



Point-of-Care Mobile Application to Guide Health Care Professionals in Conducting Substance Use Screening and Intervention: A Mixed-Methods User Experience Study

Megan A. O'Grady¹ Sandeep Kapoor^{2,3} Evan Gilmer¹ Charles J. Neighbors¹ Joseph Conigliaro^{2,3}
Nancy Kwon^{2,3} Jon Morgenstern^{2,3}

¹Health Services Research, Center on Addiction, New York, New York, United States

²Northwell Health, New Hyde Park, New York, United States

³Zucker School of Medicine at Hofstra/Northwell, Hempstead, New York, United States

Address for correspondence Megan A. O'Grady, PhD, Health Services Research, Center on Addiction, 633 Third Avenue, New York, NY 10017, United States (e-mail: mogrady@centeronaddiction.org).

ACI Open 2019;3:e1–e12.

Abstract

Background Well-documented barriers have limited the widespread, sustained adoption of screening and intervention for substance use problems in health care settings. mHealth applications may address provider-related barriers; however, there is limited research on development and user experience of such applications.

Objective This user experience study examines a provider-focused point-of-care app for substance use screening and intervention in health care settings.

Method This mixed-methods study included think-aloud tasks, task success ratings, semistructured interviews, and usability questionnaires (e.g., System Usability Scale [SUS]) to examine user experience among 12 health coaches who provide substance use services in emergency department and primary care settings.

Results The average rate of successful task completion was 94% and the mean SUS score was 76 out of 100. Qualitative data suggested the app enhanced participants' capability to complete tasks efficiently and effectively. Participants reported being satisfied with the app's features, content, screen layout, and navigation and felt it was easy to learn and could benefit patient interactions. Despite overwhelmingly positive user experience reports, there were some concerns that the app could negatively affect patient interactions due to reductions in eye contact and ability to build rapport.

Conclusion Using the "Fit between Individuals, Task, and Technology" framework to guide interpretation, overall results indicate acceptable user experience and usability for this provider-focused point-of-care mobile app for substance use screening and intervention as well as favorable potential for adoption by health care practitioners. Such mobile health technologies may help to address well-known challenges related to implementing substance use services in health care settings.

Keywords

- ▶ SBIRT
- ▶ mHealth
- ▶ screening
- ▶ substance use
- ▶ usability
- ▶ user experience

Background and Significance

Mobile devices and health applications (apps), often referred to as mobile health (mHealth) products, are popular for

accessing health information and providing a wide range of health services across medical disciplines and treatment settings.¹ mHealth apps can assist with patient management and monitoring, clinical decision support, and information

received
September 5, 2018
accepted after revision
January 14, 2019

DOI <https://doi.org/10.1055/s-0039-1684002>.
ISSN 2566-9346.

© 2019 Georg Thieme Verlag KG
Stuttgart · New York

License terms



gathering.² In particular, point-of-care mHealth apps can support providers and improve patient outcomes by allowing for more rapid, informed clinical decision making, and improved practice efficiency and knowledge.² This study describes a mixed-methods user experience study of an mHealth app for health care providers to use at point of care to help identify patient substance misuse and provide a brief counseling intervention.

Despite being a major public health issue, substance use remains under-addressed in health care settings.³ For example, a 2017 report indicates that only one in six binge drinkers in the United States are asked about alcohol use and advised to cut down by a health professional.⁴ Screening, Brief Intervention, and Referral to Treatment (SBIRT) is a model that can be used in health care settings to identify and address risky substance use.⁵ Evidence for SBIRT efficacy has been demonstrated with the strongest evidence coming from studies conducted in primary care that target reductions in alcohol use.⁶ Uptake of substance use screening and brief intervention models, like SBIRT, is recommended for health care settings by several organizations in the United States (i.e., Centers for Disease Control and Prevention; Substance Abuse and Mental Health Services Administration) and internationally (e.g., World Health Organization). Many countries have started to develop national plans and clinical guidelines to incorporate substance use screening and intervention practices in health care and community settings.^{7,8}

Unfortunately, there are well-documented barriers that have limited the widespread, sustained adoption of SBIRT in health care settings.⁹⁻¹² These factors include lack of health care provider training and comfort in addressing substance use problems, competing priorities, time constraints during health care visits, stigma related to patients with substance use problems, lack of organizational or leadership support, and limited knowledge and resources.¹¹ An answer to some of these implementation challenges has been to develop computerized, web-based, or mobile screening and brief intervention systems. Using such technologies to deliver potentially low-cost, time-saving, high-fidelity SBIRT programs may be a compelling way to increase uptake of SBIRT among health care providers.^{13,14} Reviews of these technologies suggest that they can be effective for improving substance use outcomes among patients, are feasible to use, and are acceptable to patients.^{13,15}

Better patient outcomes may be achieved by combining technology-based programs with in-person provider intervention.¹³ Several studies have shown that computer-guided, provider-delivered substance use interventions have equivalent or even better outcomes than patient self-administered computerized programs.¹⁶⁻¹⁸ However, most technology-based SBIRT programs described in the literature are patient self-administered and do not directly involve a health care provider. A few provider-focused, rather than patient-focused, mHealth apps and products for SBIRT have just very recently been developed.¹⁹⁻²¹ Like patient self-administered SBIRT programs, provider-focused point-of-care mHealth programs may help to address some of the system- and provider-related implementation barriers. For example, they may increase provider confidence and knowledge in delivering SBIRT,

improve fidelity to SBIRT procedures, and streamline SBIRT delivery. Such mHealth tools may be attractive to health care providers because streamlined SBIRT delivery could reduce workload and add time savings. SBIRT mHealth could extend the reach of the health care workforce by giving them additional information, clinical guidelines, and tools to engage patients who need substance use services.³ Further, such provider-focused mHealth programs may help to facilitate transfer of skills and knowledge learned by providers during SBIRT trainings to clinical practice, making them better able to address substance use among their patients.²¹

Despite the appeal of mHealth point-of-care tools for providers, unsystematic development and adoption can disrupt workflow resulting in increased costs, dissatisfied users, and less effective interventions.²² User experience testing is a vital standard for software systems targeting public health; however, over 95% of mHealth apps have not been adequately tested for use or uptake into clinical practice.²³ Little research on provider- or patient-focused SBIRT mHealth programs has examined user experience; there is only one published user experience study for a provider-focused point-of-care SBIRT mobile app of which we are aware. This study only very briefly described usability questionnaire results and qualitative feedback received from providers.²¹ Another study describes the acceptability of an SBIRT technology program but does not describe a robust user experience study.²⁰ In fact, most SBIRT technology-based programs have been tested in the context of randomized clinical trials examining patient outcomes.¹⁶⁻¹⁸ There has been little research conducted on SBIRT technology in clinical contexts outside of research trials, and there is a lack of studies on user experience or feasibility among providers for which the technology was built. Therefore, because providers are an important piece of the puzzle from a technology adoption standpoint, their input on user experience is essential to the production of clinically relevant provider-focused mHealth SBIRT technology.

Technology adoption models can help to provide a framework for guiding user experience studies. For example, the "Fit between Individuals, Task, and Technology" (FITT) framework suggests that adoption of technology in a health care setting is affected by the fit between the users, the technology, and the clinical tasks and processes required.²⁴ According to this model, achieving an optimal fit between these factors can improve information technology adoption. For example, the FITT framework suggests that (1) individuals must be motivated and knowledgeable about the clinical task that the technology was developed for, (2) the technology must function and perform in ways that support the clinical task, and (3) training to properly use a technology to perform the task is needed.²⁴⁻²⁷ The FITT framework is especially appealing because it can easily highlight interventions needed at the individual, task, and/or the technology level when problems are identified during a user experience study or technology implementation.²⁴ The FITT framework has been used to examine user experience and technology adoption for a variety of clinical technologies and for a range of health care topics.

Importantly, the FITT framework (see **Fig. 1**) can be easily used to examine technology adoption by applying

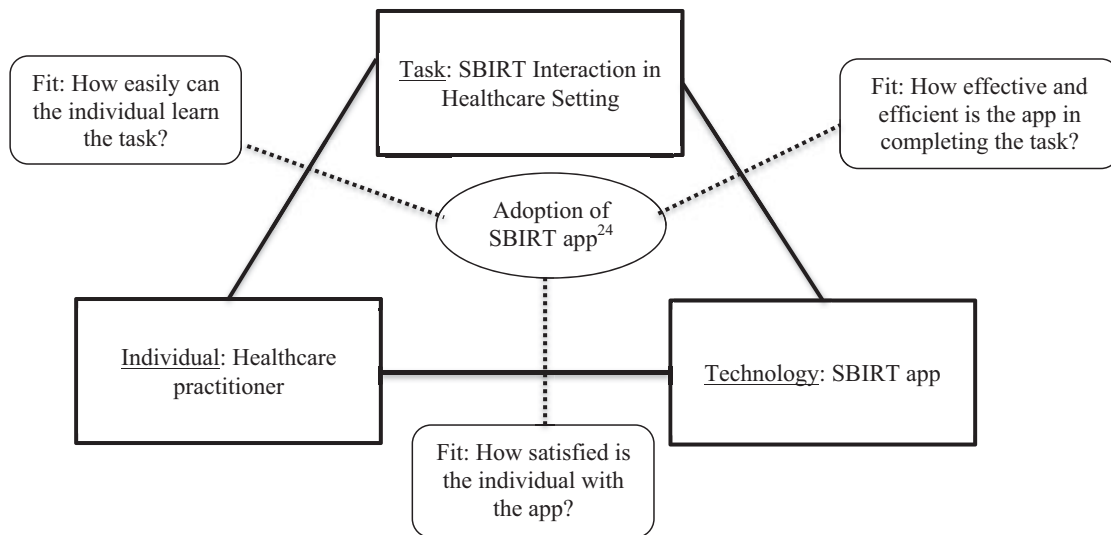


Fig. 1 FITT framework of IT adoption applied to SBIRT for health professionals app. FITT, Fit between Individuals, Task, and Technology; IT, information technology; SBIRT, Screening, Brief Intervention, and Referral to Treatment.

concepts known to affect usability and user experience, including user satisfaction, efficiency and effectiveness of the technology, as well as ability to learn to use the technology.^{27,28} For example, fit between task (i.e., SBIRT) and technology (i.e., SBIRT app) can be assessed by examining the ability of a user to complete the task accurately (e.g., app effectiveness) as well as the amount of resources the user needs to expend to complete the task (e.g., app efficiency). User satisfaction information can help to determine if there is a strong fit between the individual and the technology in a specific use context. Finally, fit between the individual and the required task can be assessed by examining the degree to which a product enables its users to learn its functions (e.g., learnability).²⁸

Objective

We developed a provider-focused point-of-care SBIRT mobile app. Using the FITT framework as a guide, we conducted a preliminary user experience study of this app. We describe the app development process and study below.

Method

App Content and Development

The SBIRT app was designed for use by providers in conjunction with patients during a health care visit. Providers are the targeted end user for the app, broadly defined to include a wide range of health care professionals that may address substance use as part of their clinical role (e.g., health coaches (HCs), physicians, nurses, social workers). It is interactive, featuring screens designed to be shown to the patient by the provider. The app is designed to assist the provider in quickly assessing the patient's level of risk due to use of substances and walk them through the steps of a brief intervention that is tailored to the patient's responses. The app contains the following sections: (1) alcohol and drug screening questions based on

validated screening tools that are administered by the provider (→Fig. 2), (2) screening results and recommendations for the provider to review (→Fig. 3), (3) feedback and educational information for the provider to share with patients that is tailored to screening results and question responses (→Fig. 4), (4) brief intervention tools (e.g., readiness ruler) for the provider to use when discussing patient motivation to change that are tailored based on the patient's response to the feedback they received (→Fig. 5), (5) goal setting tools tailored to the patient's motivation to change their substance use (→Fig. 6), and (6) a printable report for the patient that summarizes the session.

The app content and structure were developed by a team including three PhD-level clinical psychologists, one PhD-level social psychologist, and one physician. Team members had expertise in SBIRT service delivery and training, mHealth technology development, and behavioral health intervention development. This team worked closely with a designer skilled in user experience as well as a programming vendor that had

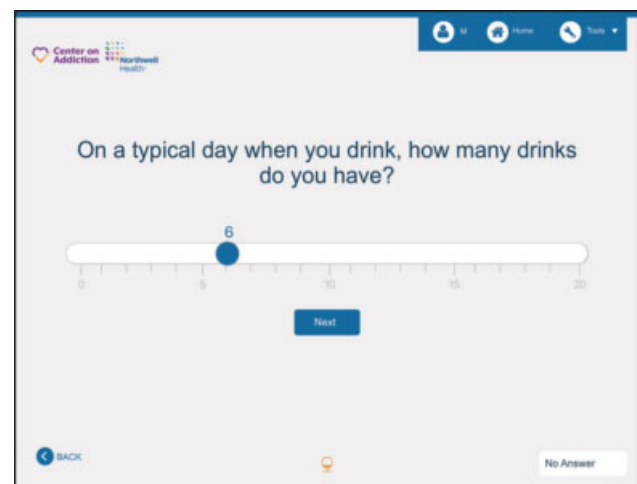


Fig. 2 Example: screening.

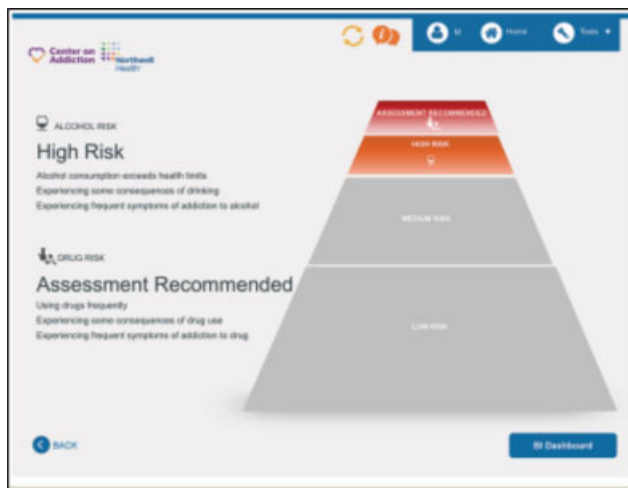


Fig. 3 Example: screening results.

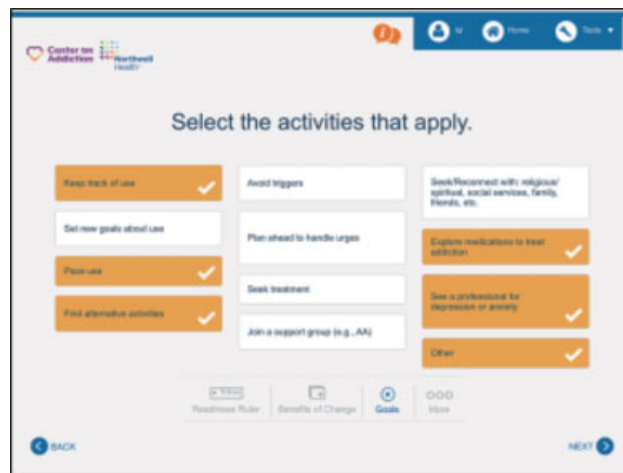


Fig. 6 Example: goal setting tool.



Fig. 4 Example: feedback and information.

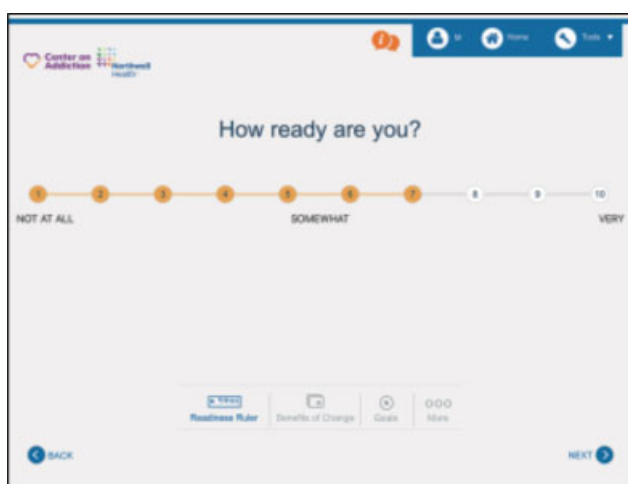


Fig. 5 Example: brief intervention tool.

experience in mHealth development. App development was iterative and done in several waves, beginning with content development, mapping algorithms, and algorithm programming. The designer created wireframes and designed screens.

Finally, the team worked with the designer and programmers to produce α versions of the app for testing. Team members tested the α versions and recommended changes. App development and testing occurred during 2015 to 2017. Extensive quality assurance testing was undertaken during this time to ensure app accuracy in calculating screening results, providing appropriate feedback screens based on screening results, and displaying proper guidance to the user.

During development, the team consulted with a three-person provider committee made up of HCs. HCs worked within a large health system to provide SBIRT to patients in primary care practices and emergency departments. The team conducted think-aloud sessions, observations, surveys, and interviews with this committee during each round of development. Adjustments and refinements were made to the app based on their feedback and a β version was finalized and used for the usability study described in this article.

Setting and Participants

This study took place in a large health system in the Mid-Atlantic region of the United States. Participants were 12 HCs who provide SBIRT services in primary care and emergency department settings. Most HCs were female (75%) and White (83%), and 8% were Hispanic or Latino. Mean age was 36 (standard deviation [SD] = 8.84). Approximately 50% had a master’s degree and the remaining had a bachelor’s degree. The average time in current role as HC was 22.5 months (SD = 19.6).

The 12 participants represented all of the HCs working in the health system at the time of the study. Despite the relatively small sample size, usability studies involving as few as 5 to 10 subjects can lead to identification of up to 80% of the surface level usability problems and be meaningful and comparable studies have used similar sample sizes.^{29–35} Both the first and second author’s institutional review boards approved the study. Participants received a meal of approximately \$20.00 value for participating.

Procedure

This convergent parallel mixed-methods study³⁶ was used to simultaneously examine qualitative and quantitative data

and consisted of activities that occurred over a several week period, including (1) a think-aloud task observation activity, (2) online usability questionnaires, and (3) semistructured interviews. We describe the full study procedure below.

Think-Aloud Task Observation Activity

Immediately following a 3-hour group app training and a 30-minute practice session, the 12 HCs were invited to participate in the study, each agreeing to participate and providing written informed consent. Participants first completed a demographics questionnaire for descriptive purposes and were given an overview of the think-aloud task observation procedures. The think-aloud method is a common approach to usability testing that enables evaluation of the ease with which a system is learned and provides insight associated with design problems.^{37,38} During this activity, participants were asked to complete 18 tasks using the SBIRT app on an Apple iPad.

Two research staff facilitated think-aloud task observation sessions with each HC. One moderated the session by asking the participant to complete 18 tasks (see further description in the Materials section) on the SBIRT app while prompting them to think aloud. The second staff used an audio device to record the HC, observed and noted HC behavior, and documented a success or fail score for each of the 18 tasks using a structured rating sheet (described below).

Online Usability Questionnaires and Semistructured Interviews

Less than 2 weeks after the initial training and think-aloud task observation session, all 12 participants were sent a link via email to access the online questionnaire designed in SurveyMonkey, of which 11 were completed (materials described below). Then, individual semistructured phone interviews were conducted until a priori thematic saturation was achieved.³⁹ Seven interviews were conducted to achieve saturation, each lasting between 9 and 20 minutes and averaging 13 minutes. We allowed time to elapse between the initial training session and the questionnaire and interview administration because we advised HCs to reflect on the app during their daily routines (e.g., how it would affect their current work) and we wanted to provide ample time for this.

Materials

Think-Aloud Task Observation Sessions

Think aloud-task observation session materials included a structured rating sheet to record HC's performance on the 18 tasks. This included tasks like entering information into the patient entry screen (e.g., age), completing alcohol screening questions, using the stop-exit-resume feature, using the readiness ruler, selecting items on the "setting goals" screen, submitting session results, and requesting a summary report. Failure to complete a task, resulting in a score of 0, occurred if the participant needed assistance from the moderator, gave up trying to solve the task, or completed the task incorrectly on the first try. Successful task completion (score = 1) was achieved if the HC completed the task

correctly on the first attempt without assistance from the moderator. Observational notes were also taken.

Online Survey

The online survey included two usability scales. The 10-item self-report version of the System Usability Scale (SUS)⁴⁰ uses a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree) and an SUS score can be calculated that ranges from 0 to 100 using a conversion formula provided by the tool authors.^{41,42} Scores under 70 indicate below average usability.⁴³ In addition to the overall score, a learnability subscale can be calculated.⁴⁴

The 19-item Computer System Usability Questionnaire (CSUQ) was also used^{42,45} and the term "app" was used to replace "system" throughout. The CSUQ uses a 7-point Likert scale (1 = strongly disagree; 7 = strongly agree). A total usability score as well as three subscales can be calculated: (1) system usefulness (e.g., satisfaction with the usefulness of the app), (2) information quality (e.g., satisfaction with the overall helpfulness and support provided by the app), and (3) interface quality (e.g., satisfaction with the quality of the interface and tools integrated within the app). While mobile-specific usability questionnaires exist (e.g., SUPR-Qm),⁴⁶ the CSUQ and SUS have been used widely in mobile app usability studies.⁴⁷ The tools better tapped into the system-based usability constructs of interest in the study rather than less applicable constructs in available mobile-specific tools.⁴⁶ Both questionnaires have shown excellent reliability and validity in previous research.⁴⁸ In the current study, appropriate internal consistency reliability was found (CSUQ $\alpha = 0.90$; SUS $\alpha = 0.73$). The CSUQ and SUS measure similar constructs⁴⁸; however, we opted to use both because each offers different usability insights. For example, the CSUQ and SUS provide different subscales of interest. In addition, the SUS, but not the CSUQ, can be calculated on a 100-point scale that importantly allows for comparison to widely accepted usability standards. Five open-ended questions were also included in the online surveys. These asked participants to detail their experience when learning to use the app, list the app's positive and negative aspects, and describe how they believed the app would affect patient interactions.

Semistructured Interviews

Semistructured interviews asked participants to reflect on their overall satisfaction with the app, app efficiency, experience in learning how to operate the app, and the app's potential impact on clinical workflow. The interviews also addressed perceived patient reaction/engagement when using the app.

Data Analytics Strategy and Framework

User experience of the app was evaluated according to the three dimensions of the FITT framework described above. As noted, we used concepts known to affect usability and user experience to operationalize each FITT dimension including effectiveness, efficiency, satisfaction, and learnability.^{27,49}

Task–Technology Fit

We used the concepts of effectiveness and efficiency to operationalize task–technology fit. In the quantitative data, we focused on effectiveness and measured this with the task–success outcomes. Task success was calculated by summing the individual scores from the 18 items and creating task–success rates (numerator = number of tasks successfully completed, denominator = total number of tasks). In the qualitative analyses, we focused on the experienced efficiency of the app.

Individual–Technology Fit

Individual–technology fit can be operationalized as satisfaction with the app. In the quantitative data, we examined the CSUQ total scale and subscale scores as well as the SUS total score. In the qualitative data, we coded for themes related to satisfaction.

Individual–Task Fit

Individual–task fit was operationalized as the app's learnability. Therefore, qualitative data were coded for learnability themes and the SUS learnability subscale was examined.

Qualitative Analysis

There were four sources of qualitative data: (1) audio recordings from the think-aloud task observation sessions, (2) audio recordings from the semistructured interviews, (3) open-ended responses from the online survey, and (4) observational notes from the think-aloud task observation sessions. Recordings from the think-aloud sessions and semistructured interviews were analyzed using rapid identification of themes from audio recordings (RITA) procedures.⁵⁰ RITA procedures allow for quick and efficient identification of themes in audio recorded qualitative data without the need for transcription. RITA involves coding for preidentified themes onto a structured coding form within prespecified time segments while listening to an audio recording. RITA procedures allow for refinement to the initial coding forms as new themes emerge or updated theme definitions are needed. In this study the coded time segments were 3 minutes.⁵⁰ Chunking audio recordings into segments facilitates establishment of reliability because it ensures that the same recording unit is used across all coders.⁵⁰

Analysis for text data from the think-aloud observational notes and open-ended survey questions used a directed content analytic approach in which codes are defined before as well as during data analysis and a systematic process of coding and identifying themes is completed. In directed content analysis, coding begins with predetermined codes, and any data that cannot be coded lead to additional codes being created.⁵¹ This analysis was completed using Atlas.ti software.⁵²

Analysis followed conventional content analysis procedures as outlined in the work of Erlingsson and Brysiewicz.⁵³ A codebook including an initial set of codes and definitions, guided by the FITT framework, was developed by three research team members with input from the first author who has expertise in qualitative research. The three team members used open coding to independently code all text

and audio data after identifying meaning units in the data. They also identified representative quotes for each code. The team members met several times to compare codes and refine the codebook after this initial coding. Additional codes not initially included were developed, defined, and added to the codebook. Next, two coders (one new coder and one of the original coders) independently reviewed and re-coded the data with the revised codebook. They met several times to review the coding together to look for agreement, missing codes, or inconsistencies. Inconsistencies in coding were reviewed and a final decision about the coding was reached by consensus between the two coders with input from the first author. In the second round of coding, all the data were combined together and organized using Atlas.ti.⁵² The coding team and the first author reviewed the final coding structure and findings together to determine how it fits within the FITT framework. The first author then identified themes from the codes and finalized representative quotes (→Table 1). To establish trustworthiness of the qualitative analysis (i.e., credibility, dependability, transferability, and confirmability) the researchers had a well-established codebook, used memos and notes to carefully document coding and analysis decisions, sought feedback on results from participants, chunked audio data into uniform segments, recruited a representative sample, and had multiple people code each transcript.

To provide greater analysis details, the team also coded for negatively and positively framed quotes within each theme (see →Table 1). However, it should be noted that positively framed codes greatly outnumbered those negatively framed (71% positive; 29% negative). There were four initial themes (efficiency, effectiveness, learnability, and satisfaction). Additional codes were developed as the analysis proceeded and at the conclusion of coding and analysis there were 13 themes: 5 for the task and technology fit dimension, 4 for the individual–technology fit dimension, and 4 for the individual–task fit dimension (see →Table 1).

Results

Task–Technology Fit

In this dimension, we were interested in effectiveness and efficiency. The average rate of successful task completion was 94%, indicative of a high level of fit between task and technology. One specific feature, the save and exit function, was mostly responsible for error when there was one, such that participants had difficulty remembering which screen area to use to access it. Qualitative data supported the task success results and suggested that the technology enhanced participants' capability to complete tasks efficiently and effectively (→Table 1). For instance, participants noted that the app provided automation of tasks they previously did by hand, reducing time spent on the task. They indicated that the app had everything they needed to do their jobs.

However, there were some concerns that integrating the technology would negatively affect patient interaction in that it would take away from building patient rapport or be less efficient overall. HCs explained that the mHealth version

Table 1 Thematic analysis coding and FITT dimension linkage

FITT dimension and theme	Theme definition	Positive [+] code quotes	Negative [-] code quotes
Task–technology			
Dialogue facilitator	Content and structure of the app incite dialogue between user and patient regarding substance use	“Using the risk pyramid is helpful for showing the patient what level they are at—interested in how its top down expansion may lead to further conversation.”	“[I] think it [iPad application] will hinder rapport building and engagement.”
Perceived patient engagement	Perceived patient response to app during the provision of SBIRT	“It will enhance my patient encounters”	‘Patient might see [screen not meant for viewing] out of the corner of their eye and become ‘turned off.’
Effectiveness	Ability of a user to complete tasks accurately and completely using the app	“This [the application] has everything we need to do our job.”	“I would not want the app to detract from that interaction because I am looking at the screen too long/too often instead of focusing on the patient.”
Efficiency	Participant is able to complete the task with minimal time and effort	“I love this feature [automatic screening calculation] because it automatically calculates the score and eliminates the paperwork making it easier to work between two different clinical sites with different processes.”	“I don’t think it will make me inefficient, but it is difficult to see whether or not this is going to be an asset or a barrier. It will probably take me longer to complete my interview with patients, personally.”
Service enhancement	Enhances delivery and/or receipt of SBIRT services	“Adds extra layer [the physical risk screen] to service relevant to [patient’s] health care visit.”	“[App] may take away from the authenticity of the service.”
Individual–technology			
Helpfulness	Perceived helpfulness of a specific feature and/or function within the app when completing a task	“Having [substance use related] information right in front of me is helpful.”	“Would rather discuss patient results in conversation moving from a macro to micro level rather than using the iPad.”
Satisfaction	Satisfactory user experience during SBIRT service provision	“The app takes into account the weekly and daily drinking limits indicating the correct level of risk that is more difficult to define using the rethinking drinking guide, so this is terrific.”	“I think it takes away from what we are doing.”
User-friendly	Level to which the app is easily operated and understood	“App was easy to navigate, and clear.”	“Having a little trouble getting the slider just right.”
User interface	Experience regarding layout, structure, and aesthetic of tools/screens, dialogue structures, icons, navigation, graphics, and multimedia	“Screens functioned and looked as expected.”	“I don’t like having to use the sliding function to answer the questions; it’s easier to just click.”
Individual–task			
Comfort	Perceived level of comfort/confidence using the app to provide SBIRT services	“It [the different app screens] was comfortable because the commands are not really complicated and the way it [the application] is set up is very convenient. As long as the content is very visible and clear I won’t have any problems presenting the screens to patients.”	“Initially I had some misunderstanding about this [question] because the original AUDIT form is different, so it [the additional question] threw me off because I’m used to only asking one question. So, this question was not comfortable for me in this [screening] process.”

(Continued)

Table 1 (Continued)

FITT dimension and theme	Theme definition	Positive [+] code quotes	Negative [-] code quotes
End user support	App provides users with help, advice, and guidance in using features and delivering SBIRT-related content	"[The app] helps you move through the steps of the BI (brief intervention) and is reminder of the tools/skills that you can use."	"I like how the standard drink tool is up there but it would be even better if we had the conversions for specific bottle sizes, like pints of rum, so if I'm using this already I don't want to use something else to calculate how many ounces are in a pint for example."
Learnability	How quickly and/or easily the user described learning to use the app	"I believe this is a great [tool] that can be helpful for seasoned professionals as well as professionals who are learning MI (motivational interviewing) and the BNI (brief negotiated interview) for the first time."	"May take some time so learn to integrate [into current routine]."
Adaptability/flexibility	Degree to which the app can accommodate the specific user style and/or technique when interacting with patients	"I appreciate the flexibility in the risk stratification for people [patients]."	"I am concerned about using this [the app] with patients and it taking away from the clinical style that I use [with the paper-based version]."

Abbreviations: FITT, Fit between Individuals, Task, and Technology; SBIRT, Screening, Brief Intervention, and Referral to Treatment.

of the intervention might reduce their ability to form a therapeutic alliance because attention may be directed away from the patient to the screen. For instance, the most commonly noted anticipated disruption between the patient and the HC concerned the ability to maintain eye contact while using the app.

In addition to initial themes related to efficiency and effectiveness, three more themes emerged in the qualitative data: (1) service enhancement, (2) dialogue facilitator, and (3) anticipated patient engagement (see ►Table 1). These themes showed that in addition to HCs seeing the app as effective, they also saw that it could enhance the services they are already providing and facilitate dialogue in a way that they perhaps were not able to do before. However, as noted above, there seemed to be apprehension on the part of a few HCs about ways the app could hinder conversation with patients.

Individual–Technology Fit

In this dimension, we focused on satisfaction with the app. The mean CSUQ total score of 5.5 (SD = 0.56) on a seven-point scale implies a high level of satisfaction with usability. Satisfaction was stable across the three subscales, with means ranging from 5.4 to 5.6, suggesting high levels of fit between the individual user and the technology (see ►Table 2). The mean total SUS score was 75.9, exceeding the industry standard of 70 that is indicative of acceptable usability (see ►Table 3).⁴³

In addition to satisfaction, three additional themes emerged (helpfulness, user friendliness, and user interface; ►Table 1). Within these themes, participants reported their impressions about the features, content, screen layout, and navigation of the app. The majority of the HCs noted how the technology could benefit their interactions with patients. However, some HCs did reveal naviga-

tion limitations on a few screens that used a slider instead of a clicking function; HCs preferred to click versus slide.

Individual–Task Fit

Individual–task fit was operationalized by examining the learnability of the app. On the SUS, 90% of participants disagreed or strongly disagreed that they needed to learn many things before they could use the application. The average learnability subscale score was 84.10, demonstrating acceptable levels of learnability⁴² and satisfactory individual–task fit.

In the qualitative analysis (►Table 1), none of the HCs indicated having trouble learning how to use the app. One HC indicated that although while they did not experience any difficulty in learning to use the app, additional practice would be helpful to learn the flow and content of the screens. Another HC noted that this app would be useful for experienced professionals already doing SBIRT as well as people who are newly learning the technique. In addition to learnability, three additional themes emerged: flexibility, comfort, and end-user support. Mixed results about the app's flexibility were found. Although some felt the app was flexible, others perceived it to be inflexible and/or rigid compared with how the SBIRT interaction would take place when not using the app. Participants felt comfortable using the app, stating that it was uncomplicated to use and provided ample user support. Specifically, participants felt the app easily walked them through the SBIRT interaction, and they reported making use of many of the available functions.

Discussion

Results of this mixed-methods study suggest that user experience and usability of a provider-focused point-of-

Table 2 CSUQ averages and percent agreement

CSUQ individual item scores	Mean	SD	Percent agree/ strongly agree
1. Overall, I am satisfied with how easy it is to use this app.	5.91	(0.70)	72.7%
2. It was simple to use this app.	6.09	(0.70)	81.8%
3. I will be able to effectively complete my work using this app.	4.64	(1.4)	36.4%
4. I will be able to complete my work using this app.	5.27	(0.9)	45.5%
5. I will be able to efficiently complete my work using this app.	4.55	(1.7)	36.4%
6. I feel comfortable using this app.	5.55	(1.3)	72.7%
7. It was easy to learn to use this app.	6.18	(0.6)	90.9%
8. I believe I will become productive quickly using this app.	4.64	(1.4)	36.4%
9. This app gives error messages that clearly tell me how to fix the problem.	4.64	(0.9)	27.3%
10. Whenever I make a mistake using the app, I recover easily and quickly.	5.00	(1.0)	45.5%
11. The information provided with this app (e.g., manual, on-screen messages) is clear.	6.09	(0.54)	90.9%
12. It is easy to find the information I needed.	5.73	(0.65)	63.6%
13. The information provided for this system (e.g., manual, on-screen messages) is easy to understand.	5.82	(0.87)	72.7%
14. The information (e.g., manual, on-screen messages) is effective in helping me complete the tasks and scenarios.	5.36	(0.92)	45.5%
15. The organization of information on the app screens is clear.	6.00	(0.45)	90.9%
16. The interface of this app is pleasant.	5.73	(0.79)	72.7%
17. I like using the interface of this app.	5.27	(1.0)	81.8%
18. This app has all the functions and capabilities I expect it to have.	5.82	(0.75)	63.6%
19. Overall, I am satisfied with this app.	5.82	(0.60)	72.7%
CSUQ subscale scores	Mean	SD	
Overall satisfaction (total scale score)	5.47	(0.56)	
System usefulness	5.37	(0.53)	
Interface quality	5.60	(0.66)	
Information quality	5.50	(0.45)	

Abbreviations: CSUQ, Computer System Usability Questionnaire; SD, standard deviation.

care SBIRT mobile app were acceptable and there is good indication that it could be adopted by health care practitioners. Using the FITT framework as a guide, results suggest that perhaps the strongest dimensions of fit were between individual and task as well as individual and technology.

It appears that there is a sound fit between individual and task given the high learnability SUS subscale score and positive qualitative feedback. Participants anticipated needing a small amount of practice before feeling fully confident in using the app. However, participants did complete a training just prior to study participation; therefore, it is unclear how providers without a several-hour training would be able to use the app. Given that several hours of in-person training may not be feasible for health care staff, future research should examine how much training is ideal for launching provider-focused technology like this SBIRT app and whether end users could train themselves using an app-based training companion that is now available with this app. In a recent acceptability study of a computer-guided

SBIRT program, providers suggested having a 30-minute training to learn to use the computer interface.²⁰ Therefore, we can conclude at this time that with just a few hours of training and practice, the app was generally easy to learn and use and that shorter training may be necessary.

Sufficient individual and technology fit was evidenced by a high SUS score that scored above accepted standards of usability, as well as high CSUQ scores. As compared with the only other published SBIRT app usability study, the app in this study scored approximately 10 points higher on the 100-point SUS scale (65.8 vs. 75.9).²¹ Therefore, while there is still room to improve, as noted by users' dissatisfaction with a specific feature, usability appears to be strong. Design and user interface factors that may have contributed to favorable usability ratings include screen and navigation simplicity and clear provision and organization of information on each screen. For example, in the screening section of the app, there are typically only one to two questions on the screen and it is very easy to navigate forward and backward. Each

Table 3 System usability subscale total, subscale, and item average scores and percent agreement

SUS items	Mean	SD	Percent agree/ strongly agree
1. I think that I would like to use this app frequently.	3.64	(0.92)	54.6%
2. I found the app unnecessarily complex.	2.09	(0.70)	0.0%
3. I thought the app was easy to use.	4.18	(0.60)	91%
4. I think that I would need the support of a technical person to be able to use this app.	1.55	(0.69)	0.0%
5. I found the various functions in this app were well integrated.	4.00	(0.0)	100%
6. I thought there was too much inconsistency in this app.	1.82	(0.75)	0.0%
7. I would imagine that most people would learn to use this app very quickly.	4.09	(0.54)	90.9%
8. I found the app very cumbersome to use.	1.45	(1.1)	27.3%
9. I felt very confident using the app.	4.09	(0.94)	81.9%
10. I needed to learn a lot of things before I could get going with this app.	1.73	(0.65)	0.0%
SUS total scale score	75.90	(11.4)	
SUS learnability subscale score (items 4 and 10)	84.10	(14.9)	

Abbreviation: SUS, System Usability Scale.

feedback screen has a singular purpose, such as to provide information about the health effects of drinking or the recommended drinking limits, rather than including all this information on the same screen. Finally, tools are easily accessible if the provider needs more guidance or information and it is easy to navigate between tools and screens. Effort spent navigating between screens was one of the factors found to reduce usability ratings in a previous SBIRT app usability study.²¹

The final dimension, task–technology fit, was somewhat less strong, though still promising. For example, there was excellent success in the task observation activity (94% success rate), with one feature (save and exit) that gave participants difficulty. Generally, participants reported that the app allowed for effective and efficient completion of SBIRT tasks. In addition, participants felt that the app could enhance the services they are already providing because of the resources it provides. However, there was a consistently emerging theme of participant concern that using technology during patient interactions might detract from the therapeutic value, turn patients off, or distract from the interaction in some way. Future research should examine how patients perceive the introduction of technology into point-of-care SBIRT interventions conducted by health care providers.

While results here suggest that the app has promise as a tool for health care practitioners, as noted in the introduction, interventions at different levels (e.g., task, individual, and technology) of the FITT model can be conducted to improve fit even more. On the individual level, to alleviate participant concerns about potential negative effects of using the technology with patients, it would be beneficial to provide more training on best practices in using technology with patients so that patient interactions remain therapeutic and consistent with the goal of the SBIRT interaction. On the technology level, addressing the features and functions that users found less easy to use (e.g., the slider function; save and exit) would increase satisfaction. As a

result of this study, we plan to revise these features as well as determine how training could address concerns about using the app with patients. There are no recommended interventions to the SBIRT task as a result of this study.

This study was preliminary and does have limitations. All HC participants in this study were already familiar with SBIRT; therefore, results are only applicable to practitioners who are familiar with SBIRT. Future research should test the app among health care providers who are novel to the SBIRT clinical practice as well as other types of health care providers. As noted in the Method section, this study only included a small sample size; therefore, we are limited in terms of being able to conduct more complex statistical analyses of the quantitative data. We did not examine patient experience and perception of the app; however, future work should address this because provider adoption and experience may be affected by the patient's reaction. In addition, future research should address providers' interest and perceived need for mHealth substance use services tools. This was outside the scope of this study, yet such research is limited and could inform future SBIRT mHealth products. Participants received training on the app prior to study activities, which could limit variability in study results. Finally, this study was conducted prior to participants fully implementing the app into their clinical workflow; therefore, their experiences and perceptions of usability could change over time. A postimplementation study should also be conducted.

Conclusion

In conclusion, this preliminary user experience and usability study suggests that a provider-focused point-of-care SBIRT mobile app has promising potential for adoption by health care practitioners. Results suggested that user satisfaction with the app was excellent, the app was easy to learn to use, and it could be effective and efficient when conducting an SBIRT interaction. Greater implementation of screening and

intervention in health care settings, along with better patient outcomes, may be achieved by adopting such technologies. Provider-focused mHealth SBIRT programs that can be used at point of care, like the one described in this study, may help in addressing some of the SBIRT system- and provider-related implementation barriers noted in the literature.

Clinical Relevance Statement

Despite being a major public health issue, substance use remains under-addressed in health care settings. User-friendly, easy to use mHealth apps may help health care providers conduct more efficient and effective substance use screening and intervention services, streamline such interactions during busy health care interactions, and increase their comfort in discussing substance use by providing support and resources. This study provides initial evidence that an app for screening and intervention provides a good user experience and may be useful for patient interactions when addressing substance use in health care settings.

Protection of Human and Animal Subjects

This study was conducted in compliance with all human subjects regulations. The study was reviewed and approved by institutional review boards of both the first and second authors.

Funding

This work was supported by the Substance Abuse and Mental Health Services Administration (SAMHSA; Grant Number 5U79TI025102). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of SAMHSA.

Conflict of Interest

None declared.

Acknowledgments

The authors wish to thank Camila Bernal, Kristen Pappacena, and Samantha Fisher for their assistance with data collection and coding for this study. We would also like to thank Laura Harrison for her assistance with facilitating app testing.

References

- Naeem F, Gire N, Xiang S, et al. Reporting and understanding the safety and adverse effect profile of mobile apps for psychosocial interventions: An update. *World J Psychiatry* 2016;6(02):187–191
- Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. *P&T* 2014;39(05):356–364
- U.S. Department of Health and Human Services (HHS); Office of the Surgeon General. Facing addiction in America: the Surgeon General's report on alcohol, drugs, and health. Washington, DC: HHS; 2016
- McKnight-Eily LR, Okoro CA, Mejia R, et al. Screening for excessive alcohol use and brief counseling of adults - 17 states and the District of Columbia, 2014. *MMWR Morb Mortal Wkly Rep* 2017; 66(12):313–319
- Sacks S, Gotham HJ, Johnson K, Padwa H, Murphy DM, Krom L. Integrating substance use disorder and health care services in an era of health reform: models, interventions, and implementation strategies. *Am J Med Res* 2016;3(01):75–124
- Álvarez-Bueno C, Rodríguez-Martín B, García-Ortiz L, Gómez-Marcos MA, Martínez-Vizcaíno V. Effectiveness of brief interventions in primary health care settings to decrease alcohol consumption by adult non-dependent drinkers: a systematic review of systematic reviews. *Prev Med* 2015;76(Suppl):S33–S38
- Glass JE, Andréasson S, Bradley KA, et al. Rethinking alcohol interventions in health care: a thematic meeting of the International Network on Brief Interventions for Alcohol & Other Drugs (INEBRIA). *Addict Sci Clin Pract* 2017;12(01):14
- SAMHSA. Systems-level implementation of Screening, Brief Intervention, and Referral to Treatment (SBIRT). Technical Assistance Publication (TAP) Series 33. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2013
- Johnson M, Jackson R, Guillaume L, Meier P, Goyder E. Barriers and facilitators to implementing screening and brief intervention for alcohol misuse: a systematic review of qualitative evidence. *J Public Health (Oxf)* 2011;33(03):412–421
- Nilsen P. Brief alcohol intervention—where to from here? Challenges remain for research and practice. *Addiction* 2010;105(06): 954–959
- Rahm AK, Boggs JM, Martin C, et al. Facilitators and barriers to implementing Screening, Brief Intervention, and Referral to Treatment (SBIRT) in primary care in integrated health care settings. *Subst Abus* 2015;36(03):281–288
- Vendetti J, Gmyrek A, Damon D, Singh M, McRee B, Del Boca F. Screening, brief intervention and referral to treatment (SBIRT): implementation barriers, facilitators and model migration. *Addiction* 2017;112(Suppl 2):23–33
- Harris SK, Knight JR. Putting the screen in screening: technology-based alcohol screening and brief interventions in medical settings. *Alcohol Res* 2014;36(01):63–79
- Marsch LA, Borodovsky JT. Technology-based interventions for preventing and treating substance use among youth. *Child Adolesc Psychiatr Clin N Am* 2016;25(04):755–768
- Bertholet N, Daepfen JB, McNeely J, Kushnir V, Cunningham JA. Smartphone application for unhealthy alcohol use: A pilot study. *Subst Abus* 2017;38(03):285–291
- Blow FC, Walton MA, Bohnert ASB, et al. A randomized controlled trial of brief interventions to reduce drug use among adults in a low-income urban emergency department: the HealthIER You study. *Addiction* 2017;112(08):1395–1405
- Bonar EE, Walton MA, Cunningham RM, et al. Computer-enhanced interventions for drug use and HIV risk in the emergency room: preliminary results on psychological precursors of behavior change. *J Subst Abuse Treat* 2014;46(01):5–14
- Cunningham RM, Chermack ST, Ehrlich PF, et al. Alcohol interventions among underage drinkers in the ED: a randomized controlled trial. *Pediatrics* 2015;136(04):e783–e793
- SBIRT [computer program]. Version 1.3. Center on Addiction and Northwell Health; 2015
- Levesque D, Umanzor C, de Aguiar E. Stage-based mobile intervention for substance use disorders in primary care: development and test of acceptability. *JMIR Med Inform* 2018;6(01):e1
- Satre DD, Ly K, Wamsley M, Curtis A, Satterfield J. A digital tool to promote alcohol and drug use Screening, Brief Intervention, and Referral to Treatment skill translation: a mobile app development and randomized controlled trial protocol. *JMIR Res Protoc* 2017;6(04):e55
- Kumar S, Nilsen WJ, Abernethy A, et al. Mobile health technology evaluation: the mHealth evidence workshop. *Am J Prev Med* 2013;45(02):228–236
- Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? *PLoS Med* 2013;10(02):e1001382
- Ammenwerth E, Iller C, Mahler C. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. *BMC Med Inform Decis Mak* 2006;6:3

- 25 Honekamp W, Ostermann H. Evaluation of a prototype health information system using the FITT framework. *Inform Prim Care* 2011;19(01):47–49
- 26 Noblin A, Shettian M, Cortelyou-Ward K, Schack Dugre J. Exploring physical therapists' perceptions of mobile application usage utilizing the FITT framework. *Inform Health Soc Care* 2017;42(02):180–193
- 27 Sheehan B, Lee Y, Rodriguez M, Tiase V, Schnall R. A comparison of usability factors of four mobile devices for accessing healthcare information by adolescents. *Appl Clin Inform* 2012;3(04):356–366
- 28 Seffah A, Kececi H, Donyaee M. QUIM: a framework for quantifying usability metrics in software quality models. Paper presented at: Second Asia-Pacific Conference on Quality Software; December 10–11, 2001; Hong Kong
- 29 Nielsen JE. Estimating the number of subjects needed for a thinking aloud test. *Int J Hum Comput Stud* 1994;41(03):385–397
- 30 Nielsen JE, Landauer TK. A mathematical model of the finding of usability problems. Paper presented at: INTERCHI'93 Conference on Human Factors in Computing Systems; April 24–29, 1993; Amsterdam
- 31 Press A, DeStio C, McCullagh L, Kapoor S, Morley J, Conigliaro J; SBIRT NY-II Team. Usability testing of a national substance use screening tool embedded in electronic health records. *JMIR Human Factors* 2016;3(02):e18
- 32 Rubin J, Chisnell D. *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. Indianapolis, IN: Wiley; 2008
- 33 Tullis T, Albert B. *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. Waltham, MA: Elsevier; 2013
- 34 Vilardaga R, Rizo J, Kientz JA, McDonell MG, Ries RK, Sobel K. User experience evaluation of a smoking cessation app in people with serious mental illness. *Nicotine Tob Res* 2016;18(05):1032–1038
- 35 Wilson V, Neilson CJ. We want it now and we want it easy: usability testing of an online health library for healthcare practitioners. *J Can Health Libr Assoc* 2014;32(02):51–59
- 36 Creswell JW, Creswell JD. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Los Angeles, CA: Sage Publications; 2017
- 37 Brinck T, Gergle D, Wood SD. *Usability for the Web: Designing Web Sites That Work*. San Francisco, CA: Morgan Kaufmann; 2001
- 38 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; 73(11–12):781–795
- 39 Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? *Qual Health Res* 2017;27(04):591–608
- 40 Brooke J. SUS: a retrospective. *J Usability Stud* 2013;8(02):29–40
- 41 Services USDoHaH. System Usability Scale. Available at: <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>. Accessed February 12, 2019
- 42 Lewis JR, Sauro J. The factor structure of the System Usability Scale. In: Kurosu M. ed. *Human Computer and Design. HCD 2009. Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer; 2009;95619:94–103
- 43 Sauro J. *A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices*. Denver, CO: Measuring Usability LLC; 2011
- 44 Borsci S, Federici S, Lauriola M. On the dimensionality of the System Usability Scale: a test of alternative measurement models. *Cogn Process* 2009;10(03):193–197
- 45 Lewis J. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *Int J Hum Comput Interact* 1995;7(01):57–78
- 46 Sauro J, Zaria P. SUPR-Qm: a questionnaire to measure the mobile app user experience. *J Usability Stud* 2017;13(01):17–37
- 47 Kortum P, Sorber M. Measuring the usability of mobile applications for phones and tablets. *Int J Hum Comput Interact* 2015;31(08):518–529
- 48 Lewis JR. Measuring perceived usability: the CSUQ, SUS, and UMUX. *Int J Hum Comput Interact* 2018;34(12):1148–1156
- 49 Bevan NJ. Extending quality in use to provide a framework for usability measurement. Paper presented at: 1st International Conference on Human Centered Design, held as Part of HCI International; 2009; San Diego, CA
- 50 Neal JW, Neal ZP, VanDyke E, Kornbluh M. Expediting the analysis of qualitative data in evaluation: a procedure for the rapid identification of themes from audio recordings (RITA). *Am J Eval* 2015;36(01): 118–132
- 51 Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005;15(09):1277–1288
- 52 Muhr T. *ATLAS.ti: The knowledge Workbench: Visual Qualitative Data, Analysis, Management, Model Building: Short User's Manual*. Berlin: Scientific Software Development; 1997
- 53 Erlingsson C, Brysiewicz P. A hands-on guide to doing content analysis. *Afr J Emerg Med* 2017;7(03):93–99