Towards Tailoring Digital Food Labels

Insights of a Smart-RCT on User-specific Interpretation of Food Composition Data

Klaus Fuchs, Timothée Barattin, Mirella Haldimann, Alexander Ilic
D-MTEC, Auto-ID Labs ETH/HSG
ETH Zurich
Zurich, ZH, Switzerland
{fuchsk, barattit, mhaldima, ailic}@ethz.ch

ABSTRACT

Front-of-pack nutrition labelling (FoPL) supports healthier food choices yet remain unmandated in most countries. Simultaneously, labels are criticized for giving standardized recommendations that overlook individual needs. To assess the potential of consumer-specific tailored labels, we thus developed and tested a tailoring logic for adapting labels to individual dietary requirements and a smartphone app that then provided tailored food labels after scanning a product's barcode. The tailoring logic was developed with dieticians, accounting for gender, age, activity, preferences, diet-related diseases. The label showed a combination of established labelling systems: Nutri-Score and Multiple Traffic Light. The application followed a smart-RCT design, randomly attributing users either with tailored or standardized labels. 33 users met the eligibility criteria for our exploratory study. We found promising evidence that tailored digital food labels are perceived as more helpful, relevant, and recommendable than current static food labels, especially in the absence of FoPL.

CCS CONCEPTS

• Information systems • Personalization • Applied computing • Consumer health • Human-centered computing • Smartphones

KEYWORDS

Digital food label, Tailoring, mHealth, Front-of-pack nutrition labelling, Food purchase behavior, Diet intervention, Nutri-Score

ACM Reference format:

Klaus Fuchs, Tobias Grundmann, Mirella Haldimann and Elgar Fleisch. 2019. Impact of Mixed Reality Food Labels on Product Selection: Insights from a User Study using Headset-mediated Food Labels at a Vending Machine. In *Proceedings of ACM MADiMa '19, October 21, 2019, Nice, France. ACM, 8 pages.* https://doi.org/ 10.1145/3347448.3357171

MADiMa '19, October 21, 2019, Nice, France

ACM ISBN 978-1-4503-6916-9/19/10…$15.00.
DOI: http://dx.doi.org/10.1145/3347448.3357171

1 Introduction

The Status Quo in Food Labelling. Diet-related non-communicable diseases (NCD's) have become the leading cause of mortality globally, accounting for more deaths than non-diet-related mortality causes combined [15]. Consequently, dietary intake has become a recognized public health priority [12]. To address the rising ubiquity of NCD's, selected national regulators have begun to introduce FoPL, most notably UK's Multiple Traffic Light (MTL) or France's Nutri-Score (NS) in France [25, 41]. In parallel, mobile Health (mHealth) technologies are establishing themselves as an inclusive, scalable and supportive conduit to health behavior change for patients as well as health-care systems [1, 3]. There exists thus the potential for the joint application of FoPL and mHealth to support food recommendations and dietary interventions [37, 43], especially relevant in regions where FoPL have not yet been mandated.

While FoPL has been shown to positively influence food choices [25, 28], few countries have managed to successfully implement them on a large scale. Multiple barriers to food label adoption exist. First, regulators are confronted with resistance from retailers, brands or manufacturers, who worry about negative economic impacts due to increased administrative and logistical efforts associated with the introduction of mandatory food labels on their products. Second, companies that offer potentially negatively labelled foods may fear declining revenues [43]. Third, societies and public debates might turn against regulators aiming to introduce food labels, since individuals may dislike being patronized by regulators about dietary consumption. Consequently, only few countries have successfully managed to introduce FoPL, including Australia with the Health Star Rating (HSR) [41], France with NS or England with the MTL label [25]. Most countries adhere to conventional nutrition labelling regulation, requiring back-of-package labelled (BoPL) information on ingredients and nutrients. Previous studies have demonstrated that especially for low-literate consumers, BoPL is significantly less effective compared to FoPL [25, 41]. As a result, today most consumers cannot yet benefit from easy-to-compare FoPL on their food choices.

Standardization compounds this issue. Printed FoPL labels have been criticized for being too generic, and in term fall short of catering to dietary needs of important socio-demographic segments [25, 39, 41]. Among consumers as well as health-care professionals, different labels such as NS, HSR or MTL are viewed differently due to their varying advantages and disadvantages. For example, the MTL is better geared toward informed consumers or consumers with specific diets (e.g. low-sugar or low-sodium) due to the selection of certain nutrition information. In contrast, the NS seems better positioned for nutrition-illiterate consumers, since its easy-to-interpret visualization does not require fundamental understanding of nutrients (macro-, micro-nutrients, minerals or vitamins) or recommended daily allowances [26, 27]. As such labels might be effective for a large part or even the majority of consumers, they still fail to consider individual dietary needs and requirements of consumers with specific needs or preferences, such as people who are disease-affected, elderly, or very (in-)active citizens.

Tailoring of food labels. With recent advances in mHealth, user-specific adaptive tailoring of health behavior interventions has become best-practice, yet has not yet been established in apps aiming to improve or support food purchasing behavior. In this study, we argue that the complex and dynamic nature of food shopping behavior calls for the development of such user-specific and just-in-time adaptive interventions (JITAI) based methods, especially for food choice support. JITAIs provide skill building (e.g., coping and planning strategies, decision-making), emotional support (e.g., encouragement), and instrumental support (e.g., feedback, reminders) and occur in an adaptive manner to facilitate support in the exact moment of need [34]. Though JITAIs can be administered through several means, advancements in smartphone technology and related wearables [24] allow for continuous in-the-moment participant monitoring and delivery of personalized coping strategies [20]. This makes mobile devices particularly well-suited for delivering JITAI interventions, and in the context of food purchasing may promise superior effectiveness as they can be triggered actively or passively in-store through mHealth.

The idea to include tailoring of FoPL as JITAI within a smartphone application extends the growing stream of diet-related mHealth research that examines diet outcomes in relation to food purchasing behavior via barcode scanning [11, 35, 37], which so far have neglected the role of tailoring user-specific needs within the application [11]. The aim of this study was hence to explore this topic through the design, implementation and validation of a mobile app 'BetterChoice' that would provide Swiss consumers with access to easy-to-understand, tailored nutritional information of barcode-scanned food items - in the absence of standardized front-of-pack nutrition labeling (FoPL) in Switzerland. We built on previous research applications [33, 43] that were designed for fully automated trials, delivering intervention remotely, and collecting individual-level data on outcomes in nutrition-labeling randomized controlled trials (RCTs). We included a tailoring logic, developed together with dieticians from the Swiss Society of Nutrition, which took into users' gender, age, physical activity, diet patterns and diet-related diseases into account before producing the food label for the respective user. The application is connected to, compares and combines several relevant food databases in Switzerland (including e.g. GS1 trustbox) with over 47'500 packaged foods. Therefore, the app can support users in the selection of healthier food choices through display of FoPL, in which the score and color-coding was calculated according to each user's dietary situation, and suggestion of substitutes with higher nutritional quality were also tailored to each user. After development, we explored the effect of tailoring digital food labels. In the following we present the development and functionality of the smartphone app and tailoring logic used for the trials and report usage statistics and common technical issues. We also report on the results of the post-hoc exploratory analysis self-reported usage intention as well as usage behavior of study participants over the intervention period.

2 System Design

The 'BetterChoice' mHealth application was designed and developed by IS and health researchers in collaboration with dieticians from the Swiss Society for Nutrition (SGE-SSN). In the following, we present the system design including i) smartphone application design, ii) digital food label design, and the iii) tailoring logic for the digital food label.

2.1 Application Design

The final versions of the app were submitted to the Swiss Apple and Android app stores. The initial version of the app was compatible with smartphones running either iOS 9 and above or Android 5. We borrowed from the design of established mHealth applications in the field of food purchase interventions [10, 11, 38], which support consumers in making healthier food choices through display of a digital food labels. Throughout several iterations with dietary experts from SGE-SSN and multiple rounds of user testing, the final design of the 'BetterChoice' was achieved. Several existing and some additional functions were identified as helpful by users or experts and mentioned below.

Barcode Scanning. First, similar to existing food label mHealth [33, 43], the app allows scanning of barcodes by activating the smartphone-integrated camera feature to capture a product's barcode-encoded Global Trade Item Number (GTIN) [6, 18].

Display of product details. Second, the app displays product details after scanning or selecting a product. Product details include the digital food label, ingredients and nutrients. Additionally, we included product images to increase the chances of visually recognizing familiar products when making a purchase decision [7, 19, 32]. Visual examination of product packages plays a key role in the search-phase of purchasing decisions, especially in the supermarket [2, 5, 8, 9]. Visuals can also improve learning, recognition and recall of product-related nutritional knowledge [4, 31]. As such, visuals could improve support for choice recommendation in fast-paced environments such as the purchasing process [28].

Recommendation of substitutes. Third, in order to support a user in making healthier product choices, users can discover healthier alternatives of higher nutritional quality within a product's category. In order to enable this feature, the database would classify each product along the recommended categorization scheme of the Swiss Public Nutrition Database [14].

Inclusion of weighted products. A lot of food items carry barcodes that encode product price, usually derived from its weight multiplied with a product's current price per kilogram (e.g. fruit, vegetables, meat, and so on). Conventional barcode-scanning mHealth apps do not correctly identify these products, due to their often uniquely generated barcode. We therefore programmed a reverse-mapping to identify the corresponding base product (2110085000005 = 'Le Gruyère Cheese') of any weighted product (e.g. barcode 2110085004959 identifies 4.95 Swiss Francs of 'Le Gruyère Cheese'). As long as product barcode labelling within retail outlets remain compliant with the GS1 GTIN standard [6] including the checksum, such a reverse-mapping yields reliable identification of weighted products. This addition appears important since weighted products often include diet-relevant data.

Missing information, products & crowdsourcing. The app also allowed users to add corresponding category information in case a scanned product was not yet categorized. This crowdsourced product-category mapping was then added to the server and made available to all users after a manual confirmative check by the experts. The app also allowed error reporting of false entries and crowdsourcing of new products that were not covered by the server's database. Users were invited to add pictures of the front and back of each product, an established approach in purchase-related mHealth applications [10, 11, 38].

Convenience. To support users in identifying healthier substitutes when shopping, the application allowed storing discovered products within an easy-to-access favorite list. From the favorite list, the user can again access the product detail pages of each stored product in order to check the digital food label, see its nutrients or ingredients or retrieve healthier substitutes.

Usability & Learning. In order to support users in familiarizing themselves with the app, an introductory tutorial regarding the app functionalities was included. Further, to enable gaining nutritional literacy, the app included helpful, educative tips regarding a scanned product's category or nutrients. In total, 130 of such texts and icons were produced by the Swiss Society for Nutrition and shown while loading the list of product substitutes including images from the server.

Languages. All text elements within the 'BetterChoice' app were translated into English and three of the four national languages of Switzerland (German, Italian, French). The application would automatically set its language to the device's default language and allow users to switch the language in the application's settings.

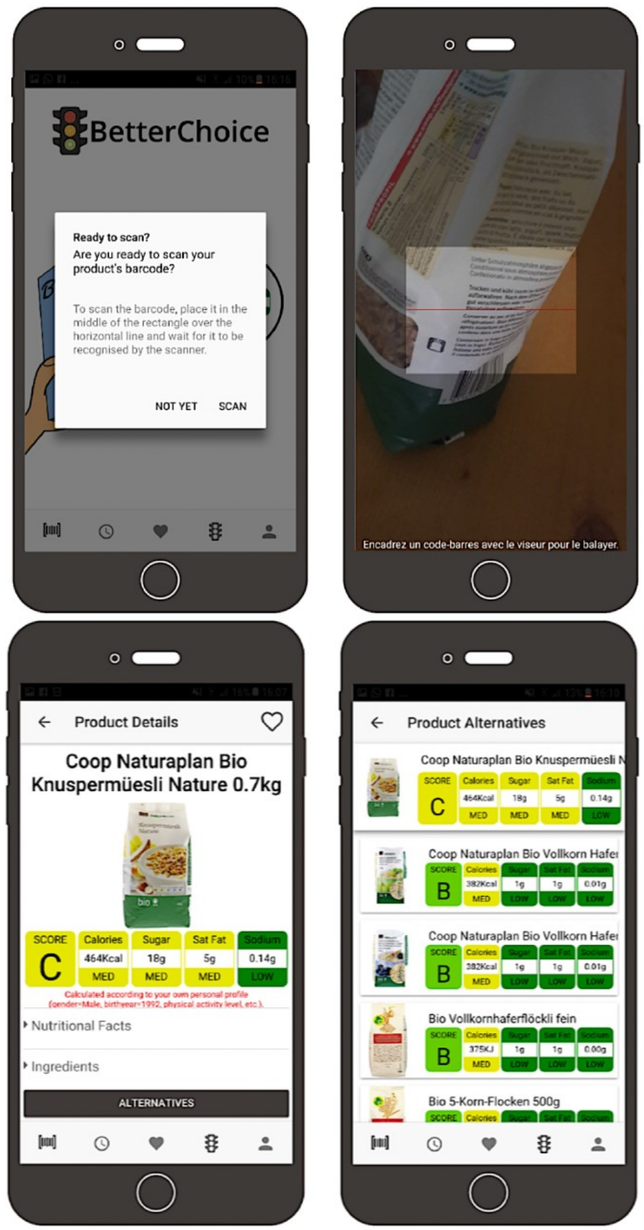

Figure 1: Functions within the 'BetterChoice' application

Data Privacy. To enable the RCT [43], the application required users to set up and maintain their user profile, from which the tailoring logic would retrieve the necessary information to calculate the tailored digital food label. The profile setup page captured valuable information necessary for identifying the dietary needs including gender, age, physical activity, diet patterns and diet-related diseases. No personally identifiable information was captured, i.e. no emails, addresses or phone numbers to minimize the risk of data breaches. The username was stored in the server backend as anonymized MD5-hash in order to allow anonymous processing of data for the post-hoc analyses.

Blocking product categories. To reduce processing of irrelevant items, we blocked not diet-related (medicine, beauty and healthcare products, tobacco, alcohol) or special food categories that ideally require advice from health-care professionals (e.g. baby food, supplements).

2.2 Digital Food Label

The digital food label within 'BetterChoice' represents a combination of the NS and MTL label [22], compounding the advantages of these two effective and established labels [35]. This decision led to multiple advantages. First, the digital food label includes the easy-to-use single-dimension Nutri-Score that allows nutrition illiterate users to identify healthy products through reducing decision complexity into one easy-to-decode color and letter scheme [25]. Second, more literate users can identify relevant nutrient information, such as amounts of sodium or sugar contained in a product. Third, this combination also allows producing a food label even in the absences of NS relevant information in the database, such as product category (e.g. beverages, added fats, cheeses and vegetable) or share of vegetables/fruit/nuts, as mandated by the Nuri-Score [26]. If such entries were missing, the 'BetterChoice' app could nonetheless compute MTL scores and asked users to specify the missing category values [38].

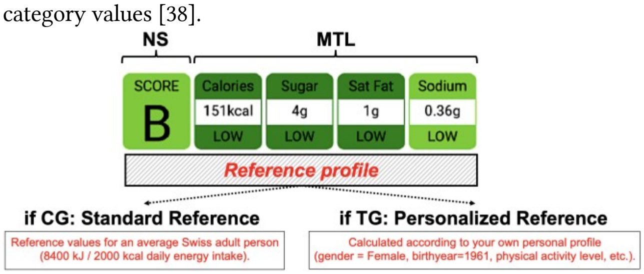

Figure 2: Digital food label within research application

The corresponding Nutri-Score, color-coding of each field and the stated reference profile of the digital food label is either tailored or standardized based on the user allocation to treatment or control group. If tailored, the label is re-calculated to the user's dietary needs according to the tailoring framework in real-time. Recipients of the standardized and tailored intervention received corresponding labels as depicted (Figure 2: red text below standardized versus tailored label). Depending on the profile, the same product can yield differently color-coded food labels, depending on the tailoring framework (Figure 3). In the given example, a 58yr-old female with average physical activity scanned the same item as an active 27yr-old man.

2.3 Tailoring Framework

The tailoring logic (table 1) was developed together with dieticians from the Swiss Society of Nutrition, and took into account a users' gender, age, physical activity, diet patterns and diet-related diseases, when generating the food label. The tailoring framework built on the NS framework [25], considering recommended daily energy and nutritional intake guidelines, and tailored the corresponding point attribution as defined by the Nutri-Score to each user. The corresponding point thresholds were increased or decreased according to the table (e.g. labels for sedentary, older users were colored in red earlier than for younger and more active users).

Figure 3: The same product can have differently tailored labels for users with different dietary needs.

In addition, the tailored labels were automatically adapted for diet-related diseases and dietary patterns in order to incentivize health-beneficial behavior, e.g. purchasing healthier products. Therefore, diabetic users who received tailored labels, were shown a 2-point malus for consumption of energy and sugar. Users with hypertension received 1-point malus for saturated fats, as well as 2-points malus for salt, but were rewarded +2 points for consuming fruit/vegetable/nuts. Also, to incentivize protein intake, vegan and vegetarian users were credited 2 additional points for protein, therefore protein-rich products appeared healthier than protein-low substitutes.

Similar to the original Nutri-Score calculation, the tailoring framework with its transformed point allocation would then add the negative points of undesired nutrient amounts and subtract positive points of health-beneficial nutrients. Afterwards, the corresponding NS were then added to the tailored label. The color coding was also adapted from the original Nutri-Score color-coding and represented the points given for each nutrient after applying the tailoring framework, also ranging from bright green (min. 0 negative points or max. 5 positive points) to bright red (max. 10 negative points or min. 0 positive points) for each nutrient. Additionally, labels for the recommended healthier substitutes were shown, and tailored for users in the treatment group. The tailoring framework was not designed with aspiration for perfectionism, but to demonstrate feasibility of user-specific tailoring for digital food labels. More work is needed to further develop the concept of the tailoring work in order to comply well with other potential dietary situations as well (e.g. micro-nutrients, cultural elements, social environments, other diseases e.g. anorexia or allergies).

Table 1: Tailoring of Digital Food Label per Person & Item

Tailoring	Energy	Sat Fat	Sugar	Salt
Gender				
Male	+12.5%	+12.5%	+12.5%	-
Female	−12.5%	−12.5%	−12.5%	-
Other	-	-	-	-
Age				
18-44	-	-	-	-
45-65	−10.0%	−10.0%	−10.0%	-
over 65	−20.0%	−20.0%	−20.0%	-
PAL				
1.4	−16.0%	−16.0%	−16.0%	−16.0%
1.5	−12.0%	−12.0%	−12.0%	−12.0%
1.6	− 8.0%	− 8.0%	− 8.0%	− 8.0%
1.7	− 4.0%	− 4.0%	− 4.0%	− 4.0%
1.8	-	-	-	-
1.9	+4.0%	+4.0%	+4.0%	+4.0%
2.0	+8.0%	+8.0%	+8.0%	+8.0%
2.1	+12.0%	+12.0%	+12.0%	+12.0%
2.2	+16.0%	+16.0%	+16.0%	+16.0%
2.3	+20.0%	+20.0%	+20.0%	+20.0%
BMI				
<18.5	+15.0%	+15.0%	+15.0%	-
[18.5, 25.0)	-	-	-	-
[25.0, 30.0)	− 5.0%	− 5.0%	− 5.0%	-
>30.0	− 10.0%	− 10.0%	− 10.0%	-
Legend: Energy (KJ), Saturated Fat (g), Sugar (g), Salt (g), all per 100g of product, PAL = Physical Activity Level				

3 Methodology

The goal of this study was the assessment of potential for tailored digital food labels on usage behavior, usage intention and performance expectation, when using a purchase-related mHealth system. The 'BetterChoice' application's key functionalities therefore included automation of the RCT including randomized segmentation of users, eligibility screening, trial intervention delivery (i.e. tailored versus standardized food labels), data collection facilitation, similar to related studies [43]. The app hence also included duplicate checks, automatic in-app logic checks, collection of informed consent, questionnaire administration, and outcome data collection from usage behavior and survey completion.

Participant recruitment. The smartphone application was made available to smartphone users in Switzerland aged 18 years or older via the Android and iOS app stores and enabled the conduct of a fully automated smartphone-delivered intervention trial. Between October 2018 and June 2019, in total 1024 users installed the 'BetterChoice' application via the iOS (601 installs) and Android (423 installs) app stores. There was no monetary incentive for users to participate in the study, nor were any large-scale promotions used for the study. In order to observe natural behavior among consumers, the users received no strict study protocol and were able to use the app when they wanted and as long as they wanted. During this time 33 participants successfully completed the study.

Intervention. As discussed, the generation of tailored digital labels followed the guidelines of actively triggered Just-in-time Adaptive Interventions (JITAI) [34]. The automated, RCT evaluated the effects of tailoring these labels [33, 43] onto proximal outcomes such as usage statistics, intention as well as performance expectation, based on the Technology Acceptance Model [29, 44]. Therefore, the application automatically collected usage data (i.e. number of product scans and number of crowdsourced new products) and structured user surveys at the end of the study period. Upon first start of the 'BetterChoice' application, a user's device was either assigned to i) treatment group (i.e. user-specific tailored food labels) [TG = 50%] or ii) control group (i.e. non-tailored standardized food labels) [CG = 50%], based on a smartphone's unique identifier. All users' assignments were static and could not be changed on the same device, also not through re-installment of the app. Therefore, even after a potential re-install of the app, the same device would be re-classified into the same group that was randomly attributed at the first installation. The generation of digital food labels within the app relied on each product's corresponding nutrient composition in the food composition database [21], which currently contains nutrition information for more than 47'500 products. Next, the application would interpret the nutritional product composition via the systematic generation of the food labels (see System design). In intervention delivery mode, nutrition labels were displayed for products, if matched successfully with an existing product in the database (the successful match rate was approximately 70%).

Eligibility checks. In order to gather meaningful results on the impact of tailored digital food labels, only eligible users were invited to participate in the study. Therefore, users were only invited to participate in the study after a minimum usage of eight product scans and having used the application on at least four separate days. In total, 33 users decided to give their consent and opted in to participate in the survey and share their usage statistics with the research team in anonymous form. The participation rate of 3.5% indicates a typical and acceptable response rate for public, non-requested invites to surveys, such as e-mail marketing which often have response rates between 1% and 5%.

Data collection. In data collection mode, eligible users were invited to share the number of scanned products, the number of crowdsourced items and end-survey questionnaire data which would be sent to the server for outcome assessment. All recorded data were transmitted via Wi-Fi or mobile networks to the trial database which was hosted on a secure remote server. Additional information of the user profile, including demographics, physical activity, diet patterns and diet-related diseases, and purchase behavior was also collected via in-app baseline questionnaire. The following ethical and security requirements were adhered to: (1) privacy-by-design: all collected data were anonymized on device and only non-personal aggregates were sent to the trial server. Therefore, no locations, no timestamps, no individual product scan logs, no addresses were stored on the server and are therefore not available for data assessment, in order to avoid processing of personalizable data. (2) Participant information statement was available to users via the app store and link throughout the trial.

Data analysis. The data analysis applied descriptive statistics and compare usage behavior (number of scans), self-reported intention to use (three items), as well as performance expectation (three items) between the two groups: TG and CG. Statistical analyses were performed in R. To perform the comparison tests, we conducted Wilcoxon-Mann-Whitney tests [13], due to the detected non-normality in the sampled population based on a Shapiro-Wilk test. All statistical tests were two-sided and tested at the 5% significance level. As large-scale studies in the field demonstrate, a sustained, significant effect on food choices from a mHealth intervention shall not be expected, as such habitual change requires more than an initial intervention [23, 36]. Therefore, the assessment of scanned food groups or mean degree of nutritional quality of products (such as NS or Nutrient Profiling Scoring Criterion values) were not calculated and compared between TG and CG, as they require further data collection, e.g. of receipts of purchased food items or diary logs, e.g. food diaries.

4 Results

4.1 Sample Description

Participants included in the study had a mean age of 38.2 (SD = 13.3) years, 48% were female, and 59% tertiary educated. There were no significant differences between the TG and CG across any of the sample describing dimensions, indicating a successful randomization for the RCT (table 2).

4.2 Usage Statistics

Over the study period and at the time point of registration up to filling in the survey, participants scanned on average 13 products and crowdsourced on average 3 products that were previously not in the database (table 3). It is observable that the treatment group on average scanned 27% and crowdsourced 94% more products compared to the control group. However, there were no significant differences between treatment and control group in terms of usage statistics, neither for the number of scanned products, nor in regard to crowdsourcing new products.

Table 2. Sample Description of Study Participants (N=33)	
Age	**mean (SD)**
years	41.43 (13.02)
Body Mass Index	**mean (SD)**
BMI in kg/m2	23.37 (3.93)
Weight	**count (%)**
Underweight (BMI <18.5)	2 (6%)
Normal (18.5 ≤ BMI < 25)	23 (70%)
Overweight (25 ≤ BMI < 30)	7 (21%)
Obese (BMI ≥ 30)	1 (3%)
Gender	**count (%)**
female	15 (45%)
male	17 (52%)
other	1 (3%)
Education	**count (%)**
to middle school	4 (12%)
to vocational or high school	10 (30%)
to BA/MA	16 (48%)
to PhD	3 (9%)
Diet Patterns	**count (%)**
Omnivore	29 (88%)
Vegan/vegetarian	4 (12)
Household Purchasing Decision	**count (%)**
No Decision	6 (18%)
Shared	15 (45%)
Sole	12 (37%)
Diet-related diseases	**count (%)**
Diabetes Type I	2 (7%)
Diabetes Type II	0 (0%)
Hypertension	4 (12%)
Physical Activity	**work, leisure**
Very Light	30%, 0%
Light	48%, 12%
Moderate	22%, 36%
Active	0%, 40%
Very Active	0%, 12%

Table 3. Usage Statistics (N=33)	
Scanning ($N_{products}$ = 429)	**mean (SD)**
All users	13.00 (6.18)
TG	14.12 (5.36)
CG	11.81 (6.76)
Crowdsourcing ($N_{crowdsourced}$ = 89)	**mean (SD)**
All users	2.70 (2.91)
TG	3.53 (3.35)
CG	1.81 (2.01)

4.3 Intention to Use

First, we compared differing perceptions about their self-reported intention to use the app, including the respective tailored or standardized label. We asked whether users in the tailored group (TG) and non-tailored group (CG) would use the app in future (1.1: $\Delta_{TG\text{-}CG}$=0.59, *P=.055*), would recommend the app to others (1.2: $\Delta_{TG\text{-}CG}$=0.80, *P=.030*), have a positive opinion about the app (1.3:

$\Delta_{TG\text{-}CG}$=0.37, *P=.40*). While the differences in the average values show a strong tendency towards favoring the tailored label, the overall intention to use (weighted average of the three items) is not significantly different. However, we did find a significant difference in terms of whether they would recommend the app (1.2) to others (friends and family) between responses in the tailored group (M=4.56, SD=0.61) and the control group (M=3.73, SD=1.06). Given the visible positive tendency for future usage (1.1) and positive opinion (1.3), and considering generally low continued usage in mHealth [30], these results suggest that tailoring may improve beneficial aspects of intention-to-use.

Table 4. Self-Reported Technology Acceptance (N=33)

Construct	**TG mean (SD)**	**CG mean (SD)**	***P*-value**
Intention to Use	4.29 (0.63)	3.70 (1.01)	.16
1.1: I intend to keep using the app during my next shopping trips over the next weeks.	4.00 (0.61)	3.41 (0.91)	.055[A]
1.2: I will recommend the app to my friends, because I think they should try it out.	4.56 (0.61)	3.76 (1.06)	**.030**[*]
1.3: I have a very positive opinion/perception about the app.	4.31 (0.68)	3.94 (1.06)	.40
Performance Expectancy	4.10 (0.94)	3.38 (0.94)	.078[A]
2.1 This app supports me in my struggle to identify healthy products among the many products available today.	3.94 (0.90)	3.53 (0.85)	.18
2.2 This app helps me in assessing my dietary intake from my grocery purchases.	4.19 (1.01)	3.31 (0.91)	**.039***
2.3 This app gives me recommendations that are very relevant to my personal lifestyle.	4.18 (0.81)	3.29 (1.07)	**.018***

*: significant on *P<.05*, [A]: significant on *P<.1*
Agreement Likert scale for items [1=very low ; 5=very high]

4.4 Performance Expectancy

We then examined whether groups had differing expectations regarding app performance. We asked users in both groups whether they thought the app was helpful to identify healthier products (2.1: $\Delta_{TG\text{-}CG}$=0.41, *P=.18*), supporting in assessing dietary intake (2.2: $\Delta_{TG\text{-}CG}$=0.88, *P=.039*), and whether the app provides relevant recommendations in regard to their personal, individual life-style (2.3: $\Delta_{TG\text{-}CG}$=0.89, *P=.018*). While again, the differences in the average values show a strong tendency towards favoring the tailored label, the overall performance expectation (weighted average of the three items) is not significantly different. However, we did find significant differences between the groups in their perceived performance expectation towards assessing dietary intake (2.2) and perceived personal relevance of the recommendation to one's personal lifestyle (2.3). These findings suggest that tailoring has significant effects on performance expectation via perceived quality of dietary assessment and perceived personal relevance.

5 Discussion

In this paper we developed, implemented and assessed a novel tailoring framework allowing for personalization of digital food labels based on user-specific dietary needs within a purchase-related mHealth application. This exploratory study was conducted as a fully automated RCT across 33 users who fulfilled the eligibility criteria, agreed to anonymously sharing their usage behavior and successfully completed the study survey. The study aimed to observe whether tailored labels have the potential to overcome the criticized disadvantages of current purchase-related mHealth and FoPL labels [25, 39, 41].

Benefits of study protocol include the observable advantages of tailored labels over standardized labels across the intention to use and performance expectancy items assessed in this study. Albeit not always significant, tailored labels scored higher on every single dimension when compared to their static counterparts. Also, the values for all dimension were significantly higher than neutral (Likert scale for all items ranked from 1 to 5). This indicates that tailored digital food labels indeed have potential to feature increased expected performance as well as intention-to-use. More specifically, this study shows that especially in terms of perceived personal relevance and perceived helpfulness in assessing one's dietary intake, tailored labels score significantly higher when compared to standardized labels. Both represent very relevant advantages in the design of purchase-related health interventions. In addition, the study demonstrated that there was a significant increase in the intention to recommend the application when receiving tailored labels. In conclusion, even despite the fact that there was no observable significant difference in actual label usage or crowdsourcing of new product items, the findings suggest the hypothesized potential of improving at least some important aspects of behavior change: intention-to-use and performance expectancy.

Strengths of this study include its rigorous RCT design with strict assignment of either tailored or standardized labels and device-based, privacy-preserving duplicate checks, automatic in-app logic checks, collection of informed consent, questionnaire administration, and outcome data collection from usage behavior and survey completion. To our knowledge, this study was the first nutrition labeling RCT in a real-world setting that assessed the potential of user-specific digital food labels, with a successful randomized assignment of tailored vs standardized nutrition labels for packaged products in any store across an entire country.

The findings of this study ought to be considered with certain limitations. The sample represents a first limitation. In this study, we were able to attract users from all sorts of socio-demographic and biometric segments. This is especially worth mentioning, as similar, even large-scale studies on purchase-related mHealth seem to suffer under much stronger biases towards female users (e.g. 88% female users) [36], or to healthy users [45]. When compared to Switzerland's societal distribution of the body-mass-index [17], the study sample and population have similar characteristics (Underweight: CH (2%) vs. Study (6%), Normal: CH (54%) vs. Study (70%), Overweight: CH (31%) vs. Study (21%), Obese: CH (13%) vs Study(3%)). Nonetheless, further studies could aim for a stratified sampling strategy to increase the amount of non-healthy, older and less nutritionally literate segments of society. The sufficient yet low sample size (N = 33) may represent another closely related limitation for external validity and may have played a role in limiting the significance of some of the statistical comparisons conducted. It is also worth mentioning the low retention rate in this context. As described earlier, 4% of all users decided to participate in the optional study. To increase such figures future studies may thus opt for a more invasive study design. In any case, we recommend optimizing and testing the user conversion rate for future studies on purchase-related mHealth. Another limitation deals with the fact that we selected individual established concepts (intention-to-use, performance expectancy) to gauge potential for tailored digital food labels which we saw as especially relevant in the initial stage of a JITAI behavioral change (BC) intervention. Future study designs could assess users' compliance and intervention efficacy. This limitation is closely linked to the decision to maximize user privacy and not to store detailed product scan logs. In future studies, such data would yield interesting insights into food categories and mid-term effects of tailored digital food labels on scanning behavior and food choice [42]. Another important limitation deals with barcode identification. Approximately 70% of barcode scans were successfully identified in the server backend. Even though, data of 47'500 products were available for the study, there are still thousands of products missing. This problem is compounded when users frequently buy groceries abroad, as can be the case for users who live or work in border regions. Practitioners and researchers should be aware of the effort involved in preparing and maintaining such a country-wide product database, as new product continuously enter the retail landscape. A final limitation, we believe warrants additional attention is the way we designed our intervention labels. As discussed previously, much effort was put into establishing an effective tailoring design. Nonetheless it remains unclear how label aspects may affect such interventions. A future line of research might be interested to compare such label aspects, by for example creating multiple treatment groups and tailoring towards even further aspects, e.g. cultural aspects or food literacy.

6 Conclusion

This study adds to the body of research in meaningful ways. First, the development of a tailoring framework for personalization of digital food labels has not been suggested before but represents a promising purchase-intervention even in the absence of FoPL. Tailored labels add to the body of purchase-related mHealth and yield relevant advantages over standardized labels such as perceived relevance, perceived helpfulness in identifying healthier alternatives and intention to recommend to others. Specifically, this study extends the important work on FoPL such as Nutri-Score or NCPS by proposing individualized, tailored labels, calculated via a scalable, novel tailoring framework that compares user profiles to product profile in real-time. Also, the reverse-mapping enabled barcode-scanning of weighted products, which previously was not possible due to the uniqueness of their price encoding. Given the relatively high compliance of Swiss retailers with the GS1 GTIN standard, reverse-mapping price-encoded barcodes to their base products represents a valuable extension to current purchase-related and diet-related mHealth applications. Especially since weighted products include processed cheese and meat as well as fresh fruit and vegetable, their impact on diets cannot be ignored.

In future works on this field, we plan to extend the findings of this study by assessing usage behavior across a longer observation period. Naturally, we aim to increase the larger sample size through a broader public campaign together with the SGE-SSN in the context of a future study on the 'BetterChoice' application. In addition, we suggest including assessment of the degree of nutritional quality of the scanned products and categories within a future study protocol, e.g. with the Nutri-Score points or NPSC points to see if tailored labels lead to an increased uptake of healthy food products. Future work also should assess automated input of or extension of the number of tailoring variables for digital food labels. Researchers could include dynamic tailoring of digital food labels based on recent logs from a food-logging application or purchase data from loyalty cards [18, 40]. Also, automated tracking of recent physical activity, genome-sequencing or blood tests could support the tailoring approaches. Last, but not least, the inclusion of user-specific food allergy information to tailoring of food labels seems promising, as the purchase mHealth could directly inform the user of relevant allergens in each food item [16].

As a final note, the introduction and increased usage of tailored food labels may also represent a novel approach to increasing transparency between products' ingredients and consumers. There exists consistent evidence, that the introduction of transparency measures such as FoPL can lead to healthier product reformulation by manufacturers [43]. Therefore, tailored digital food labelling could have an additional, indirect health-beneficial effect besides individual improvement of food choices via food reformulation [43], which could not be tested in the current study. Tailored, digital food labels might therefore, similarly to conventional FoPL, work to affect the diets of populations by creating an incentive for food manufacturers to improve the nutritional profile of their products.

REFERENCES

[1] Agarwal, R. et al. 2010. Research Commentary —The Digital Transformation of Healthcare: Current Status and the Road Ahead. *Information Systems Research*. 21, 4 (Dec. 2010), 796–809. DOI:https://doi.org/10.1287/isre.1100.0327.

[2] Andrews, J.C. et al. 2000. Are Some Comparative Nutrition Claims Misleading? The Role of Nutrition Knowledge, Ad Claim Type and Disclosure Conditions. *Journal of Advertising*. 29, 3 (2000), 29–42. DOI:https://doi.org/10.1080/00913367.2000.10673615.

[3] Beaglehole, R. et al. 2011. Priority actions for the non-communicable disease crisis. *The Lancet*. 377, 9775 (Apr. 2011), 1438–1447. DOI:https://doi.org/10.1016/S0140-6736(11)60393-0.

[4] Bergstein, J. and Erdelyi, M. 2008. Recognition hypermnesia: how to get it. *Memory (Hove, England)*. 16, 7 (2008), 689–702. DOI:https://doi.org/10.1080/09658210802169095.

[5] Breugelmans, E. et al. 2006. The effects of shelf display on online grocery choices. *KUL Working Paper No. MO 0607*.

[6] Brock, D.L. 2001. White Paper: Integrating the Electronic Product Code (EPC) and the Global Trade Item Number (GTIN). *MIT Auto-ID Center*. February 1 (2001), 1–25.

[7] ChanLin, L.-J. et al. 2003. Web-based instruction in learning nutrition. *Journal of Instructional Psychology*. 30, 1 (2003), 12–21.

[8] Clement, J. et al. 2013. Understanding consumers' in-store visual perception: The influence of package design features on visual attention. *Journal of Retailing and Consumer Services*. 20, 2 (2013), 234–239. DOI:https://doi.org/10.1016/j.jretconser.2013.01.003.

[9] Clement, J. 2007. Visual influence on in-store buying decisions: an eye-track experiment on the visual influence of packaging design. *Journal of Marketing Management*. 23, 9–10 (2007), 917–928. DOI:https://doi.org/10.1362/026725707X250395.

[10] Codecheck: Produkte checken und gesund einkaufen: 2016. *http://www.codecheck.info/*. Accessed: 2016-04-04.

[11] Dunford, E. et al. 2014. FoodSwitch: A Mobile Phone App to Enable Consumers to Make Healthier Food Choices and Crowdsourcing of National Food Composition Data. *JMIR mHealth and uHealth*. 2, 3 (2014), e37. DOI:https://doi.org/10.2196/mhealth.3230.

[12] FAO and WHO 2018. The Nutrition Challenge - Food System Solutions. *World Health Organization: Food and nutrition policies*. (2018).

[13] Fay, M.P. and Proschan, M. a. 2010. Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics Surveys*. 4, 0 (2010), 1–39.

[14] Federal Food Safety and Veterinary Office (FSVO). Swiss Food Composition Database (SFCD): 2016. *www.naehrwertdaten.ch*. Accessed: 2016-04-04.

[15] Forouzanfar, M.H. et al. 2015. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*. 386, 10010 (2015), 2287–2323. DOI:https://doi.org/10.1016/S0140-6736(15)00128-2.

[16] Frey, R.M. et al. 2016. Universal Food Allergy Number. *Proceedings of the 6th International Conference on the Internet of Things - IoT'16* (2016).

[17] FSVO 2017. Body Mass Index (BMI) in Switzerland 2014/2015. *Fachinformation Ernährung*. (2017), 8–10.

[18] Fuchs, K. et al. 2018. Automation of Data Collection Techniques for Recording Food Intake: a Review of Publicly Available and Well-Adopted Diet Apps. *2018 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, South Korea* (2018), 58–65.

[19] Fuchs, K.L. et al. 2016. Swiss FoodQuiz: Inducing Nutritional Knowledge via a Visual Learning based Serious Game. *ECIS 2016 Proceedings* (2016).

[20] Goldstein, S.P. et al. 2017. Return of the JITAI: Applying a Just-in-Time Adaptive Intervention Framework to the Development of m-Health Solutions for Addictive Behaviors. (2017). DOI:https://doi.org/10.1007/s12529-016-9627-y.

[21] GS1 Switzerland: Trustbox Database: 2018. *https://www.trustbox.swiss/*. Accessed: 2018-08-28.

[22] Hamlin, R. and McNeill, L. 2016. Does the Australasian "health star rating" front of pack nutritional label system work? *Nutrients*. (2016). DOI:https://doi.org/10.3390/nu8060327.

[23] Harrington, R.A. et al. 2019. A Pilot Randomized Controlled Trial of a Digital Intervention Aimed at Improving Food Purchasing Behavior: The Front-of-Pack Food Labels Impact on Consumer Choice Study. *JMIR Formative Research*. (2019). DOI:https://doi.org/10.2196/formative.9910.

[24] Hassannejad, H. et al. 2017. Automatic diet monitoring: a review of computer vision and wearable sensor-based methods. *International Journal of Food Sciences and Nutrition*. 68, 6 (2017), 656–670. DOI:https://doi.org/10.1080/09637486.2017.1283683.

[25] Julia, C. et al. 2017. Perception of different formats of front-of-pack nutrition labels according to sociodemographic, lifestyle and dietary factors in a French population: Cross-sectional study among the NutriNet-Santé cohort participants. *BMJ Open*. (2017). DOI:https://doi.org/10.1136/bmjopen-2017-016108.

[26] Julia, C. and Hercberg, S. 2017. Development of a new front-of-pack nutrition label in France: the five-colour Nutri-Score. *Public Health Panorama*. 3, 4 (2017), 712–725.

[27] Julia, C. and Hercberg, S. 2017. Nutri-Score: Evidence of the effectiveness of the French front-of-pack nutrition label. *Ernahrungs Umschau*. 64, 12 (2017), 181–187. DOI:https://doi.org/10.4455/eu.2017.048.

[28] Kalnikaite, V. et al. 2013. Decision-making in the aisles: Informing, overwhelming or nudging supermarket shoppers? *Personal and Ubiquitous Computing*. 17, 6 (2013), 1247–1259. DOI:https://doi.org/10.1007/s00779-012-0589-z.

[29] Kaushik, A.K. and Rahman, Z. 2015. An alternative model of self-service retail technology adoption. *Journal of Services Marketing*. 29, 5 (2015), 406–420. DOI:https://doi.org/10.1108/JSM-08-2014-0276.

[30] König, L.M. et al. 2018. Describing the process of adopting nutrition and fitness apps: Behavior stage model approach. *Journal of Medical Internet Research*. 20, 3 (2018). DOI:https://doi.org/10.2196/mhealth.8261.

[31] Kroll, J.F. and Potter, M.C. 1984. Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning and Verbal Behavior*. 23, 1 (1984), 39–66. DOI:https://doi.org/10.1016/S0022-5371(84)90499-7.

[32] Lu, Z.-L. et al. 2011. Visual perceptual learning. *Neurobiology of learning and memory*. 95, 2 (2011), 145–51. DOI:https://doi.org/10.1016/j.nlm.2010.09.010.

[33] Mhurchu, C.N. et al. 2019. Effects of interpretive front-of-pack nutrition labels on consumer food purchases: A randomized controlled trial. *Obesity Research & Clinical Practice*. (2019).

[34] Nahum-shani, I. et al. 2015. Building health behavior models to guide the developmnet of just-in-time adaptive interventions: A pragmatic framework. *Health Psychology*. 34, Supplement (2015), 1209–1219. DOI:https://doi.org/10.1037/hea0000306.

[35] Ni Mhurchu, C. et al. 2018. Do nutrition labels influence healthier food choices? Analysis of label viewing behaviour and subsequent food purchases in a labelling intervention trial. *Appetite*. (2018). DOI:https://doi.org/10.1016/j.appet.2017.11.105.

[36] Ni Mhurchu, C. et al. 2018. Do nutrition labels influence healthier food choices? Analysis of label viewing behaviour and subsequent food purchases in a labelling intervention trial. *Appetite*. 121, (Feb. 2018), 360–365. DOI:https://doi.org/10.1016/J.APPET.2017.11.105.

[37] O'Flaherty, M. et al. 2016. Exploring potential mortality reductions in 9 European countries by improving diet and lifestyle: A modelling approach. *International Journal of Cardiology*. 207, (2016), 286–291. DOI:https://doi.org/10.1016/j.ijcard.2016.01.147.

[38] Openfoodfacts Mobile Applications: 2019. *https://world.openfoodfacts.org*. Accessed: 2019-03-30.

[39] Roberto, C. a. et al. 2012. Evaluation of Consumer Understanding of Different Front-of-Package Nutrition Labels, 2010–2011. *Preventing Chronic Disease*. 9, 1 (2012), 2010–2011. DOI:https://doi.org/10.5888/pcd9.120015.

[40] Steele, R. 2015. An Overview of the State of the Art of Automated Capture of Dietary Intake Information. *Critical Reviews in Food Science and Nutrition*. 55, 13 (2015), 1929–1938. DOI:https://doi.org/10.1080/10408398.2013.765828.

[41] Talati, Z. et al. 2016. Do health claims and front-of-pack labels lead to a positivity bias in unhealthy foods? *Nutrients*. (2016). DOI:https://doi.org/10.3390/nu8120787.

[42] Van't Riet, J. 2013. Sales effects of product health information at points of purchase: A systematic review. *Public Health Nutrition*.

[43] Volkova, E. et al. 2016. "Smart" RCTs: Development of a Smartphone App for Fully Automated Nutrition-Labeling Intervention Trials. *JMIR mHealth and uHealth*. 4, 1 (2016), e23. DOI:https://doi.org/10.2196/mhealth.5219.

[44] Wang, W. and Benbasat, I. Trust in and Adoption of Online Recommendation. 6, 3, 72–101.

[45] Williamson, D.A. et al. 2006. Two-year internet-based randomized controlled trial for weight loss in African-American girls. *Obesity (Silver Spring, Md.)*. 14, 7 (2006), 1231–1243. DOI:https://doi.org/10.1038/oby.2006.140.