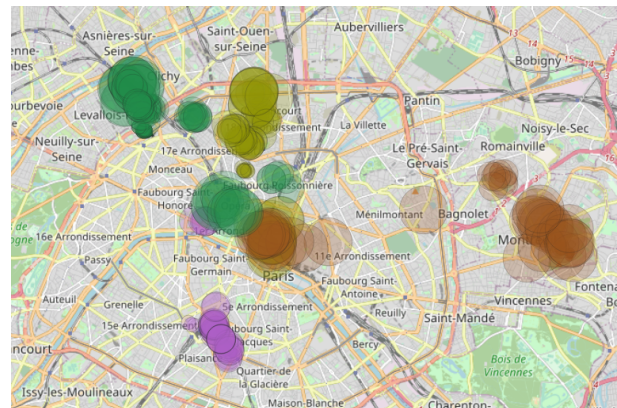
Figure 2. Fixation point when fast zooming out for the first task. Each colour is for a different stage on the task 1 to localise for the same participant.

Some people try to position themselves carefully on the map. Once they have found a landmark they return to the original point with a zoom in in order to place the point correctly on their mental map (Figure 3).
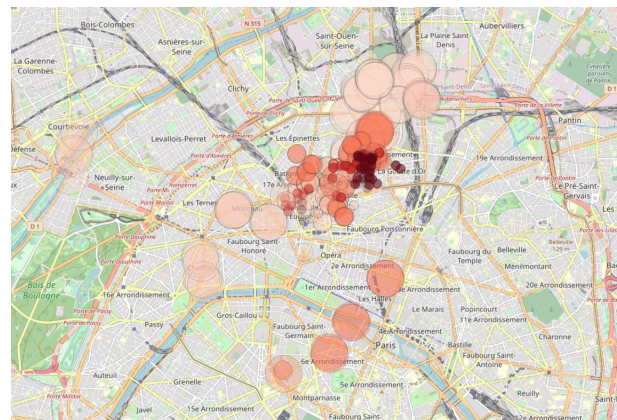
Figure 3. Fixation during a 're-zoom'. The colour shade represents the order of the fixations. The lighter colour represents the first points

Participants who already have a detailed mental map of the area they are in, look at several potential landmarks. However, with gaze data only, it is not clear whether they are trying to triangulate the landmarks in order to find their bearings or whether they are moving from one landmark to another until they find one they know.
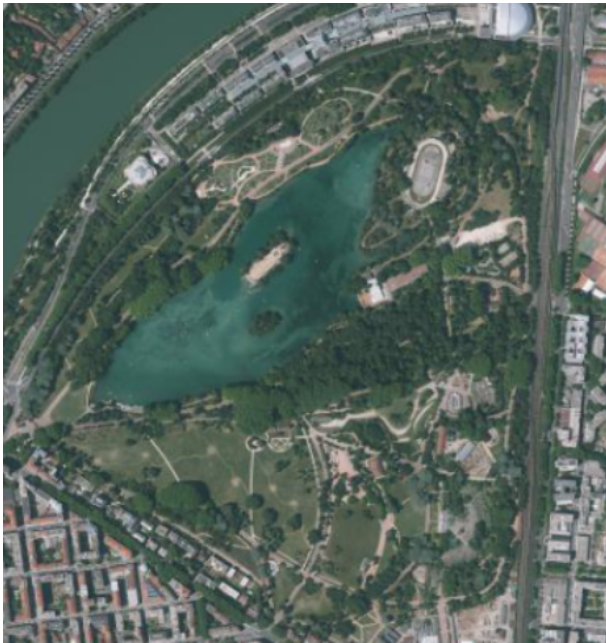
Figure 4. The aerial image the participants had to locate in Lyon, France during task 2 and first location.

In the case of some candidates, it is possible to see a difference in reaction for the four locations of task 1. Indeed, one of the points is located outside Paris with few well-known landmarks. Participants who do not know the area will tend to look for more context to find their way around as shown in the Figure 5
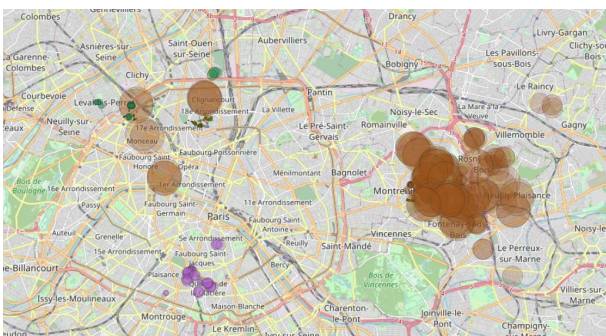
Figure 5. difference for a candidate for the four locations in task 1. Each colour represents a location.

In task 2 the hypothesis to be tested is whether candidates use the visual landmarks visible in pictures to locate themselves on a map. One of the two stages in the task is a view of the "Parc de la tête d'or" in the city of Lyon (Figure 4. Results show that candidates find their way by searching

landmarks. But must of them use only one landmark, e.g. the Rhone river, and focus on it. They see a green space and/or a river and focus on it without looking for another visual element that could be used as an anchor (Figure 6). The participants are not that expert in this area. There is also the question of the candidates' lack of a sense of scale, to correlate the dimensions in the image and in the map.
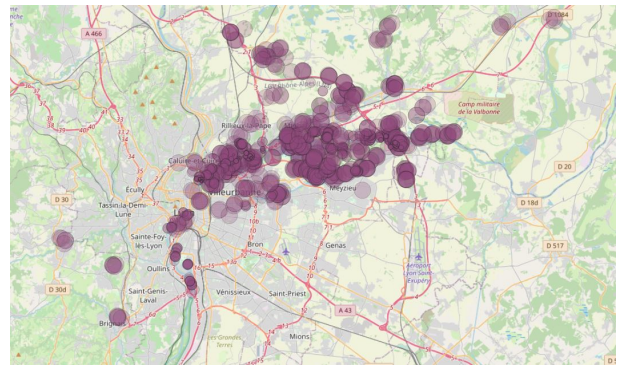
Figure 6. Fixation points when searching for a park outside Lyon for one participant.

In the second stage of task 2, an aerial image of a particular building with a north/south-oriented railway track is used (Figure 7). Users are able to find the location by using this orientation of the railway track. Not all candidates use the strategy of the oriented railway from the beginning. They first look for a track and then realize the vertical orientation of the track. Once they have found the railway line, they look for a building or a triangular road shape. As they go along they refine their search space with precise landmarks. In addition, many candidates asked if the aerial image was correctly oriented N/S after they had started to look for railway tracks, confirming the gaze data. This task also shows that large linear objects such as the Rhone river, or the railway, are prime landmarks.

Figure 7. Image of Dijon, France used during the task 2 and second location

For tasks 3 and 4, the current protocol does not allow for a visual analysis of the interpretation of the textual instructions. Indeed, it is not possible to know whether a person

is looking in a direction other than the one indicated because they have misunderstood the indications or because they wanted to have more spatial context.

The lack of a sense of scale is also found in tasks 3 and 4. When given an indication such as "just south of the city centre" some participants were really looking south of the city in stage 1 of task 3. However, in task 4, stage 2, the instruction was given in relation to a smaller object "west of the Lyon confluence", and the participants had difficulties to find the indicated area.

4.2 Zoom and pan

We also analysed the gaze behaviour during a pan or a zoom interaction. We define a zoom action as starting when the user starts to change scale and to consider progressive zooms, the zoom action is finished when the user stops zooming for one second. Similarly, panning starts when the coordinates change at a fixed scale. It is considered finished when the centre is fixed for more than one second.

Staring at the mouse is one of the zooming behaviours that appears several times during this experiment (Figure 8).
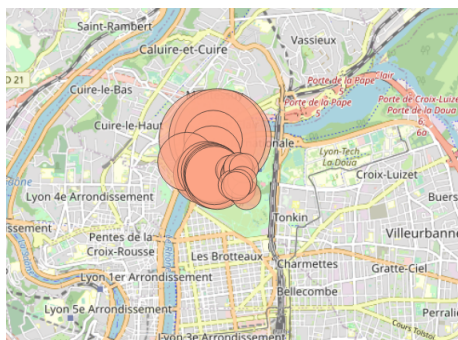

Figure 8. Fixed fixation point when zooming, which corresponds to the user staring the mouse cursor

When looking at linear landmarks, such as the railway track, some participants use the opposite behaviour by fixing the landmark during the pan (so gaze is moving with the map) (Figure 9).

4.3 Quantitative analysis

We are also interested in the influence of the four tasks in the behaviour of the participant, both the gaze behaviour represented by the number of fixations per task, and the interaction behaviour that can be represented by the number and frequency of pan and zoom actions. We conducted several ANOVA tests, summarized in Table 2 to analyse this influence. The ANOVA tests show that the frequency of zooming and panning actions is not influenced by the type of task requested (respectively $p = 0.169$ and $p = 0.123$), while the number of panning and zooming actions is (respectively $p = 1.26e - 06$ and $p = 5.04e - 10$). This shows that the behaviour of users on pan and zoom
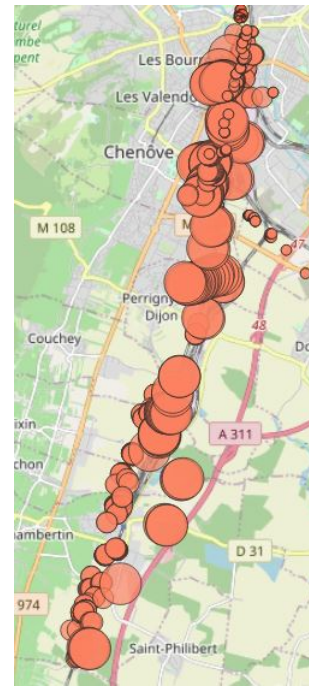
Figure 9. Fixation points remain focused on the railway when panning.

Table 2. p-values of the ANOVA tests on the influence of the task on the number of fixations, pan, and zoom actions.

	fixation	pan	zoom
1	**0.00673**	0.488	0.917
2	**0.0264**	**0.00415**	**0.00235**
3	0.9	0.292	0.207
4	0.235	0.827	0.604
all	**0.0005**	**1.26e-06**	**5.04e-10**

is similar but differs on the time spent on a stage. An ANOVA test on the influence of the task on time shows a significant correlation ($p = 1.6e - 05$), which confirms that some tasks are longer to solve than others.

The analysis of the influence of the map backgrounds is done with the data from stage 2 on task 1. In particular, we are focusing on the visibility of the railway line at the scale of the city. The majority of participants used the railway line as a landmark. However, the railway is less visible on the city scale on Google Maps than on OSM. It is thus possible to see whether a person managed to locate the railway easily if he zoomed directly on a railway or not. The fixation points above zoom level 16, located near the different railways were selected and analyzed (Figure 10).

The number of fixations in the areas near the railway tracks do not follow a normal distribution, so a Kruskal-Wallis test is run to analyze the effect of the map on the fixations, wich appears to be moderate ($p = 0.118$).
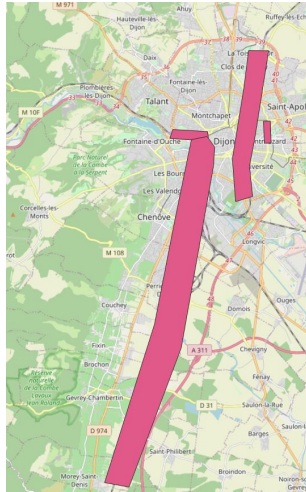
Figure 10. railway buffer zone

5 Conclusions and future research

To conclude, this experiment shows that eye-tracking is an interesting method to study the behaviour of users of pan-scalar maps. The analysis of the gaze behaviour confirms the use of landmarks in locating tasks (Touya et al., 2023), and we identified different types of gaze behaviour during a zoom or a pan action.

However, we think that the experiment dataset can be analysed further, in particular the gaze behaviour during the tasks with textual instructions. But we also would like to use more the vector data, as we did for the railway line (Figure 10), to more closely analyse the gaze behaviour. Furthermore, we think that the experiment protocol has limitations that need to be addressed in further studies. In particular, we do not know why participants are looking at a specific location of the map, and the eye-tracker used is sometimes not precise enough to distinguish between several landmarks. In this study, we were only interested in pan-scalar maps, but it would be interesting to adopt a similar approach to study the gaze behaviour of users of vario-scale maps (van Oosterom et al., 2014).

Finally, our use of eye-tracking data only captures where visual attention is focused, but peripheral vision certainly plays an important role in the success of the tasks given in the experiment (Rosenholtz et al., 2012). Further research should focus on the role played by peripheral vision during the multi-scale exploration of pan-scalar maps.

References

Cybulski, P. and Horbiński, T.: User Experience in Using Graphical User Interfaces of Web Maps, ISPRS International Journal of Geo-Information, 9, 412, https://doi.org/10.3390/ijgi9070412, 2020.

Dumont, M., Touya, G., and Duchêne, C.: Designing multi-scale maps: lessons learned from existing prac-
tices, International Journal of Cartography, 6, 121–151, https://doi.org/10.1080/23729333.2020.1717832, 2020.

Gruget, M., Touya, G., and Muehlenhaus, I.: Missing the city for buildings? A critical review of pan-scalar map generalization and design in contemporary zoomable maps, International Journal of Cartography, 0, 1–31, https://doi.org/10.1080/23729333.2022.2153467, 2023.

Keskin, M. and Kettunen, P.: Potential of eye-tracking for interactive geovisual exploration aided by machine learning, International Journal of Cartography, 0, 1–23, https://doi.org/10.1080/23729333.2022.2150379, 2023.

Keskin, M., Krassanakis, V., and Çöltekin, A.: Visual Attention and Recognition Differences Based on Expertise in a Map Reading and Memorability Study, ISPRS International Journal of Geo-Information, 12, 21, https://doi.org/10.3390/ijgi12010021, 2023.

Kiefer, P., Giannopoulos, I., Raubal, M., and Duchowski, A.: Eye tracking for spatial research: Cognition, computation, challenges, Spatial Cognition & Computation, 17, 1–19, https://doi.org/10.1080/13875868.2016.1254634, 2017.

Krassanakis, V. and Cybulski, P.: Eye Tracking Research in Cartography: Looking into the Future, ISPRS International Journal of Geo-Information, 10, 411, https://doi.org/10.3390/ijgi10060411, 2021.

Ooms, K., Dupont, L., and Lapon, L.: Mixing Methods and Triangulating Results to Study the Influence of Panning on Map Users' Attentive Behaviour, The Cartographic Journal, 54, 196–213, https://doi.org/10.1080/00087041.2016.1213517, 2017.

Rosenholtz, R., Huang, J., and Ehinger, K.: Rethinking the Role of Top-Down Attention in Vision: Effects Attributable to a Lossy Representation in Peripheral Vision, Frontiers in Psychology, 3, https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00013, 2012.

Skopeliti, A. and Stamou, L.: Online Map Services: Contemporary Cartography or a New Cartographic Culture?, ISPRS International Journal of Geo-Information, 8, 215, https://doi.org/10.3390/ijgi8050215, 2019.

Touya, G., Gruget, M., and Muehlenhaus, I.: Where Am I Now? Modelling Disorientation in Pan-Scalar Maps, ISPRS International Journal of Geo-Information, 12, 62, https://doi.org/10.3390/ijgi12020062, number: 2 Publisher: Multidisciplinary Digital Publishing Institute, 2023.

van Oosterom, P., Meijers, M., Stoter, J., and Šuba, R.: Data Structures for Continuous Generalisation: tGAP and SSC, in: Abstracting Geographic Information in a Data Rich World, edited by Burghardt, D., Duchêne, C., and Mackaness, W., Lecture Notes in Geoinformation and Cartography, pp. 83–117, Springer International Publishing, https://doi.org/10.1007/978-3-319-00203-3_4, 2014.

Yarbus, A. L.: Eye Movements and Vision, New York, NY, USA, plenum press edn., 1967.

Zagermann, J., Pfeil, U., and Reiterer, H.: Measuring Cognitive Load using Eye Tracking Technology in Visual Computing, in: Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization, BELIV '16, pp. 78–85, Association for Computing Machinery, New York, NY, USA, https://doi.org/10.1145/2993901.2993908, 2016.

Çöltekin, A., Heil, B., Garlandini, S., and Fabrikant, S. I.: Evaluating the Effectiveness of Interactive Map Interface Designs: A Case Study Integrating Usability Metrics with Eye-Movement Analysis, Cartography and Geographic Information Science, 36, 5–17, https://doi.org/10.1559/152304009787340197, 2009.