



Published in final edited form as:

AIDS Educ Prev. 2021 August ; 33(4): 312–324. doi:10.1521/aeap.2021.33.4.312.

Usability Evaluation of the mLab App for Improving Home HIV Testing Behaviors in Youth at Risk of HIV Infection

Gabriella Flynn, MEd¹, Thomas Scherr, PhD², Robert Garofalo, MD, MPH³, Lisa M. Kuhns, PhD³, Brady Bushover, MPH⁴, Nathanael Nash, MSW³, Rindcy Davis, PhD⁵, Rebecca Schnall, PhD, MPH, RN-BC, FAAN⁴

¹University of South Florida, College of Public Health, 4202 E. Fowler Ave, Tampa, FL 33620, USA

²Department of Chemistry, Vanderbilt University, 2201 West End Ave, Nashville, Tennessee 37235, USA

³Division of Adolescent Medicine, Lurie Children's Hospital, & Northwestern University, Feinberg School of Medicine, Department of Pediatrics, 1440 N. Dayton Street, Chicago, Illinois 60642 USA

⁴Division of Scholarship and Research, Columbia University School of Nursing, 560 West 168th Street, New York, New York 10032, USA

⁵H. Sergievsky Center, Columbia University Medical Center, 630 West 168th Street, New York, New York 10032, USA

Abstract

Improving access to HIV testing among youth at high-risk is essential for reaching those who are most at risk for HIV and least likely to access healthcare services. This study evaluates the usability of mLab, an app with image-processing feature that analyzes photos of OraQuick HIV self-tests and provides real-time, personalized feedback. mLab includes HIV prevention information, testing reminders and instructions. It was developed through iterative feedback with a youth advisory board (N=8). The final design underwent heuristic (N=5) and end-user testing (N=20). Participants completed think-aloud protocols with use-case scenarios. Experts rated mLab following Nielsen's heuristic checklist. End-users used the Health Information Technology Usability Evaluation Scale. While there were some usability problems, overall study participants found mLab useful and user-friendly. This study provides important insights into using a mobile app with imaging algorithm for interpreting HIV test results with the ultimate goal of improving HIV testing and prevention in populations at high-risk.

Keywords

HIV prevention; mHealth; usability evaluation; YMSM

Mobile health (mHealth) is an increasingly popular platform for delivery of health interventions, especially those targeting young adults, since smartphones are a favored communication tool among this population (Lenhart, Ling, Campbell, & Purcell, 2015; Pew Research Center, 2018; Sheehan, Lee, Rodriguez, Tiase, & Schnall, 2012; Silver, 2019; Ybarra et al., 2017). Additionally, mHealth applications can be useful for reaching marginalized groups, like young men who have sex with men (YMSM) who may be less likely to access care in person for fear of stigma or discrimination (Beach et al., 2018; Catalani, Philbrick, Fraser, Mechael, & Israelski, 2013; Devi et al., 2015; Rossman & Macapagal, 2017). This is especially important as YMSM have high incidence of HIV in the United States (US). To illustrate, in 2018, adult and adolescent MSM made up 69% of all the new HIV diagnoses in the US and its dependent areas (Dailey et al., 2020).

HIV at-home testing interventions may help reduce barriers to accessing HIV testing and related services. However, there are concerns among MSM that while completing HIV self-testing at home they could misinterpret their results or perform the test incorrectly (Legrand et al., 2017; Schnall, John, & Carballo-Dieiguez, 2016; Zhong et al., 2017). mHealth interventions serve as a promising avenue for overcoming these HIV testing barriers and increasing testing rates in YMSM (LeGrand, Muessig, Horvath, Rosengreen, & Hightow-Weidman, 2017). In the past decade, there has been an increase in mHealth interventions targeting HIV prevention (Catalani et al., 2013; Forrest et al., 2015). However, these interventions are often not developed using a rigorous design process to ensure usability (Bevan, Carter, Earthy, Geis, & Harker, 2016; Davis, Gardner, & Schnall, 2020; Nguyen et al., 2019).

Given the importance of testing in the HIV care continuum, apps promoting HIV testing, particularly among youth at high-risk of HIV, like YMSM, are an important area for developing health interventions. However, a recent review of over 100 eHealth interventions found less than half of the interventions were designed for HIV negative individuals, and less than a quarter of the interventions linked people to care (Maloney, Bratcher, Wilkerson, & Sullivan, 2020). Additionally, only 37% of interventions were designed for young adults, illustrating the need for HIV prevention apps for YMSM, our target population (Maloney et al., 2020). Furthermore, the number of usability studies evaluating e-Health HIV interventions focusing on at-home HIV testing and linkage to care among YMSM is also limited. A comprehensive review of usability evaluation methods in eHealth HIV interventions yielded only 28 studies that assessed eHealth usability, and none of the interventions specifically focused on both HIV at-home testing and linkage to HIV information and care (Davis, Gardner, & Schnall, 2020). This small number of interventions, with none centering on HIV at-home testing and linkage to HIV information and care, helps illustrate that the lack of eHealth HIV prevention studies focusing on testing and linkage to and the need for rigorous usability testing.

In this paper, we report on the user-centered design process to improve the usability of the mLab App (Jaspers, 2009; Schnall et al., 2016; Tan, Liu, & Bishu, 2009). Conducting usability studies of an iterative design process is essential for ensuring the usability of an HIV prevention app (Hayat & Ramdani, 2020; Kushniruk, Senathirajah, & Borycki, 2017).

Methods

The mLab App consists of educational HIV prevention information, automated testing reminders, instructions for at-home HIV testing, and an image processing algorithm. This image processing algorithm enables end-users to upload photographs of their OraQuick at-home HIV test result and receive real-time feedback displaying results on a subsequent window with information personalized to users' test results. Before the usability evaluation, a youth advisory board (YAB) comprised of eight community members representative of our target population reviewed initial iterations of the App developed by our research team and made modification recommendations. The final iteration of the mLab App designed by the YAB was tested during the usability evaluation with informatics experts and potential end-users (i.e. YMSM) in individual research sessions. The experts performed a think-aloud protocol while completing a set of tasks known as "use-cases" during their individual research sessions. The use-cases were used to help evaluators find usability problems and categorize and rate the severity of the issues. The experts' review also provided a means for comparing the experts' experience using the App with those of the end-users, creating a more comprehensive evaluation. Experts rated the mLab App and the tasks they completed following Nielsen's heuristic checklist. End-users also participated in the same think-aloud protocol and use-case scenario as the heuristic evaluators to find and rate usability problems during their individual research sessions. End-users rated the usability of the tasks they completed in the use-case scenario using the Health Information Technology Usability Evaluation Scale (Schnall, Cho, & Liu, 2018).

Study activities took place from January-August 2019. The FDA issued an Investigational Device Exemption to use the mLab App since the image-processing algorithm is authorized for investigational purposes only. The Institutional Review Boards at Lurie Children's Hospital and Columbia University approved all study activities; all participants consented before participating in research activities and were compensated for their participation.

Youth advisory board

Sample selection.—The YAB was comprised of eight YMSM. They were recruited from the Mailman School of Public Health or were former study participants of Columbia University School of Nursing (CUSON) research studies. All YAB members were 18- to 29-year-old MSM, reflecting the demographic of the study population. The YAB meetings took place at CUSON.

Procedures.—Members of the YAB met four times from January-March 2019. At each meeting, the research coordinator demonstrated use of the mLab App and each of the mLab's components while the YAB provided feedback. The YAB sessions were audio recorded, and the coordinator recorded field notes. After each meeting, the software developer and graphic designer worked together to update the App based on suggestions from the YAB. Following each update, YAB members re-convened to review the revised version of the App.

Heuristic evaluation

Sample selection.—Five experts in human-computer interaction evaluated the mLab App. The sample was chosen based on Nielsen’s recommendation to include three-to-five experts to achieve saturation in identifying usability flaws (Nielsen & Molich, 1990). The heuristic evaluators were faculty and IT Directors from the New York-Presbyterian Hospital or CUSON with an average of 17 years of informatics experience (range 7–35 years). These heuristic evaluators were invited to join because of their expertise in IT and human-computer interactions. All sessions took place on the Columbia University Medical Center campus.

Procedures.—Heuristic evaluators participated in usability study visits and completed a think-aloud protocol (Bertini, Gabrielli, & Kimani, 2006). Evaluators were first given a mLab App login and asked to complete an OraQuick HIV home test. Experts were then asked to complete the use-cases tasks. The use-case tasks were: watch the OraQuick video in the App, take the OraQuick test, read test results and enter their interpretation of the visual results into the App, upload a picture of the OraQuick test into the App, receive results from the mLab App algorithm, check for nearby services, contact the mLab App team to ask a question and check the inbox, check the testing history and timeline, and finally watch videos on PEP and PrEP in the “learn” section of the app. The Heuristic evaluators’ OraQuick tests provided data to test the App’s image upload and interpretation feature. As evaluators completed use-case tasks, they described what they were thinking, feeling, seeing, and doing. Morae software captured all verbal comments and movements of what participants did on their phones while completing the use-case scenarios. Morae software was used by the research team to review the recordings and analyze task completion to assess if the App was usable and if users could complete the principal functions of mLab App. The research team members (GF, BB) conducting the study visits also took field notes. The goal of having heuristic evaluators complete a think-aloud protocol was to compare their process and experience with those of the end-users.

Instruments.—After completing the use-cases, usability experts completed a demographic survey and Nielsen’s heuristic checklist for usable interface design (Bertini et al., 2006). The completion of the use-case tasks was compared to Nielsen’s heuristic checklist and rated on a scale from 0–4, where 0 indicates no usability problem, the App is ready for use, and 4 being usability catastrophe meaning this problem must be immediately addressed (Bertini et al., 2006).

Data Analysis.—The mean severity scores were calculated for each of the 10 heuristic principles following Nielsen’s heuristic checklist identifying the usability problems found during the completion of the use-case tasks (Mack & Nielsen, 1994). Morae recordings of the heuristic evaluation session and evaluator comments were reviewed by research team members (GF, BB). The recordings were organized by mLab’s central tasks (e.g. clicking on start timer, uploading images to mLab App, getting results on the results page, and checking for nearby services) and analyzed based on the overall task success. Task success rate was included in this heuristic evaluation to provide a more comprehensive understanding of the severity of the usability problems. Task success rate was evaluated based on the participant’s ability to complete the assigned task without assistance from the research team member (GF,

BB) administering the visit. The tasks were categorized based on the Morae software's task completion categories. A task was classified as "completed with ease" if the participant did not require guidance. If participants asked the research team members (GF, BB) questions on how to complete the assigned task, they were labeled as "completed with difficulty." Only participants who could not complete the task, even with the research team members' (GF, BB) help, were labeled as "failed to complete." These different task ratings along with the Morae recordings and completion of Nielsen's heuristic checklist help to identify central areas of usability concern for improving the App.

End-user testing

Sample selection.—Eligibility criteria included: 18–29 years of age, assigned male sex at birth of any current gender identification, understand and read English, sexually active and at risk for HIV infection per Centers for Disease Control and Prevention (CDC) guidance (e.g., recent anal sex with men), smartphone ownership, and self-report being HIV-negative or unknown status (Centers for Disease Control and Prevention, 2019). A sample size of 20 participants was determined based on prior studies' evidence of adequacy in identifying usability issues (Faulkner, 2003). Recruitment occurred through Instagram posts, posting flyers at community-based organizations, and recruiting from a REDCap database of past volunteers who expressed interest in research with CUSON.

Procedures.—During usability study visits, participants first completed a self-administered electronic survey which collected information on the participants' technology use, health literacy, PrEP and PEP exposure, and HIV testing. After the first survey, participants completed an OraQuick test and followed the same use-case tasks as the heuristic evaluators. Like the heuristic evaluators, they were recorded using Morae software and followed the same think-aloud protocol for easy comparison.

Instruments.—Following the think-aloud protocol, participants completed the Health Information Technology Usability Evaluation Scale (Health ITUES). The Health ITUES is a 20 item Likert scale question, allowing alterations at an item level to match the specific task/expectation and health IT system while keeping the construct level standardized with higher scores indicating better App usability (Schnall et al., 2018). Health ITUES were used for this analysis as it is an instrument which has been validated for use in mHealth technologies (Schnall et al., 2018). The sociodemographic characteristics or the testing history of end-users were not linked to their subscore.

Data analysis.—The study team reviewed and organized each session's Morae recordings, analyzing the task success rate of the defined tasks which were also completed by the heuristic evaluators (Table 3). The task success rate was evaluated based on the user's ability to complete the assigned task without guidance from the research team member conducting the visit. Two research team members (GF, BB) also reviewed the Morae recordings of the end-users and identified common usability concerns, positive feedback, and areas of improvement.

Results: Youth Advisory Board

The YAB provided feedback and created the final version of the mLAB App that heuristic evaluators and end-users tested. Their feedback primarily centered around graphic design enhancements and content development in the App.

Graphic design enhancements.

YAB members provided iterative feedback on the user interface of the initial App developed at each of their four meetings. A consistent point of feedback was the design of the login screen, particularly the color scheme and font. Additionally, the testing timer was revised with each iteration. The YAB recommended the testing timer be visible but not overshadowing in order to not make participants feel nervous. As seen from Figure 1 below, the timer's design went through several iterations ending with a small timer appearing on the upper right-hand corner of the App.

Youth advisory board app development: Content development.

The YAB also contributed to the App's "learn" section. The YAB suggested including more information about HIV risk, treatment, biomedical prevention, and improving the interface by including visual graphics. Based on these recommendations, additional revisions were made to expand the content and the interface was updated to include a search function, collapsible details, and incorporation of infographics, links to articles and videos on the content areas.

Results: Heuristic Evaluator

All heuristic evaluators were cisgender females (n=5). Each visit took approximately two hours to complete. Mean severity scores were based on Nielsen's 10 usability heuristics and ranged from 0.2–2.6 (Table 1) (Nielsen, 1992; Nielsen & Molich, 1990). The heuristic evaluators also engaged in think-aloud procedures during their heuristic evaluation to complement their mean severity scores and to help identify usability problems. The heuristic evaluation mean severity scores were used to categorize the usability problems found in the App and helped to rate the severity of the issue.

Usability Problem: User Control and Freedom

The heuristic principle identified with the highest (i.e., worst) mean severity score was "user control and freedom" (Mean 2.6, SD 1.1). An example of poor user control and freedom in the App is the testing timer. Initially, the testing timer lived locally within the user's browser, creating the unintended consequence of requiring users to keep their device's browser open to the App. If the participant exited the App, the timer stopped, not allowing participants to read their results within accurate time windows.

Usability Problem: Visibility of System Status

Another high mean severity score was "visibility of system status" (Mean 2.2, SD 1.1). Experts had trouble uploading images of their OraQuick tests to the App for the algorithm to

read. Many found the language difficult to understand and were unsure if they uploaded the pictures correctly or if the image was acceptable.

Usability Problem: Help and Documentation

In terms of the usability factor “help and documentation” (Mean 1.4, SD 0.9), the connection was unclear between how the user would use the App in conjunction with the OraQuick tests. One evaluator stated:

When we look at some of the standard heuristic principles, the two, the two that I think have the most issue are...well, there are three because there is the mapping, the mental model between the actual test kit and mLab, there is the recovery from errors, and then navigation is not as clear as it needs to be.

Results: End-users

Twenty end-users participated (10 Chicago; 10 New York City). The mean age was 24.25, ages ranging between 18–29 years. The end-users were 55% White, 20% Black or African American, 15% Asian or Asian American, and 20% multiracial. 45% of participants completed some college, and 35% had an undergraduate degree.

End-user Health-ITUES Usability Problems

The Health-ITUES results are listed in Table 2. The Health-ITUES scores allow for broad identification of usability issues. The scale that scored the highest (i.e. best) for the end-users was perceived ease of use, capturing the user-system interaction (Schnall et al., 2018). The scale with the lowest (i.e. worst) score was the user control scale, corroborating the heuristic evaluators’ ratings in their heuristic evaluation mean severity scores. The sociodemographic characteristics or the testing history of end-users were not linked to their subscore.

End-user Qualitative Feedback

In addition to the Health-ITUES, we used the Morae audio recordings for our qualitative analysis. Most end-users found the service-locator feature very useful. One participant stated:

I think the finding service thing is really useful though because I can’t tell you how many times I’ve been (looking), maybe not just for HIV stuff, but just for like healthcare information and stuff. It’s good to have an easy portal.

Many of the participants also found the results feature reassuring, providing an added layer of confirmation. One end-user mentioned:

It’d be a really useful tool. If I was ever to do another HIV home test, I would use the [mLab] app, especially to take a photo of the result and to get the second confirmation.

End-user Task Completion

In addition to end-user feedback, four primary tasks were coded and analyzed for success rate (Table 3). These were the same four tasks that were analyzed during the heuristic evaluation and were selected because they are the four central tasks of the App that lead to other app procedures. Without completing these tasks, users could not complete the remaining tasks in the use-case scenarios. These tasks were assessed for completion success rate to understand the usability of the App and identify potential usability problem areas for the App developers to address before testing the efficacy of the App in a randomized controlled trial.

Of the 25 participants (end-user, n=20; heuristic experts, n=5), 92% were able to check for nearby services with ease, without prompting from research team members (GF, BB) administering the study visit. Furthermore, 84% started the timer with ease, while 16% completed the task with some difficulty. Additionally, 60% of the total participants uploaded the image to the App, while 12% failed to complete this task. One end-user expressed frustration after the timer reset when he exited the App:

Fix that timer guys...Because if that timer was at 5 minutes left and then you did something and then I came back, that'll just mess everything up...you need to make it 'stupid-proof' and this is one the things that you need to do to make it stupid-proof.

Results: App Updates from Testing Feedback

This feedback from Health-ITUES rating, think-aloud protocol, task success rate, and heuristic mean severity scores together provide the App developers with a details on features needing improvement. This usability evaluation's informed updates to the App. One update included moving the OraQuick testing timer to the webserver, which allows participants to navigate through the App and exit it without interfering with the timer. To improve "help and documentation," an overlay walk-through was created assisting users the first three times they enter mLab, and is also available anytime through a button on the home screen. The overlay helps to bridge the gap between the OraQuick device and the mLab Application, showing participants how to use the App in conjunction with the OraQuick test. Additionally, several end-users and heuristic evaluators described the image uploading language for the App as ambiguous, which may explain why 28% of participants reported difficulty uploading their OraQuick test image. Based on heuristic and end-user suggestions, the image uploading directions were updated to be more explicit. However, this was changed to follow the FDA's guidance which required us to include additional details even though this did not align with our participants' preferences.

Discussion

Using technology to provide support for HIV self-testing has the potential to increase HIV testing uptake among YMSM (LeGrand et al., 2017). This paper focuses on a multi-method, user-centered design approach to evaluate the usability of the mLab App. Although the findings yielded feedback on specific areas needing improvement, overall, participants

found mLab highly usable receiving high scores on validated usability scales. Participants commented on how they enjoyed using the App, and the secondary confirmation the imaging algorithm provided with future action steps. This feedback shows that the mLab App met the study's overarching goal to develop an HIV prevention app that could help users better understand their HIV results and link people to HIV information and services.

Our study is distinct in the field of mHealth usability evaluations. The study population represents an underrepresented group of end-users -- young sexual minorities at high-risk for acquiring HIV in their lifetime (Rasberry et al., 2018; Wood, Lee, Barg, Castillo, & Dowshen, 2017). Additionally, despite the extensive use of apps by youth, few apps are designed or evaluated by YMSM (Zapata, Fernández-Alemán, Idri, & Toval, 2015). In a 2015 review of 285 HIV and AIDS smartphone apps, 7% of apps were developed for MSM (Sullivan, Jones, Kishore, & Stephenson, 2015). While technology-based interventions promoting at-home HIV testing have increased, the interventions do not include all the features available in mLab (LeGrand et al., 2017). For example, one study used video conferencing to provide pre and post-test counseling and support for at-home testing (Maksut, Eaton, Siembida, Driffin, & Baldwin, 2016). Another study used an app to prompt a follow-up call for post-test counseling and linkage to care once the home-test was opened (Wray, Chan, Simpanen, & Operario, 2017). A study in Seattle and Atlanta focused on prevention recommendations, ordering HIV self-tests and condoms, reminders, and service locators (Sullivan et al., 2017). However, none of the interventions provided HIV testing information, linkage to care, and result interpretation in one central mobile health application, which has been achieved through the development of the mLab App.

The innovation of this approach is grounded in the image-processing algorithm and the multi-method, user-focused usability evaluation process. This is highlighted by a recent comprehensive literature review of usability evaluations of eHealth HIV interventions, which yielded 128 full-text articles from 2005–2019 (Davis et al., 2020). However, only 28 articles described usability evaluations, with just five studies conducting think-aloud protocols with end-users and experts and only five studies using use-case scenarios (Davis et al., 2020). Only two studies addressed HIV at-home testing in the field of HIV prevention (Davis et al., 2020). Both studies only used one method of evaluation, questionnaires, and only one study used a validated questionnaire, the System Usability Scale, illustrating a lack in the literature on current mix-methods eHealth usability evaluations (Davis et al., 2020; Sullivan et al., 2017). The mLab usability study expands the current literature by providing an example of a multi-method evaluation of an innovative HIV testing intervention using validated questionnaires, think-aloud protocols, and use-case scenarios. This study provides a potential framework for future multi-method eHealth usability studies to follow. In addition to providing a framework for future evaluations, the innovative imaging algorithm while currently being used for YMSM in HIV at-home testing purposes, could be tried in various other populations at risk of HIV infection and has the potential to also expand beyond HIV testing and move into reading and processing other at-home medical tests.

Limitations

A limitation of our evaluation was that our sample was primarily recruited from a database that included participants who had either expressed interest in future studies or had participated in past research with CUSON. Since most participants had previously expressed interest in research, they may be more interested in HIV prevention and health management than their peers. Additionally, all heuristic evaluators were assigned female sex at birth, not representative of our target population. Since the heuristic evaluators are not representative of our target population, their experience interfacing with the App could be different as the App was targeted towards another population. Despite these limitations, we were able to identify areas of improvement to refine the mLab App further and test the imaging algorithm.

Conclusion

The feedback from the YAB, end-users, and expert evaluators led to graphic design enhancements and updates to both the content and functional features of the mLab App. Both informatics experts and end-users described the App as helpful in organizing at-home testing, aesthetically pleasing, and easy to use. Feedback from this usability evaluation informed the refinement of the mLab App, which is currently being tested in a multi-site, 12-month efficacy trial. This study expands the current literature on usability evaluation of at-home HIV testing apps by having both end-users and heuristic evaluators use validated measures to evaluate a new, comprehensive HIV testing and prevention App. This usability evaluation provides essential insights into using imaging algorithms to help users understand their results in both the field of HIV testing and other medical testing fields and has the potential to centralize HIV self-testing.

Acknowledgments

The authors would like to thank the study staff of the Division of Adolescent Medicine at Lurie Children's Hospital and the Columbia University School of Nursing for their help in the execution of this usability evaluation.

The research was supported by National Institute of Mental Health, Grant/Award Number: R01MH118151 and National Institute of Nursing Research, Grant/Award Number: K24NR018621. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Beach LB, Greene GJ, Lindeman P, Johnson AK, Adames CN, Thomann M, Wahington P, & Phillips G (2018). Barriers and facilitators to seeking HIV Services in Chicago among young men who have sex with men: Perspectives of HIV service providers. *AIDS Patient Care and STDs*, 32(11), 468–476. [PubMed: 30398956]
- Bertini E, Gabrielli S, & Kimani S (2006). Appropriating and assessing heuristics for mobile computing. Paper presented at the Advanced Visual Interfaces International Working Conference, Venezia, Italy.
- Bevan N, Carter J, Earthy J, Geis T, & Harker S (2016). New ISO Standards for Usability, Usability Reports and Usability Measures. Paper presented at the Proceedings, Part I, of the 18th International Conference on Human-Computer Interaction. Theory, Design, Development and Practice - Volume 9731.

- Catalani C, Philbrick W, Fraser H, Mechael P, & Israelski DM (2013). mHealth for HIV Treatment & Prevention: A Systematic Review of the Literature. *Open AIDS J*, 7, 17–41. [PubMed: 24133558]
- Centers for Disease Control and Prevention. HIV Risk Behaviors. (2019). Retrieved from <https://www.cdc.gov/hiv/risk/estimates/riskbehaviors.html>
- Dailey A, Gant Z, Johnson S, Li J, Wang S, Hawkins D, ... Hernandez A (2020). HIV Surveillance Report, 2018 (Updated). Retrieved from <http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html>
- Davis R, Gardner J, & Schnall R (2020). A Review of Usability Evaluation Methods and Their Use for Testing eHealth HIV Interventions. *Curr HIV/AIDS Rep*, 17(3), 203–218. [PubMed: 32390078]
- Devi BR, Syed-Abdul S, Kumar A, Iqbal U, Nguyen PA, Li YC, & Jian WS (2015). mHealth: An updated systematic review with a focus on HIV/AIDS and tuberculosis long term management using mobile phones. *Comput Methods Programs Biomed*, 122(2), 257–265. [PubMed: 26304621]
- Faulkner L (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments and Computers*, 35(3), 379–383.
- Forrest JI, Wiens M, Kanters S, Nsanzimana S, Lester RT, & Mills EJ (2015). Mobile health applications for HIV prevention and care in Africa. *Curr Opin HIV AIDS*, 10(6), 464–471. [PubMed: 26352394]
- Hayat SN, & Ramdani F (2020). A comparative analysis of usability evaluation methods of academic mobile application: are four methods better? Paper presented at the Proceedings of the 5th International Conference on Sustainable Information Engineering and Technology, Malang, Indonesia.
- Jaspers MW (2009). A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. *Int J Med Inform*, 78(5), 340–353. [PubMed: 19046928]
- Kushniruk A, Senathirajah Y, & Borycki E (2017). Towards a Usability and Error “Safety Net”: A Multi-Phased Multi-Method Approach to Ensuring System Usability and Safety. *Stud Health Technol Inform*, 245, 763–767. [PubMed: 29295201]
- LeGrand S, Muessig K, Horvath K, Rosengreen A, & Hightow-Weidman L (2017). Using Technology to Support HIV Self-Testing Among Men Who Have Sex with Men. *Current Opinion in HIV and AIDS*, 12(5), 425–431. [PubMed: 28617712]
- Lenhart A, Ling R, Campbell S, & Purcell K (2015). Teens and Mobile Phones. Pew Research Center. Retrieved from <https://www.pewinternet.org/2010/04/20/teens-and-mobile-phones/>
- Mack R, & Nielsen J (1994). Usability inspection methods. New York, NY: Wiley & Sons.
- Maksut JL, Eaton LA, Siembida EJ, Driffin DD, & Baldwin R (2016). A Test of Concept Study of At-Home, Self-Administered HIV Testing With Web-Based Peer Counseling Via Video Chat for Men Who Have Sex With Men. *JMIR Public Health Surveill*, 2(2), e170. [PubMed: 27974287]
- Maloney KM, Bratcher A, Wilkerson R, & Sullivan PS (2020). Electronic and other new media technology interventions for HIV care and prevention: a systematic review. *J Int AIDS Soc*, 23(1), e25439. [PubMed: 31909896]
- Nguyen LH, Tran BX, Rocha LEC, Nguyen HLT, Yang C, Latkin CA, Throson A, & Strömdahl S (2019). A Systematic Review of eHealth Interventions Addressing HIV/STI Prevention Among Men Who Have Sex With Men. *AIDS and Behavior*, 23(9), 2253–2272. [PubMed: 31401741]
- Nielsen J (1992). Finding usability problems through heuristic evaluation. Paper presented at the CHI '92 SIGCHI Conference on Human factors in Computing Systems.
- Nielsen J, & Molich R (1990). Heuristic evaluation of user interfaces. Paper presented at the SIGCHI conference on Human factors in computing systems, ACM.
- Pew Research Center. (2018). Smartphone access nearly ubiquitous among teens, while having a home computer varies by income. Retrieved from https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/pi_2018-05-31_teenstech_0-04/
- Rasberry C, Lowry R, Johns M, Robin L, Dunville R, Pampati S, ... Balaji A (2018). Sexual Risk Behavior Differences Among Sexual Minority High School Students — United States, 2015 and 2017. *Morbidity and Mortality Weekly Report*. Retrieved from

- Rossmann K, Salamanca P, & Macapagal K (2017). A Qualitative Study Examining Young Adults' Experiences of Disclosure and Nondisclosure of LGBTQ Identity to Health Care Providers. *Journal of Homosexuality*, 64(10), 1390–1410. [PubMed: 28459379]
- Schnall R, Cho H, & Liu J (2018). Health Information Technology Usability Evaluation Scale (Health-ITUES) for Usability Assessment of Mobile Health Technology: Validation Study. *JMIR Mhealth Uhealth*, 6(1), e4. [PubMed: 29305343]
- Schnall R, John RM, & Carballo-Diequez A (2016). Do high-risk young adults use the HIV self-test appropriately? Observations from a think-aloud study. *AIDS and Behavior*, 20(4), 939–948. [PubMed: 26518679]
- Schnall R, Rojas M, Bakken S, Brown W, Carballo-Diequez A, Carry M, ... Travers J (2016). A user-centered model for designing consumer mobile health (mHealth) applications (apps). *J Biomed Inform*, 60, 243–251. [PubMed: 26903153]
- Sheehan B, Lee Y, Rodriguez M, Tiase V, & Schnall R (2012). A comparison of usability factors of four mobile devices for accessing healthcare information by adolescents. *Appl Clin Inform*, 3(4), 356–366. [PubMed: 23227134]
- Silver L (2019). Smartphone ownership is growing rapidly around the world, but not always equally. Pew Research Center. Retrieved from <https://www.pewresearch.org/global/2019/02/05/in-emerging-economies-smartphone-adoption-has-grown-more-quickly-among-younger-generations/>
- Sullivan PS, Driggers R, Stekler JD, Siegler A, Goldenberg T, McDougal SJ, ... Stephenson R (2017). Usability and Acceptability of a Mobile Comprehensive HIV Prevention App for Men Who Have Sex With Men: A Pilot Study. *JMIR Mhealth Uhealth*, 5(3), e26. [PubMed: 28279949]
- Sullivan PS, Jones J, Kishore N, & Stephenson R (2015). The Roles of Technology in Primary HIV Prevention for Men Who Have Sex with Men. *Curr HIV/AIDS Rep*, 12(4), 481–488. [PubMed: 26519083]
- Tan WS, Liu D, & Bishu R (2009). Web evaluation: Heuristic evaluation vs. user testing. *International Journal of Industrial Ergonomics*, 39, 621–627.
- Wood SM, Lee S, Barg FK, Castillo M, & Dowshen N (2017). Young Transgender Women's Attitudes Toward HIV Pre-exposure Prophylaxis. *J Adolesc Health*, 60(5), 549–555. [PubMed: 28132744]
- Wray T, Chan PA, Simpanen E, & Operario D (2017). eTEST: Developing a Smart Home HIV Testing Kit that Enables Active, Real-Time Follow-Up and Referral After Testing. *JMIR Mhealth Uhealth*, 5(5), e62. [PubMed: 28483744]
- Ybarra ML, Prescott TL, Phillips GL, Bull SS, Parsons JT, & Mustanski B (2017). Pilot RCT Results of an mHealth HIV Prevention Program for Sexual Minority Male Adolescents. *Pediatrics*, 140(1).
- Zapata BC, Fernández-Alemán JL, Idri A, & Toval A (2015). Empirical Studies on Usability of mHealth Apps: A Systematic Literature Review. *Journal of Medical Systems*, 39(2), 1. [PubMed: 25600193]
- Zhong F, Tang W, Cheng W, Lin P, Wu Q, Cai Y, Tang S, Fan L, Zhao Y, Chen X, Mao J, Meng G, Tucker JD, & Xu H (2017). Acceptability and feasibility of a social entrepreneurship testing model to promote HIV self-testing and linkage to care among men who have sex with men. *HIV medicine*, 18(5), 376–382. [PubMed: 27601301]

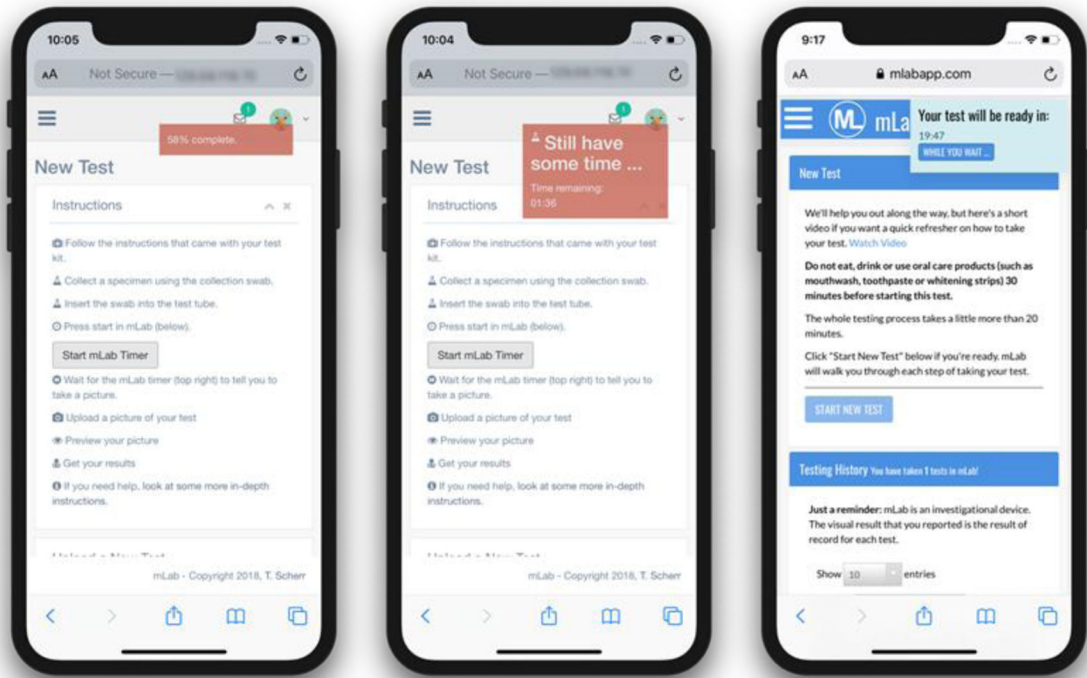


Figure 1.
mLab App Timer Developments

Table 1Heuristic Evaluation Mean Severity Scores^a

Heuristic Principle	Mean	SD
User control and freedom	2.6	(1.1)
Visibility of system status	2.2	(1.1)
Match between system and the real world	1.4	(1.1)
Error prevention	1.4	(1.1)
Help and documentation	1.4	(0.9)
Consistency and standards	1.2	(1.0)
Flexibility and efficiency of use	1.2	(1.3)
Recognition rather than recall	1.0	(1.0)
Help users recognize, diagnose, and recover from errors ^b	0.75	(0.9)
Aesthetic and minimalist design	0.2	(0.4)

^aRating score from 0=best to 4=worst; no usability problem (0), cosmetic problem only (1), minor usability problem (2), major usability problem (3), usability catastrophe (4).

^b4 out of 5 evaluators answered the question

Table 2Health-ITUES Scores (n=20)^a

Scale	Mean	(SD)
Impact	4.2	(1.2)
Perceived usefulness	4.1	(1.2)
Perceived ease of use	4.4	(0.9)
User control	3.6	(1.2)
Overall health-ITUES score	4.1	(1.1)

^aRating is based off a five-point Likert scale with five being the highest or best score.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3Task Success Rate (n=25)^a

Task	Percentage completed with ease	Percentage completed with difficulty	Percentage failed to complete
Click on start timer	84	16	0
Upload image to mLab app	60	28	12
Get your results on the results page	80	4	16
Check for nearby services	92	0	8

^aThe task success rate percentages include attempts by both heuristic evaluators and end-users

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript