# Assignment 3 (Phase 4)

IC1007 HT17-1

Human-computer Interaction: Principles and Design

**Group 11**

961021-6849 Emelie Tham
960707-4987 Mila Grancharova
931002-1440 Pournami Krishnan Krishnakumari
940920-7850 Siddhant Gupta

## Table of Contents

# 1 Introduction

An evaluation has been conducted on the "My Diet" app prototype by a team of four evaluators. The evaluation method used is described in section 2. We performed 2 iterations of testing and revising the prototype. Section 3 presents the first iteration, conducted on two test subjects. Section 4 presents the second iteration, also conducted on two subjects. For each iteration, we present the current state of the prototype (if any changes have been made), the test, the results and suggestions for improvements. In section 5, the final results base on both iterations are presented, along with recommendations for improvements for the finished system. In section 6, we reflect upon the evaluation.

# 2 Method

This evaluation is based on the testing method for lo-fi paper prototypes described by Marc Rettig (Rettig, Prototyping for Tiny Fingers, 1995). The method consists of three stages, preparing for the test, conducting the test, and evaluating the results.

The preparation phase starts off with selecting users to conduct the test on. These people should represent the intended user group of the product. The test subjects should make up as wide a variety as possible of the user group. Next, the test scenarios are prepared. The scenarios are tasks that the user should perform using the prototype. These tasks don't need to be many, as the prototype has limitations, but broad enough to allow conclusions about the design to be drawn from them. The last part of the preparation phase is to practice. The test team practices performing the tasks they decided upon and make sure that they all know what to do in a real testing situation.

The tests are performed with one user at a time. When conducting the test, the test team, which consists of four people, divide themselves into different roles. The roles are greeter, facilitator, computer and observer. The greeter's job is to welcome the user. The facilitator instructs the user with the tasks they should perform and encourages them express what they are thinking while performing them. The computer moves parts of the interface around when the user interacts with the prototype, simulating the finished product. The observer takes notes of the problems that arise, as well as any solution they can think of, without speaking in order not to affect the user.

After the testing is done, the results are evaluated. The evaluation team organizes and prioritizes the notes taken by the observer during the tests. The notes set the base for a written report of test results. This report is then discussed and the team agrees on what changes to make. A revisited prototype is then created, following the changes agreed on. This ends one iteration of testing.

This evaluation differs somewhat from the method described above because the prototype is computerized, rather than the paper prototype Retting's method was designed for. This means that there is no person acting the role of computer during the tests.

Retting does not explain how exactly the prioritization of usability problems (noted by the observer) should be done. In this evaluation, we first rate the problems by severity and then give suggestions for improvement for the problems we rated highly. The severity rating system is loosely based on Nielsen's severity rating for heuristic evaluation *(Nielsen, Severity Ratings for Usability Problems, 1995)*. The system is based on rating 1-4:

"**1** = Cosmetic problem only: need not be fixed unless extra time is available on project"
"**2** = Minor usability problem: fixing this should be given low priority"
"**3** = Major usability problem: important to fix, so should be given high priority"
"**4** = Usability catastrophe: imperative to fix this before product can be released"

The number 1-4 is decided on by considering three factors; **frequency** (how often does the problem arise), **impact** (how easy is it for the users to overcome the problem) and **persistence** (how bothered will the users be by the problem).

Nielsen's original system also includes a 0 rating which means that it is not a usability problem at all. That is not applicable to this evaluation since any problem found by a user is a usability problem. Therefore the 0 rating is not used here.

# 3 First Iteration

In the first iteration of testing, the prototype was just as given by the designers (a ppt). Two tests with similar setup and the same test scenarios were performed on two subjects. In section 3.1, the tasks given to the subjects are presented in the order they were given. It also describes the setup of both tests. Section 3.2 describes the first test along with what results came out of it. Section 3.3 describes the second test and its results. In section 3.4, we give the conclusions that we arrived on after a discussion of the test results, as well as suggestions for improvement.

## 3.1 Setup and Test Scenario

The setup of the test was as described in the methodology in section 2. The environment is shown in Image 1.

Image 1: Test Subject testing the app prototype.

1. Create an account and set it up.
2. Check out details about “bagel melt” and go back to “My Diet” page.
3. Edit diet and try to replace “cooked chicken” with another meal.
4. Edit diet and try to replace ”bagel melt” with another meal.
5. Add a meal and go back to “My Diet” page.
6. Find the monthly analysis.
7. Check out your profile.
8. Log out.

## 3.2 First test

### 3.2.1 Subject

In the first iteration, the prototype was tested on a female elite swimmer in her early twenties. The test subject has experience with using android and iPhone applications, but has not used diet apps before.

### 3.2.2 Execution

The test was performed on 2017-10-12 and lasted for around 30 minutes. Another 30 minutes were used by the evaluation group to discuss and prioritize the problems that were mentioned by the test subject.

### 3.2.3 Results

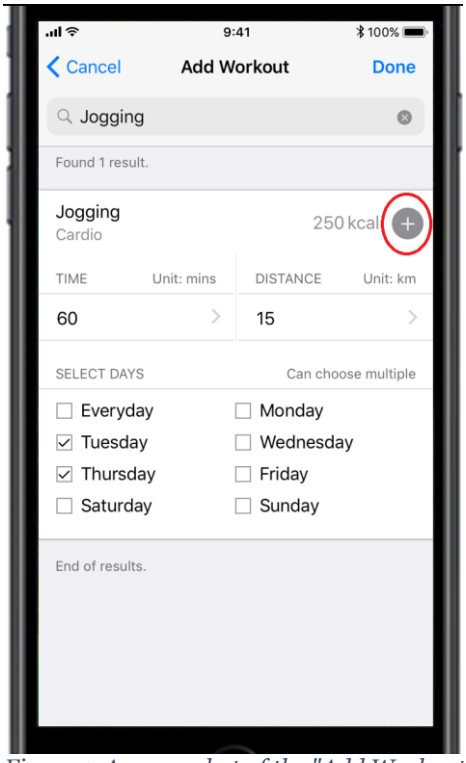


*Figure 1. A screenshot of the "Add Workout" screen.*

1. Create an account and set it up.
   - The user could not figure out how to add a workout without the help of the facilitator. She could not connect the action of adding a workout with the add button marked red in Figure 1.
   - The user wanted to have more options on the "Food Preference" page, such as options for supplements and other dietary preferences.

2. Check out details about "bagel melt" and go back to "My Diet" page.
   The user had no problems performing this task.

3. Edit diet and try to replace "cooked chicken" with another meal.
   - The user assumed that the "Edit Diet" option meant changing the meal's quantity, removing or adding meals, but not replacing the meal.

4. Edit diet and try to replace "bagel melt" with another meal.
   - The user wanted to have more recommended meals to choose from.

5. Add a meal and go back to "My Diet" page.
   The user had no problems performing this task.

6. Find the monthly analysis.
   The user had no problems performing this task.

7. Check out your profile.
   The user had no problems performing this task.

8. Log out.
   The user had no problems performing this task.

## 3.3 Second test

Section 3.3.1 describes the subject used in this test. Section 3.3.2 goes through the details of how the test was executed. In section 3.3.3, the raw results of the test are presented in order of appearance. Each issue the subject found is presented under the task they were performing at the time.

### 3.3.1 Subject

The subject is a female football player in her early twenties. She has a lot of experience using android and iPhone apps, as well as different diet apps in particular. She speaks English well.

### 3.3.2 Execution

The test was performed on 2017-10-12. The test took about 20 minutes to perform. The results were then discussed and prioritized by the evaluation team over the next 30 minutes. Another hour was then spent discussing improvements based on the results of both test in this iteration.

### 3.3.3 Results

1. Create an account and set it up.
   - The user didn't understand that she had to click the plus-button, shown in Figure 1, before she could click "Done".

2. Check out details about "bagel melt" and go back to "My Diet" page.
   - When looking at the main page, "My Diet", the subjects said: "Did I add all these meals?" She didn't understand that the app provided her with exactly what to eat. She thought that she has entered all meals she had, and the app only calculated if she was getting the right amount of all nutrients.

3. Edit diet and try to replace "cooked chicken" with another meal.
   - She thought the "Edit meal" function was frustrating. It didn't behave as she expected. She expected to be able to customize the meal from there.

4. Edit diet and try to replace ”bagel melt” with another meal.
   - The user found that it was frustrating to have to click the replace button several times until she found something she liked to eat.
   - If she found a replacement she liked but then continued asking for replacements to see what else she could have, she wouldn't be able to get back to the thing she liked.

5. Add a meal and go back to “My Diet” page.
   - The user found the way to add a meal counter-intuitive. She tried clicking in the meal itself and looking for a button there.
   - After a few seconds, she managed to find the “Add meal” option.
   - She thought that “Add meal” was not customized enough. She didn't like the search engine and would rather have typed in what nutrients her meal contained manually.

6. Find the monthly analysis.
   The user had no problems performing this task.

7. Check out your profile.
   The user had no problems performing this task.

8. Log out.
   The user had no problems performing this task.

### 3.4 Usability Problems and Suggestions for Improvement

Based on the results of the two tests conducted in this iteration the usability problems that were found are presented and improvements are suggested. They appear in order of the severity of the problem. The severity is rated by the evaluation team. Nielsen's severity rating system is used as a guideline for measuring the problem's severity *(Nielsen, Severity Ratings for Usability Problems, 1995)*:

“1 = Cosmetic problem only: need not be fixed unless extra time is available on project”
“2 = Minor usability problem: fixing this should be given low priority”
“3 = Major usability problem: important to fix, so should be given high priority”
“4 = Usability catastrophe: imperative to fix this before product can be released”

More detail on the severity rating system used is given in the methodology in section 2.

#### 3.4.1 Severity rating 4

- **Problem**: The problem concerns the user's understanding of the purpose of the app. We assume that the user knows that this is a diet app which should help them maintain their shape during the off-season, before they see the interface. What is lacking is something which tells the user that a diet which they should follow will be generated based on their specifications.
  **Possible solution**: It is suggested that this information is presented before a new account is set up, so that the user understands what they are looking at.

- **Problem**: Neither of the test subjects understood the purpose of the plus sign on the “Add workout” page and therefore failed to add a workout at first. It is unclear that the calories by the plus sign are dependent on the user's input and that they have to mark the plus sign and then click “Done” in order to add the workout.
  **Possible solution:** To make its purpose more obvious, the plus symbol should be changed to something like “Add this workout”. It should also be placed underneath the workout which is currently being created to show that it depends on it.

#### 3.4.2 Severity rating 3

- **Problem:** In the “Food preference” page of the setup, there is nowhere for the user to input *specific* foods they don't like.

**Possible solution:** Apart from the allergies entry, there should be a place to enter specific foods the user doesn't want in their diet, even if they are not allergic to it.

- **Problem:** Both users found the name "Edit diet" misleading and apparently didn't read the information on the page explaining what the edit function does.
  **Possible solution:** The function's name should be changed to something more self-explanatory, such as "Replace a meal".

- **Problem:** In the "Edit meal" function, the user doesn't get an overview of what replacements are offered and has to click repeatedly until they find something they like.
  **Possible solution:** The "Edit meal" function should show the different replacement options for a meal in a menu.

- **Problem:** The user can not add a custom meal with exact amount of nutrients if they know the amounts.
  **Possible solution:** There should be an option for users to add a custom meal, not only searching for a standard one. In the customized meal, the user can, if they want, add what quantities of different nutrients they had.

# 4 Second Iteration

In the second iteration, two tests with similar setup and the same tasks were performed on two subjects. In section 4.1, the tasks given to the subjects are presented in the order they were given. It also describes the setup of both tests. Section 4.2 describes the first test along with what results came out of it. Section 4.3 describes the second test. In section 4.4, we give the conclusions that we arrived on after a discussion of the test results, as well as suggestions for improvement.

In the second iteration, the same prototype was used as in the first iteration, except some changes based on the findings of the first iteration, presented in section 3.4. Not all issues that were found in the first iteration were fixed in the revisited prototype, but the most important ones (according to the given rating) were. The changes in the prototype are shown in Figures 2-6.

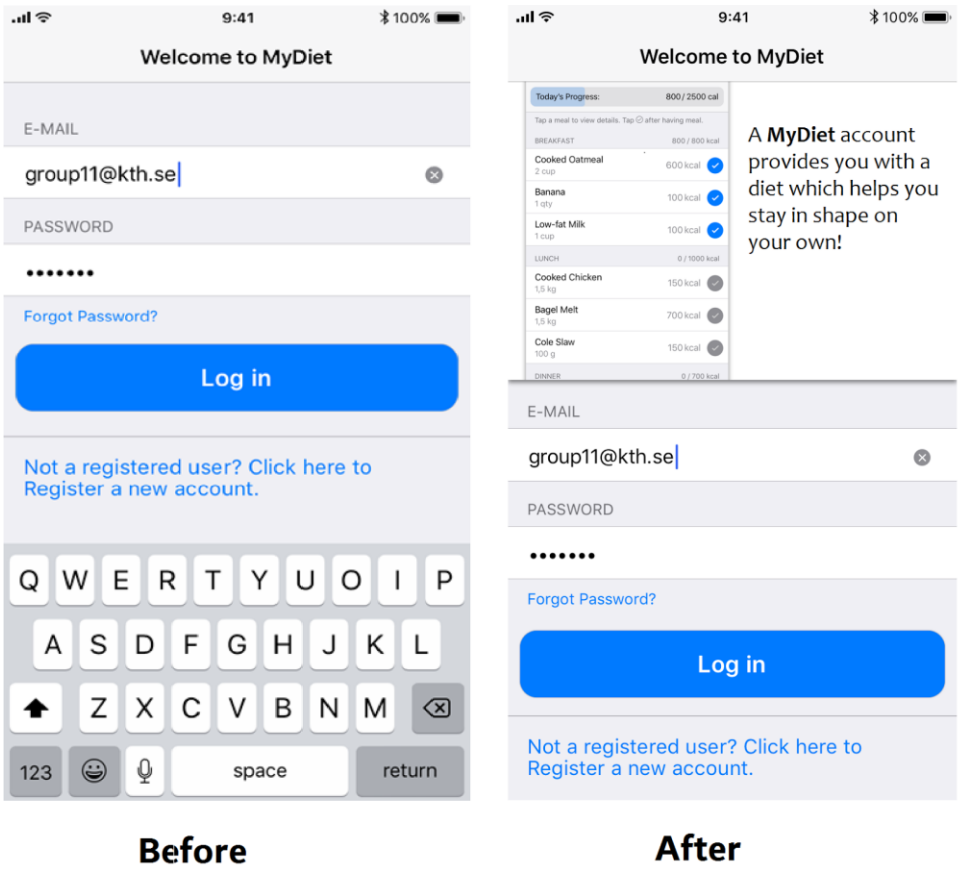


*Figure 2. Changes to the front page.*

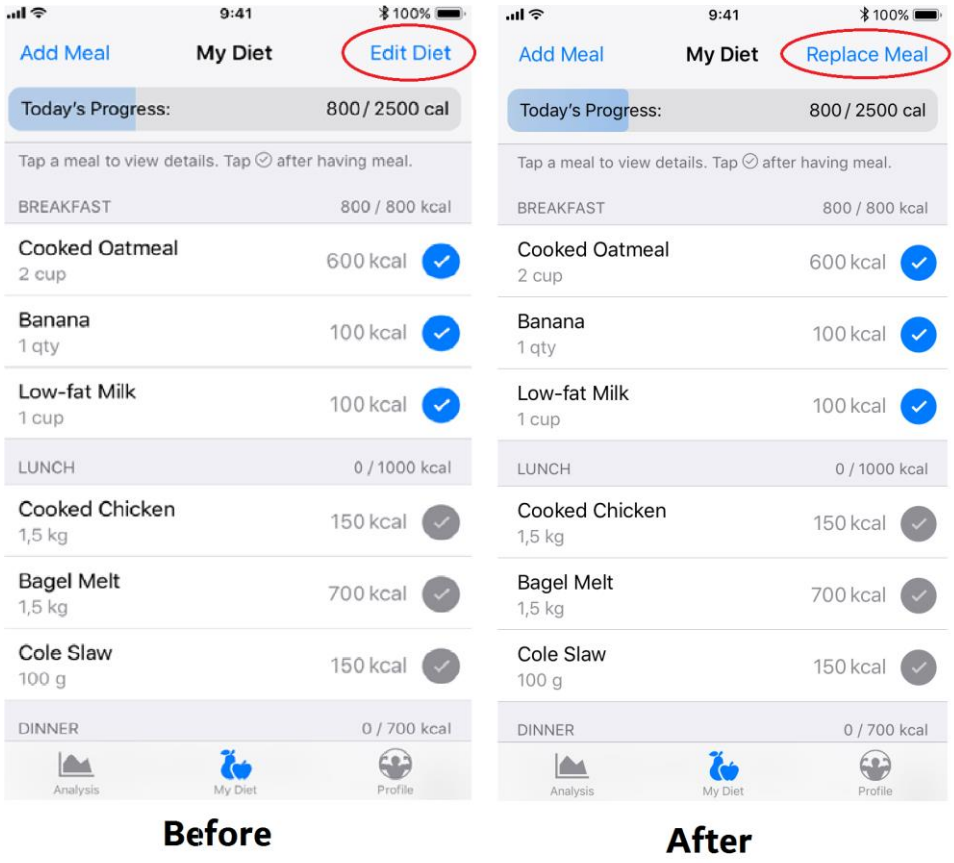


*Figure 3. A more self-explanatory name.*

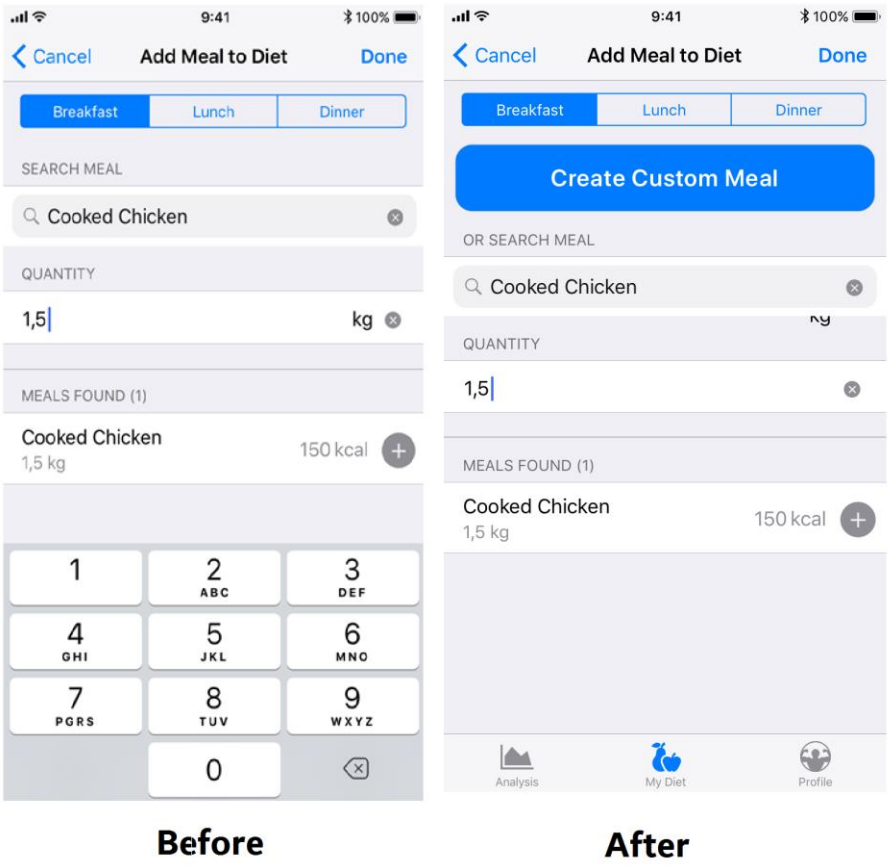


*Figure 4. Custom meal function added.*

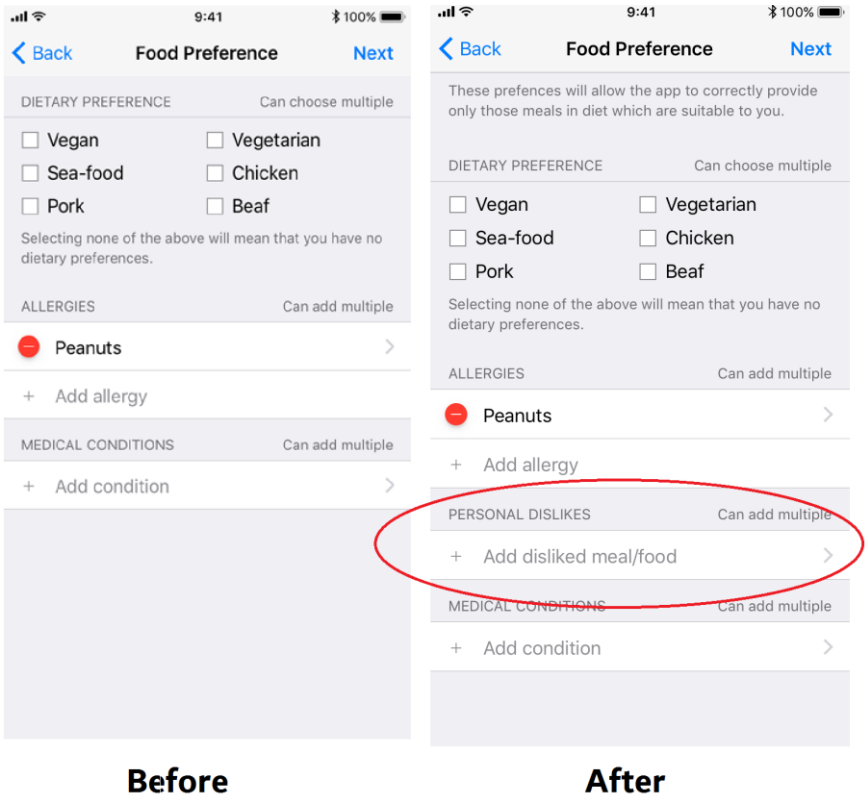


*Figure 5. Option for inputting foods the user dislikes added.*

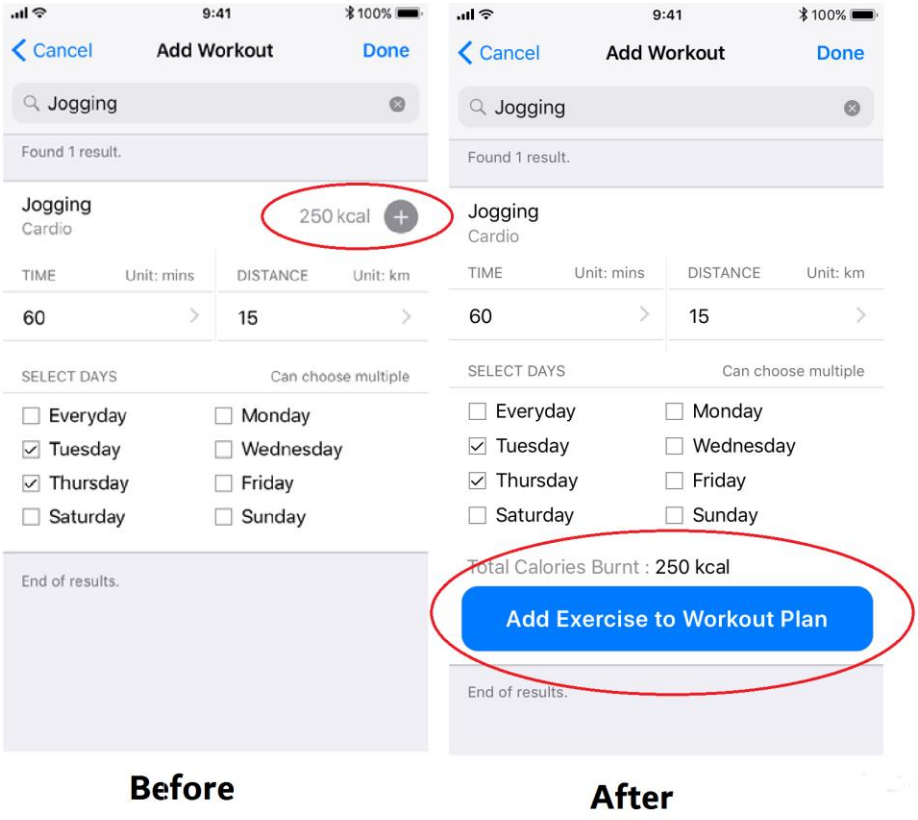


*Figure 6. Unclear icon changed to something more explanatory.*

## 4.1 Setup and Test Scenarios

The test setup was not changed between iteration. The environment in which the tests were conducted is shown in Figure 1. The test scenarios are tasks given to the user to perform during the test. These were not changed between iteration. The reason is that we want to see if the changes made from the first to the second iteration will give different results with the new test subjects. To do an as fair assessment as possible on this, we kept all other factors as possible the same in both iterations (except of course the test subjects, since the old test subjects would have known what to expect).

Test scenarios given to the subjects:

1. Create an account and set it up.
2. Check out details about “bagel melt” and go back to “My Diet” page.
3. Edit diet and try to replace “cooked chicken” with another meal.
4. Edit diet and try to replace ”bagel melt” with another meal.
5. Add a meal and go back to “My Diet” page.
6. Find the monthly analysis.
7. Check out your profile.
8. Log out

## 4.2 First test

Section 4.2.1 describes the subject used in this test. Section 4.2.2 walks through the details of how the test was executed. In section 4.2.3, the raw results of the test are presented in order of appearance. Each issue the subject found is presented under the task they were performing at the time.

### 4.2.1 Subject

The subject is a female badminton player in her early twenties. She has experience using android and iPhone apps, but does not use any diet apps.

### 4.2.2 Execution

The test was performed on 2017-10-13 and lasted for around 20 minutes. Another 30 minutes were used by the evaluation group to discuss and prioritize the problems that were mentioned by the test subject.

### 4.2.3 Results

1. Create an account and set it up.
   The user had no problems executing the task.

2. Check out details about “bagel melt” and go back to “My Diet” page.
   The user had no problems executing the task.

3. Edit diet and try to replace “cooked chicken” with another meal.
   - The user mistook the edit button (marked red in Figure 7) for a refresh button so it took her a few seconds to find the button and edit her meal.

4. Edit diet and try to replace ”bagel melt” with another meal.
   The user had no problems executing the task.

5. Add a meal and go back to “My Diet” page.
   - The user wanted to be able to add a meal categorized as a “Snack” - e.g. a banana or a protein shake. She argued that she does not only eat “Breakfast”, “Lunch” and “Dinner” but also have snacks in between meals.

6. Find the monthly analysis.
   - The user wanted the analysis to also include an analysis on her weight so that she can see/check that the diet is working.
   - The user wanted the “Today’s Analysis” to show nutrition in more detail, e.g. fat should be displayed as unsaturated and saturated fat.

7. Check out your profile.
   The user had no problems executing this task.

8. Log out.
   The user had no problems executing this task.

## 4.3 Second test

Section 4.3.1 describes the subject used in this test. Section 4.3.2 walks through the details of how the test was executed. In section 4.3.3, the raw results of the test are presented in order of appearance. Each issue the subject found is presented under the task they were performing at the time.

### 4.3.1 Subject

The subject is a male runner in his late twenties. He has a lot of experience using android and iPhone apps, but does currently not use any diet apps.

### 4.3.2 Execution

The test was performed on 2017-10-13. The test took about 15 minutes to perform. The results were then discussed and prioritized by the evaluation team over the next 30 minutes. Another hour was then spent discussing improvements based on the results of both test in this iteration.

### 4.3.3 Results

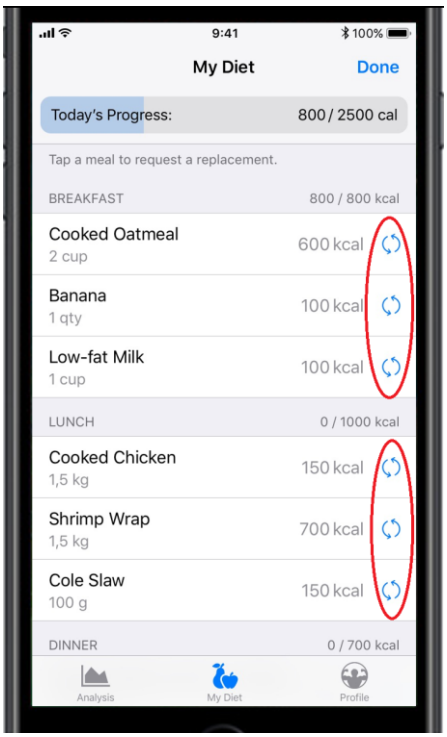


*Figure 7. A screenshot of the "My Diet" editing screen.*

1. Create an account and set it up.
   The user had no problems performing this task.

2. Check out details about “bagel melt” and go back to “My Diet” page.
   The user had no problems performing this task.

3. Edit diet and try to replace “cooked chicken” with another meal
   The user had no problems performing this task.

4. Edit diet and try to replace ”bagel melt” with another meal.
   The user had no problems performing this task.

5. Add a meal and go back to “My Diet” page.
   - The user pointed out that there is no option to remove a meal if he is not having it or replacing it with a custom meal. After the facilitator explained that he could simply leave it un-ticked, he said that it would make him nervous leaving it there. He would rather remove it.

6. Find the monthly analysis.
   - Wished for more detail in the monthly analysis view, such as how they did on different nutrients through the month.

7. Check out your profile.
   The user had no problems performing this task.

8. Log out.
   The user had no problems performing this task.

## 4.4 Suggestions for Improvement

Based on the results of the two tests conducted in this iteration, thee usability problems that were found are presented and improvements are suggested. They appear in order of the severity to the problem. The severity is rated by the evaluation team. Nielsen’s severity rating system is used as a guideline for measuring the problem’s severity *(Nielsen, Severity Ratings for Usability Problems, 1995)*:

“1 = Cosmetic problem only: need not be fixed unless extra time is available on project”
“2 = Minor usability problem: fixing this should be given low priority”
“3 = Major usability problem: important to fix, so should be given high priority”
“4 = Usability catastrophe: imperative to fix this before product can be released”

More detail on the severity rating system used is given in the methodology in section 2.

### 4.4.1 Severity rating 4

- **Problem:** The test subjects wanted to check their weight progress. The app however lacks the function of tracking the user’s change in weight, which means that the user’s have no way - other than by noting down their own weight to themselves- of checking if the diet is working or not. This also means that the app has no way of knowing if the diet is working or not. In other words, the app lacks the function of receiving feedback on how effective the diet is and as a result might lead to the user not being able to maintain their weight. Thereby, this is considered a problem with severity rating 4.
  **Possible Solution**: The app should include an option that allows the user to input the current state of their weight - this could be something that the app notifies the user to do weekly, daily or monthly.

### 4.4.2 Severity Rating 3

- **Problem**: One of the test subjects found it difficult to add a meal as a “Snack” as there were no options that allowed the subject to do so. Since snacks (e.g. protein shake or a banana) are a part of an athlete's diet, the severity rate for this problem is 3.

**Possible Solution**: The app should include an option that allows the user to input their meal as a "Snack".

### 4.4.3 Severity Rating 2

- **Problem:** When editing the diet, the test subjects did not directly or intuitively click the edit button (see Figure 7) - most of the subjects clicked directly on the meal instead of the button. Some of the test subjects even mistook the edit button as a refresh button. This means that the icon does not serve its purpose and instead confuses the user. For this reason, it is given a severity rating of 2.
  **Possible Solution:** Remove the edit button completely and instead allow the users to click directly on the object of interest to edit it.

- **Problem:** When the user wants to analyse their nutrition, they are not able to attain more detailed information. E.g. one of the test subjects wanted to check how much saturated fat was consumed but was unable to do so due to the fact that the nutrition analysis lacked that specific information. This, however, does not hinder the user from using the application correctly and thus, it is given a severity rating of 2.
  **Possible Solution**: Add more detailed nutrition information and categorize the sub-nutrition into their respective nutrition type.

### 4.4.4 Severity Rating 1

- **Problem:** One of the test subjects found it annoying that a meal which was not eaten (e.g. if the user ate something else instead and added it to the diet) was left unticked for that day. The test subject understood that the calories for that meal will not be calculated towards the daily total, but didn't like that they could not remove the meal from the screen. This isn't a big usability problem since I did not stop them from using the product with ease. Therefore, it is given a low severity rating.
  **Possible Solution:** Make it possible to remove meal from screen, possible also to restore it.

# 5 Conclusion

Several problems that were brought up by the test subjects were due to the design of the app not being able to completely satisfy the target group's needs. Other problems were concerning the design not making it clear to the user what the intended use of all functions is and how to use them. By iteratively evaluating the prototype, the evaluation team was able to pinpoint different usability problems and also provide suggested solutions to them.

The first iteration of the evaluation showed that the design was not able to fully convey the app's purpose to the test subjects. The evaluation team thereby purposed that the app should have an introductory text before the user can fill in the needed forms. It also brought to light other problems such as the "Edit diet" function being misleading. The exact problems that were found in the first iteration, along with their suggested solutions were presented in section 3.4.

By the second iteration, some improvements were made to the prototype. These changes were based on the results of the first iteration and can be seen in section 4. Introducing these improvements seemingly shifted the user's focus to different usability problems, suggesting that we had taken steps to solving the first problems. Comparing to the first iteration, the test subjects were now commenting on problems concerning the app not satisfying their needs, rather than problems with understanding functions.

In the second iteration, the test subjects mentioned that the app lacked information that they would have liked to have access to (e.g. information on specific nutrition, information on their weight progress, etc). All the problems found in the second iteration, along with suggested solutions, were presented in section 4.4.

In conclusion, after iteratively evaluating the app, it has become more usable for the intended target group and can to a certain degree cover their needs. However, many elements of the app could still be

improved. When the prototype was improved after the first iteration, that allowed the users to find different, possibly subtler, problems in the second iteration. To cover more usability issues, more iterations should be conducted and with more test subjects.

# 6 Self-Critical Reflection

Our task has been tricky in the sense that we are both the designers and the evaluators of our product. In this reflection, we discuss our efforts as designers and as evaluators.

As designers, our methodology heavily depended on rapid prototyping. Since the beginning, we bounced many ideas and concepts, but the idea we penned down a basic design. This process was time-consuming and difficult as it meant many hours of work were trashed once we moved to a better idea or solution.

For brainstorming, we all first went separate ways, as to gather as many distinct ideas as possible and then came together to pin down the most pressing ones. This worked well as by the time we discussed, we understood what is important. On the other hand, this approach means that we may have missed some great ideas which were pushed away by the time of discussion, because each one of us had already formed some understanding of what is important and what is not. Altogether, we think that the method we used was not perfect but quite well suited for our time constraints, as it saved us a lot of time. We still had plenty of ideas to dismiss and it took a lot of time.

As for our design approach, we went with very subtle colours and clean look as we understood early-on that the application will contain a-lot of text. The application design took inspiration from Apple iPhone applications. This decision made user on-boarding easier as they were already familiar with button placements. Some of our original design decision didn't yield planned results and ultimately had to be updated, like providing detailed explanation on each page, making links and buttons more visible and understandable, allowing users more freedom to choose, etc.

The most difficult part of performing the evaluation was to step out of the shoes of the designers and make a test plan which looks objectively at the product. We realized that we had to test a broad part of the interaction and not just the parts we (having designed it) were most concerned about. This lead to positive results as we found out that the problems which the users found with the interaction were not the things we felt most insecure about, but some of the things we thought were good. The positive thing about this is that the design team will through this evaluation realize completely new problems with their design, meaning that the evaluation has been meaningful. The most striking example of this is when the second test subject in the first iteration didn't understand what the app actually does (generates a diet). We, as designers, could not have imagined that this problem would occur since our entire work was based around that idea.

Since the evaluation brought up many usability problems that we as designers had not foreseen, it was successful. The improvements on the prototype after the first iteration seemed to work, as the users did not point out the same problems in the second iteration. This also suggests that the methodology of the evaluation worked. For a more comprehensive evaluation, more iterations should be conducted and on more users. Unfortunately, we did not have time for more. In a future evaluation project, we should work more efficiently, possible splitting the team up somehow to conduct more tests.

# References

Nielsen, J. (1995, January 1). *Severity Ratings for Usability Problems*. Retrieved from Nielsen Norman Group: https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/

Retting, M. (1994, April). Prototyping for Tiny Fingers. Communications of the ACM, Vol 37, No 4.