**Summative Evaluation Report**

1. **Hypothesis, Tasks and Metrics**

The Hypothesis have been adjusted in this iteration according to plenum feedback from iteration 4.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Hypothesis** | **Metrics** | **Test Task** | |
| 1 | Users can save events within the details view of an event. The users do not make more than 2 errors in average for the given task. (H1: ) | Count the number of errors the user is doing while performing the task. | **Test Task 3:** You want to save that event you have shared. Please save the events in “My events”. |
| 2 | Users do find the CTA for sharing an event and it takes not more than 10 seconds to complete the task.  (H2: ) | Count the time until the user clicks the sharing CTA. | **Test Task 2:** Tomorrow you will be going on a date, and you are interested into finding coffee shops near your area. Search for events with “Cafe”. As you want to ask your date if this place is ok, you are trying to share this event. Please try to find the share button |
| 3 | Users can remove the event from “My Events List”. As we are assuming that this CTA is easy to find, we assume the users do not make more than 2 errors in average for the given task.  (H3: ) | Count the number of errors the user is doing while trying to remove the saved event. | **Test Task 4:** Now your friend gave you feedback and she does not like the event. Please try to remove the event from your Saved Events-List. |

1. **Data Analytics – Results**
   1. **Descriptive Data of test participants**

For the Evaluation Testing 11 Participants took part. However, for one user it was found out during the test to be not an appropriate candidate. Due to language barrier from the user the decision was to remove the entry for the evaluation. Therefore, the following evaluation was done with 10 users as we excluded the test results for this mentioned user.

During the evaluation, 10 participants with test users from the target user group were tested. 5 female (50%) and 5 male (50%) participants took part of the evaluation. The age of the participants was between 19 and 38 years with an average value of . In addition, 7 participants (70%) were students, whereas 3 (30%) were employees.

* 1. **Total Task Duration**

Measuring the task-time gives an indication of the time-based efficiency of the service [1]. The time [s] is measured according to the time the participant takes to complete a task. [2] The completion time was measured for task 2. The average user needed four times longer to complete the task compared to what we expected . Since time is parametrical, we use the t-test for the statistical analysis. As the boxplot in Figure 1 shows, there is no outlier for the time within the data.



Figure 1: No outlier in the completion time for task 2

Furthermore, as the Shapirow-Wilk test shows, the time distribution does not significantly deviate from a normal distribution (). Thus, the requirements for the t-test are fulfilled. As the t-test shows, H1 cannot be confirmed. The users do not need significantly less than 11seconds performing task 2 (). In contrary, the t-test in the opposite direction shows, that users do need significantly more than 10s to perform the task ().

* 1. **Error Rate**

The average number of errors for task 3 is lower than what expected (. However, the Wilcoxon test showed that participants do not make significantly less than 3 mistakes performing task 3 (). Thus, the result is not consistent with H1.

The average number of errors for task 4 is also lower than what expected (. Again, the H3 cannot be confirmed, as the Wilcoxon test showed that participants do not make significantly less than 3 mistakes performing task 4 ().

“Measuring the Error Rate gives often the “Why” behind longer task times.” [3] Therefore, we counted the number of errors each test user performed, and we noted down for task 3 what the participant made inaccurately to achieve the required task.

As additional information during the evaluation, 9 out of 10 test users did not use the search field but started scrolling down immediately. Moreover, 4 out of 10 participants did not understand the icons used for a coffee shop in the “Map” overview from the service.

1. **Discussion and Conclusion**

Hypothesis 1:

For the first hypothesis, the results from the evaluation showed that the theory had to be rejected. The number of errors was lower than expected when saving an event. However, the Wilcoxon test showed, that this difference was not significant. That could be because of the test setup. First, the users had to share the event. There is a function provided by the standard system, that allows users to save the link. However, this is not the intention of “saving” an event. Another reason could be that the hypothesis is too strict, allowing only two mistakes. As all users were novel to the app, they were still exploring the different sections.

Hypothesis 2:

The time required to share an event using the “Map” option showed the second hypothesis to be rejected. In average the participants needed 40.6 sec to fulfil the task. Therefore being 30.6 seconds in average over the proposed time. This could be that while searching for the desired event five out of 10 participants still took time to explore the events and find one that “best suited them” before sharing it with someone. Upon these cases the testers did not interfere to prevent a disruption in the user experience during the task. One solution would be to separate the tasks. Thus, one task would be searching for an event on the map. The second task, then would focus on the share function and is measured as soon as the users have opened an event, they are interested in.

Hypothesis 3:

The third hypothesis, for which the participants had to remove their saved event from “My Events” was rejected as well. The Wilcoxon test showed that the users did not make significantly less mistakes than expected. Participants had problems understanding how to remove one event once it has been added to the personal events list. 6 out of 10 participants had problems once they were in the “My events” page not knowing where to click to remove the event. They expected, to be able to directly remove it from “My Events”, and not only after clicking on the event. Another reason was, that the users often scroll to the bottom of the event, overlooking the remove button. This could be improved by adding additional buttons to the event cards on “My Events”, or by making the button in the event detail page more salient.

1. **References**

[1] *Usability Metrics*. (2020). Usability Geek. Abgerufen am 7. Februar 2022, von <https://usabilitygeek.com/usability-metrics-a-guide-to-quantify-system-usability/>

[2] Jeff Sauro. (2011). *10 Things To Know About Task Times*. Measuring U. Abgerufen am 7. Februar 2022, von <https://measuringu.com/task-times/>

[3] Jeff Sauro. (2012). *Measuring Errors in the User Experience*. Measuring U. Abgerufen am 7. Februar 2022, von https://measuringu.com/errors-ux/