Nursing Informatics 2018
A.K. Rotegård et al. (Eds.)

doi:10.3233/978-1-61499-872-3-106

Usability Testing of a mHealth App to Support Self-Management of HIV-Associated Non-AIDS Related Symptoms

Samantha Stonbraker[a], Hwayoung Cho[a], Gabriella Hermosi[a], Adrienne Pichon[a], Rebecca Schnall[a]

[a] School of Nursing, Columbia University, New York, New York, USA

Abstract

Mobile health (mHealth) applications (apps) are uniquely poised to offer the information that persons living with HIV (PLWH) need to manage the symptoms associated with their chronic condition. The purpose of this study was to assess the usability of a mHealth app designed to help PLWH self-manage the symptoms associated with their HIV and HIV-associated non-AIDS (HANA) conditions. We conducted a heuristic evaluation with five experts in informatics and end-user testing with 20 PLWH. End-users completed the PSSUQ and Health-ITUES validated measures of system usability. Mean severity scores for the 10-item heuristic checklist completed by experts ranged from 0.4–2.4. End-users gave the system high scores on the PSSUQ and Health-ITUES usability measures (mean 2.23±0.83 and 4.24±0.62 respectively). Results indicated the system is usable and will be ready for future efficacy testing after incorporation of recommended feedback.

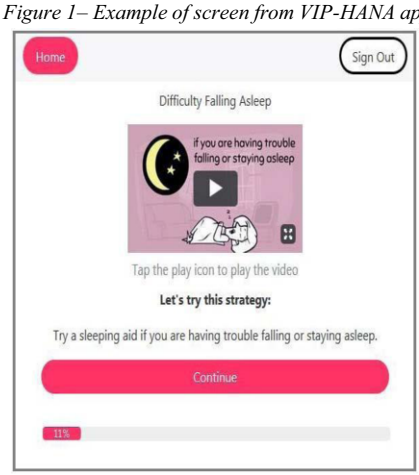
Keywords:
Mobile Health, HIV/AIDS, self-management

Introduction

Mobile health (mHealth) technologies have the potential to improve self-management activities of individuals living with chronic conditions, including those living with Human Immunodeficiency Virus (HIV) [1, 2]. In the developed world, HIV is categorized as a chronic condition, since persons living with HIV (PLWH) are able to live with the illness for decades [3]. Unfortunately, even with optimal antiretroviral therapy (ART) adherence, over time HIV infection may increase the risk of comorbid conditions such as cardiovascular disease, liver disease, diabetes, and asthma among others [4, 5], often referred to as HIV-associated non-AIDS (HANA) conditions. HANA conditions are more common among older individuals even after adjusting for age and other known risk factors [6]. Therefore, the future of self-management for PLWH, particularly those living with HANA conditions, will require patients to self-manage symptoms associated with these conditions [7]. mHealth technologies are well suited to provide relevant self-management information as mobile technology has a nearly ubiquitous presence globally [1]. Additionally, to expand access to relevant information, the content in mHealth apps can be provided in a way that those with low health literacy are able to understand [8].

The Video Information Provider (VIP) for HANA conditions (VIP-HANA) app was designed to assist PLWH self-manage their symptom experience by providing self-care strategies. Through the VIP-HANA app, PLWH are able to indicate which of nine included HANA conditions they have. Then, by completing sessions in the app, users are able to track the associated symptoms that they experience (of 28 possible symptoms), rank the severity of those symptoms, and receive self-care strategies that can help to ameliorate them. Additional functions allow users to review their symptom experience over time, set and change their avatar (character who guides participants through the app), set weekly reminders to complete sessions in the app, and email the history of their symptoms and associated strategies to themselves. The self-care strategies presented in the app were derived through a process described elsewhere [9] but briefly, we conducted a large online survey and then had a group of experts review and agree upon the most appropriate, effective, and safe self-care strategies for improving symptom outcomes in PLWH. Figure 1 provides an example of how the app would appear during a session while providing a strategy to PLWH.

Figure 1– Example of screen from VIP-HANA app



For any new system to have true value and impact among its intended audience, a critical first step is to establish its usability [10]. The goals of usability testing are to identify potential problems with using the system, improve system design, and increase the likelihood of technology acceptance among end-users. To achieve these goals, usability testing can include a combination of methods, such as including both experts and target end-users during assessments, to comprehensively identify usability concerns and generate ideas for system improvement [11; 12]. The purpose of this study was to assess the usability of the VIP-HANA app by identifying violations of usability principles as well as to establish recommendations for improvement of the system from the perspectives of both informatics experts and a sample of target end-users. Findings were used to inform further refinement of the app to be incorporated prior to efficacy testing.

Methods

Usability testing of the VIP-HANA app included two components. First, a heuristic evaluation was conducted with experts in informatics and second, end-user testing was conducted with PLWH who also had at least one HANA condition.

Heuristic Evaluation

Sample

Five informatics experts were recruited as usability evaluators in accordance with Nielsen's recommendation to include three to five evaluators in usability testing, as it is unlikely to collect additional information with additional evaluations [13]. Experts were faculty from New York-Presbyterian Hospital / Columbia University School of Nursing and doctoral students from Columbia University Department of Biomedical Informatics who have at least a master's degree in Human Computer Interaction or Informatics and had published in the field of informatics.

Procedures

Usability experts were provided with a Beta version of the VIP-HANA app and were asked to complete a session in the app using user scenarios that represented the main functions of the system and to think-aloud while doing so [14]. Participants were asked to describe what they were thinking, seeing and doing as they completed a session using the provided tasks. All participants verbally stated comments and on screen movements were recorded using Morae software™ [15], which allows the researcher to complete an in-depth analysis of the testing session.

Instruments

Following their use of the app, usability experts completed an online Heuristic Evaluation Checklist created by Bright et al. [16], which assess how a system adheres to Nielsen's 10 usability principles [13] and was administered through Qualtrics, an online survey software. Each question on this 10-item checklist is scored on a 5-point Likert scale ranging from 0 (not a usability problem) to 4 (usability catastrophe).

Data Analysis

Morae recordings and evaluators' comments regarding usability problems were reviewed and analyzed by two members of the research team to identify areas of usability concern that could be targets for improvement. Most commonly cited concerns were discussed and any discrepancies were deliberated until consensus was reached. Mean severity scores were calculated for each of the 10 usability heuristics [17].

End-user testing

Sample

Eligibility criteria for participation was: 18 years of age or older, English-speaking, living with HIV, diagnosed with at least one of 9 possible HANA conditions, and have experienced at least 3 of 28 possible associated symptoms in the past 7 days. The cognitive state of each participant was also assessed with a shortened version of the Mini Mental State Examination (MMSE) and any participant who provided all "acceptable" responses was eligible [18]. Exclusion criteria included any "unacceptable" response on the MMSE, current enrollment in another mobile app or text messaging study for PLWH, or being pregnant. Potential participants were contacted using a research database and were initially introduced to the study. Interested individuals were then screened for eligibility and a more thorough explanation of study procedures was provided. Recruitment continued until twenty individuals agreed to participate in usability testing. The rationale for including 20 participants was based on past usability research which indicated that 95% of usability problems in a system can be identified with 20 users [19].

Procedures

During usability testing, participants were first provided a brief explanation of the app and its functionalities and were then given a list of tasks to complete (Table 1). Tasks representing the main features of the app were selected to ensure that any usability concerns associated with these functions would be identified. Participants were then asked to speak aloud their thoughts while they worked through the provided tasks to enable researchers to comprehensively and reliably identify usability concerns [20]. Morae software™ [15] was used to record both participants voices and on-screen movements. A member of the research team was present during sessions to provide guidance when a participant was unable to move through a task independently and to take notes. Following their use of the app, participants completed a survey using Qualtrics survey management software.

Table 1– List of Tasks for End-users to Complete During Testing Sessions

1. Log into VIP-HANA web-application using provided username and password
2. Begin the first session, answer questions related to the symptoms you have been experiencing
3. Play a video of the recommended strategies for the symptoms you reported.
4. Review your history of the most recent symptoms/ strategies
5. Email the history to yourself
6. Log out
7. Log back in
8. Review the "How our app works" section
9. Set a weekly reminder
10. Change your avatar
11. Complete a second session
12. Log out

Instruments

The survey included demographic and technology use questions, a health literacy assessment, and two validated measures of usability. Health literacy levels were assessed using the Newest Vital Sign (NVS), in which participants are provided a nutrition label and then asked to answer a series of six questions regarding the information contained in the label. Scores range from 0 to 6 where a score of 0 or 1 indicates a "high likelihood" of, 2 or 3 indicates the "possibility" of, and 4-6 indicates likely to have "adequate" health literacy [21].The first usability measure was self-reported ease of use which was measured with the Health Information Technology Usability Evaluation Scale (Health-ITUES) [22; 23]. This 20-item tool is designed to support customization at the item level to match the specific task/expectation and health IT system while retaining standardization at the construct level and has been demonstrated to be useful for evaluating the usability of mHealth technology [24]. It is scored on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) where higher scores indicate a system that is easier to use. The second usability measure was the short version of the Post Study System Usability Questionnaire (PSSUQ) which is a 16-item survey that assesses users' perceived satisfaction with a system [25]. Scoring is based on a 7-point Likert scale ranging from 1 (strongly agree) to 7 (strongly disagree) including a neutral midpoint. Lower scores indicate more perceived satisfaction with the system.

Data Analysis

The study team reviewed transcripts of recordings of user sessions and notes taken by a study team member during the usability sessions to identify "critical incidents" [26], which were defined as events that indicate positive or negative attributes of usability. Incidents and other comments relevant to usability were then summarized to identify common usability concerns. STATA SE 14 (StataCorp 2015, College Station, TX) was used to analyze the survey data.

Results

Heuristic testing

Overall, severity scores ranged from 0.4 and 2.4 on the 10 items on the heuristic checklist, where scores closest to 0 indicate a more usable product. Mean severity scores for each of the usability heuristics are presented in Table 2. The area identified as the most in need of improvement was "user control and freedom," where evaluators identified that the app did not have sufficient ability to move forward and backwards. The next most identified area for improvement was, "help and documentation" for which evaluators identified the lack of a help function as a predominant concern. Additionally, "match between the system and real world" was an area of concern as evaluators cited too advanced of vocabulary, visibility of menu items, and comprehensibility of videos as issues that might impede system usability.

Experts recommended that the content of videos be audio-recorded to assist end-users who have low literacy. Additionally, having a Spanish language option was suggested to improve reach of the app. Beyond the heuristic checklist, evaluators provided positive feedback regarding the user experience. For instance, one participant said, "I can see the home button right there and it's simple as signing out right there, so I can always exit. It's very minimal, simple, color is very aesthetically pleasing."

Table 2– Mean Severity Scores of Heuristic Testing

Nielsen's Heuristic Checklist	Mean (SD)
1. Visibility of system status	0.8 (0.84)
2. Match between system and real world	1.2 (0.84)
3. User control and freedom	2.4 (0.89)
4. Consistency and standards	0.4 (0.89)
5. Help users recognize, diagnose, and recover from errors	0.4 (0.89)
6. Error prevention	0.4 (0.89)
7. Recognition rather than recall	0.6 (0.89)
8. Flexibility and efficiency of use	0.8 (1.10)
9. Aesthetic and minimalist design	0.6 (0.89)
10. Help and Documentation	1.4 (0.89)

End-user testing

The mean age of end-users was 54.4 years (SD = 1.71) and they were mostly female (60%) and African American (85%). Half of participants (50%) had less than a high school education while 35% had a high school education and 15% had education beyond high school. Based on NVS health literacy assessment, almost all (95%) had either high likelihood or some possibility of low health literacy. Of participants' technology use, 40% reported daily computer use, 25% reporting using a computer several times per week, 5% indicated computer use several times per month while the remaining 30% reported using a computer either once a month, less often, or never (Table 3). The most frequently used mobile device was an Android phone (70% of participants) and a vast majority (90%) reported using a cell phone several times a day.

Table 3– End-users' Technology Use

Technology	n	(%)
Frequency of Computer Use		
Several times per day	7	(35)
Once a day	1	(5)
Several times per week	5	(25)
Several times per month	1	(5)
Once a month or less often	4	(20)
Never	2	(10)
Most Frequently Used Mobile Device		
Android Phone	14	(70)
iPhone	4	(20)
Tablet or Netbook	2	(10)
Frequency of Cell Phone Use		
Several times a day	18	(90)
Once a day	2	(10)
When Started Using a Mobile Device		
In the past six months	2	(10)
In the past year	3	(15)
In the past two years	15	(75)

End-users reported high usability scores on the PSSUQ and Health-ITUES usability measures (mean 2.23 ± 0.83 and 4.24 ± 0.62 respectively) (Table 4).

Table 4– Usability Measures

Measure	Mean	SD
PSSUQ Overall	2.23	0.83
System Quality	2.17	0.75
Information Quality	2.34	0.90
Interface Quality	2.12	0.91
Health-ITUES Overall	4.24	0.62
Quality of Work Life	4.47	0.85
Perceived Usefulness	4.19	0.80
Perceived Ease of Use	4.27	0.64
User Control	4.10	0.68

End-users provided numerous positive comments regarding the usability of the VIP-HANA app and were generally able to move through sessions without difficulty. Overall, participants liked the layout of the app. For example, one participant commented, "It is a good layout, simple, it is filled in nice. Everything is compact and you can see what you are doing." The process of moving through symptom questions and recommended strategies was described as easy by most participants. As one said, "the questions are nice and clear to me, like a kid could understand it." Participants also reported a positive user experience and some further described the app as fun. For example, one participant mentioned, "it [using the app] can be fun, makes me more like, not afraid of it." Participants also mentioned liking some of the additional features of the app, such as being able to select an avatar from a variety of options, although the steps required to change the avatar within the app were confusing to most participants.

End-users had more difficulty completing advanced functions of the app, such as being able to change their avatar, enlarge videos, email themselves their symptom reports, and to set a weekly reminder. It was determined that the system could refined to make these tasks more accessible by more clearly identifying the functions of each component of the app. For instance, several participants were confused by the "home" button and the functionalities that might be available if "home" were selected. Participants commented that words such as "menu" would more accurately represent the functionality of the button. Contrary to feedback received from experts

in the heuristic evaluation, end-users did not notice or comment on the lack of a back button. Additionally, rather than identifying the need for a help feature as experts did, end-users indicated that app features need to be clearly marked. As one participant said, "Put it [what you want the user to do] there plain and simple, just say it." Participants also indicated that they might not have understood the vocabulary used and many were unable to read words such as "fatigue." Many participants had difficulty identifying the buttons or links to click on, citing size "it was kinda small," or simply did not know what icons, such as the "expand video" icon (⛶), were for. Additionally, participants were not clear on what the progress bar was or meant and recommended clearly labeling it with "% completed" or simply, progress" to make its function more obvious. Throughout sessions, participants emphasized that clear, simple language was critical and that large fonts and bright colors were useful to help them identify needed information.

Discussion

Evidence-based methods, such as mHealth apps, to help PLWH manage their symptoms are an important component of effective self-management and may lead to improved quality of life [4; 27]. In this study, we conducted usability testing of a mHealth app designed to help PLWH manage the symptoms associated with their HANA conditions with both informatics experts and target end-users to identify usability concerns and provide recommendations for further refinement of the VIP-HANA app. Although heuristic evaluators and end-users had different recommendations for specific features in need of adjustment, overall, they both gave the system high scores on validated usability scales and numerous positive comments on app features were provided. Key areas where the app could be improved, such as the identification of functionalities in need of clearer labels and determining places where intended functionality was not well understood were identified and shared with the software developer to refine the app.

Considering that target end-users were either novice technology users or individuals with low health literacy, findings from this study have numerous design implications. First, using a mHealth app is not an intuitive process for users who may not have much experience using a computer or mobile device [28]. System designers must consider this when designing layouts and other visual aspects of apps and focus on developing designs that work for first-time app users. Additionally, the vocabulary used to describe features and content of the app needs to match the literacy level of target end-users. In our study, paticipants demonstrated limited vocabularies and exhibited very literal understanding of words, for instance, in part of the app there is a section on current and older reports; one participant thought she should click on the tab that said "older" as she considered herself an older adult. Therefore, clearly defining words such as "avatar" and clearly indicating where to access certain functions would also be useful. Additionally large, easy to click on links and icons that are brightly colored were highly desired by participants in this study. These implications are similar to those identified in an Institute of Medicine report, where writing actionable content, displaying content clearly, and organizing and simplifying were among the primary recommendations [8]. These considerations are increasingly important to enable those with low health literacy, such as older adults, to access widely used mHealth technologies meant to improve health [29].

One limitation of this study was that end-users were recruited from a database that included participants who had participated in past research. This may have altered results as participants may have been more active information seekers, have more interest in their health and self-management, and have more time to seek out studies than their peers. An additional limitation was that at least one member of the research team was present during each of the usability sessions, and despite attempts to create a comfortable, private, and welcoming research space, social desirability bias may have influenced results [30]. Lastly, we conducted usability testing on a computer to enable more effective analysis with Morae recordings. Thus, participant's comfort, or lack thereof, may not be an accurate representation of their ability to use the app on their smartphone. Despite this limitation, the Beta version used in the online platform mirrored the way the app appears on a mobile device thus, we were able to observe and record while experts and end-users moved through the different features that are available in the VIP-HANA app. Through our analysis, we were also able to identify areas where usability was a potential concern and from those findings, were able to comprehensively establish ways to further refine the app to make it more usable

Conclusions

Both informatics experts and end-users rated the VIP-HANA app as a highly usable system. Additionally, end-users found it easy to work through sessions of the app and provided positive feedback about the layout as well as other system functions. Notably, experts and end-users had different feedback regarding essential components for usability. All participants provided useful recommendations to further improve the app. Results will inform researchers, designers, and software developers about considerations that may make mHealth apps more accessible and useful to end-users with low health literacy. Also, findings from this study will enable the study team and programmers to further develop and refine the app before moving into the efficacy trial.

Acknowledgements

This research was funded by the Video Information Provider for HIV-Associated Non-AIDS (VIP-HANA) Symptoms grant (R01 NR015737), PI: Rebecca Schnall, funded by the National Institute of Nursing Research, National Institutes of Health. Samantha Stonbraker was supported by the Reducing Health Disparities Through Informatics (RHeaDI) training grant (T32 NR007969), funded by the National Institute of Nursing Research, National Institutes of Health.

References

[1] G.D. Kirk, S.S. Himelhoch, R.P. Westergaard, and C.G. Beckwith, Using mobile health technology to improve HIV care for persons living with HIV and substance abuse, *AIDS Research and Treatment* **2013** (2013).

[2] R. Schnall, S. Bakken, M. Rojas, J. Travers, and A. Carballo-Dieguez, mHealth technology as a persuasive tool for treatment, care and management of persons living with HIV, *AIDS and Behavior* **19** (2015), 81-89.

[3] S. Deeks, S. Lewin, and D. Havlir, The end of AIDS: HIV infection as a chronic disease, *The Lancet* **382** (2013), 1525-1533.

[4] M. Greene, A.C. Justice, H.W. Lampiris, and V. Valcour, Management of human immunodeficiency virus infection in advanced age, *JAMA* **309** (2013), 1397-1405.

[5] W.C. McCormick, Summary report from the human immunodeficiency virus and aging consensus project, *Journal of the American Geriatrics Society* **60** (2012), 974-979.

[6] M. John, The clinical implications of HIV infection and aging, *Oral Diseases* **22** (2016), 79-86.

[7] D. Swendeman, B.L. Ingram, and M.J. Rotheram-Borus, Common elements in self-management of HIV and other chronic illnesses: an integrative framework, *AIDS Care* **21** (2009), 1321-1334.

[8] J. Broderick, T. Devine, E. Langhans, A.J. Lemerise, S. Lier, and L. Harris, Designing health literate mobile apps, in, Institute of Medicine of the National Academies, 2014.

[9] S. Iribarren, K. Siegel, S. Hirshfield, S. Olender, J. Voss, J. Krongold, H. Luft, and R. Schnall, Self-Management Strategies for Coping with Adverse Symptoms in Persons Living with HIV with HIV Associated Non-AIDS Conditions, *AIDS and Behavior* (2017), 1-11.

[10] L. Goldberg, B. Lide, S. Lowry, H.A. Massett, T. O'Connell, J. Preece, W. Quesenbery, and B. Shneiderman, Usability and accessibility in consumer health informatics: current trends and future challenges, *American Journal of Preventive Medicine* **40** (2011), S187-S197.

[11] M.W. Jaspers, A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence, *International Journal of Medical Informatics* **78** (2009), 340-353.

[12] W.-s. Tan, D. Liu, and R. Bishu, Web evaluation: Heuristic evaluation vs. user testing, *International Journal of Industrial Ergonomics* **39** (2009), 621-627.

[13] J. Nielsen and R. Molich, Heuristic evaluation of user interfaces, in: *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 1990, pp. 249-256.

[14] E. Bertini, S. Gabrielli, and S. Kimani, Appropriating and assessing heuristics for mobile computing, in: *Proceedings of the working conference on Advanced visual interfaces*, ACM, 2006, pp. 119-126.

[15] TechSmith, Morae Usability and Web Site Testing, (1995).

[16] T.J. Bright, S. Bakken, and S.B. Johnson, Heuristic evaluation of eNote: an electronic notes system, in: *AMIA Annual Symposium Proceedings*, American Medical Informatics Association, 2006, p. 864.

[17] R.L. Mack and J. Nielsen, *Usability inspection methods*, Wiley & Sons New York, NY, 1994.

[18] M. Folstein, S. Folstein, and P. McHugh, "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician, *Journal of Psychiatry Research* **12** (1975), 189-198.

[19] L. Faulkner, Beyond the five-user assumption: Benefits of increased sample sizes in usability testing, *Behavior Research Methods* **35** (2003), 379-383.

[20] R. Jääskeläinen, Think-aloud protocol, *Handbook of translation studies* **1** (2010), 371-374.

[21] B.D. Weiss, M.Z. Mays, W. Martz, K.M. Castro, D.A. DeWalt, M.P. Pignone, J. Mockbee, and F.A. Hale, Quick assessment of literacy in primary care: The Newest Vital Sign, *Annals of Family Medicine* **3** (2005), 514-522.

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

[23] R. Schnall, H. Cho, and J. Liu, Validation of the Health Information Technology Usability Evaluation Scale (Health-ITUES) for Usability Assessment of Mobile Health Technology, *JMIR mHealth uHealth* (Under Review).

[24] W. Brown, P.-Y. Yen, M. Rojas, and R. Schnall, Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology, *Journal of Biomedical Informatics* **46** (2013), 1080-1087.

[25] J.R. Lewis, Psychometric evaluation of the PSSUQ using data from five years of usability studies, *International Journal of Human-Computer Interaction* **14** (2002), 463-488.

[26] B.-E. Andersson and S.-G. Nilsson, Studies in the reliability and validity of the critical incident technique, *Journal of Applied Psychology* **48** (1964), 398.

[27] H. Cho, S. Iribarren, and R. Schnall, Technology-Mediated Interventions and Quality of Life for Persons Living with HIV/AIDS, *Applied Clinical Informatics* **8** (2017), 348-368.

[28] B. Sheehan, Y. Lee, M. Rodriguez, V. Tiase, and R. Schnall, A comparison of usability factors of four mobile devices for accessing healthcare information by adolescents, *Applied Clinical Informatics* **3** (2012), 356-366.

[29] H. Levy, A.T. Janke, and K.M. Langa, Health literacy and the digital divide among older Americans, *Journal of General Internal Medicine* **30** (2015), 284-289.

[30] A. Bowling, Mode of questionnaire administration can have serious effects on data quality, *Journal of Public Health* **27** (2005), 281-291.

Address for correspondence

Corresponding author: Samantha Stonbraker
Columbia University School of Nursing
Email: Slb2201@cumc.columbia.edu
630 West 168th Street, Mail Code 6, New York, NY 10032.
Phone: 720-880-8194