Original Paper

Development of an Intervention Targeting Multiple Health Behaviors Among High School Students: Participatory Design Study Using Heuristic Evaluation and Usability Testing

Ulrika Müssener[1], PhD; Kristin Thomas[1], PhD; Catharina Linderoth[1], MSc; Marie Löf[1,2], PhD; Katarina Åsberg[1], MSc; Pontus Henriksson[1], PhD; Marcus Bendtsen[1], PhD

[1]Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden
[2]Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden

Corresponding Author:
Ulrika Müssener, PhD
Department of Health, Medicine and Caring Sciences
Linköping University
Sandbäcksgatan 1
Linköping, 58183
Sweden
Phone: 46 732702426
Email: ulrika.mussener@liu.se

Abstract

Background: Mobile electronic platforms provide exciting possibilities for health behavior promotion. For instance, they can promote smoking cessation, moderate alcohol consumption, healthy eating, and physical activity. Young adults in Sweden are proficient in the use of technology, having been exposed to computers, smartphones, and the internet from an early age. However, with the high availability of mobile health (mHealth) interventions of varying quality, it is critical to optimize the usability of mHealth interventions to ensure long-term use of these health promotion interventions.

Objective: This study aims to investigate the usability of an mHealth intervention (LIFE4YOUth) targeting health behaviors among high school students through heuristic evaluation and usability testing.

Methods: A preliminary version of the LIFE4YOUth mHealth intervention, which was aimed at promoting healthy eating, physical activity, smoking cessation, and nonrisky drinking among high school students, was developed in early 2019. We completed a total of 15 heuristic evaluations and 5 usability tests to evaluate the usability of the mHealth intervention prototype to improve its functioning, content, and design.

Results: Heuristic evaluation from a total of 15 experts (10 employees and 5 university students, both women and men, aged 18-25 years) revealed that the major usability problems and the worst ratings, a total of 17 problems termed *usability catastrophes*, concerned shortcomings in displaying easy-to-understand information to the users or technical errors. The results of the usability testing including 5 high school students (both girls and boys, aged 15-18 years) showed that the design, quality, and quantity of content in the intervention may impact the users' level of engagement. Poor functionality was considered a major barrier to usability. Of the 5 participants, one rated the LIFE4YOUth intervention as poor, 2 rated as average, and 2 assessed it as good, according to the System Usability Scale.

Conclusions: High school students have high expectations of digital products. If an mHealth intervention does not offer optimal functions, they may cease to use it. Optimizing the usability of mHealth interventions is a critical step in the development process. Heuristic evaluation and usability testing in this study provided valuable knowledge about the prototype from a user's perspective. The findings may lead to the development of similar interventions targeting the high school population.

XSL•FO
RenderX

Introduction

Health Behaviors Among Young People

Chronic diseases are the leading cause of death and disability worldwide. Globally, up to an estimated 80% of cases of heart disease, stroke, and type 2 diabetes and more than 30% of cancers can be prevented by reducing smoking, harmful alcohol use, improving diet, and engaging in regular physical activity [1,2]. Previous research has shown that multiple risk behaviors increase the risk of chronic disease and all-cause mortality, more so than the combined effects of single behaviors [3-5]. Health behaviors typically emerge during adolescence, track into adulthood, and commonly co-occur [6-8]. Therefore, adolescence is a critical age to intervene and interrupt a trajectory toward poor adult health [9-11]. Evidently, effective and evidence-based health behavior promotion interventions are needed.

Health Behavior Promotion Among Youths Through Multiple Mobile Health Interventions

Mobile platforms provide exciting possibilities for the promotion of health behaviors through mobile health (mHealth) interventions. Previous research has shown that interventions targeting multiple health behaviors at the same time might be effective in improving the general lifestyle among adults [5,12], with less evidence among adolescents [13]. A meta-analysis [14] examined the effectiveness of text message–based interventions for tobacco and alcohol cessation in a young adult population. Only 5 of the 14 studies reported significant differences between groups of substance use behavior outcomes, and the included randomized controlled trials (RCTs) lacked details regarding the intervention content. A more recent systematic review and meta-analysis [15] investigated the effectiveness of school-based eHealth interventions, defined as interventions delivered via the internet, computers, mobile technology, or telehealth, to prevent multiple health risk behaviors among adolescents. A total of 22 publications were included, assessing 16 interventions that targeted 2 or more of the following behaviors: alcohol use, smoking, diet, physical activity, screen time and sitting, or sleep. Only short-term effects were found for improving physical activity, screen time, and fruit and vegetable intake, and all trials were considered to be of low quality. There was limited evidence of their effect on alcohol and smoking. Although the selection criteria in the meta-analysis included studies with intervention components delivered via mobile technology, no studies of mHealth-only interventions were identified [15].

The Development of mHealth Interventions

Young adults are referred to as digital natives; many are proficient in the use of technology, having been exposed to computers, smartphones, and the internet from an early age. Indeed, young adults have the highest level of smartphone ownership among all age groups [16]. It is critical to optimize the usability of interventions targeting this age group; if they do not enjoy the program, they may disengage. Several reviews have shown the feasibility, acceptability, and efficacy of digital interventions for behavior change [17,18] and other health interventions, such as disease self-management [19-21] and medication adherence [22,23]. Previous systematic reviews on texting and mobile app interventions emphasized the urgent need to examine the development processes of mHealth interventions [23,24].

In a participatory design approach, research is undertaken with, rather than on, people, allowing researchers to gain an understanding of context-specific requirements and challenges [25]. Heuristic evaluation and usability testing are commonly used to support the development and refinement of the effectiveness of mHealth interventions. During heuristic evaluation, trained evaluators review an intervention to find usability problems, assign them to a specific category of heuristics, and ascribe a severity rating to provide distinct usability information [26,27]. Heuristics are often used to identify usability issues, such as problems with unclear functions, confusing navigation, and consistency issues [27-29]. Usability tests consist of a human-computer interaction and refer to evaluating an intervention by testing it with potential end users with the goal of identifying understandability, ease of learning, and attractiveness and to determine participants' satisfaction with the intervention. Usability testing provides developers with feedback about what does and does not work in the intervention and determines whether the features in the interventions are acceptable to and feasible for users and also determines what can be improved [30-33].

Objectives

This study aims to investigate the usability of an mHealth intervention (LIFE4YOUth) targeting health risk behaviors among high school students through heuristic evaluation and usability testing.

Methods

Procedures

A preliminary version of the LIFE4YOUth mHealth intervention targeting health risk behaviors among high school students was developed in early 2019. LIFE4YOUth is one of 7 multiple mHealth interventions in the MoBILE (mHealth–Multiple Lifestyle Behaviors) research program (funded by Forte 2018-01410, principal investigator: ML) aimed at promoting healthy eating, physical activity, smoking cessation, and nonrisky drinking in 7 different populations in the health care system. The intervention includes information about health behaviors, tips on behavior change strategies, and activities. The formative research process of developing a novel multiple mHealth intervention is described in detail in a study protocol elsewhere [34]. This study reports on findings from the first stage of the formative research process: heuristic evaluation and usability testing. The aim is to investigate the usability of a prototype app in terms of function, content, and design.

Setting, Participants, and Recruitment

Heuristic Evaluation

Recruitment of participants for the heuristic evaluation was undertaken by members of the research team through paper advertisements (posters) in public areas at Linköping University, Sweden. The inclusion criteria for the heuristic evaluation were

university students and employees, both women and men, aged 18 to 25 years, at the Faculty of Medicine and Health Sciences at Linköping University who were willing to participate and owned a mobile phone. Participants showed their interest by contacting the research leader via email or telephone.

Usability Tests

School staff at 5 high schools selected for convenience in Östergötland, Sweden, were contacted via email and informed about the research project. Students from all schools were invited to participate in the usability testing. The recruitment of participants was performed by school staff through paper advertisements (posters and leaflets), digital advertisements (student email and school website), and information disseminated in the classrooms. The inclusion criteria for the usability tests were high school students, both female and male, aged 15 to 18 years, willing to participate and owning a mobile phone. High school students showed their interest by contacting the research leader via email or telephone.

Data Collection

Informed Consent

All participants provided written informed consent before participation in all study procedures (heuristic evaluation and usability tests).

Heuristic Evaluations

A total of 15 experts (10 employees and 5 university students) were recruited. For the heuristic evaluation, a set of 10 heuristics published by Nielsen [26] was used to evaluate the LIFE4YOUth prototype. The heuristics for usability evaluation according to Nielsen are listed in Textbox 1.

We selected Nielsen's 10 heuristics because they have been thoroughly tested and are quick and easy to apply. When applying heuristic evaluation, participants evaluate an app to find usability problems, assign them to a specific category of heuristics, and assign a severity rating [26,27]. All participants were invited to a brief training session (approximately 45 min) conducted by a research assistant (CL), to receive instructions on the main principles of heuristic evaluation, and to learn to use the heuristics to evaluate the intervention [27]. The participants were sent a link to a high-fidelity prototype [35] of the intervention, including the actual software start page, menu page, and the 4 intervention modules (alcohol, smoking, physical activity, and diet). Each participant was asked to identify usability problems independently in a given protocol (Multimedia Appendix 1). Participants were asked to identify a problem, describe it, identify the relevant heuristic for the problem (eg, visibility of system status or match between the system and the real world), and give it a severity rating [26,27,29].

The procedure itself was a two-part process: the participants first familiarized themselves with the system and its usage with reference to the materials and training provided by the assistant. Then, they independently applied the 10 heuristics, as given in Textbox 1. The participants detected a usability problem, assigned each problem to a violation of a heuristic, and described the problem in their own words. Then, participants assigned severity scores to each problem by using the severity rating factors of impact presented in Textbox 2.

The introduction took place in the beginning of May 2019 in a conference room at Linköping University, Sweden. The evaluations were performed wherever the participants preferred and were sent to the research assistant in a prepaid envelope. After 10 days, a reminder was sent by a text message. Heuristic evaluations from all participants (n=15) were gathered during the last week of May 2019.

Textbox 1. Heuristics for usability evaluation according to Nielsen.

1. Visibility of system status: The system should always keep users informed about what is going on through appropriate feedback within reasonable time

2. Match between system and the real world: The system should speak the users' language, with words, phrases, and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order

3. User control and freedom: Users often choose system functions by mistake and will need a clearly marked *emergency exit* to leave the unwanted state without having to go through an extended dialog. Support undo and redo

4. Consistency and standards: Users should not have to wonder whether different words, situations, or actions mean the same thing

5. Error prevention: Even better than good error messages is a careful design, which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action

6. Recognition rather than recall: Minimize users' memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialog to another. Instructions for the use of the system should be visible or easily retrievable whenever appropriate

7. Flexibility and efficiency of use: Accelerators—unseen by the novice user—may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions

8. Aesthetic and minimalist design: Dialogs should not contain information that is irrelevant or rarely needed. Every extra unit of information in a dialog competes with the relevant units of information and diminishes their relative visibility

9. Help users recognize, diagnose, and recover from errors: Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution

10. Help and documentation: Although it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large

XSL•FO
RenderX

Textbox 2. Scale for severity rating according to Nielsen.

> 0: Not a usability problem
>
> 1: Cosmetic problem only—need not to be fixed unless extra time is available
>
> 2: Minor usability problem—fixing this should be given low priority
>
> 3: Major usability problem—important to fix, should be given high priority
>
> 4: Usability catastrophe—imperative to fix this before product can be released

Usability Tests

The usability test consisted of a human-computer interaction evaluation, which focused on the perceptions and performance of users in a laboratory setting. It entailed the completion of 30 goal-oriented tasks by targeted end users [30,32]. In total, 5 usability tests were completed with high school students. The tests were carried out in June 2019 in a medical informatics laboratory room and were run on an iPhone. The participants went through a 60-min session, during which all interactions with the intervention were recorded using video. The same prototype, as in the heuristic evaluation, was used. The test manager gave the participants printouts with the 30 tasks.

Examples of tasks used in the usability testing are described in Textbox 3.

The participants were asked to complete the 30 tasks, and the test manager asked the participants to explain their actions as they performed them using a think-aloud method [36,37]. An observer (the research leader) noted potential issues as the given tasks were performed by the participants. After completing the session, the participants completed a paper version of the System Usability Scale (SUS), providing a global view of their subjective assessment of usability based on 10 questions [38]. SUS has been shown to be reliable and valid with a variety of different technologies and users.

Textbox 3. Examples of tasks in the usability testing.

> - Go to the start page and explain what you think of the layout
> - Where can you find support to drink less?
> - Log what you drink. Explain how you interpret the graphs for the alcohol consumption you logged
> - You know that people feel healthier after exercising and want to find out more about getting started. How would you look for this information?
> - If you want to plan a physical activity, how do you find that activity, and specify your level of involvement?
> - You can set personal goals for your eating habits under the "Timeline" tab. Explain how to set your goal and what the graph shows you
> - What information is there under the "Risks" tab in the Diet module? Does it give you a good overview?
> - Go to the Smoking module. What do you think about the scope of the information provided on the first page?
> - If you want to know more about the benefits of quitting smoking, how would you go about this?

Data Analyses

Heuristic Evaluation

The 10 heuristics were pooled, and the identified problems were categorized as major issues [26-29]. The focus of the analysis was to identify usability problems and critical issues and to explore whether any functions in the intervention performed better than others. A master list was compiled to collect all the described usability problems, duplicate problems were removed to enable analyses, and the list was verified by 2 of the authors (UM and CL) for accuracy and to ensure validity and prevent bias in the analysis process. Descriptive statistics were used to summarize the heuristic violations and associated severity scores. The severity rating scale from 0 (not a problem) to 4 (usability catastrophe) is shown in Textbox 2.

Usability Tests

The usability tests were recorded and transcribed verbatim by a professional transcription company. All transcripts were checked and validated by the first author (UM). Analysis of the video recordings was inspired by program theory development using an inductive approach, taking both verbal and visible conduct into consideration [39]. Transcripts were analyzed thematically using an iterative coding procedure. The focus of the analysis was on the features of the intervention that needed to be redesigned with regard to function, content, and design. Overall, 2 authors (UM and CL) individually read the transcriptions and viewed the video recordings to acquire a comprehensive understanding. The categories were identified using an iterative process of reading and rereading the transcripts. Patterns were searched for, and usability issues were coded into categories. The first coding was initiated by the first author. Next, the coding was presented and discussed between the 2 authors (UM and CL), and boundaries for coding were established jointly [39].

The analysis of the SUS score was conducted according to the scoring strategy of Brooke [40], with the score for each item ranging from 0 to 4. The score from positively worded items (1, 3, 5, 7, and 9) is calculated as the scale position minus 1. For the negatively worded items (2, 4, 6, 8, and 10), the score is calculated as 5 minus the scale position. The sum of the scores is then multiplied by 2.5 to obtain the overall value of SUS ranging from 0 to 100. The average SUS scores were used to

identify average satisfaction [40]. According to Bangor et al [38], the average SUS score of approximately 70 can be interpreted as good or acceptable [38].

Results

Heuristic Evaluation

The heuristic evaluation resulted in a total of 121 usability problems and 131 heuristic violations reported by 15 participants. The usability problems identified through heuristic evaluation are summarized and presented by place of occurrence (eg, alcohol, diet, physical activity, and smoking modules as well as the start page), number of heuristic violations, and mean severity ratings in Figure 1. The alcohol module generated the maximum number of usability problems and problem descriptions (n=49), followed by the diet module (n=39), the

physical activity module (n=13), and the smoking cessation module (n=8). The start page also had usability problems (n=12). The average severity ratings ranged from 2.1 to 2.8 (on a scale of 0-4). The start page module had the highest severity rating, that is, 2.8.

The line in Figure 1 shows the mean severity ratings for usability problems on a scale of 0 to 4, where 0=none, 1=cosmetic problems, 2=minor problems, 3=major problems, and 4=usability catastrophe.

Of the 10 types of heuristic violations depicted in Table 1, categories 2 (*Match between system and the real world*) and 4 (*Consistency and standards*) dominated at 33 and 32, respectively. The heuristic categories 1 (*Visibility of system status*; n=4) and 10 (*Help and documentation*; n=3) had the lowest violations across all views.

Figure 1. Number of usability problems identified presented by place of occurrence, counts of heuristic violations identified by the 15 participants.
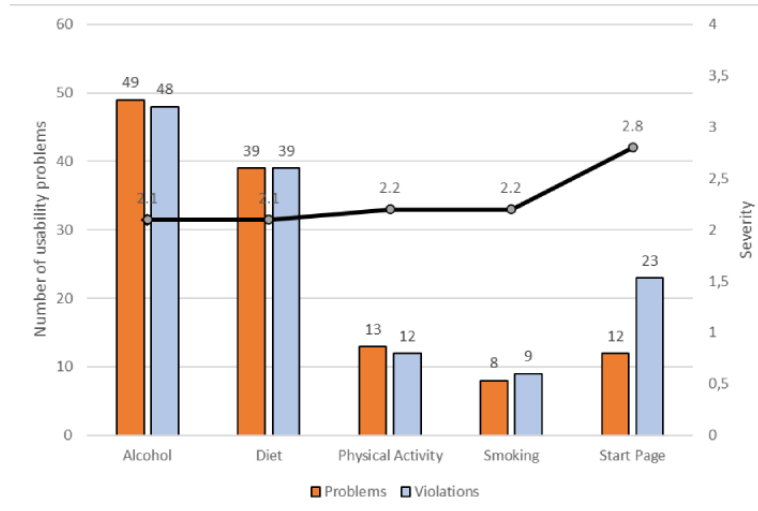

Table 1. Presentation of the frequencies of 10 heuristic violations sorted by Nielsen's heuristics reported by the 15 participants.

Violations	Number of participants
Visibility	4
Match	33
User control	6
Consistency	32
Error	8
Recognition	8
Flexibility	21
Aesthetics	10
Recover	5
Help	3

Most severity ratings were in the major severity category, indicating that fixing the problem should be given high priority (severity rating 3). The problems categorized with the highest severity rating were in the alcohol and diet modules. Most of the minor usability problems, that is, fixing those problems should be given low priority (severity rating 2), were also in the alcohol and diet modules. There were also a total of 17

problems reported as usability catastrophes across all modules, indicating the imperative to fix the problem before the product can be released (severity rating 4). Figure 2 shows severity ratings across system views.

When analyzing the nature of usability problems, the heuristic evaluation revealed that major usability problems and

XSL•FO
RenderX

catastrophic ratings concerned shortcomings in displaying easy-to-understand information to the users or technical errors. Examples of these types of usability problems provided by participants are as follows:

> *Navigation unclear. There's a constant mixture of headings [in the various modules], unclear headlining, different selectable functions, sometimes with hidden text. You don't know where you are* [in the intervention]

> *If the diagram's x- and y-axes don't have any labels, you don't understand because the table headings are unclear and don't stand alone.*

> *In the figures it's unclear which direction the scales go in, you don't understand the colors, it's confusing.*

> *Several choices of wording, and the concepts are difficult and complicated. Think about having simple, consistent wording and simpler language.*

> *It's not possible to save all the choices you've entered, you can't change your choices. The planning disappears when you browse through the tabs.*

> *You can't go back easily. In other words, it's not possible to find your way "home" easily, there's no home button.*

Figure 2. Number of severity ratings (minor, major, and catastrophic) for each module as reported by the 15 participants.
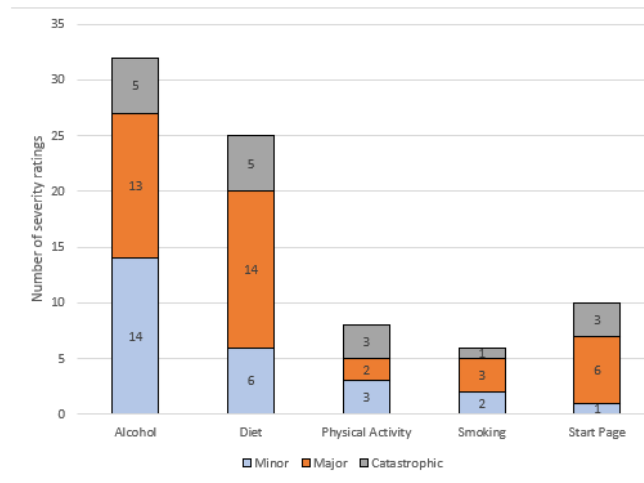
Usability Tests

A total of 5 high school students, 3 adolescent girls and 2 adolescent boys, participated in the usability testing. The analysis of the data from the transcripts revealed 3 main categories of barriers that limited usability: (1) design, (2) content, (3) and functionality.

Design

Design refers to how the intervention is perceived visually. The data showed that the design of the prototype influenced engagement and interest among users, thereby motivating or hindering user engagement with the intervention. Design comprises both aesthetics, that is, how attractive the intervention is perceived to be, and clarity, that is, how the intervention is structured and presented. For example, participants described how clear headlines and appealing infographics were integral to usability. Moreover, participants highlighted the use of clearly presented text, figures, and tables as a way to improve the usability of the intervention. In addition, characteristics of the intervention such as the use of attractive pictures, symbols, and videos were described as further improving usability:

> *And then I wonder if you could make it a bit more fun or something /.../ to have some background color other than white.* [Adolescent 2]

> *Perhaps it would be good to have images? I don't think it captures the user's interest much [without images].* [Adolescent 1]

Regarding the font size, participants preferred a larger font size and suggested clearly defined headlines:

> */.../ an intervention that looks a bit half-hearted and where things aren't, like, centered or whatever. Perhaps it looks less professional than it actually is, and so you don't trust it as much /.../ because you've put so much effort into it being scientifically correct, so you could put a bit of effort into it looking good or, like, fixing it.* [Adolescent 4]

> */.../ and bigger, clearer headings. I think bullet points are good, easy to read /.../ I don't think you get particularly excited using it.* [Adolescent 1]

Content

The second category refers to both the quality and quantity of intervention *content*. Problems relating to quantity were described as too much text, with large blocks of text making the information difficult to process, which subsequently limits the usability:

> */.../ there's a lot of text. It gets a bit, like, too much text to manage to read through everything. /.../ you don't want there to be too much text. It should be, like, quite quick and easy.* [Adolescent 5]

> *You can quite easily get tired with a lot of text.* [Adolescent 3]

> *Sometimes less is more /.../ if it's simpler, it's easier.* [Adolescent 4]

The quality of the content referred to the terminology whereby language was perceived as too heavy, too complicated, and difficult to comprehend. Indeed, words, phrases, and concepts were perceived as unfamiliar and not tailored to the target age group. Thus, engaging with the content was too taxing and limited the usability:

> *What on earth is "moderate level" and what on earth is "strenuous level"?* [Adolescent 4]
>
> *"Dietary index"? What does "dietary index" mean?* [Adolescent 1]
>
> *Very strange words. People don't use them at all!* [Adolescent 3]

Functionality

The third category *functionality* referred to the need for the app to be effortless to navigate, to be quick to use, and to have a logical flow, according to participants who wanted an easy-to-use interface. Poor functionality was considered a major usability barrier. Participants described that usability was limited when navigation was complex and included multiple modules or functions. For example, participants stated that scrolling to find information required too much effort and the intervention needed to be easier to grasp:

> *I think fewer stages would be good.* [Adolescent 1]
>
> *Yes, because otherwise perhaps it's a bit so-so, that you go in and first you can click there, and then click there /.../ in other words, it gets too much.* [Adolescent 2]

Participants stated that time was precious and they did not want to spend time navigating unnecessarily, such as entering what they had been eating or how many activity minutes they had participated in. Thus, engagement with the app must be effective and targeted to facilitate usability:

> *You should be able to get an overview extremely quickly.* [Adolescent 4]

> *Like, how many times I ate fruit or berries last week? Yes, I might have eaten fruit once. I might have eaten it ten times, I might not have eaten it at all. I don't know. It was a bit difficult.* [Adolescent 5]

Participants also described the importance of a logical flow between different components of the app. In addition, the features that guided or prompted navigation of the intervention could facilitate usability:

> *You need a bit more help orientating yourself, where you are. Or some kind of main menu that comes up, and then you can tap on alcohol and after you've tapped on that subheadings appear. Then you can choose between them.* [Adolescent 4]
>
> *Maybe [it would help if] everything is in categories instead, and you tap, and maybe then it appears. Not showing everything there from the start ... instead, you can go into what you're interested in.* [Adolescent 1]

Furthermore, relating to the logical flow within the intervention, there was a desire for consistency between the different modules to improve the usability of the intervention:

> *Because maybe you can't have the same subheadings for everything, but that they're still very consistent, that there are reminders and text messages, then they should be in all [modules] so you can get familiar with it and find things.* [Adolescent 1]

The SUS Score

The results of the SUS scores are given in Table 2. As shown in this table, the analysis of the SUS score identified that the intervention was rated with an average score of 66.6. According to Bangor et al [38], an average SUS score below 70 indicates that the system has shortcomings that need to be addressed. Of the 5 participants, one rated the LIFE4YOUth app as poor, 2 as average, and 2 assessed it as good.

Table 2. Result of the System Usability Scale.

Questions	P1[a]	P2	P3	P4	P5	Average
1. I think that I would like to use this app	4	4	3	2	3	3.2
2. I found the app unnecessarily complex	1	0	2	1	1	1
3. I found the app easy to use	1	3	3	2	2	2.2
4. I think I would need support from a technical person to use this app	4	4	3	3	4	3.6
5. I found the various functions in this app were well integrated	2	3	4	2	2	2.6
6. I thought there was too much inconsistency in this app	1	2	3	2	2	2
7. I would imagine that most people would learn to use this app very quickly	0	3	4	3	4	2.8
8. I found the app very cumbersome to use	1	3	3	2	4	2.6
9. I felt very confident using the app	1	2	3	1	4	2.2
10. I needed to learn a lot of things before I could start using this app	4	4	3	4	4	3.8
SUS[b] score (sum×2.5) maximum 100	47.5	70	77.5	55	75	66.6

[a]P1-5: person 1-5.

[b]SUS: System Usability Scale.

Discussion

Principal Findings

As described in our previous protocol study for a participatory design [34], the prompt expansion of device capability presents many challenges for developers of mHealth interventions, especially when designing interventions that aim to affect multiple individual lifestyle behaviors [25]. This study investigated the usability of the LIFE4YOUth intervention targeting health behaviors among high school students. Previous studies have suggested using a combination of different usability methods to provide insights to developers about potential usability problems [41]. This paper reports data from the first stage of a formative research process [42], which included heuristic evaluation and usability testing.

The heuristic evaluation revealed that the major usability problems and the catastrophic ratings concerned shortcomings such as information display and comprehensive information, meaning that the intervention needed to speak the users' language with consistent information that appears in a natural and logical order for the users. The results from the usability testing showed that design (aesthetics and clarity), content (quality and quantity), and functionality (effortless and logic flow) enabled usability. The findings of this study are consistent with those of previous research, which found that participants wanted features that reduced the amount of time and effort required from them [43]. The amount of data participants are expected to enter into self-monitoring apps should be carefully considered in future intervention development [44] because the results showed that frustration with a large quantity of data is one of the most common complaints of users and results in apps being deleted entirely [45].

Engagement was an issue closely related to usability. For instance, participants explained that when the intervention did not appear to have a logical flow, they would quickly cease to use it. Previous research has stressed that engagement refers both to how a user interacts with technology and their emotional response to it [46]. In a study investigating usability barriers and enablers for interventions targeting harmful drinking in young adults, participants stated that positive experiences of usability made them engage more with the intervention and made them more likely to keep using it [43]. In a think-aloud study among adult smokers and drinkers, users revealed their choice of smoking cessation or alcohol reduction apps to be influenced by the apps immediate look and feel, social proof, and titles judged to be realistic and relevant [47,48]. Individuals seem more motivated to engage with and process information more thoroughly if the message is personally relevant and meaningful [49]. Theoretical models of user engagement propose that an individual's characteristics and personal circumstances may influence their user experiences of digital interventions [50]. Engagement is an ongoing issue for mHealth intervention development. Low login rates and limited use of interventions are issues consistently reported in the literature, and higher engagement through logins and repeated use is associated with better effects of the intervention [51].

According to previous research, the most important factors during the design process are flexibility and responsiveness to the input and feedback of the target audience [52]. Optimizing usability for mHealth interventions is a critical step in the development process. Young adults, the age group with the highest use of mHealth interventions, have high expectations of digital products. If an intervention does not have optimal usability, engagement will be compromised [46]. We agree with previous researchers who conclude that it is not good enough to have an evidence-based intervention per se, but rather, aspects such as delivery, design, aesthetics, usability, and functionality need to be carefully considered [43]. As there is a high demand for interventions that target skilled technology users, such as high school students, the LIFE4YOUth mHealth intervention needs to be iteratively refined and improved and will, thereafter, be pilot tested. An RCT will be conducted to determine the efficacy of the intervention at a later stage. If effective in the subsequent RCT, the program has the potential to be implemented nationally through school health services. According to the World Health Organization, education can play an important role in health promotion of the youth [53]. School multidisciplinary teams provide good accessibility for high school students, and therefore, a natural setting for attempting to endorse health behaviors for as many high school students as possible [15].

Strengths and Limitations

A limitation of the study is the small and partly nonrepresentative sample that highlights the need for caution when interpreting the results. The results cannot be used for far-reaching conclusions.

Combining heuristic evaluation and usability testing is a strength of this study. Heuristic evaluation provided insights to developers about potential usability problems, particularly in terms of identifying problems with user interface usability. The results from the heuristic evaluation were also used as inspiration to create tasks applied in the usability tests. Usability tests provided knowledge regarding whether specific tasks could be performed in the sequences of actions they were designed for to give direct input into how real users used the system.

The heuristic evaluations were performed wherever the participants preferred, for example, in the participant's home. This was done to facilitate participation and to optimize that the participants felt no time pressure. Hence, the validity of the data could not be controlled for. This study was not conducted to identify every usability problem with the mHealth intervention but instead to show how heuristic evaluation and usability testing with a small number of users could identify a large proportion of usability problems and assist in making significant improvements to an mHealth intervention targeting multiple health behaviors. The methods used were valuable in identifying not only major areas and themes that needed modification but also smaller, easily fixed problems that users encountered.

Conclusions

Through participatory design using heuristic evaluation and usability testing, this study resulted in in-depth knowledge

regarding the aspects of intervention content and structure that end users (eg, high school students) considered important. This knowledge can be used when designing an mHealth intervention targeting multiple health behaviors. In summary, heuristic evaluation showed that the major usability problems and the catastrophic ratings concerned information display and language use. Usability testing showed that design (aesthetics and clarity), content (quality and quantity), and functionality (effortless and logic flow) enabled usability. This knowledge is valuable in guiding further development of a final version of the novel multiple mHealth intervention program LIFE4YOUth.

Acknowledgments

Conflicts of Interest

MB owns a private company that develops and distributes lifestyle interventions for use in health care settings.

Multimedia Appendix 1

Protocol used for participants to report usability problems tied to the place of occurrence, usability problem, problem description, heuristics violated, and severity rating scoring.
[DOCX File , 13 KB-Multimedia Appendix 1]

References

1. Gore FM, Bloem PJ, Patton G, Ferguson J, Joseph V, Coffey C, et al. Global burden of disease in young people aged 10-24 years: a systematic analysis. Lancet 2011 Jun 18;377(9783):2093-2102. [doi: 10.1016/S0140-6736(11)60512-6] [Medline: 21652063]

2. Mokdad AH, Forouzanfar MH, Daoud F, Mokdad AA, El Bcheraoui C, Moradi-Lakeh M, et al. Global burden of diseases, injuries, and risk factors for young people's health during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2016 Jun 11;387(10036):2383-2401. [doi: 10.1016/S0140-6736(16)00648-6] [Medline: 27174305]

3. Meader N, King K, Wright K, Graham HM, Petticrew M, Power C, et al. Multiple risk behavior interventions: meta-analyses of RCTs. Am J Prev Med 2017 Jul;53(1):e19-e30 [FREE Full text] [doi: 10.1016/j.amepre.2017.01.032] [Medline: 28258777]

4. Schuit A, van Loon AM, Tijhuis M, Ocké M. Clustering of lifestyle risk factors in a general adult population. Prev Med 2002 Sep;35(3):219-224. [doi: 10.1006/pmed.2002.1064] [Medline: 12202063]

5. de Vries H, Kremers S, Smeets T, Reubsaet A. Clustering of diet, physical activity and smoking and a general willingness to change. Psychol Health 2008;23(3):265-278. [doi: 10.1080/14768320701349107] [Medline: 25160478]

6. Simmonds M, Llewellyn A, Owen CG, Woolacott N. Predicting adult obesity from childhood obesity: a systematic review and meta-analysis. Obes Rev 2016 Feb;17(2):95-107. [doi: 10.1111/obr.12334] [Medline: 26696565]

7. Hall WD, Patton G, Stockings E, Weier M, Lynskey M, Morley KI, et al. Why young people's substance use matters for global health. Lancet Psychiatry 2016 Mar;3(3):265-279. [doi: 10.1016/S2215-0366(16)00013-4] [Medline: 26905482]

8. Stockings E, Hall WD, Lynskey M, Morley KI, Reavley N, Strang J, et al. Prevention, early intervention, harm reduction, and treatment of substance use in young people. Lancet Psychiatry 2016 Mar;3(3):280-296. [doi: 10.1016/S2215-0366(16)00002-X] [Medline: 26905481]

9. Patton GC, Coffey C, Sawyer SM, Viner RM, Haller DM, Bose K, et al. Global patterns of mortality in young people: a systematic analysis of population health data. Lancet 2009 Sep 12;374(9693):881-892 [FREE Full text] [doi: 10.1016/S0140-6736(09)60741-8] [Medline: 19748397]

10. Catalano RF, Fagan AA, Gavin LE, Greenberg MT, Irwin CE, Ross DA, et al. Worldwide application of prevention science in adolescent health. Lancet 2012 Apr 28;379(9826):1653-1664 [FREE Full text] [doi: 10.1016/S0140-6736(12)60238-4] [Medline: 22538180]

11. Global Action Plan for the Prevention and Control of Noncommunicable Diseases-2020. World Health Organization. 2013. URL: https://apps.who.int/iris/bitstream/handle/10665/94384/9789241506236_eng.pdf?sequence=1 [accessed 2020-09-09]

12. Prochaska JJ, Spring B, Nigg CR. Multiple health behavior change research: an introduction and overview. Prev Med 2008 Mar;46(3):181-188 [FREE Full text] [doi: 10.1016/j.ypmed.2008.02.001] [Medline: 18319098]

13. Hale DR, Fitzgerald-Yau N, Viner RM. A systematic review of effective interventions for reducing multiple health risk behaviors in adolescence. Am J Public Health 2014 May;104(5):e19-e41. [doi: 10.2105/AJPH.2014.301874] [Medline: 24625172]

14. Mason M, Ola B, Zaharakis N, Zhang J. Text messaging interventions for adolescent and young adult substance use: a meta-analysis. Prev Sci 2015 Feb;16(2):181-188. [doi: 10.1007/s11121-014-0498-7] [Medline: 24930386]

15. Kraag G, Zeegers MP, Kok G, Hosman C, Abu-Saad HH. School programs targeting stress management in children and adolescents: a meta-analysis. J School Psychol 2006 Dec;44(6):449-472. [doi: 10.1016/j.jsp.2006.07.001]

XSL·FO
RenderX

16. Suffoletto B. Text message behavioral interventions: from here to where? Curr Opin Psychol 2016 Feb 1;9:16-21 [FREE Full text] [doi: 10.1016/j.copsyc.2015.09.012] [Medline: 26665157]

17. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]

18. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. J Med Internet Res 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]

19. Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-centered ehealth interventions for children, adolescents, and adults with sickle cell disease: systematic review. J Med Internet Res 2018 Jul 19;20(7):e10940 [FREE Full text] [doi: 10.2196/10940] [Medline: 30026178]

20. Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, et al. Apps and adolescents: a systematic review of adolescents' use of mobile phone and tablet apps that support personal management of their chronic or long-term physical conditions. J Med Internet Res 2015 Dec 23;17(12):e287 [FREE Full text] [doi: 10.2196/jmir.5043] [Medline: 26701961]

21. Ramsey WA, Heidelberg RE, Gilbert AM, Heneghan MB, Badawy SM, Alberts NM. eHealth and mHealth interventions in pediatric cancer: a systematic review of interventions across the cancer continuum. Psychooncology 2020 Jan;29(1):17-37. [doi: 10.1002/pon.5280] [Medline: 31692183]

22. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]

23. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text messaging and mobile phone apps as interventions to improve adherence in adolescents with chronic health conditions: a systematic review. JMIR Mhealth Uhealth 2017 May 15;5(5):e66 [FREE Full text] [doi: 10.2196/mhealth.7798] [Medline: 28506955]

24. Ricci-Cabello I, Bobrow K, Islam SM, Chow CK, Maddison R, Whittaker R, et al. Examining development processes for text messaging interventions to prevent cardiovascular disease: systematic literature review. JMIR Mhealth Uhealth 2019 Mar 29;7(3):e12191 [FREE Full text] [doi: 10.2196/12191] [Medline: 30924790]

25. Maar MA, Yeates K, Toth Z, Barron M, Boesch L, Hua-Stewart D, et al. Unpacking the black box: a formative research approach to the development of theory-driven, evidence-based, and culturally safe text messages in mobile health interventions. JMIR Mhealth Uhealth 2016 Jan 22;4(1):e10 [FREE Full text] [doi: 10.2196/mhealth.4994] [Medline: 26800712]

26. Heuristic evaluation. In: Usability Inspection Methods. New York, USA: John Wiley & Sons; 1994.

27. Heuristic Evaluations and Expert Reviews. Usability. 2019. URL: https://www.usability.gov/howto-and-tools/methods/heuristic-evaluation.html [accessed 2020-09-09]

28. Georgsson M, Staggers N, Weir C. A modified user-oriented heuristic evaluation of a mobile health system for diabetes self-management support. Comput Inform Nurs 2016 Feb;34(2):77-84 [FREE Full text] [doi: 10.1097/CIN.0000000000000209] [Medline: 26657618]

29. Baumel A, Muench F. Heuristic evaluation of ehealth interventions: establishing standards that relate to the therapeutic process perspective. JMIR Ment Health 2016 Jan 13;3(1):e5 [FREE Full text] [doi: 10.2196/mental.4563] [Medline: 26764209]

30. Usability Engineering. In: Usability Engineering. San Francisco, CA: Academic Press; 1993.

31. Hinchliffe A, Mummery WK. Applying usability testing techniques to improve a health promotion website. Health Promot J Austr 2008 Apr;19(1):29-35. [doi: 10.1071/he08029] [Medline: 18481929]

32. Running a Usability Test. Usability. 2019. URL: https://www.usability.gov/how-to-andtools/methods/running-usability-tests.html [accessed 2020-09-09]

33. Ehrler F, Weinhold T, Joe J, Lovis C, Blondon K. A mobile app (bedside mobility) to support nurses' tasks at the patient's bedside: usability study. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e57 [FREE Full text] [doi: 10.2196/mhealth.9079] [Medline: 29563074]

34. Müssener U, Löf M, Bendtsen P, Bendtsen M. Using mobile devices to deliver lifestyle interventions targeting at-risk high school students: protocol for a participatory design study. JMIR Res Protoc 2020 Jan 6;9(1):e14588 [FREE Full text] [doi: 10.2196/14588] [Medline: 31904576]

35. Walker M, Takayama L, Landay JA. High-Fidelity or Low-Fidelity, Paper or Computer? Choosing Attributes when Testing Web Prototypes. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2016 Presented at: HFES'16; September 19-23, 2016; Hilton, Washington. [doi: 10.1177/154193120204600513]

36. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]

37. Charters E. The use of think-aloud methods in qualitative research an introduction to think-aloud methods. Brock Educ J 2003 Jul 1;12(2):-. [doi: 10.26522/brocked.v12i2.38]

38. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]

39. Video in Qualitative Research. New York, USA: Sage Publishing; 2010.

40. John B. SUS: A quick and dirty usability scale. In: Jordan P, editor. Usability Evaluation In Industry. London, UK: Taylor & Francis Group; 2016.

41. Thyvalikakath TP, Monaco V, Thambuganipalle H, Schleyer T. Comparative study of heuristic evaluation and usability testing methods. Stud Health Technol Inform 2009;143:322-327 [FREE Full text] [Medline: 19380955]

42. Brown KG, Gerhardt MW. Formative evaluation: an integrative practice model and case study. Personnel Psychol 2002 Dec;55(4):951-983. [doi: 10.1111/j.1744-6570.2002.tb00137.x]

43. Milward J, Deluca P, Drummond C, Watson R, Dunne J, Kimergård A. Usability testing of the branch smartphone app designed to reduce harmful drinking in young adults. JMIR Mhealth Uhealth 2017 Aug 8;5(8):e109 [FREE Full text] [doi: 10.2196/mhealth.7836] [Medline: 28790022]

44. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]

45. Brauer C. AppDynamics. 2014. URL: https://www.appdynamics.com/media/uploaded-files/1425406960/app-attention-span-research-report-1.pdf [accessed 2019-11-18]

46. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to internet-delivered health behavior change interventions aimed at adolescents or young adults: a systematic review. Health Educ Behav 2011 Feb;38(1):49-62. [doi: 10.1177/1090198110372878] [Medline: 21189422]

47. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]

48. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]

49. Alkhaldi G, Hamilton FL, Lau R, Webster R, Michie S, Murray E. The effectiveness of prompts to promote engagement with digital interventions: a systematic review. J Med Internet Res 2016 Jan 8;18(1):e6 [FREE Full text] [doi: 10.2196/jmir.4790] [Medline: 26747176]

50. Short CE, James EL, Rebar AL, Duncan MJ, Courneya KS, Plotnikoff RC, et al. Designing more engaging computer-tailored physical activity behaviour change interventions for breast cancer survivors: lessons from the iMove More for Life study. Support Care Cancer 2017 Nov;25(11):3569-3585. [doi: 10.1007/s00520-017-3786-5] [Medline: 28624949]

51. Hutton HE, Wilson LM, Apelberg BJ, Tang EA, Odelola O, Bass EB, et al. A systematic review of randomized controlled trials: web-based interventions for smoking cessation among adolescents, college students, and adults. Nicotine Tob Res 2011 Apr;13(4):227-238. [doi: 10.1093/ntr/ntq252] [Medline: 21350042]

52. Abroms LC, Whittaker R, Free C, van Alstyne J, Schindler-Ruwisch JM. Developing and pretesting a text messaging program for health behavior change: recommended steps. JMIR Mhealth Uhealth 2015 Dec 21;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]

53. New Horizons for Health Through Mobile Technologies. World Health Organization. 2011. URL: https://apps.who.int/iris/bitstream/handle/10665/44607/9789241564250_eng.pdf [accessed 2019-11-14]

Abbreviations

mHealth: mobile health
RCT: randomized controlled trial
SUS: System Usability Scale

XSL•FO
RenderX