Feedback on Shopping Receipt Data Through a Mobile App: A Pilot Study

Martin Lurz
Technical University of Munich
martin.lurz@tum.de

Alexander Esterbauer
Technical University of Munich
alex.esterbauer@tum.de

Markus Böhm
University of Applied Sciences Landshut
markus.boehm@haw-landshut.de

Helmut Krcmar
Technical University of Munich
helmut.krcmar@tum.de

Abstract

Mobile applications have become popular tools for supporting healthy nutrition behavior. Current tools are primarily based on the detailed tracking of a user's direct consumption, thus intervening only during or even after a user has eaten something. With increasing home office hours (especially during the COVID-19 pandemic), people are eating more often at home, which has also led to a decrease in fruit and vegetable consumption. Therefore, we aim to support people in the food-shopping process. We suggest a mobile application that helps people reflect on their purchases and tries to nudge users toward healthier product choices. We conducted a pilot study with 31 participants who used the application for two consecutive weeks. During this time, we observed a decrease in the caloric values per 100 g of purchases. Furthermore, we collected positive feedback on the app regarding acceptance, usability, and user experience.

Keywords: mHealth, Nutrition, Shopping, App, Nudging.

1. Introduction

The number of people with overweight and obesity continues to increase (Abelson & Kennedy, 2004). One major cause for this is seen in the increase of highly processed foods and the associated changes in lifestyle habits and eating habits (Monteiro et al., 2010). In addition, studies show that the consumption of fruits and vegetables has been decreasing in recent years (Jordan et al., 2021; Siegel, 2019). Therefore, this increase in unhealthy nutrition combined with the lack of physical activity and the lack of knowledge about a healthy diet is among the main factors contributing to weight gain in the population (Mesas et al., 2012). Because obesity has been determined to be a risk factor for so-called noncommunicable diseases such as certain cancers and heart disease (Hruby et al., 2016), interventions are necessary.

Nutrition-tracking and weight loss apps have become increasingly popular, with thousands of new health apps being added to the different stores every year (The IQVIA Institute, 2021). These mobile health (mHealth) apps for nutrition provide a method of capturing and influencing dietary patterns. However, studies have shown that monitoring food intake is considered annoying for users in the long run. Users initially display a high level of motivation to enter, track, and receive information about their eating behavior (Hilliard et al., 2014). However, after a short time, the high interaction effort demotivates users to continue tracking their food intake or it seems like too much additional time effort to them, which is why they skip tracking quick snacks, for example (König et al., 2020). Despite every new development in nutrition tracking such as image recognition, current apps still require strong user interaction (Hingle & Patrick, 2016). Thus, alternative tracking methods that minimize required user interaction are needed.

Furthermore, dietary behavior has changed considerably in recent years. The COVID-19 pandemic not only caused a strong increase in home office hours but was also expected to increase malnutrition (Balanzá-Martínez et al., 2020). These expectations have been shown to be true for a nonnegligible part of society (Visser et al., 2020). Therefore, we aimed to support people regarding their shopping behavior. Thus, to minimize required user input while supporting users in reflecting on their shopping behavior, we suggest a shopping receipt–based app that nudges the user to purchase healthy products. This app reduces user interaction of scanning shopping receipts and provides instantaneous feedback on the single products.

In this practice-based research paper, our research question is as follows:

RQ: Can ratings of one's shopping receipt be used to nudge people to healthier purchases?

URI: https://hdl.handle.net/10125/102985
978-0-9981331-6-4

2. Background

Documentation of nutrition has long been applied in the fields of nutritional sciences. With the advent of the smartphone, research on nutrition apps has also grown in the field of Information Systems (IS), among others.

2.1. Nutrition-Tracking

In the area of nutrition sciences, different methods have been established for recording a person's nutrition and nutritional behavior (Johansson, 2014). One of the most popular methods is the so-called nutrition diary. Here, the consumed food is recorded immediately before or after consumption, including the exact weight. This has the advantage that the participants usually still know exactly what they consumed but requires recording several times per day. However, such detailed logging is often negatively perceived by users (Lee et al., 2017). Other methods, such as the Food Frequency Questionnaire or the 24-h Recall, assess over a longer time period per session and thus require less logging effort. However, one drawback of these methods is the fact that the lists of entered consumed food items are incomplete (Thompson & Subar, 2017). Regardless of the recording method, it is important to note that errors may occur with all of the prior mentioned methods (MacIntyre, 2009), as data are often falsified consciously or unconsciously during tracking (Calvert et al., 1997; Macdiarmid & Blundell, 1997).

Different attempts to automate the tracking process have become subject to research in the past few years. The recording of products by means of barcodes has become a widespread option (Byrd-Bredbenner & Bredbenner, 2010). However, it is limited to pre¬packaged food. Other, more novel attempts range from picture-based meal recording (Bedri et al., 2020; Freitas et al., 2020; Jia et al., 2019) to estimating food and consumption values via chewing or swallowing sensors (Hussain et al., 2018; Sazonov et al., 2008). However, it should be noted that these approaches are usually not suitable for daily use because they either require special sensors on the neck for chewing detection (Hussain et al., 2018) or—in the case of a video camera—often raise privacy concerns for the users (Doulah et al., 2020). In addition, most studies have described tests conducted only in laboratory environments, with limitations such as participants being not allowed to talk (Sazonov et al., 2008), a limited variety of food items (Freitas et al., 2020), or simply detecting food versus nonfood activities without identifying the single items or their quantity (Bedri et al., 2020; Jia et al., 2019). Thus, until passive tracking solutions are reliable companions in day-to-day use, simplified tracking methods through the user could be helpful for improving long-term use.

2.2. Food Item Rating

One of the most well-known ways of rating a food is based on the calories and macronutrients contained per 100 g, in the form of the mandatory nutrition tables and ingredient lists on food items. This method is supposed to help shoppers make healthier choices (Grunert & Wills, 2007). Accordingly, consumers can compare foods based on nutritional information, ingredient information, and other factors such as serving size. Despite these positive aspects, there are many challenges to the information approach. Research on consumer behavior indicates that although consumers understand the basics of nutrition labels, they may be overwhelmed when faced with too much and too complex information (Kalnikaitė et al., 2013). Comprehending and applying nutrition information to food choices is a mentally demanding task that many consumers find excessively difficult and tedious (Guthrie et al., 2015). This high mental demand can lead to decision bias (Kalnikaitė et al., 2013). Thus, the goal of such label approaches should be to minimize the cognitive demand required to process information about the nutritional content of foods and to evaluate a product as healthy or unhealthy (Thorndike et al., 2012).

A more simplified way of rating food items is through scores, which are intended to make it easier for consumers to assess and compare products. One of the most famous examples of food scores is the so-called Nutri-Score (Chantal et al., 2017). This score is intended to improve the comparability of food items of one category and is composed of an ascending 5-point scale, ranging from A to E, which informs the consumers about possible negative effects on their health when buying groceries. The basis of the Nutri-Score is the respective ingredients of a product. In other words, this means that the higher the proportion of vegetables, fruits, proteins, or nuts, the better the score. In contrast, ingredients with high energy content, sugar, saturated fatty acids, and a high salt or sodium quantity have a negative effect on the rating. Consumers react positively to this measure, and with its help, they also resort to healthier foods from the same category (De Temmerman et al., 2021). The Nutri-Score is currently being used in several European countries, including France, Belgium, Germany, Spain, the Netherlands, and Luxembourg, on official packaging (Foodwatch, 2020).

2.3. Nutrition Apps

Apps that aim to support users in adopting a healthier nutrition behavior, often simply referred to as nutrition apps, have gained popularity, especially on smartphones, where these apps can be used anywhere at any time (Cho, 2016; König et al., 2018). Although numerous convenient nutrition apps have been developed in the past few years (Franco et al., 2016), their structure is mostly similar. To support nutritional diets, these apps are usually equipped with functions such as food recommendations for healthy cooking, nutritional information, self-monitoring, disease prevention, and various other functions (Holzmann et al., 2017). To provide personalized and correct feedback, users must log their meals on a daily basis into a digital nutrition diary (Leipold et al., 2018). Although the market is rapidly growing, only a small number of these apps are successful (Cho, 2016). One reason might be the aforementioned required regular logging that users often report as burdensome (Hauptmann et al., 2021; Lee et al., 2017). Such perceptions might cause low adherence rates in users (Helander et al., 2014), with high dropout rates starting in the first few days (Thompson & Subar, 2017).

With regard to nutrition apps focusing on the food-shopping process, different approaches have been made to improve the healthfulness of people's purchases. For example, López et al. (2017) presented an app that generates shopping lists for users. However, we were able to identify only one publication investigating nutrition apps based on the analysis of shopping receipts (Sainz-De-Abajo et al., 2020), in which the evaluation was focused solely on elderly people. Thus, no conclusion regarding the acceptance and its effects on (young) adults could be made.

Our analysis of commercially available apps in the stores for iOS and Android showed that the most popular solutions seem to be mainly aimed at weight loss and use a nutrition diary approach to collect data, such as Lifesum (lifesum.com), myfitnesspal (myfitnesspal.com), or YAZIO (yazio.com). Although some allow barcode scanning to quickly add packaged foods, each product has to be scanned individually, thus still being cumbersome if many products have to be scanned. The same applies to apps targeted toward scanning products while shopping such as Yuka (yuka.io) or the Open Food Facts app (openfoodfacts.org). None of the currently available apps are targeted at providing nutrition support via scanning of shopping receipts, which would allow tracking many products in minimal time. Although different receipt-scanning apps exists, such as MrReceipt (mrreceipt.com) or Receipt Scanner (easy-expense.com), these are targeted at expense control and they do not provide nutritional feedback.

3. Concept

The main goal of this study was to develop a nutrition app that would enable users to quickly access information on their shopping behavior to be able to make more informed choices toward a healthier diet.

3.1. Process

To develop a useful technical artifact, we applied the Design Science Research Approach (DSRA) as described by Hevner et al. (2004), to our research. DSRA defines three major keystones that must be considered during development: knowledge base, environment, and IS research.

To define our environment, we looked at the current challenges of nutrition apps and found that major areas for improvement included the topics of high dropout rates, mainly due to the time-consuming dietary recording mechanism, and the problem of over- and underreporting of quantities consumed.

Following this, we reviewed the published scientific literature to gain insights into the current state of the knowledge base. We collected information on prior mobile apps aimed at nutrition (behavior). We evaluated how data gathering/nutrition tracking was implanted by investigating characteristics, such as accuracy, the timing of recording, and frequency of recording. With regard to feedback, we investigated different approaches to evaluating nutrition protocols. We could identify three different approaches based on the level of detail of the available information: assessment based on individual foods (amount consumed and frequency), overall nutrition (relationships between dishes and meals), and nutrient and energy content (composition in terms of macro- and micronutrients).

Building on the two insight areas, we then defined basic functionalities for a potential mobile application, followed by a develop/evaluate cycle.

3.2. Technical Features

To make the tracking of purchases comfortable, we included the feature of taking a photo of the shopping receipt. The taken picture is then analyzed using the Azure Form Recognizer service (Microsoft, 2022), which returns the content as text elements.

To retrieve the nutritional values of packaged products, we retrieved data from the Open Food Facts database (Open Food Facts, 2022), which is a crowdsourced database of food from around the world. Introduced in 2012, it has grown rapidly to more than 1 million products.

To give users a feeling about their overall purchase, an average score is calculated per shopping receipt. Therefore, only products that have a Nutri-Score given are considered. Then, the Nutri-Score of each product is added. Products purchased more than once are counted accordingly. For products that were purchased in grams, an attempt is made to infer the number based on the standardized portion size. For example, an apple has a standard weight of 178 g according to the German food database Bundeslebensmittelschlüssel (Max Rubner-Institute, 2021). With a purchased quantity of 1.1 kg, six times the rating of the apple is included in the calculation. The accumulated Nutri-Score number is divided by the number of accumulated products. This results in a number that lies between 1 and 5 and, according to the breakdown at the beginning of this chapter, leads to the overall Nutri-Score rating. To obtain a clear score, results with decimal places are mathematically rounded up or down.

3.3. User Interface

In developing the user interface, we followed the guidelines of Kalnikaitė et al. (2013) to display nutrition information in a simple and easy-to-understand manner and used default design elements to enable users to quickly navigate through the app without having to learn new concepts. Furthermore, we included the Nutri-Score as an easy way to evaluate purchases. Because, according to Kalnikaitė et al. (2013), shoppers blame the size and position of traditional food labels for paying less attention to nutritional values, we aimed to display the Nutri-Score in a clear and prominent position. This was done to enable users to quickly obtain a rough idea of a food's healthiness.

When opening the app for the first time, the user is first welcomed by a tutorial that provides a brief introduction to the added value of the application, which steps are necessary to scan an invoice correctly, what to look out for, what the Nutri-Score is, and how it is to be interpreted. This is to give the user quick insight into the functions and instructions on how to use the app.

As seen in Figure 1 (left), the home screen of the app shows information about the user's general purchasing behavior. A graphic visualizes the percentage distribution of the ratings of the individual products of the purchase and compares it with the distribution of the purchase before. The color distribution suggests the healthier purchase and also indicates how the current purchase compares with the last one. To obtain trends over a longer period, purchases made in the past 2 months can be compared with one another to see how the purchasing behavior has developed. This is to ensure that long-term nutritional goals develop and that users are thus tied to the app in the long term.

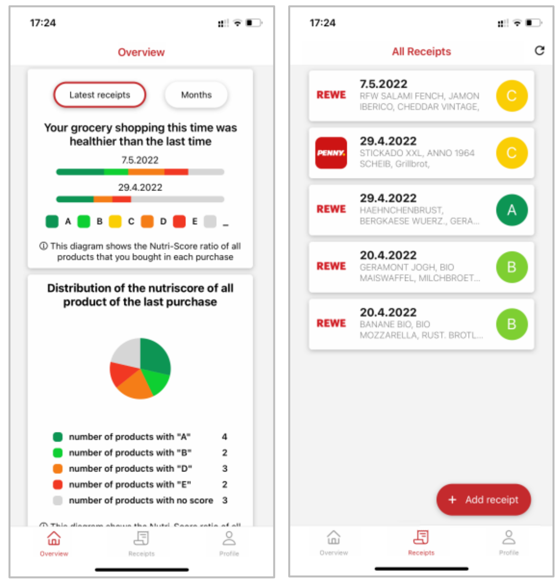

Figure 1: Home screen (left) and shopping receipt overview (right).

On the receipts screen, an overview of all scanned receipts is given, as shown in Figure 1 (right). A new receipt can be added through the "add receipt" button by either taking a picture of a receipt or selecting a picture from one's gallery.

When clicking on one scanned receipt, the entries are listed with their quantity, name, and the Nutri-Score, as shown in Figure 2 (left). In addition, the linked product is also displayed below the product name to make the user aware of which product was recognized with the entry. This allows the user to quickly identify whether a product was incorrectly linked to the entry and in that case connect the correct product.

By tapping on an entry, the user gets an overview of the food's exact nutritional values, as shown in Figure 2 (right). In addition to the calorific value of a product, the prototype also shows the fat, salt, and sugar content contained in the product and overcomes, for example, the trick of food companies to replace fat with high sugar content or vice versa. As Becker et al. (2016) showed that color-coded icons can lead to higher awareness, an ample system is used for this section.

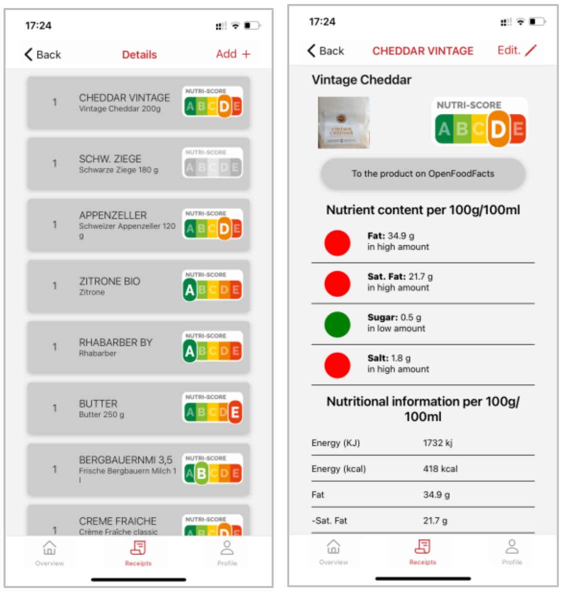

Figure 2: Product list of a shopping receipt (left) and detailed view for a single product (right).

4. Evaluation

To evaluate the impact of such an application on the quality of food selection and the suitability for automated tracking, we conducted a pilot user study that allowed us to investigate how users would apply the app in their normal everyday life.

4.1. Methodology

Our study consisted of four main stages: recruiting participants, a prequestionnaire, a 2-week app-testing period, and a postquestionnaire, as shown in the study flow in Figure 3. Each participant was assigned an individual study ID at the beginning, which was assessed in the questionnaires as well as in the app to ensure completion of each participant.

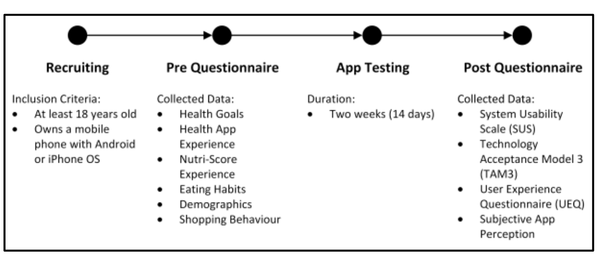

Figure 3: Pilot study procedure.

We recruited participants through the social media/social circles of the researchers. This allowed for a younger demographic in order to complement the research of Sainz-De-Abajo et al. (2020). To achieve a normal distribution, we did not exclude participants based on age or gender. The only inclusion criterion was that participants had to be at least 18 years old. A pre¬selection was not made, because the scope of the feedback should not be influenced.

The prequestionnaire administered at the beginning of the study primarily aimed to obtain background information about the participants. Thus, we assessed the following control variables as they might influence app usage, as well as later assessed the perceptions and acceptance of the participants: demographic data such as age, gender, height, weight, and the highest level of education; prior knowledge of health app usage; and eating habits and healthiness of food choices (Lau et al., 1986). Furthermore, we also assessed the shopping behavior along with information on how many meals were eaten at home.

After completing the questionnaire, participants received a link to a website where the app could be obtained as well as instructions on installing it on the end device. Apple TestFlight was used for iOS devices, and owners of Android devices received access to an apk-file for a manual installation.

During the app-testing period, participants were asked to scan their shopping receipts for 2 weeks. We chose this period to ensure that each participant would scan multiple receipts. During this phase, we tracked how long it took participants to scan receipts, how many products were purchased and automatically linked, as well as how many had to be linked manually by the user. The total number of calories and unrecognized items were also gathered. In addition, during the 2 weeks, participants were asked to fill out a 24-hour recall protocol in which they entered everything they had eaten that day. We included this test to evaluate whether scanned shopping receipts line up with the consumed calories per person (with respect to how many meals are eaten not at home).

After 2 weeks, a postquestionnaire was sent out. It included questions on usability through the System Usability Scale (SUS) (Brooke, 1996); app perception using the items Perceived Ease of Use (PEOU), Perceived Usefulness (PU), Behavioral Intention (BI), and Attitude Toward Using (AT) from the Technology Acceptance Model (TAM) (Davis, 1989); and user experience through the User Experience Questionnaire (UEQ) (Schrepp et al., 2014).

Before the pilot user study started, we carried out a pretest with three individuals to identify errors in the application and obtain suggestions for improvement.

4.2. Participants

Overall, we included 31 participants who had registered for the study during the recruitment period and completed the study by filling out the postquestionnaire. As shown in Table 1, 64.5% were male and 35.5% were female. The mean (M) age of the participants was 27.32 years with a median age of 25.0 years. Of the participants, 77.4% had a bachelor's degree or higher, 16.1% had completed an apprenticeship, and 6.5% had achieved a high school degree. Furthermore, 35.48% of the participants had used a nutrition-tracking app before, but only one person was using one during the study.

Table 1: Participant statistics.

		Participants	
Amount		31	
Gender		Male	20
		Female	11
		Mean	SD
Age (years)		27.32	7.18
BMI		23.89	3.32
Perceived importance of health		15.45	1.89
Percentage of consumed food bought in the supermarket		77.42	19.14
Duration of food-shopping trip (minutes)		26.87	12.09
		Yes	No
Used nutrition apps before		35.48%	64.52%
Currently using a nutrition app		3.23%	96.77%
Knows the Nutri-Score		74.19%	25.81%
Food Shopping	Once a Week	35.48%	
	2–3 Times	54.84%	
	4–6 Times	9.68%	
	Daily	0%	
Living Situation	Single	51.61%	
	With roommates / parents	25.81%	
	With partner	16.13%	
	With partner and child(ren)	6.45%	

4.3. Results

First, we wanted to determine whether the calculated calories of the shopping receipts reflected the actual intake. Thus, we investigated if calories of the 24-hour recall protocol were in line with the overall bought calories during the study period. As it is not possible to determine who has eaten which product in a multiperson household, only single-person households were considered in the evaluation. Furthermore, for a more realistic calculation, we included values from the prequestionnaire such as meals being eaten out during the week as well as the percentage of food purchased at the supermarket. An average of 93% (M = 0.934, SD = 0.47) of the calories consumed at home could be calculated by scanning the receipts. Compared with the total calories ingested, without considering whether the calories were consumed at home or not, 69% (M = 0.692, SD = 0.32) were covered by the receipts.

Table 2: Results regarding usability, technology acceptance, and user experience

	Mean	SD
SUS (0–100)	78.39	13.06
TAM		
PEOU (1–7)	6.09	1.04
PU (1–7)	4.69	1.10
BI (1–7)	4.23	1.76
AT (1–7)	4.69	1.31
UEQ		
Attractiveness (-3 to +3)	1.387	0.92
Perspicuity (-3 to +3)	2.202	0.41
Efficiency (-3 to +3)	1.379	0.93
Dependability (-3 to +3)	1.129	0.51
Stimulation (-3 to +3)	1.137	0.77
Novelty (-3 to +3)	0.847	0.50

Regarding the tested prototype, the usability of the systems was well received, with an average score of 78, indicating good usability and corresponding to grade B (Sauro & Lewis, 2016). There were no statistically significant differences in characteristics such as gender between the subgroups. Differences in the usability ratings between people who shop frequently or only about once a week were only small. Regarding acceptance, the mean of PEOU, as shown in Table 2, indicates that the participants regarded the application as easy to use (mean 6.09), which was also consistently reported (standard deviation 1.04). In addition, other items of the TAM such as PU showed a positive indication as well as a positive attitude regarding use. The lowest values were given for BI, which was rated only slightly above undecided.

When evaluating the UEQ, we found that all six scales were in the range of values considered positive. As shown in Figure 4, the app performed above average on five of the six scales. In terms of Perspicuity, the prototype even scored excellently and was on this scale among the best 10% of the results. Only Dependability scored just below average. Attractiveness, Efficiency, Stimulation, and Novelty were all above average and thus performed better than 50% of the results in the benchmark.

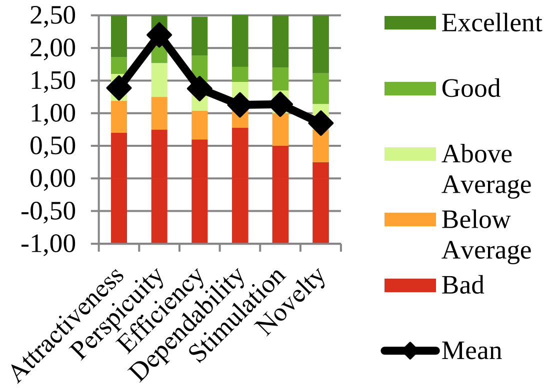

Figure 4: UEQ performance benchmark results.

Apart from the validated scales of the postquestionnaire, we also analyzed app usage during the 2-week testing period. Overall, during the study period, 171 receipts from 15 different supermarket chains were scanned, including the biggest supermarket chains in Germany such as Edeka, Rewe, Aldi, and Lidl, as well as smaller organic supermarkets such as a Denns. A total of 1195 products were linked to an associated product on the receipts. On average, participants scanned five receipts and linked 35 products. Although participants stated that they found the Nutri-Score evaluation very useful, it was noticed that in many of the products purchased, not enough information was given through the Open Food Facts data to correctly calculate the Nutri-Score. As a result, 28.76% of the products lacked the Nutri-Score information and thus could not be included in the calculation of the overall receipt score. In the total of 1195 linked items, 851 had an assigned Nutri-Score, with the remaining products missing a specification.

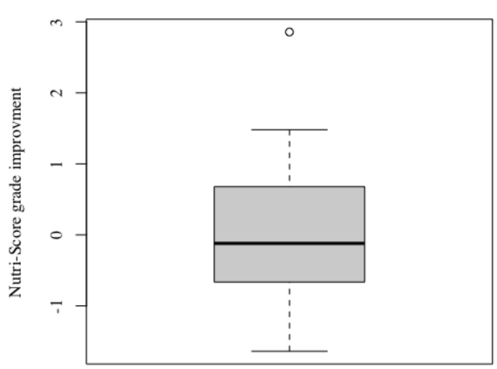

Figure 5: Boxplot of the difference in the Nutri-Score mean value between the two periods.

Finally, we wanted to determine if eating behavior values changed during the testing period. Therefore, we looked at the average Nutri-Score as well as caloric values. For these analyses, we again split up the testing period into two parts, as done for the app usage.

As shown in Figure 5, our evaluations on the Nutri-Score could not detect any significant improvement in the receipts' Nutri-Score between the two periods. Neither the time of usage, importance of health, nor duration of shopping time showed any significant correlations. The mean of all receipt scores changed from 2.22 on average in the first period to 2.29 in the second period. Both the mean and median scores showed little change between the two periods. Although some participants improved their grocery shopping scores, many also worsened, and therefore, no trend can be measured.

With regard to caloric values, slight improvements were observed for calories per 100 g. However, no significant correlations could be found, neither across time of use nor with duration of shopping time. Nevertheless, the average calories per 100 g decreased on average by 20 kcal from 200 in the first period to 180 in the second, as shown in Figure 6.

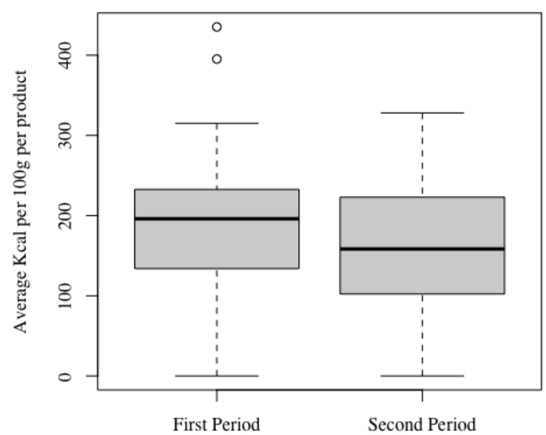

Figure 6: Comparison of the average kcal per 100g between the two periods.

5. Discussion

Based on the presented results, we found the approach to be a valid instrument for nutrition assessment. Of course, this pilot study focused on mostly single households; thus, the calorie calculation for families might be more imprecise than the calorie intake of single household members, as certain products may be consumed by only some members of the household. Nevertheless, this app might help nudge the purchasers in the households to select healthier options.

With regard to perception and acceptance, measured with the SUS, TAM, and UEQ, most of the participants in the user study seemed to have an above average like of the prototype. Participants quickly found their way around and did not need technical support in using the application. Overall, the application appears to perform well across all UEQ and SUS criteria. This becomes even more clear when comparing our usability rating to that of commercial apps. Based on the results of Ferrara et al. (2019), we see that the proposed solution performed better than four of the seven evaluated apps. Because this was the first user evaluation of the technical artifact, we regard this as a positive sign for the general idea.

Although the number of manual interactions per receipt decreased significantly throughout the user study, it is surprising that in the second period, on average, still only 30% of entries could be automatically assigned to a product. Because, according to Chu et al. (2010), shopping behavior is often repeated, we expected a higher value. One explanation seems to be that the duration of the user study, as it was limited to 2 weeks and thus was possibly not long enough to sufficiently track the repetitive shopping behavior of the participants. Furthermore, the results might also be attributed to the fact that the letter recognition of the app did not always completely recognize the characters of an entry.

With regard to nutrition behavior, we did not detect any improvements in the Nutri-Score level. However, there might be multiple reasons for this finding. First, the Nutri-Score evaluation from the Open Food Facts database is missing for some products and thus cannot be included in the evaluation of the receipt. Second, not all products currently provide a Nutri-Score on their packaging. Therefore, it might have been harder for some participants to compare products and easily find healthy alternatives.

However, we did observe a 10% average reduction in the energy density of shopped items (200 kcal/100 g to 180 kcal/100 g) during the 2-week study pilot. We see this as a promising sign that the app has a positive nudging effect of users' purchase behavior.

Based on what we have learned from the pilot study, we suggest the following design recommendations for future nutrition applications using shopping receipts:

1. Allow data improvement.
   Because of the large number of different foods that can be purchased, it should be possible to edit information about a product so that users can add any missing data needed to calculate a Nutri-Score value.
2. Don't make it too simple.
   Although precise macro- and micronutrient values might be complicated for some users, they should still be available to allow comparison on a more precise level than just a score, especially as scores might not be available for all products.
3. Provide precise recommendations.
   Nutritional values are difficult to compare, either due to complicated nutrition tables or missing labels such as the Nutri-Score. Therefore, users need precise guidance and/or product recommendations to improve.

Of course, our findings are not without limitations. First, the sample of the evaluation was not representative of society in general, in terms of either gender or age. The gender distribution was not balanced, and most participants were younger than 30 years; thus, perceptions and usage may differ among other demographic groups. In addition, this pilot study focused on mostly single households; therefore, we cannot assume that the same effects would apply if the app were used in a multiperson household. Other limitations include user behavior during the study, which might deviate from regular behavior. For the questionnaires, known influencing factors such as confirmation bias and response bias might apply.

This work contributes to the research on the automation of food intake monitoring, as the evaluation of the user study showed that the designed prototype can calculate a close estimate of the calories consumed at home. In addition, we contribute to the IS research in the area of development of health behavior change support systems by providing design recommendations concluded in our pilot study regarding the presented system, with a low effort that encourages the user to eat healthier in the long term by assessing the individual's shopping behavior. Although we did not observe statistically significant results in dietary change over the 2 weeks, the results of the pilot study on this approach show a reduction in the calorie density of the purchased items. Furthermore, participants in the user study appreciated the food assessment through the receipt-scanning process and had an interest in using the application over the long term.

6. Conclusion

In this paper, we presented the design of and evaluated a nutrition app that requires only little input from users by analyzing pictures of shopping receipts. from the pilot study, we not only provided a proof of concept regarding the validity of using shopping receipts as a base to track users' nutrition but also showed that people found the app easy to use. Our research question was whether such an application could improve users' shopping behavior toward more healthy products. Although no positive change regarding the Nutri-Score could be detected, we found a 10% decrease in the average caloric density between the first and the second period of app testing. Thus, we see potential in this approach, especially if identified challenges, such as the missing Nutri-Score values for some products, are resolved.

For future work, we suggest expanding the app with more precise suggestions for users on how to improve their shopping behavior to determine whether such a suggestion could improve the impact of the app. Furthermore, the next study should include more participants and expand over a longer period to be able to gain more stable results and get insights regarding the usage behavior over a longer time frame.

7. References

Abelson, P., & Kennedy, D. (2004). The obesity epidemic [Editorial]. *Science*, *304*, 1413.

Balanzá-Martínez, V., Atienza-Carbonell, B., Kapczinski, F., & De Boni, R. B. (2020). Lifestyle behaviours during the COVID-19 - time to connect. *Acta psychiatrica Scandinavica*, *141*(5), 399-400.

Becker, M. W., Sundar, R. P., Bello, N., Alzahabi, R., Weatherspoon, L., & Bix, L. (2016). Assessing attentional prioritization of front-of-pack nutrition labels using change detection. *Applied Ergonomics*, *54*, 90-99.

Bedri, A., Li, D., Khurana, R., Bhuwalka, K., & Goel, M. (2020). Fitbyte: Automatic diet monitoring in unconstrained situations using multimodal sensing on eyeglasses. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,

Brooke, J. (1996). SUS: A quick and dirty usability scale. In P. W. Jordan (Ed.), *Usability evaluation in industry* (pp. 189–194). CRC Press.

Byrd-Bredbenner, C., & Bredbenner, C. A. (2010). Assessing the home food environment nutrient supply using mobile barcode (Universal Product Code) scanning technology. *Nutrition & Food Science*

Calvert, C., Cade, J., Barrett, J. H., Woodhouse, A., & Group, U. S. (1997). Using cross-check questions to address the problem of mis-reporting of specific food groups on Food Frequency Questionnaires. *European Journal of Clinical Nutrition*, *51*(10), 708-712.

Chantal, J., Hercberg, S., & World Health Organization. (2017). Development of a new front-of-pack nutrition label in France: the five-colour Nutri-Score. *Public Health Panorama*, *3*(04), 712-725.

Cho, J. (2016). The impact of post-adoption beliefs on the continued use of health apps. *International journal of medical informatics*, *87*, 75-83.

Chu, J., Arce-Urriza, M., Cebollada-Calvo, J.-J., & Chintagunta, P. K. (2010). An Empirical Analysis of Shopping Behavior Across Online and Offline Channels for Grocery Products: The Moderating Effects of Household and Product Characteristics. *Journal of Interactive Marketing*, *24*(4), 251-268.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

De Temmerman, J., Heeremans, E., Slabbinck, H., & Vermeir, I. (2021). The impact of the Nutri-Score nutrition label on perceived healthiness and purchase intentions. *Appetite*, *157*, 104995.

Doulah, A., Ghosh, T., Hossain, D., Imtiaz, M. H., & Sazonov, E. (2020). "Automatic ingestion monitor version 2"–a novel wearable device for automatic food intake detection and passive capture of food images. *IEEE Journal of Biomedical and Health Informatics*, *25*(2), 568-576.

Ferrara, G., Kim, J., Lin, S., Hua, J., & Seto, E. (2019). A Focused Review of Smartphone Diet-Tracking Apps: Usability, Functionality, Coherence With Behavior Change Theory, and Comparative Validity of Nutrient Intake and Energy Estimates. *JMIR Mhealth Uhealth*, *7*(5), e9232.

Foodwatch. (2020). *Nutri-Score - A 5-Colour Nutrition label to help people shop more healthily*. Foodwatch. https://www.foodwatch.org/en/campaigns/nutri-score/

Franco, R. Z., Fallaize, R., Lovegrove, J. A., & Hwang, F. (2016). Popular nutrition-related mobile apps: a feature assessment. *JMIR mHealth and uHealth*, *4*(3), e85.

Freitas, C. N., Cordeiro, F. R., & Macario, V. (2020). Myfood: A food segmentation and classification system to aid nutritional monitoring. 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI),

Grunert, K. G., & Wills, J. M. (2007). A review of European research on consumer response to nutrition information on food labels. *Journal of Public Health*, *15*(5), 385-399.

Guthrie, J., Mancino, L., & Lin, C.-T. J. (2015). Nudging Consumers toward Better Food Choices: Policy Approaches to Changing Food Consumption Behaviors. *Psychology & Marketing*, *32*(5), 501-511.

Hauptmann, H., Leipold, N., Madenach, M., Wintergerst, M., Lurz, M., Groh, G., . . . Krcmar, H. (2021). Effects and challenges of using a nutrition assistance system: results of a long-term mixed-method study. *User Modeling and User-Adapted Interaction*

Helander, E., Kaipainen, K., Korhonen, I., & Wansink, B. (2014). Factors related to sustained use of a free mobile app for dietary self-monitoring with photography and peer feedback: retrospective cohort study. *Journal of medical Internet research*, *16*(4), e109.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 75-105.

Hilliard, M., Hahn, A., Ridge, A., Eakin, M., & Riekert, K. (2014). User Preferences and Design Recommendations for an mHealth App to Promote Cystic Fibrosis Self-Management. *JMIR Mhealth Uhealth 2014;2(4):e44*

Hingle, M., & Patrick, H. (2016). There Are Thousands of Apps for That: Navigating Mobile Technology for Nutrition Education and Behavior. *Journal of Nutrition Education and Behavior*, *48*(3), 213-218.e211.

Holzmann, S. L., Pröll, K., Hauner, H., & Holzapfel, C. (2017). Nutrition apps: Quality and limitations. An explorative investigation on the basis of selected apps. *Ernaehrungs Umschau International*, *64*(5), 80-89.

Hruby, A., Manson, J. E., Qi, L., Malik, V. S., Rimm, E. B., Sun, Q., . . . Hu, F. B. (2016). Determinants and Consequences of Obesity. *American Journal of Public Health*, *106*(9), 1656-1662.

Hussain, G., Javed, K., Cho, J., & Yi, J. (2018). Food intake detection and classification using a necklace-type piezoelectric wearable sensor system. *IEICE TRANSACTIONS on Information and Systems*, *101*(11), 2795-2807.

Jia, W., Li, Y., Qu, R., Baranowski, T., Burke, L. E., Zhang, H., . . . Mao, Z.-H. (2019). Automatic food detection in egocentric images using artificial intelligence technology. *Public Health Nutrition*, *22*(7), 1168-1179.

Johansson, G. (2014). *Nutritional epidemiology: Dietary assessments: use, design concepts, biologica markers, pitfalls and validation*. Halmstad University Press

Jordan, I., Keding, G. B., Stosius, L., Hawrysz, I., Janiszewska, K., & Heil, E. A. (2021). Changes in Vegetable Consumption in Times of COVID-19—First Findings From an International Civil Science Project [Original Research]. *Frontiers in Nutrition*, *8*

Kalnikaitė, V., Bird, J., & Rogers, Y. (2013). Decision-making in the aisles: informing, overwhelming or nudging supermarket shoppers? *Personal and Ubiquitous Computing*, *17*(6), 1247-1259.

König, L. M., Attig, C., Franke, T., & Renner, B. (2020). Barriers to and facilitators for using nutrition apps: a scoping review and conceptual framework. *JMIR mHealth and uHealth*

König, L. M., Sproesser, G., Schupp, H. T., & Renner, B. (2018). Describing the process of adopting nutrition and fitness apps: behavior stage model approach. *JMIR mHealth and uHealth*, *6*(3), e55.

Lau, R. R., Hartman, K. A., & Ware, J. E. (1986). Health as a value: Methodological and theoretical considerations. *Health psychology*, *5*(1), 25.

Lee, J.-E., Song, S., Ahn, J. S., Kim, Y., & Lee, J. E. (2017). Use of a mobile application for self-monitoring dietary intake: Feasibility test and an intervention study. *Nutrients*, *9*(7), 748.

Leipold, N., Madenach, M., Schäfer, H., Lurz, M., Groh, G., Böhm, M., . . . Krcmar, H. (2018). Nutrilize a Personalized Nutrition Recommender System: an Enable Study.

López, D., Torres, M., Vélez, J., Grullon, J., Negrón, E., Pérez, C. M., & Palacios, C. (2017). Development and evaluation of a nutritional smartphone application for making smart and healthy choices in grocery shopping. *Healthcare Informatics Research*, *23*(1), 16-24.

Macdiarmid, J. I., & Blundell, J. E. (1997). Dietary under-reporting: what people say about recording their food intake. *European Journal of Clinical Nutrition*, *51*(3), 199-200.

MacIntyre, U. E. (2009). Measuring food intake. *Introduction to Human Nutrition*, *2*, 238-275.

Max Rubner-Institute. (2021). *Food Analysis and the German Nutrient Database*. https://www.mri.bund.de/en/institutes/safety-and-quality-of-fruit-and-vegetables/research-areas/food-analysis-and-german-nutrient-database/

Mesas, A. E., Guallar-Castillón, P., León-Muñoz, L. M., Graciani, A., López-García, E., Gutiérrez-Fisac, J. L., . . . Rodríguez-Artalejo, F. (2012). Obesity-Related Eating Behaviors Are Associated with Low Physical Activity and Poor Diet Quality in Spain. *The Journal of Nutrition*, *142*(7), 1321-1328.

Microsoft. (2022). *Azure Form Recognizer*. https://azure.microsoft.com/en-us/services/form-recognizer/

Monteiro, C. A., Levy, R. B., Claro, R. M., de Castro, I. R. R., & Cannon, G. (2010). Increasing consumption of ultra-processed foods and likely impact on human health: evidence from Brazil. *Public Health Nutrition*, *14*(1), 5-13.

Open Food Facts. (2022). *Open Food Facts - Germany*. https://de-en.openfoodfacts.org/

Sainz-De-Abajo, B., García-Alonso, J. M., Berrocal-Olmeda, J. J., Laso-Mangas, S., & Torre-Díez, I. D. L. (2020). FoodScan: Food Monitoring App by Scanning the Groceries Receipts. *IEEE Access*, *8*, 227915-227924.

Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann

Sazonov, E., Schuckers, S., Lopez-Meyer, P., Makeyev, O., Sazonova, N., Melanson, E. L., & Neuman, M. (2008). Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior. *Physiological measurement*, *29*(5), 525.

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2014). Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios. In A. Marcus (Ed.), *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience. DUXU 2014* (pp. 383-392). Springer.

Siegel, K. R. (2019). Insufficient Consumption of Fruits and Vegetables among Individuals 15 Years and Older in 28 Low- and Middle-Income Countries: What Can Be Done? *The Journal of Nutrition*, *149*(7), 1105-1106.

The IQVIA Institute. (2021). *Digital Health Trends 2021*. https://www.iqvia.com/insights/the-iqvia-institute/reports/digital-health-trends-2021

Thompson, F. E., & Subar, A. F. (2017). Chapter 1 - Dietary Assessment Methodology. In A. M. Coulston, C. J. Boushey, M. G. Ferruzzi, & L. M. Delahanty (Eds.), *Nutrition in the Prevention and Treatment of Disease (Fourth Edition)* (pp. 5-48). Academic Press.

Thorndike, A. N., Sonnenberg, L., Riis, J., Barraclough, S., & Levy, D. E. (2012). A 2-Phase Labeling and Choice Architecture Intervention to Improve Healthy Food and Beverage Choices. *American Journal of Public Health*, *102*(3), 527-533.

Visser, M., Schaap, L. A., & Wijnhoven, H. A. H. (2020). Self-Reported Impact of the COVID-19 Pandemic on Nutrition and Physical Activity Behaviour in Dutch Older Adults Living Independently. *Nutrients*, *12*(12), 3708.