

# Map Evaluation under COVID-19 restrictions: A new visual approach based on think aloud interviews

Martin Knura <sup>a,\*</sup>, Jochen Schiewe <sup>a</sup>

<sup>a</sup> Lab for Geoinformation and Geovisualization (g2lab), HafenCity University Hamburg, Germany, [martin.knura@hcu-hamburg.de](mailto:martin.knura@hcu-hamburg.de), [jochen.schiewe@hcu-hamburg.de](mailto:jochen.schiewe@hcu-hamburg.de)

\* Corresponding author

**Abstract:** With the beginning of the COVID-19 pandemic, the execution of eye-tracking user studies in indoor environments was no longer possible, and remote and contactless substitutes are needed. With this paper, we want to introduce an alternative method to eye tracking, completely feasible under COVID-19 restrictions. Our main technique are think aloud interviews, where participants constantly verbalize their thoughts as they move through a test. We record the screen and the mouse movements during the interviews, and analyse both the statements and the mouse positions afterwards. With this information, we can encode the approximate map position of the user's attention for each second of the interview. This allows us to use the same visual methods as for eye-tracking studies, like attention maps or trajectory maps. We implement our method conducting a user study with 21 participants to identify user behaviour while solving high-level interpretation tasks, and with the results of this study, we can show that our new method provides a useful substitute for eye-tracking user studies.

**Keywords:** Map Evaluation, Cognition in Geovisualization, Visual Analytics, Usability Testing

## 1. Introduction

The use of eye tracking is a well-established method in cartographic research to evaluate and optimize the usability of maps and other information visualizations (Kiefer et al. 2017). These eye-tracking user studies usually take place in indoor laboratories, making use of the advantages of stationary systems, e.g. straightforward calibration and high accuracy. But with the beginning of the COVID-19 pandemic, the execution of eye tracking studies in indoor environments was no longer possible. Although there are also head-mounted devices for outdoor environments available, entirely contactless and therefore remote substitutes for eye tracking studies are needed. Alternatives discussed in the research community either lack of accuracy and precision (e.g. webcam-based eye tracking), focus on slightly different aspects of map usability (e.g. mouse tracking), or relinquish the direct measurement of visual attention (e.g. web-based or postal surveys).

With this paper, we want to propose a new approach completely feasible under COVID-19 restrictions, combining classic research techniques from social sciences with methods from UX research and visual analytics. The paper is structured as followed: after the literature review (section 2), we introduce our method design (section 3) and the application for our recent study (section 4). After that, we discuss advantages and limitations of our method compared to eye-tracking (section 5), and conclude the paper (section 6).

## 2. Related Work

Reading a map to acquire spatial information is a complex visual procedure, as we perceive the space through the eyes, reason about our task and, if needed, perform a visual search for more information. Eye-tracking can measure the user's visual attention doing this, and therefore has become a popular research method to understand spatial cognition. Kiefer et al. (2017) provide a general overview about eye-tracking research in the domains of spatial cognition and GIS, while Krassanakis and Cybulski (2019) focus on cartography-related eye-tracking studies between 2009 and 2018. In their most recent work (Krassanakis & Cybulski 2021), the authors discuss the online shift of eye-tracking experimentation as one of the main future perspectives. While online eye-tracking gives the opportunity to expand the overall number of participants, the most important problems they identified are that the experimental process is not fully supervised by the experimental operator, and the accuracy and precision of the data are not adequate.

For analysing eye-tracking data, Andrienko et al. (2012) provide guidelines for method selection depending on the tasks at hand. Other authors develop approaches for the comparison of movement sequences between users or user groups, e.g. based on geometric vectors (Dewhurst et al. 2012) or visualized as a graph (Dolezalova & Popelka 2016).

The same visualizations can also be applied to other methods. As a less expensive alternative to eye-tracking,

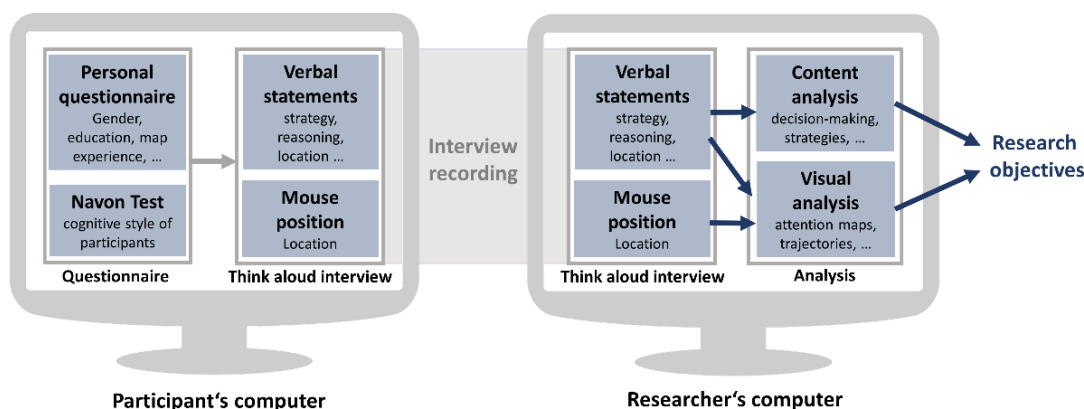


Figure 1. Study design.

Krassanakis and Kesidis (2020) proposed a toolbox based on computer mouse tracking, which can be mainly used to analyse interactive visual search experiments.

### 3. Method design

With the ongoing COVID-19 pandemic, researchers have to adapt their experiments to the regulations of contact restrictions, which means that most studies have to be conducted completely remote. In our case, the overall design of the study remains the same as planned and contains of three parts: a personal questionnaire, a test on the participants' cognitive style and the actual experiment. For the first two parts, the online shift is straightforward: Instead of completing a questionnaire just before the experiment, we send the questions to the participants, and invite them to perform a Navon test to define their cognitive style, using the open source software library PsyToolkit<sup>1</sup>. In the following, we want to introduce our approach to conduct the main part of our study as a substitute for eye-tracking experiments.

#### 3.1 Think Aloud Interviews

Our main technique to evaluate the user's visual attention while reading maps are think aloud interviews (Eccles & Aarsal 2017). During think aloud interviews, participants constantly verbalize their thoughts as they move through the test and solve given tasks. The statements are recorded, and analysed later on, using methods from content analysis.

Think aloud interviews can be subdivided in two types: While the thoughts in a *concurrent* think aloud are verbalized *during* the task-solving process, participants in a *retrospective* think aloud describe their line of thoughts *after* solving the task. As a result, concurrent think aloud is much more demanding to the participants. In contrast, participants in retrospective think aloud tend to justify their decision instead of recapping their initial thoughts (Häder 2015).

The interviews are held via the web-based video conference platform Zoom, which has a function to record the session. In addition to the verbal descriptions, the participants are asked to use their computer mouse as

a pointer during the interviews, providing a second indicator for identifying the exact location on the map the participants focus on. To track the mouse symbol, the participants were asked to share their screens during the interviews, and the verbal statements and the whole screen are recorded, providing the map and the mouse symbol for further analysis. Videos of the participants itself were not recorded.

#### 3.2 Analysis

We analyse the recorded interviews in two ways (see figure 1). First, we conduct a structured analysis of the think aloud statements of the participants using interview encoding techniques from social sciences, focusing on reasoning and decision-making while solving the tasks. Second, we manually encode the approximate map position the participant focuses on for each second during the interview, using the verbal description and the mouse symbol as indicators. To reduce the quantity of possible map positions, we use predefined map areas for position encoding. These areas can be derived from the basemap (e.g. grid lines, map categories), calculated using the visualized data (e.g. cluster areas, Voronoi polygons), or defined by the research question (e.g. focus areas). Alternatively, a pre study could provide the data for the predefined map areas.

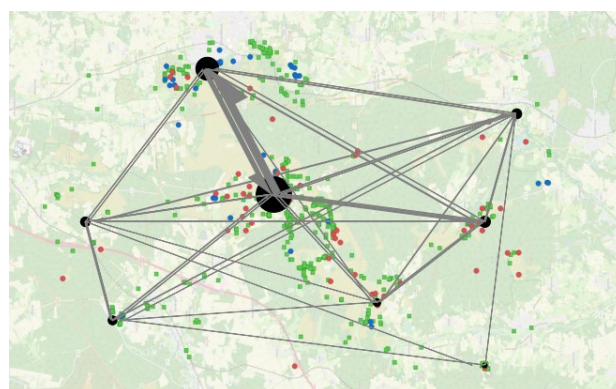


Figure 2. Attention map using eight predefined point clusters.

<sup>1</sup> <https://www.psychotoolkit.org/experiment-library/navon.html>

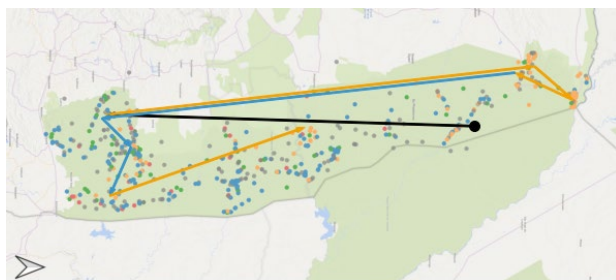


Figure 3. Trajectory path of a participant using predefined point clusters. Black dot is the first cluster identified, arrow color represents the class the participants is focusing while transitioning

This detailed position encoding process<sup>2</sup> enables researchers the possibility to use the same visual methods as for eye movement studies. Eye tracking data usually consists of four components: user identifier, time, position and fixation duration (Andrienko et al. 2012). With our method, we can provide user identifier and time, encode the approximate position, and use the time of user-area interaction as a substitute for fixation duration. With this data, we can generate visualizations known from the visual analytics, e.g. summary maps of the spatial distribution of the participants' attention (see Figure 2), or map displays of trajectories (see Figure 3).

Furthermore, we can arrange the subjects of participant's statements along a timeline and visualize them as a function from the moment the task was started, as in Figure 4. With path similarity analysis, different behaviours between the participants can also be identified, as well as different strategies through path comparison as shown in Figure 5.

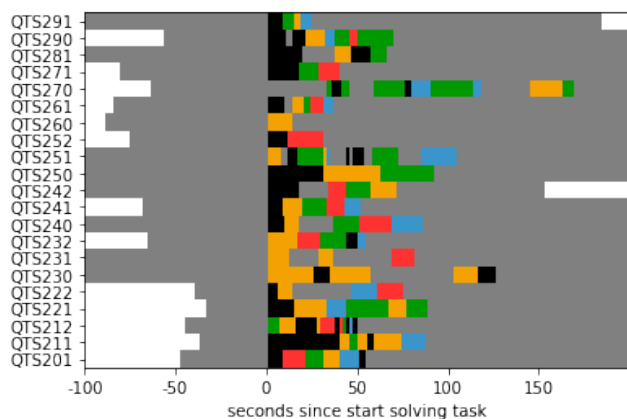


Figure 4. Temporal view of different categories described as a function of time since starting the task. Color represents the respective colored point category on the map. All other statements (regarding multi-colored point clusters, background map, legend, etc.) in grey.

<sup>2</sup> Template for position encoding and related python code for visualizations presented in this paper can be found here: <https://gitlab.com/g2lab/think-aloud-visualizations>

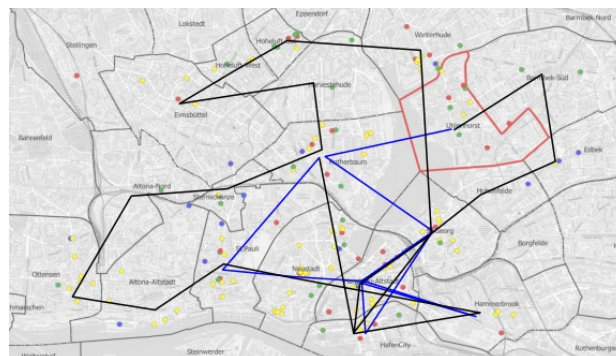


Figure 5. Comparison of trajectories of two participants searching for districts with similar clusters as the red district.

#### 4. Application: Understanding user behaviour

As a first application, we implemented our new method conducting a study to analyse the user behaviour while solving high-level interpretation tasks on point data. First, we perform a postal pre-study with 25 participants to

- identify relevant interpretation tasks for point data sets
- define point cluster for the map position encoding process
- define guideline questions for the interviews

We then conduct think aloud interviews as described above with 21 participants – none of them already taking part in the pre-study – to

- identify pattern in user behaviour while solving high-level interpretation tasks
- investigate the influence of map complexity and point data set complexity on the users' solving strategies

The participants could decide themselves when to start the tasks. Although we guided the interviews by predefined questions to make the interviews more comparable, we waited with our first question until the participants – in their opinion – seem to finish the respective task. Therefore, the method can be described as a mixture of *concurrent* and *retrospective* think aloud.

In the following, we want to present the findings obtained with our new approach while analysing the results of a task which is classified as a relation-seeking task according to the typology of Andrienko and Andrienko (2006). For this task, the participants have to find city districts in Hamburg that have a similar price level of Italian restaurants than the reference district, using the map which is shown in the background of Figure 5 (without the trajectories, the red border of the reference district was also not shown in the initial task).

To get an overview about the different task-solving strategies of the participants, we create an attention map, a summary map of the spatial distribution of all participants' attention (see Figure 6). Although we can



identify the districts which were focused on the most, it is not possible to differentiate between different strategies due to the high number of different paths used. For further investigation, we need to triangulate between the findings of the content and the visual analysis.



Figure 6. Attention map for the 20 districts in the inner city of Hamburg. Point and arrow sizes represent the number of times a participant relates to the respective district or transition.

As their first step during the task-solving process, all participants searched for the reference district and described the price level at hand. The content analysis of the participants' statements revealed that they used four different comparison criteria for the following relation-seeking, with the majority using two or more of them:

- the overall number of points in the district (named 13 times)
- the point proportion between the different price level classes (11)
- the mean price level (10)
- the absence of the highest price level class (10)

As we are interested in the different user strategies to find similar districts, we compare the trajectory paths (see Figure 8 in Appendix) and the corresponding path lengths (see Figure 9 in Appendix) of all users. With this method, there are two groups of participants who stand out against the others: First, there are participants who use many short paths, mostly between neighbouring districts, while searching for similarities (*stepwise strategy*; see also the black path in Figure 5). The other group is much less strategic, as they use longer paths and return to the reference district several times during the search (*star strategy*). Between the other participants, we identified two more clusters using the MultiMatch algorithm for position similarity (Dewhurst et al. 2012): One cluster tends to focus on the central clusters in the south (see also the blue path in Figure 5), while the other moves more towards the top of the map. To visualize the different strategies, we created cumulated flow maps for each cluster (see Figure 7).

In summary, our method allows us to identify different comparison criteria the participants used qualitatively, and to distinguish between different search strategies quantitatively. With these findings, we can answer our research question regarding the user behaviour while

solving this high-level interpretation tasks. Nevertheless, we found no evidence that specific comparison criteria and search strategies correlated.



Figure 7. Cumulated flow maps for each search strategy cluster: stepwise (top left), star (top right), central cluster focus (bottom left), north cluster focus (bottom right).

## 5. Discussion

We introduce our new method as a contactless and remote substitute for eye-tracking studies, combining classic research techniques from social sciences with UX research and visual analytics. In the following section, we want to discuss the advantages and limitations of our method.

### 5.1 Advantages

As we combine different research approaches, we can also utilise the advantages of these methods. The analysis of the verbal statements during the task-solving process provides us detailed insights into the reasoning of the participants, and complements the path analysis revealing the visual processing. Therefore, we could identify different comparison criteria through analysing the statements, which cannot be derived from path analysis. In contrast, the participants did not describe their detailed search strategies explicitly, but we could retrace them using visual analytics.

While the different research approaches provide us complementary findings for different research questions regarding the task-solving process, we also utilise the method variety to support the validity of the respective methods. Using predefined interview guideline questions, we make sure that the statements of the participants are easier to compare and contain the necessary topics related to our research questions. At the same time, the mouse cursor provides us a second indicator for understanding the participants' reasoning during the interview, as it increases the precision of identifying the exact area in focus.

Another key advantage of our method originates in the circumstances of the COVID-19 pandemic. As we had to conduct a remote user study, we were forced to develop a method which is robust against differences in technical setups. In contrast to webcam- or eye-tracking, a system

calibration is not necessary within our method, and screen size or lighting has no influence on the outcome, as we encode the map positions manually and in relation to the map instead of pixelwise. In general, only a microphone and a stable internet connection is required to participate in our study, so the overall costs for the software and hardware configuration are much lower than for comparable eye-tracking setups.

## 5.2 Limitations

As described in section 3.2, an essential part of our method is the manual encoding of the map positions the participants are focusing on. This is also one of the most important limitations of our method: First, our method is less scalable than eye-tracking, as every minute of the interview requires approximately four minutes of encoding time. Second, we introduce an encoding bias in our study. The setup of our method provides a high flexibility regarding the technical requirements for the study participants, but the downside is the lower precision of the visual encoding results. While we can encode the approximate position of user-area interaction for every second during the interview, we rely on two indicators: the verbal statements of the participants, and the position of their mouse cursors. As long as both indicators are consistent to each other, the encoding accuracy depends primarily on the validity of the predefined map areas, and their respective borders. If the mouse cursor is moving contrary to the statements given by the participants, the encoder has to decide between the two indicators. Although we provide guidelines for encoding and prioritize the statements over the mouse cursor in most of the conflicts, we cannot erase all of these uncertainties during the encoding process.

Using a predefined set of map areas for the encoding process is another source of uncertainty in our method. Although the predefined areas can be optimized to answer a specific research question, or derived from a pre-study, a bias is introduced. Regarding the visual analysis, the number of different areas has a big impact on the interpretation of the result, while the exact positions in the visualizations are predefined and not derived from the gaze position data.

Compared to eye-tracking, another limitation of our approach is a constraint in measuring success: As viewing time heavily depends on the users' capabilities to express their thoughts coherent, we cannot use this factor as a dependent variable in *concurrent* think aloud interviews. Alternatively, viewing time can be used in a fully *retrospective* think aloud study, with the drawback of losing all information of the initial task-solving strategy.

## 6. Conclusion

During the COVID-19 pandemic, researchers have to adapt their experiments to the regulations of contact restrictions. In this paper, we introduced a new approach for a contactless and remote substitute for eye-tracking studies. We combine think aloud interviews, a classic

research method from social science, with the methods of UX research and visual analytics by manually encoding the approximate map position of user-area interaction for every second of the recorded interviews.

We implemented our method conducting a user study with 21 participants, analysing user behaviour while solving complex interpretation task. With the results of our application, we can show that our new method provides a useful substitute for eye tracking user studies under contact restrictions, leading to reasonable and satisfying visual results, and with the addition of gaining insights into the task-solving strategies by analysing the content of the participants' verbal statements. Compared to eye-tracking, the major limitations of our methods are the lower accuracy regarding the exact gaze position on the map, and the higher time and work effort due to the manually encoding process.

As a next step, we want to improve our method by conducting a user study for further evaluation. In this study, we plan to use a combination of eye-tracking and think aloud interviews to oppose the visual results and findings of both methods.

## 7. Acknowledgements

This research is part of the project *Improvement of task-oriented visual interpretation of VGI point data* (TOVIP), founded by the *German Research Foundation* (DFG) within the SPP 1894 *Volunteered Geographic Information: Interpretation, Visualisation and Social Computing*.

## 8. References

- Andrienko, N. & Andrienko, G. (2006). Exploratory Analysis of Spatial and Temporal Data. A Systematic Approach. Springer. <https://doi.org/10.1007/3-540-31190-4>
- 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. <https://doi.org/10.1109/TVCG.2012.276>
- Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R. & Holmqvist, K. (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behavior Research Methods*, 44, 1079–1100. <https://doi.org/10.3758/s13428-012-0212-2>
- Dolezalova, J. & Popelka, S. (2016). ScanGraph: A novel scanpath comparison method using visualization of graph cliques. *Journal of Eye Movement Research*, 9(4): 5, 1-13.
- Eccles, D. & Aarsal, G. (2017). The think aloud method: what is it and how do I use it? *Qualitative Research in Sport, Exercise and Health*, 9(4), 514-531. <https://doi.org/10.1080/2159676X.2017.1331501>

- Häder, M. (2006). Empirische Sozialforschung – eine Einführung. Wiesbaden, VS Verlag für Sozialwissenschaften. DOI: 10.1007/978-3-531-19675-6
- Kiefer, P., Giannopoulos, I., Raubal, M. & Duchowski, A. (2017). Eye tracking for spatial research: Cognition, computation, challenges. *Spatial Cognition & Computation*, 17(1-2), 1-19. <https://doi.org/10.1080/13875868.2016.1254634>
- Krassanakis, V. & Cybulski, P. (2019). A review on eye movement analysis in map reading process: the status of the last decade. *Geodesy and Cartography, Committee*

*on Geodesy of the Polish Academy of Sciences*, 68 (1), 191-209.

- Krassanakis, V. & Cybulski, P. (2021). Eye Tracking Research in Cartography: Looking into the Future. *ISPRS Int. Journal of Geo-Information*, 10, 411. <https://doi.org/10.3390/ijgi10060411>

- Krassanakis, V. & Kesidis, A. L. (2020). MatMouse: A Mouse Movements Tracking and Analysis Toolbox for Visual Search Experiments. *Multimodal Technologies and Interaction*, 4(4), 83. <https://doi.org/10.3390/mti4040083>

## 9. Appendix

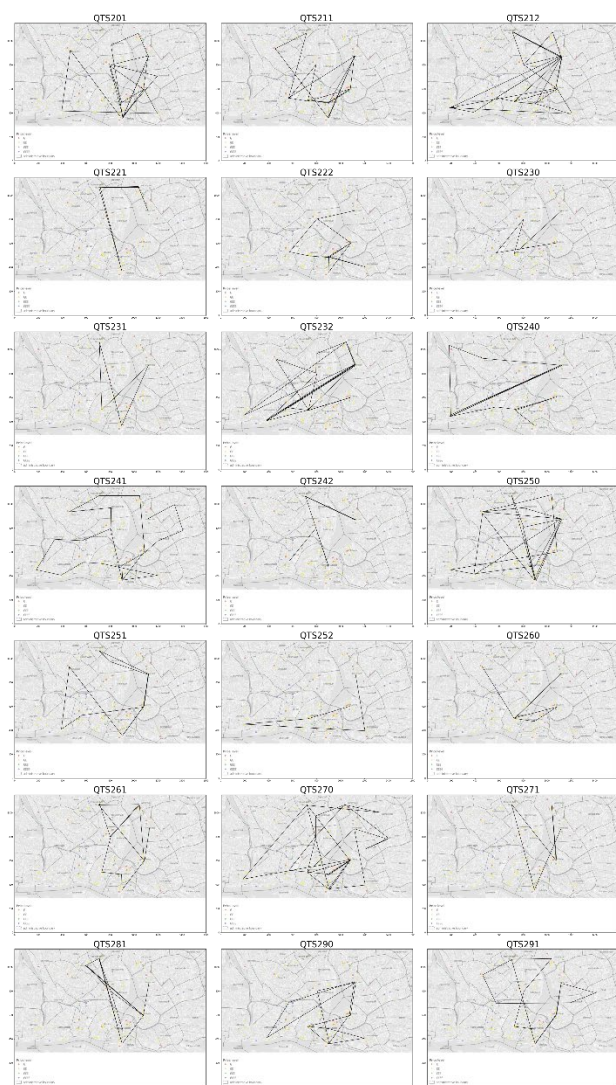


Figure 8. Trajectory paths of each participant.

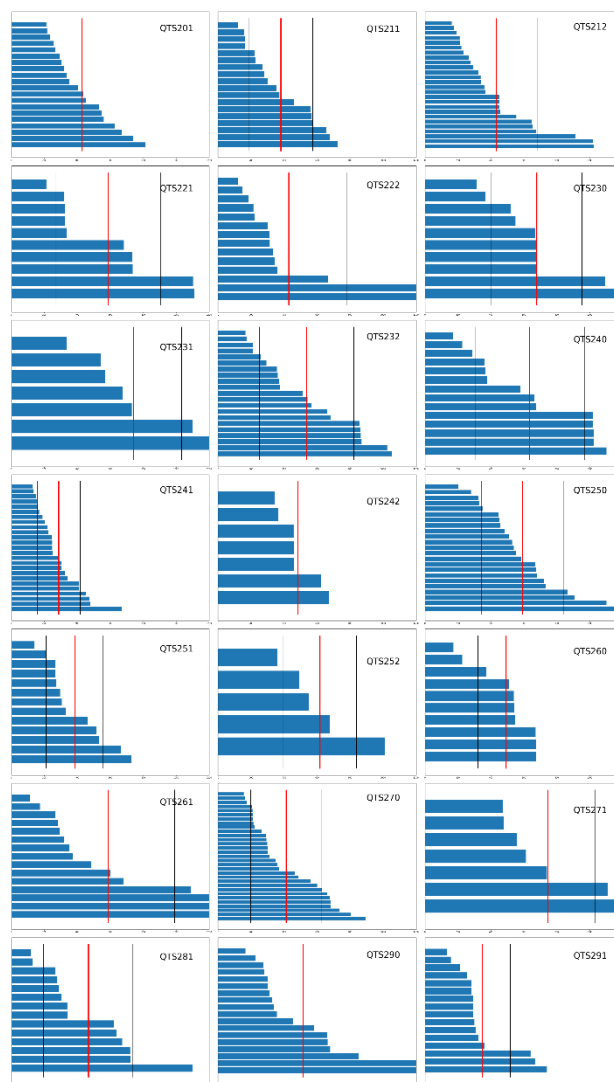


Figure 9. Sorted path lengths in map pixels of each participant corresponding to the trajectory paths in Figure 7. Red line indicates the mean path length.