Quantitative Evaluation

1.1 Hypotheses

The following hypotheses were made:

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

2. Subjects complete tasks 1-7 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].

1.2 Tasks

The following tasks should be conducted by the user:

1. Task: Search for a recipe
   Take a look into your pantry. Search for a recipe with your ingredients: pumpkin, onions, garlic, salad. Open the recipe and pretend you are cooking it now.

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

3. Task: Like a recipe
   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. Task: The pumpkin is somehow lost…
   You are running out of pumpkin. Try to remove it from your pantry.

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

6. Task: You bought a tomato
   You went shopping for tomatoes. Add "tomato" to your digital pantry within the Foodiyo App.

7. Task: Watching a recipe video & reading the recipe steps
   Check out the pumpkin soup recipe again. Read the recipe steps and play the recipe video.

1.3 Metrics

The following metrics are measured:

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]*

1.4 User Test Setup:

***Figure 3**: User test setup of 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: video recording & questionnaire

1.5 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 1-7 without any advice [1,3]. After each task they have to answer a “single each question” (SEQ) to measure the task level satisfaction.

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

2. Results

2.1 Demographic Data

A total of 10 people took part in the user testing. 7 men and 3 women were acquired. The test participants were between 21 and 32 years old (*MW* = 27, *SD* = 3.92). The majority of the subjects were students. Based on their profession, for example as IT specialist or software developer, as well as marketing manager, it can be concluded that at least some of the test persons have experience with certain software tools.

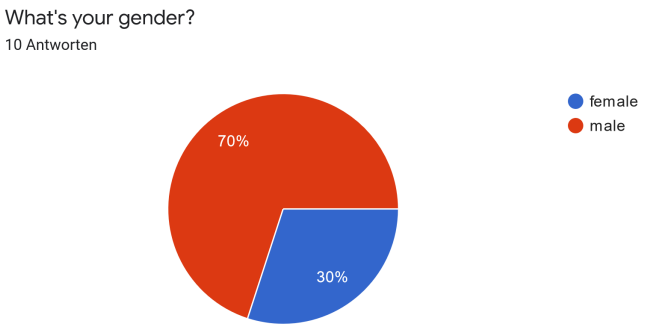

***Figure 4**: Gender distribution*

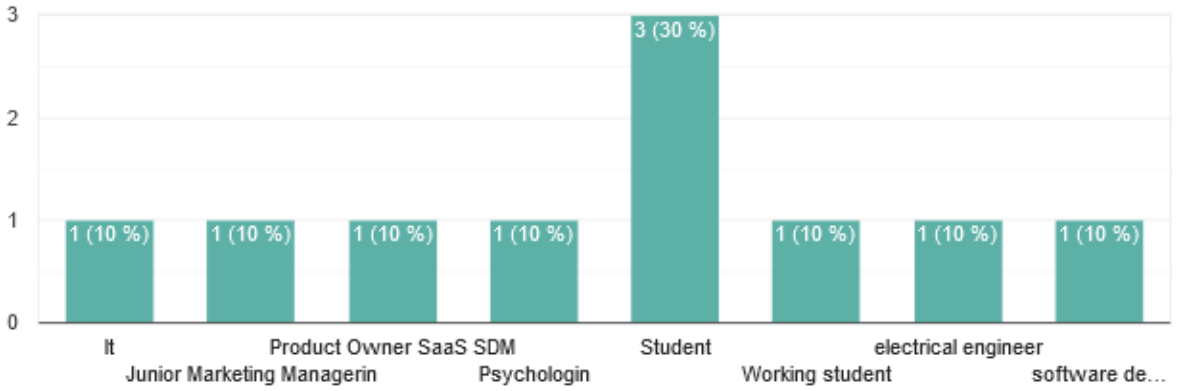

***Figure 5**: Reported professions of the test participants*

2.2 Observations during the user tests

The subjects had to conduct 7 tasks in total. The following observations were made:

- Task 1: Search for a recipe

  There was confusion about how to add the ingredients. It took some time for the user to open the category containers. For some users it also took some

time to find the button “save & get recipe suggestions”. Highlighting this button with a different color would have been helpful. One of the subjects wanted to click on another recipe within the home section to add it to the digital cookbook within the foodiyo app. However, this function is not implemented yet.

- Task 2: Avoid dishes that contain meat

  It was unclear to the user if the setup steps have to be conducted every time the user wants to change the diet type.

- Task 3: Like a recipe

  Most of the users stated that the process for this task is simple and easy to follow. Some of the users wanted to like other recipes, e.g. the couscous recipe instead of the pumpkin soup recipe. However, this feature is not fully implemented.
  For one of the subjects the cookbook icon in the navigation bar was confusing. He expected a larger “heart” icon in the navigation bar. Also a confirmation after adding the ingredient to the favorites would have been nice. Furthermore, there were confusions about where to find the favorite recipe.

- Task 4: The pumpkin is somehow lost…

  The function of removing an ingredient from the digital pantry was considered as very intuitive by most of the users. During two user tests the subjects did not click on the “save & update recipes” button at the end of the task. Highlighting this button would have been more supportive to successfully complete the task.
  Two users were asking if it is possible to also set the amount of ingredients, e.g. there is ½ pumpkin left.

- Task 5: I want to cook the pumpkin soup again!

  One of the users wanted to open the recipe from the digital cookbook. However the recipe was not linked (function not fully implemented). Choosing the pumpkin recipe is only possible from the home menu. There were confusions about deselecting the ingredient. An “X” or a hint on the ingredient symbol would be helpful to know that it can be removed from the recipe. Right now this feature seems to be a bit too hidden.

Moreover, one of the subjects was confused about the button “remove ingredient” on the overlay screen (figure 4). He would expect to also have a button “replace ingredient”.

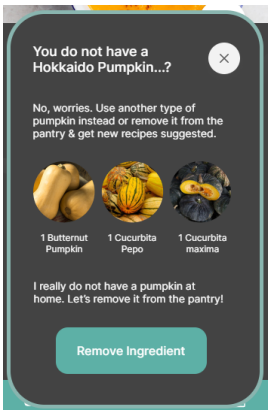

***Figure 6**: Overlay Screen for replacing the hokkaido pumpkin by another one.*

- Task 6: You bought a tomato

  Some subjects tried to click on the different sections (e.g. fruit section) to add an ingredient. The process of the task was unclear. A dropdown menu with input support probably would have been helpful. Also it took some time for users to find the add button. Highlighting this button would have been helpful. Some users also mentioned that an individual plus-button for each category section would have been even more logical.

  Most of the users would have liked to get suggestions for the predefined ingredients when entering the first letters of the ingredient. So, while entering the new ingredient an input support would have been helpful [10], e.g. a dropdown menu or suggestion box for ingredients (figure 5). One of the users mentioned that being able to set the amount of the ingredients would also have been important.

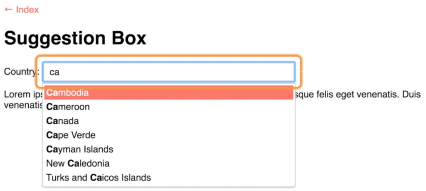

***Figure 7**: Example of a suggestion box for supporting the user while entering a new ingredient [10]*

- Task 7: Watching a recipe video & reading the recipe steps

  For most of the users the play button was not well visible. Highlighting this button would have been necessary.

- Further Feedback

  In general most of the users liked the feature to get recipes suggested based on the ingredients they have got at home. Also they appreciate the feature of selecting alternative ingredients.
  Having a digital shopping list to know what to buy for certain recipes would also be a very helpful feature. One user suggested also entering their own recipes and sharing them with friends would have been awesome. Furthermore, the nutrition information for the dishes and a filter option based on nutrition values would have been great.

  Some users stated that entering all the ingredients they have got at home would be very time consuming.

2.3 Errors

Errors can be unintended actions, slips or mistakes that a user makes while attempting a task. During the user tests the number of errors the participants made were counted for each task [9].

Descriptive Statistics

	Valid	Missing	Mean	Std. Deviation	Minimum	Maximum
Error	10	0	2.100	1.197	0.000	4.000

***Figure 8**: Descriptive statistics of the user errors*

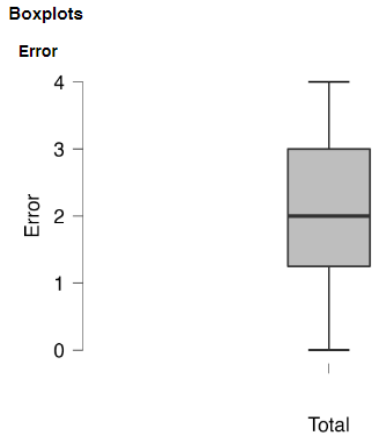

***Figure 9**: Boxplot of the user errors*

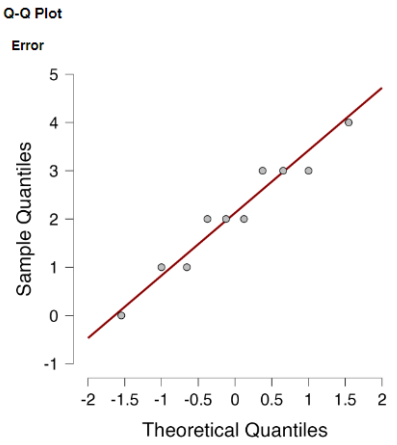

***Figure 10**: Q-Q Plot of the user errors*

H0: Subjects do make more than 4 errors in total when conducting the tasks.
H1: Subjects do not make more than 4 errors in total when conducting the tasks.

→ *M* = 2.10, *SD* = *1.97*
→ H0 can be rejected.

2.4 Completion Rate

Mean effectiveness for all subjects:

$$Effectiveness = \frac{67}{70} \times 100\% = 95,71\%$$

H0: Subjects complete tasks 1-7 with a completion rate of 77% or less.
H1: Subjects complete tasks 1-7 with a completion rate of 78%.

→ *Effectiveness* = 95.71%
→ H0 can be rejected.

2.5 System Usability Score (SUS)

Descriptive Statistics

	Valid	Missing	Mean	Std. Deviation	Minimum	Maximum
SUS Score	10	0	86.750	10.675	70.000	100.000

***Figure 11**: Descriptive statistics of the System Usability Score (SUS)*

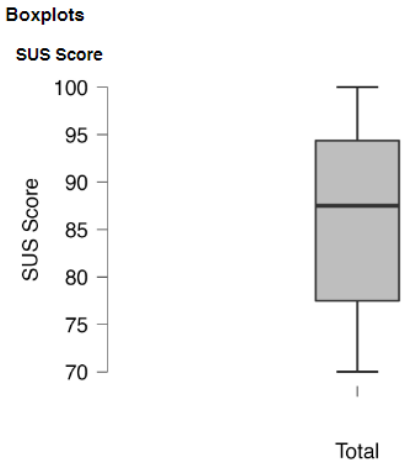

***Figure 12**: Boxplot of the System Usability Score (SUS)*

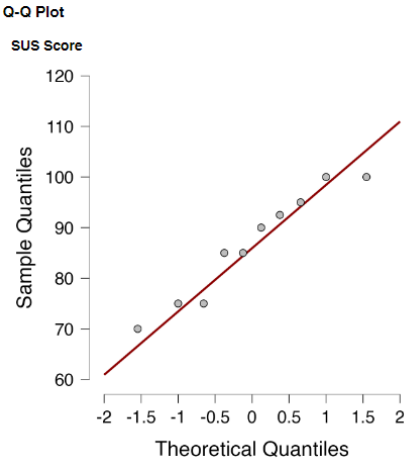

***Figure 13**: Q-Q Plot of the System Usability Score (SUS)*

H0: Subjects do not exceed a SUS-Score of 70 regarding the complete application.
H1: Subjects do exceed a SUS-Score of 70 regarding the complete application.

→ *M* = 86.75, *SD* = 10,68
→ H0 can be rejected.

An SUS score of 86.75 was achieved. This value reflects an acceptability rating of "acceptable" and an adjective rating of "excellent". Thus, the value exceeds the minimum value of 70 expected in the hypothesis, which means that the app has no critical usability deficits [6].

2.6 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. The Single Ease Question, (SEQ) a single-question post-task questionnaire measures users' perception of usability based on the last attempted task. Since the task is still fresh in the participants' minds, their answer provides a useful assessment of the experience for that particular task [1,2,8,9].

Descriptive Statistics

	Valid	Missing	Mean	Std. Deviation	Variance	Minimum	Maximum
Task1	10	0	6.200	0.789	0.622	5.000	7.000
Task2	10	0	6.900	0.316	0.100	6.000	7.000
Task3	10	0	6.300	0.823	0.678	5.000	7.000
Task4	10	0	6.400	0.516	0.267	6.000	7.000
Task5	10	0	6.000	1.247	1.556	3.000	7.000
Task6	10	0	5.400	1.265	1.600	3.000	7.000
Task7	10	0	5.200	1.687	2.844	1.000	7.000

***Figure 14**: Descriptive statistics of the tasks 1 -7*

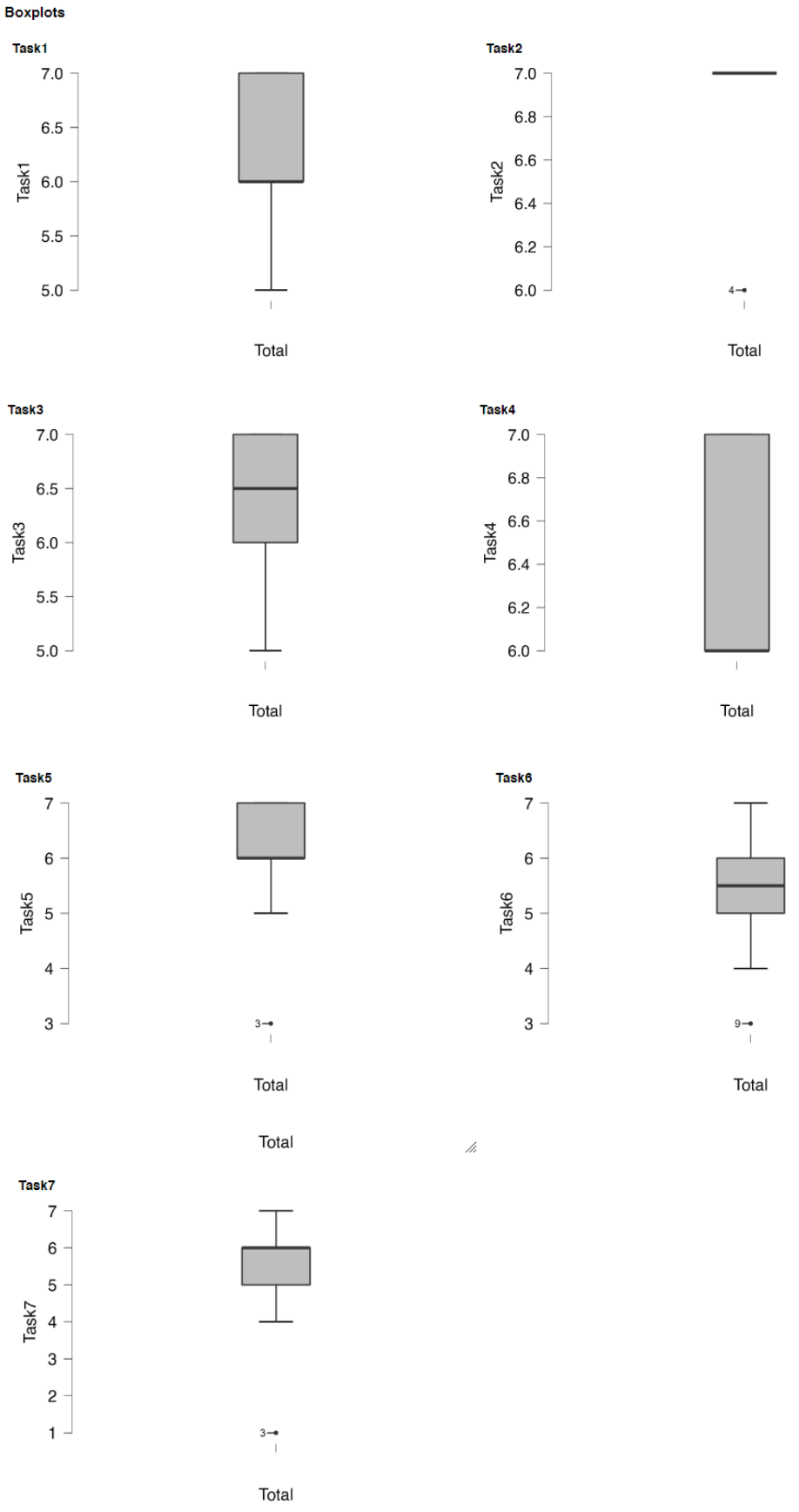

***Figure 15**: Boxplots of the tasks 1 -7*

H0: The average Task Level Satisfaction does not exceed a value of 5.
H1: The average Task Level Satisfaction does exceed a value of 5.

→ M = 6,06
→ H0 can be rejected.

The mean Task Level Satisfaction across all test subjects and tasks reaches a value of 6.06. Thus, the app is rated as easy to use [9]. The range of values extends from 5.2 to 6.9, which means that even the lowest mean value of the individual test subjects exceeds the minimum value of 5 expected in the hypothesis.

3. Discussion

The Foodiyo App was perceived as very intuitive, which is also confirmed by the excellent SUS score. The appealing design in particular was repeatedly praised. The concept of the app in general was also rated as very helpful and useful. Users liked the possibility of entering existing ingredients and receiving recipes based on them. In addition, the option that alternative ingredients are suggested was rated as very useful. The categories for categorizing ingredients were also helpful. According to the data analysis, almost all processes were successfully completed without any major problems and very few mistakes were made. The ease of use and good usability are reflected in the results of the task level satisfaction.

However, some confusion arose especially regarding the small range of functions, since some features were not yet fully implemented. In addition, some buttons were only found after some time or were not pressed at all. This could be optimized by an improved ergonomic design.

In addition, the criticism that the input of ingredients is currently still very time-consuming is quite justified. This could possibly be remedied as the quality of speech recognition progresses. Furthermore, the introduction of the function of being able to create recipes yourself and share recipes with friends could be added. Since healthy nutrition is becoming an increasingly important part of everyday life for many at the moment, the display of nutrition values would also be quite advantageous.

Implications for the ergonomic design of the app
For the future development of the app, for example, a more conspicuous design of buttons could be considered, as these were often not immediately recognized. Regarding the design of the cookbook button, an enlargement of the heart icon or the mere use of a heart button in the nav bar could be discussed, as it is already used in some other apps. There should also be a short confirmation in the form of a popup about the successful addition of the recipe to the favorites.

Another important feature that could be implemented is the ability to set amounts already when adding ingredients. The function to remove ingredients from the recipe should be indicated in a more prominent way, by an additional "x", for example. For the addition of ingredients directly in the pantry, a different placement or marking of the plus button should be considered, and an auto-complete function in the form of a drop-down menu should be introduced when manually entering ingredients. The play button of the cooking video should also be made more conspicuous, possibly in a different color.

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

[10] “Suggestion Boxes”, https://pressbooks.library.ryerson.ca/wafd/chapter/suggestion-boxes/, access on 27.01.2022