Eye tracking experiment

Introduction

The eye-tracking study was carried out on a city map of Salzburg, on which the urban green areas of the parks were shown in dark green and the recreation grounds in light green. The most important green areas were labeled with the corresponding names. In addition, there is a legend in the bottom right-hand corner of the stimulus so that participants can distinguish between the different areas.

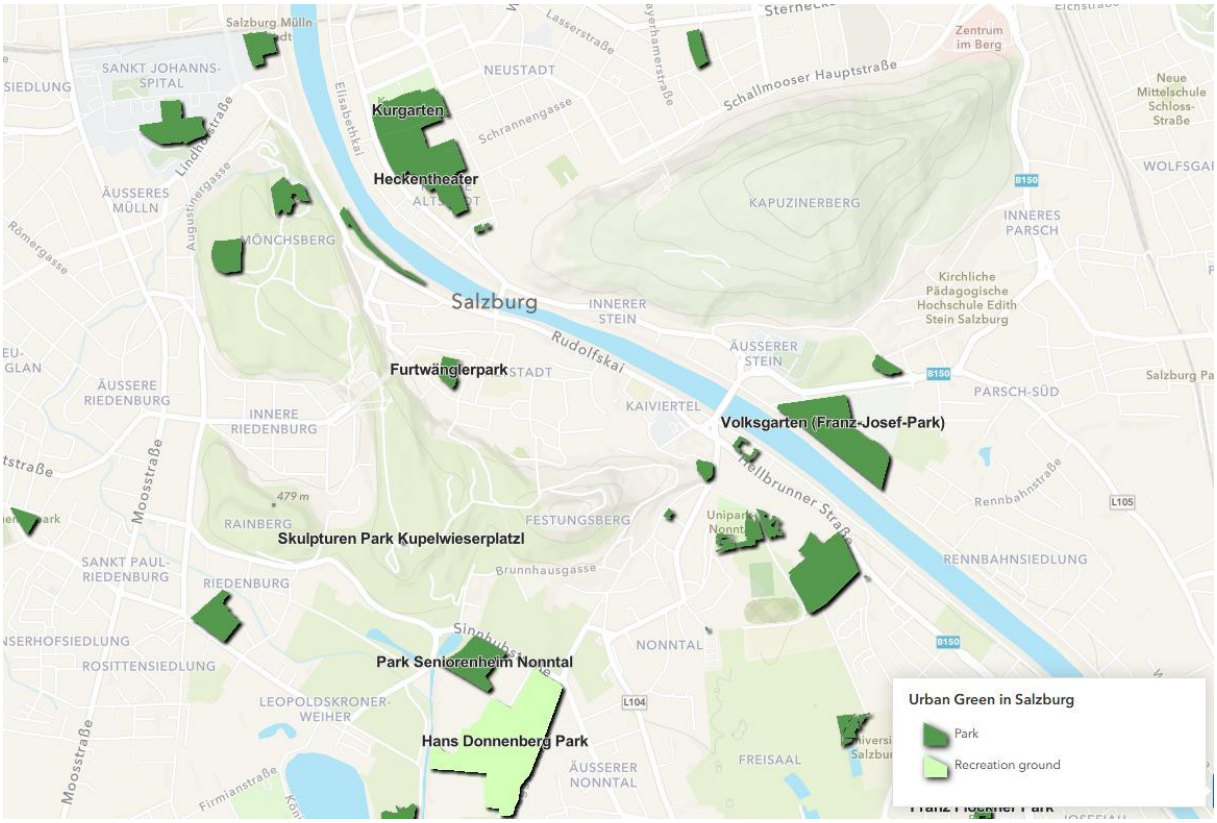

Figure 1. *Map of Salzburg showing urban green in the form of parks and recreation grounds.*

The aim of this study was to test the effectiveness and user-friendliness of this cartographical output. A special focus was placed on the use of the labels and the legend. In this context, the participants had to complete three different tasks, which were either clickable or included a survey question. The first task was a training task in which the participants were asked to get an overview of the content of the map. After that, the participants were asked to weigh up the different sizes of the features shown and choose the green area consisting of the largest area. The last question was aimed at filtering between the two feature types in this map and then clicking on a recreation ground.

Table 1. *Average accuracy data of the participants.*

E-T Data Grade	Eye-tracker Log	E-T Data Integrity	Gaze On Screen	E-T Sampling Rate	"Gaze vs Click" Accuracy
Perfect	OK	100%	100%	31 Hz	83 %
Very Good	OK	100%	94%	33 Hz	N/A
Good	OK	100%	88%	31 Hz	76 %
Average	OK	100%	79%	71 Hz	71 %
Very Good	OK	96%	100%	29 Hz	82 %

In total, the study was conducted by five participants aged between 23 and 28. There were two male and three female participants. The overall quality grade of the participants ranged from *good* (3) to *perfect* (6). This resulted in a *very good* average eye tracking data grade and an average "Gaze vs. Click" accuracy of 70 %.

Results

Analysis of the survey questions

In order to better understand the results of the eye tracking, it is important to link these to the answers to the survey questions. There were two survey questions, the first one being a text-based question and the second one a multiple-choice question.

Table 2. *Survey results. Question 1 asked about the features shown on the map and question 2 asked about the biggest green area shown on the map.*

ID	Age	Gender	Question 1	Question 2
1	28	female	Parks and recreational grounds	Hans Donnenberg Park
2	23	female	Salzburg	Volksgarten
3	28	male	Parks polygons	Volksgarten
4	23	male	Heckentheather	Kurgarten
5	26	female	Building	Hans Donnenberg Park

When analyzing the results of the first question, the first thing that stands out is that only one person used the terminology described in the legend. This leads to the assumption that the other participants did not pay as much attention to the legend on the map as they did to the general features.

(a) Heatmap of all participants for Question 1

(b) Heatmap of participant 5 for Question 1

Figure 2. *Heatmap with all participants (a) and participant 5 (b) for the first question.*

When looking at the overall heatmap for the first question with all participants, it becomes clear that there are many different hotspots that mainly focus on certain features of the map. At the same time, one hotspot can be localized to the map legend. This hotspot differs when looking at the individual heatmaps, as can be seen in figure 2 (b), where the participant did not take the legend into account when looking at the map. This becomes even clearer as soon as an area of interest (AOI) is drawn on the legend. Analyzing the AOI metrics per participant, it becomes apparent that only participant 1 fixated on the legend, which resulted in the most accurate answer.

For the second survey question, the participants chose different answers. This result was to be expected, as the selected areas appear to be roughly the same size, although the largest green area shown is *Hans Donnenberg Park*.

(a) Heatmap of all participants for Question 2

(b) Heatmap of participant 5 for Question 2

Figure 3. *Heatmap with all participants (a) and participant 5 (b) for the second question.*

The heatmap for question 2, showing all participants, illustrates that different features were considered while focusing on the largest area. Nonetheless, the densest hotspot can be seen near the correct

answer of *Hans Donnenberg Park*. At the same time, little to no hotspots were found near the other two answer options, *Kurgarten* and *Volksgarten*. This raises the question of why three out of five participants chose a different answer from *Hans Donnenberg Park,* even though the largest hotspot was identified there. Analyzing the individual heatmaps as shown in figure 3 (b) leads to similar questions, since the only relevant hotspot of labeled green spaces is found at *Hans Donnenberg Park* and not in the other parks.

Table 3. *Overall statistics of all participants for the AOIs for the second question.*

AOI	Fixations [total count]	Fixations [ms]	Gazes [total count]	Gazes [ms]
Hans Donnenberg Park	16	1209	295	5980
Kurgarten	3	549	50	1107
Volksgarten	0		11	252

The results of the heatmaps can be validated when compared to some of the overall statistics for the AOIs drawn. The AOIs were drawn according to the three identified answers to the survey question. When looking at the overall statistics, it is clear that the *Hans Donnenberg Park* AOI was fixated the most and for the longest time. While *Kurgarten* was also slightly fixated, *Volksgarten* was not fixated at all. Only a few gazes were detected here, with the total time looking being significantly lower (252 ms) than at *Hans Donnenberg Park* (5980 ms).

Consequently, it is difficult to link the answers to the second survey question with the results of the heatmaps and AOI metrics shown. It is quite unclear why the majority of the participants chose a different answer from the proven hotspot of *Hans Donnenberg Park*. This could indicate that the quality of the map and in particular the labeling of the highlighted features is insufficient.

Analysis of the clickable question

In addition to the two survey questions, the participants were asked to identify the location of a recreation ground on the map and click on its feature. In doing so, the participants had to understand the legend to know which features were shown on the map.

(a) Heatmap of all participants for the clickable question

(b) Heatmap of participant 5 for the clickable question

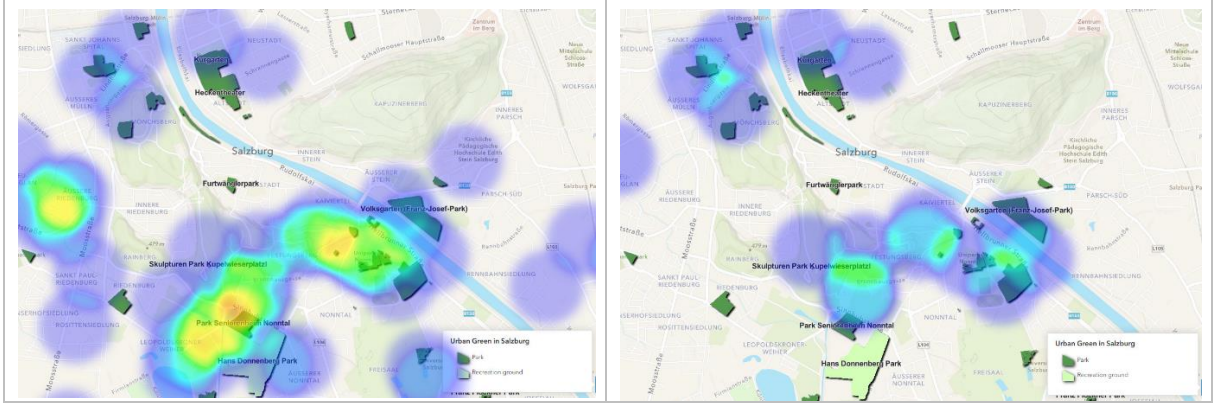

Figure 4. *Heatmap with all participants (a) and participant 5 (b) for the clickable question.*

The heatmap of all participants shows a clear indication of a hotspot near to the area of the recreation ground which had to be clicked during this task. At the same time, however, it becomes clear that other areas were also highly viewed, especially by individual participants, as shown in figure 4 (b). In this context, it makes sense that the legend was viewed, since it contributes to successfully clicking on the right area.

Table 4. *AOI statistics per participant for the clickable question.*

ID	Visits [total count]	Click [total count]	Click time [ms]	"Gaze vs. Click" accuracy
1	1	1	2343	83
2	4	0	0	0
3	0	1	3817	69
4	0	0	0	0
5	0	1	5753	72

Evaluating the AOI statistics per participant clarifies the results of the individual heatmaps. The count of the visits to the AOI shows that two participants (3 & 5) who clicked on the correct answer did not actively visit the AOI. This could be directly related to the "Gaze vs. Click" accuracy, since this value was lower for participant 3 and 5 than for participant 1. No click could be detected for participant 2 and 4, although participant 2 visited the *Hans Donnenberg Park* AOI the most. A possible explanation for this could be that the two participants were the only mobile users in this survey.

Conclusion

In conclusion, the results were largely satisfying and demonstrated the effectiveness and usability of the map. It became clear how the participants interacted with the cartographic output and were able to complete the tasks. Regarding the legend and labeling, not all participants used these two features effectively, which sometimes led to different answers.

At the same time, some results were unexpected and could be related to insufficient accuracy of eye tracking or cartographic output. Especially the lack of fixations, gazes and hotspots on features that participants considered as the largest green area was a difficult result to interpret. Similarly, it was unexpected that no fixations were detected by some participants on the last question, even though they clicked on the correct feature. This may indicate that eye tracking with desktop and mobile devices is not accurate enough for some tasks, although most of the results of this study were satisfactory.