Quantitative Evaluation - Hypotheses, Test Cases, Metrics

Hypotheses:

1. Subjects do not make more than 4 errors in total when conducting the tasks [1,2,7,9].

2. Subjects complete tasks 1-5 with a completion rate of 78% [2,5].

3. Subjects do exceed a SUS-Score of 70 regarding the complete application [1,2,6].

4. The average Task Level Satisfaction does exceed a value of 5 [9].

Test Cases:

1. Take a look into your pantry. Search for a recipe with your ingredients: pumpkin, onions, garlic, salad.

2. You try to avoid dishes that contain meat. Open the app and set that you want to receive only vegetarian recipe suggestions.

3. Check the main screen and look for a recipe with pumpkin as a main ingredient. Add this recipe to your favorites. Where do you find it now?

4. You are running out of pumpkin. Try to remove it from your pantry.

5. You want to cook the pumpkin soup again. Is there a possibility to replace the hokkaido pumpkin?

(settings of the tasks before have to be set again in the next task e.g. set vegetarian diet also in tasks 3-5)

Metrics:

1. **Number of errors**

   → wrong clicks, too many clicks, wrong tasks executed [1-3,9,10].

2. **Completion Rate** to measure effectiveness of the application

   The completion rate is calculated by assigning a binary value of “1” if the test participant manages to complete a task and “0” if he/she does not.The effectiveness can be calculated by measuring the completion rate [2,5].

$$Effectiveness = \frac{number\ of\ Task\ completed\ successfully}{total\ number\ of\ task\ undertaken} \times 100\%$$

3. **Overall SUS-Score** for measure the usability of the application [1]

4. **Task Level Satisfaction**
   After users attempt a task they will be given a questionnaire so as to measure how difficult that task was. For this, the SEQ (Single Ease Question) will be used [1,2,8,9].

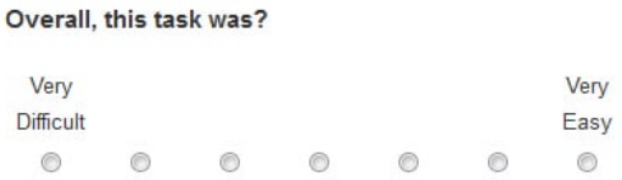

***Figure 2**: SEQ (Single Ease Question) which will be used to measure the task level satisfaction [2]*

Quantitative Evaluation - Setting

User Test Setup:

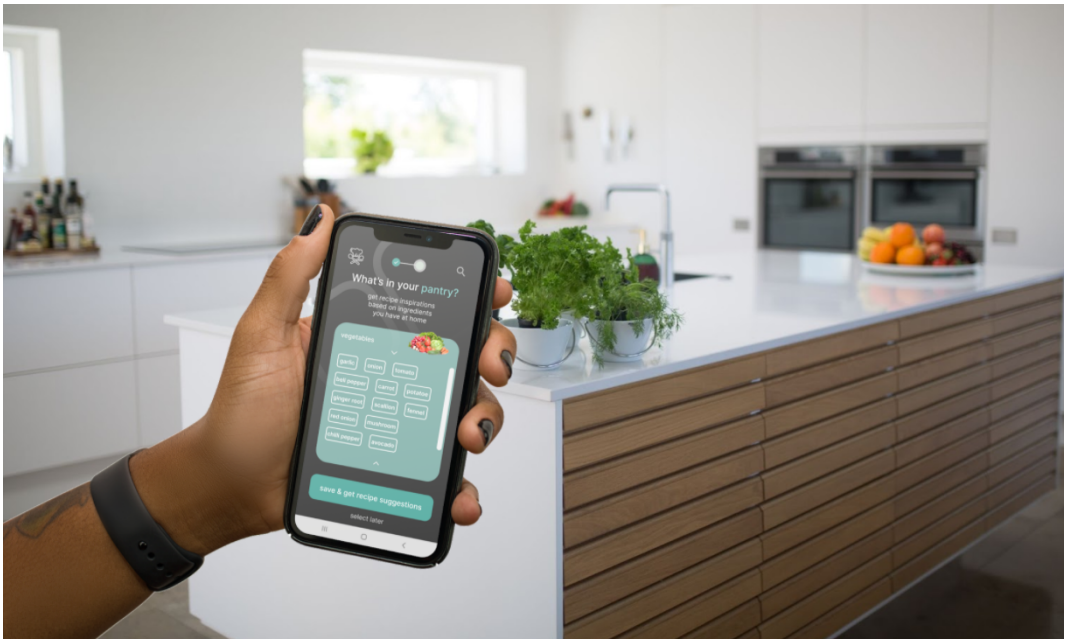

***Figure 2**: User Test Setup: User will be in the kitchen, checking his pantry for ingredients and meanwhile testing the app FOODIYO.*

- Data recording with smartphone camera (screen and voice recording) [4].
- Online or live setting due to covid-pandemic [1,4].
- Used methods: thinking aloud method, video recording & questionnaire

Procedure:

1. Subjects are welcomed and briefly introduced to the topic of the study

2. Permission for recording is obtained, screen and voice is recorded [1,4].

3. Task description is shown via screen sharing or paper (online vs. live).

4. Subjects conduct task 1-5 without any advice (app is started again before every task) by using the thinking aloud method [1,3].

5. Afterwards subjects are introduced to fill out the SUS questionnaire with regard to the usability of the whole application [1].

6. After the study: recordings are analyzed in terms of errors, completion rate and SUS as well as SEQ score and compared with the hypotheses [1].

References

[1] Nielsen, Jakob. (1993). Usability Engineering. chapters 6.7, 6.8, 7.1, 7.2

[2] Sauro, Jeff. (2010). A Practical Guide to Measuring Usability.

[3] Lewis, Clayton. (1982). Using the "Thinking-aloud" Method in Cognitive Interface Design. S.1-6

[4] Mackay, Wendy & Janecek, Paul. (2000). Video artifacts for design: Bridging the gap between abstraction and detail. 10.1145/347642.347666.

[5] “10 Things To Know About Completion Rates”, https://measuringu.com/completion-rates/, access on 18.01.2022

[6] Aaron Bangor, Philip T. Kortum & James T. Miller (2008) An Empirical Evaluation of the System Usability Scale, International Journal of Human–Computer Interaction, 24:6, 574-594, DOI: 10.1080/10447310802205776

[7] Don, Norman (2013). The Design of Everyday Things

[8] “Efficiency”, http://ui-designer.net/usability/efficiency.htm, access on 17.01.2022

[9] “Usability Metrics – A Guide To Quantify The Usability Of Any System”, https://usabilitygeek.com/usability-metrics-a-guide-to-quantify-system-usability/, access on 17.01.2022