AIDS Behav
DOI 10.1007/s10461-014-0818-8

ORIGINAL PAPER

What Should the Ideal HIV Self-Test Look Like? A Usability Study of Test Prototypes in Unsupervised HIV Self-Testing in Kenya, Malawi, and South Africa

Roger B. Peck · Jeanette M. Lim · Heidi van Rooyen · Wanjiru Mukoma · Lignet Chepuka · Pooja Bansil · Lucia C. Knight · Nelly Muturi · Ellen Chirwa · Arthur M. Lee · Jeff D. Wellhausen · Olivia Tulloch · Miriam Taegtmeyer

Abstract HIV self-testing (HIVST) is increasingly being sought and offered globally, yet there is limited information about the test features that will be required for an HIV self-test to be easy to use, acceptable to users, and feasible for manufacturers to produce. We conducted formative usability research with participants who were naïve to HIVST using five prototypes in Kenya, Malawi, and South Africa. The tests selected ranged from early-stage prototypes to commercially ready products and had a diverse set of features. A total of 150 lay users were video-recorded conducting unsupervised self-testing and interviewed to understand their opinions of the test. Participants did not receive a test result, but interpreted standardized result panels. This study demonstrated that users will refer to the instructions included with the test, but these can be confusing or difficult to follow. Errors were common, with less than 25 % of participants conducting all steps correctly and 47.3 % of participants performing multiple errors, particularly in sample collection and transfer. Participants also had difficulty interpreting results. To overcome these issues, the ideal HIV self-test requires pictorial instructions that are easy to understand, simple sample collection with integrated test components, fewer steps, and results that are easy to interpret.

Keywords HIV self-test · Target product profile · Usability · Sub-Saharan Africa · Prototypes

R. B. Peck (✉) · J. M. Lim · P. Bansil · A. M. Lee · J. D. Wellhausen
PATH, 2201 Westlake Avenue, Suite 200, Seattle, WA 98121, USA
e-mail: rpeck@path.org

H. van Rooyen
Human Sciences Research Council, Durban, South Africa

W. Mukoma · N. Muturi
Liverpool VCT, Care, and Treatment, Nairobi, Kenya

L. Chepuka · E. Chirwa
Kazumu College of Nursing, Blantyre, Malawi

L. C. Knight
School of Public Health, University of the Western Cape, Bellville, South Africa

O. Tulloch · M. Taegtmeyer
Liverpool School of Tropical Medicine, Liverpool, UK

Introduction

Despite progress in the global scale-up of HIV testing and counselling (HTC), significant gaps in access to universal HIV testing remain [1]. Through a combination of convenience, confidentiality, and privacy, HIV self-testing (HIVST) has the potential to reduce this gap and reach populations that are poorly served by available testing programs. Limited evidence suggests that HIVST is feasible and preferred to facility-based testing, but that the accuracy of self-testing varies [2].

Most published studies on self-testing come from high-income countries; few have systematically examined HIVST among lay users. A review of 11 studies across different contexts suggested HIVST has high acceptability in most instances and increases the reach of HIV testing services with users capable of providing accurate results [3]. Data indicating user preference for oral or blood rapid tests are weak; however, users generally have more difficulty performing blood-based rapid tests [4, 5]. In Singapore, of 350 participants using fingerstick rapid tests, almost 90 % found the kit and instructions easy to use and

Published online: 20 June 2014

understand. However, 85 % failed to perform all steps correctly, especially blood sampling, and 56 % received invalid results [6]. Another study in Spain with 313 participants conducting unsupervised HIVST using whole-blood with adapted instructions found that only 8 % obtained an invalid self-test, with 1.1 % of positive results interpreted as negative [7]. High levels of accuracy (99.2 %) were obtained in Malawian communities from unsupervised oral HIVST with the use of illustrated instructions, after a brief product demonstration [8].

In general, most rapid diagnostic tests for HIVST are derived from products for professional use in assisted testing environments, for example, by modifying the labelling, packaging, and/or instructions. There is only one test kit approved by the United States Food and Drug Administration for HIVST in the United States [9] and few products are available for HIVST outside of the United States. The First International Symposium on HIV Self-Testing issued a clear call for new self-testing products and for further research on market demand and target product profile (TPP) for self-test kits [1]. A detailed TPP describes the minimally acceptable and ideal product characteristics and specifications for target population, performance, specimen type and volume, storage conditions, time to results, nature of results, and other parameters [10, 11]. The ASSURED criteria (Affordable, Sensitive, Specific, User-friendly, Rapid and robust, Equipment-free, Delivered to those who need it) [12] set out basic test requirements, but how a test achieves those criteria is dependent on the features included by the manufacturer and their usability in the hands of the intended users. There are limited data available on HIVST prototypes.

Currently there is no broadly available TPP in existence for HIVST. Usability testing is required to identify use-related problems and hazards of test features in order to address these early in the development process [13]. To address this gap, we report findings from a formative usability study with lay users involving a range of HIV self-test prototypes, including rapid oral and fingerstick test kits. The overall aim was to create a TPP for use in developing high-quality HIVST options for target populations.

Methods

We used a mixed method approach to assess the usability of five HIV self-test kit prototypes among lay users. Data-collection techniques comprised video observation of lay users conducting unsupervised self-testing, quantitative participant observation checklist, exit questionnaires, and qualitative interviews.

Test Prototypes and Instructions for Use

The study objective was to evaluate the usability of a wide variety of test features suitable for HIV self-test kits. Therefore, we identified existing HIV rapid tests and prototype tests with features potentially suitable for HIVST through interaction with manufacturers to ensure a range of characteristics (Table 1). We do not name prototypes but reference these as O1 (oral test), FS1, FS2, FS3, and FS4 (fingerstick tests). Only non-functional tests incapable of developing control and test signals were used. Simple, pictorial test instructions appropriate for lay users were developed. These were reviewed in country and a simple accompanying text was translated and back-translated to ensure accurate versions in local languages.

Timeline and Settings

The selected prototypes were evaluated between March and September 2013 in Kenya, Malawi, and South Africa—countries with generalized HIV epidemics, commitment to universal access to HIV testing, and emerging markets for HIVST. The setting differed by country: users in Kenya conducted testing at four urban voluntary counselling and testing sites, in Malawi at clinics in two districts, and in South Africa at an office in a community-based setting. Each study site had clear referral pathways to ensure that those who wanted to learn their HIV status were able to voluntarily access HTC on-site or nearby. Each site provided a private testing space with two complete test kits (components and instructions), pencil, and clock. The study provided additional identical lancets and generic blade-style lancets for participants to use, if needed.

Participant Selection

A sample size of ten people per test kit from potential target populations in each country were chosen to detect over 80 % of the usability problems with a device [10, 14]. We purposively sampled the population to include participants with a range of age, sex, education (as a proxy for literacy), geographic location, and socioeconomic characteristics. Participants were purposively recruited through community-based mobilizers, provider referral, or study staff who completed the informed consent process. Participants who had ever self-tested for HIV were excluded, but those who had previously received a provider-initiated HIV test or who had conducted another self-test (e.g. pregnancy, blood sugar monitoring) were included. Verification of prior self-testing for HIV relied on self-reports.

Table 1 Characteristics of study test prototypes

Test type	O1	FS1	FS2	FS3	FS4
Sample type	Oral fluid	Fingerstick	Fingerstick	Fingerstick	Fingerstick
Instructions	Single page, one-sided print, graphic+English+local language	Single page, one-sided print, graphic+English+local language	Single page, one-sided print, graphic+English+local language	Single page, two-sided print, graphic+English+local language	Single page, one-sided print, graphic+English+local language
Lancet	N/A	Single-use, button-activated lancet	Single-use, touch-sensitive lancet	Single-use, touch-sensitive lancet	Integrated single-use, touch-sensitive lancet
Sample collection	Oral swab using test device	Dropper pipette	Capillary tube	Capillary tube	Direct to capillary tube which is integrated into test device
Sample volume	N/A	2 drops (approximately 25–30 μL)	5 μL	30 μL	12.5 μL
Sample transfer	Integrated sample collection and transfer device	Count blood drops from pipette directly onto test	Capillary tube integrated into test device	Blood in capillary tube added to test liquid; mixture added to test device	Push test arm to enable blood flow into test device sample port
Test liquid	Premeasured; test added directly to test liquid	Premeasured	Premeasured; test added directly to test liquid	Three different steps involved: count drops, add blood to test liquid, add rest of test liquid	Count drops
Number of components	4	7	7	8	6
Waiting time for results to develop (min)	20	15	15	0	15
Platform technology	Lateral flow strip test	Lateral flow strip test	Lateral flow strip test	Flow-through test	Lateral flow strip test
Results format	Control line, test line	Control line, HIV-1 test line, HIV-2 test line	Control line, test line (device does not label lines)	Control symbol, test symbol (control and test symbols differ from each other)	Control line, test line

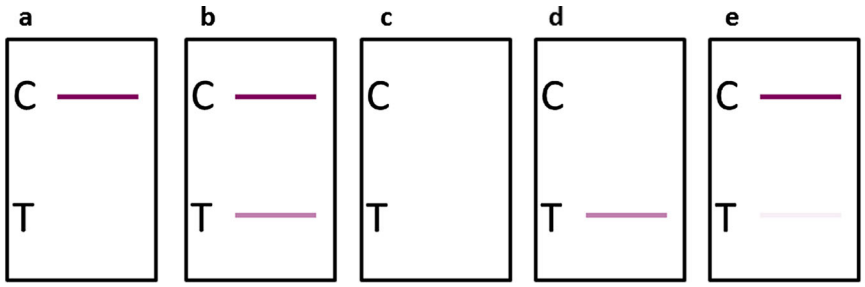

Fig. 1 Example of standardized test panel for results interpretation

a: Negative test result - Control line (C) present and test line (T) absent.
b: Positive test result - Control line (C) and test line (T) present.
c: Invalid test result - Control line (C) and test line (T) absent.
d: Invalid test result - Control line (C) absent and test line (T) present.
e: Weak positive test result - Control line (C) present and test line (T) present, but significantly weaker color intensity.

Procedures

Real-time video recording was used to provide a record of the procedure while creating the conditions of unsupervised HIVST [15]. Participants were instructed to perform the HIV self-test with the materials supplied. No additional instructions or coaching was provided. If participants requested help, they were asked to conduct the test to the best of their ability, without assistance. Study staff observed individuals via live video and recorded performance of each step on a predetermined standardized checklist. Observers also made brief field notes of significant observations. A second staff member observed the video, reviewed the checklist, and resolved any discrepant observations [16].

Non-functional tests incapable of developing control and test signals were used in the procedure. As a separate step, after the participant completed the procedure, each participant was asked to interpret a standardized test panel specific to the prototype to assess the participant's ability to interpret positive, negative, and invalid test results (Fig. 1). Interpreting test results, specifically faint lines can be difficult. Test kit instructions state that users should interpret any visible line as positive. Therefore, tests with weak positive lines were included in the test results panel to see how lay users managed this task.

After the completion of the results panels, users were asked to participate in an interviewer-administered questionnaire and a semi-structured interview to explore their opinions about the test and its features, counselling, and potential scenarios for buying and using self-tests [17].

Analysis

Observation checklist data (including field notes) and closed-ended questions from the exit interview were entered into Excel. The quantitative data were analyzed for frequencies. The qualitative interview data were transcribed and, when necessary, translated into English. A framework approach was used for analysis of the qualitative data with a common coding framework created to include major themes, categories, and concepts [18]. Interviews were coded and managed in QSR Nvivo 10 software. The framework was revised throughout the analysis, new themes were identified, and linked themes collapsed together. Findings from video observation, exit questionnaires, and qualitative interviews were triangulated to ensure trustworthiness of the data.

Ethical approval to conduct the study was obtained in Kenya, Malawi, and South Africa.

Results

The study included 150 participants from Kenya (n = 49), Malawi (n = 47), and South Africa (n = 54), resulting in 33 participants using an oral test (O1) and 117 using a fingerstick test (FS1 = 29, FS2 = 29, FS3 = 30, FS4 = 29) (Table 2). Errors were common. Less than 25 % of all users correctly performed all steps (39.4 % O1, 24.1 % FS1, 20.7 % FS2, 13.3 % FS3, and 24.1 % FS4), and 47.3 % of participants conducted multiple errors. Data have been analyzed to identify test attributes that performed well or need improvement and can be used to inform TPP development.

Instructions

Video observation revealed that 88 % of participants reviewed the instructions before, during, and after testing. Many referred to them step-by-step. For each test, most

Table 2 Participant demographics

	O1 (n = 33) n (%)	FS1 (n = 29) n (%)	FS2 (n = 29) n (%)	FS3 (n = 30) n (%)	FS4 (n = 29) n (%)
Age					
Mean (standard deviation)	31.9 (11.6)	34.6 (13.2)	32.7 (9.7)	34.5 (12.4)	30.0 (10.9)
Median	27	33	31	30	27
Range	19–61	18–68	19–54	20–63	18–60
Sex					
Female	16 (48.5)	13 (44.8)	18 (62.1)	17 (56.7)	16 (55.2)
Level of education					
None to completed primary	8 (24.2)	8 (27.6)	8 (27.6)	5 (16.7)	8 (27.6)
Some or completed secondary	15 (45.5)	16 (55.2)	17 (58.6)	17 (56.7)	12 (41.4)
Diploma/certificate or college degree	10 (30.3)	5 (17.2)	4 (13.8)	8 (26.7)	9 (31.0)
Occupation					
Employed	19 (57.6)	16 (55.2)	20 (69.0)	18 (60.0)	16 (55.2)
Unemployed	11 (33.3)	11 (37.9)	7 (24.1)	9 (30.0)	10 (34.5)
Student	3 (9.1)	2 (6.9)	2 (6.9)	3 (10.0)	3 (10.3)
Used a self-test before					
Yes (diabetes, pregnancy)	2 (6.1)	2 (6.9)	4 (13.8)	3 (10.0)	2 (6.9)

Table 3 Participants conducting steps correctly

	O1 (n = 33) n (%)	FS1 (n = 29) n (%)	FS2 (n = 29) n (%)	FS3 (n = 30) n (%)	FS4 (n = 29) n (%)
Number of test kits used					
Used one test kit to conduct the test	19 (57.6)	21 (72.4)	22 (75.9)	18 (60.0)	24(82.8)[a]
Test instructions					
Reviewed instructions	29 (87.9)	26 (89.7)	29 (100.0)	25 (83.3)	23 (79.3)
Oral sample collection					
Collected oral sample correctly	13 (39.4)	N/A	N/A	N/A	N/A
Fingerstick sampling					
Used lancet correctly	N/A	17 (58.6)	16 (55.2)	21 (70.0)	17 (58.6)
Number of participants who pricked finger only one time	N/A	13 (44.8)	18 (62.1)	17 (56.7)	18 (62.1)
Number of times pricked self (range)	N/A	0–5+	1–4	1–5	1–3
Collected sample correctly	N/A	14 (48.3)	17 (58.6)	9 (30.0)	15 (51.7)
Sample transfer					
Transferred sample correctly	N/A	7 (24.1)	N/A	17 (56.7)	12 (41.4)
Test liquid					
Added test liquid correctly	28 (84.8)	16 (55.2)	17 (58.6)	6 (20.0)	12 (41.4)
Timing					
Timed results correctly	17 (51.5)	5 (17.2)	6 (20.7)	12 (40.0)	8 (27.6)
Correctly conducted all steps	13 (39.4)	7 (24.1)	6 (20.7)	4 (13.3)	7 (24.1)
Conducted more than one error	5 (15.2)	17 (58.6)	9 (31.0)	23 (76.7)	17 (58.6)

[a] Due to limited tests available for use, most participants were only offered one test kit

participants (61.3–92.3 %, depending on the test) used only one kit (Table 3). Of those who opened a second test kit, some did so to use another lancet, and others redid the test, stating that they wanted to repeat incorrect steps. The inclusion of pictorial and written instructions was considered helpful by the respondents.

"They're good, especially because of the photographs…" (Kenya, Female, FS4)

Local translations were also valued.

"You cannot say this to somebody who has never gone to school … unless you put it in [their] mother tongue, I bet this is only meant for people who are educated." (Kenya, Female, FS4)

Qualitative interviews revealed that some participants felt instructions pertaining to specific test components or procedures were unclear. Test components not described in the test instructions, such as bandages and desiccant sachets, caused confusion and errors. Errors included participants adding the desiccant sachet contents to the test liquid or onto the sample collection device (n = 3), and recording time on the bandage instead of the space provided (n = 1).

"[The step] was a bit confusing, because at first I didn't know if I should remove the [cap] on the test or pour over it. Because there is no instruction for that step. Even the picture doesn't show." (South Africa, Male, FS3)

Oral Swabbing

Less than half (39.4 %, n = 13/33) of the participants collected the oral sample correctly. Errors included using the wrong end of the collection device (n = 2); swabbing only one side of the mouth or only the upper or lower gum (n = 11); swabbing the incorrect area of the mouth (n = 9), such as teeth, tongue, or cheek; or swabbing incorrectly (e.g., like brushing teeth). Some dipped the swab in the test liquid before swabbing (n = 3). Participants explained that their confusion was generally due to unclear instructions, misunderstanding language, and complicated procedures.

"Eh It was difficult because I couldn't understand the word "swab," I couldn't understand what was required of me…" (South Africa, Male, O1)
"They were not difficult, except step number eight which is about swabbing of the gums, it was difficult to follow the instructions and the picture wasn't all that clear." (Malawi, Male, O1)

Fingerstick Sampling and Sample Transfer to Device

Only 60.7 % (n = 71/117) of participants were able to prick their finger with the lancet and produce a drop of blood. Videos showed participants examining the lancet in detail. Others unintentionally triggered the lancet before pricking (n = 20), and 43.6 % of participants pricked themselves more than once. Some utilized a second lancet from another test kit or provided separately (n = 47), including 24 who opted for the generic blade-style lancet; two were observed reaching for personal items (e.g., safety pin or razor blade) to try to prick themselves. A few participants (n = 6) were observed conducting the process with greater ease with a second lancet. Field note observations showed that some participants indicated a great deal of frustration with pricking their finger (including two abandoning the test after failing). Few participants (n = 3) expressed reluctance or fear of pricking.

Exit interviews revealed that participants did not know how to use the triggered (touch sensitive or button) lancets.

"It was difficult to use the lancet from the testing device, even though the picture was showing how to do it… I didn't figure [it] out." (Malawi, Male, FS4)

Some participants felt that the provided instructions alone were inadequate to guide them through the steps.

"I had a problem on how to use the lancet such that I pricked the wrong finger, we need to be taught how to use it otherwise people will be just pricking themselves several times." (Malawi, Female, FS2)
"That thing for pricking the finger [lancet], it doesn't show that there is something inside…that should be reviewed a little so that even if you can't see it, you can tell that there is something that can prick you inside." (Kenya, Female, FS2)

All fingerstick tests required the participant to collect blood into a pipette or capillary. Only 47 % (n = 55/117) were able do this according to instructions. Tests with a separate blood collection device (FS1 and FS3) resulted in 39 % (n = 23/59) of participants correctly collecting and transferring the sample to the test. Participants had difficulties with both the mechanics of using a pipette and a capillary tube. Some applied the sample directly to the test (n = 8), either after not being able to collect blood or because of incorrectly interpreting the instructions. Tests with sample collection integrated into the test cassette (devices FS2 and FS4) had better results, with 55.2 % (n = 32/58) of participants applying blood directly to the test correctly.

Test Liquid

All of the tests in this study utilized a test liquid. The method for combining sample with test liquid varied. With a test that involved three different steps for adding test liquid (FS3), only 20.0 % (n = 6/30) of participants did this correctly. In contrast, when the test was added to the premeasured test liquid in one step (O1, FS2), this was performed correctly by 72.6 % of the participants

Table 4 Participants correctly identifying test result panels

	O1 (n = 33) n (%)	FS1 (n = 29) n (%)	FS2 (n = 29) n (%)	FS3 (n = 30) n (%)	FS4 (n = 28)[a] n (%)
Test Result 1—negative					
Negative	27 (81.8)	27 (93.1)	14 (48.3)	26 (86.7)	25 (89.3)
Positive	2 (6.1)	1 (3.4)	2 (6.9)	2 (6.7)	–
Invalid	3 (9.1)	1(3.4)	13 (44.8)	2 (6.7)	2 (7.1)
Don't know	1 (3.0)	–	–	–	1 (3.6)
Test Result 2—invalid (no lines)					
Negative	6 (18.2)	4 (13.8)	6 (20.7)	3 (10.0)	2 (7.4)
Positive	1 (3.0)	1 (3.4)	3 (10.3)	–	1 (3.7)
Invalid	26 (78.8)	24 (82.8)	20 (69.0)	27 (90.0)	23 (85.2)
Don't know	–	–	–	–	1 (3.7)
Test Result 3—positive					
Negative	4 (12.1)	7 (24.1)	1 (3.4)	1 (3.3)	1 (3.6)
Positive	27 (81.8)	18 (62.1)	26 (89.7)	29 (96.7)	26 (92.9)
Invalid	2 (6.1)	4 (13.8)	2 (6.9)	–	–
Don't know	–	–	–	–	1 (3.6)
Test Result 4—invalid (no control lines)					
Negative	3 (9.1)	6 (20.7)	13 (44.8)	1 (3.3)	8 (28.6)
Positive	7 (21.2)	3 (10.3)	3 (10.3)	2 (6.7)	2 (7.1)
Invalid	22 (66.7)	20 (69.0)	12 (41.4)	27 (90.0)	17 (60.7)
Don't know	1 (3.0)	–	1 (3.4)	–	1 (3.6)
Test Result 5—weak positive					
Negative	17 (51.5)	10 (34.5)[b]	14 (48.3)		21 (75.0)
Positive	15 (45.5)	16 (55.2)[b]	7 (24.1)		2 (7.1)
Invalid	1 (3.0)	3 (10.3)[b]	8 (27.6)		4 (14.3)
Don't know	–	–	–		1 (3.6)
Test Result 6—invalid					
Negative		4 (13.8)			
Positive		2 (6.9)			
Invalid		22 (75.9)			
Don't know		1 (3.4)			
Test Result 7—positive					
Negative		7 (24.1)			
Positive		21 (72.4)			
Invalid		1 (3.4)			
Don't know		–			
Test Result 8—invalid					
Negative		2 (6.9)			
Positive		9 (31.0)			
Invalid		18 (62.1)			
Don't know		–			

[a] One participant did not complete test result interpretation. Also, one additional participant did not give response to Test Result 2

[b] FS1 Test Result 5 is positive rather than weak positive

(n = 45/62). Opening test liquid containers proved problematic. Two participants used a personal item or a generic lancet to cut or puncture sealed containers. Containers with pop-open tops presented challenges when participants attempted to unscrew (n = 6), resulting in five instances of spillage. For the FS2 that delivered test liquid automatically by pushing the test into a sealed cap (thereby eliminating the need to open the container or count drops), 41.4 % (n = 12/29) of participants were observed not pushing the test far enough to break the seal and adequately immerse it in the liquid.

Participants found that the steps involving the test fluid could be complex.

"Oh this watery like substance?…I saw it, but I got confused as to what shall I use it for." (South Africa, Male, FS1)

Timing

Most participants (69.3 %) did not wait for the required time to read results, including a number in the exit interviews who explained that they did not do this as they knew they would not receive test results in this study. Participants who did time their test generally used the clock provided in the testing space; however, a few used their watch or mobile phones (n = 3).

Results Interpretation

Most participants identified negative and strong positive results correctly (Table 4). Negative test results were correctly interpreted 79.9 % of the time (n = 119/149). Strong positive results were correctly interpreted 78.7 % of the time (n = 163/207). Weak positive results with faint test lines were correctly interpreted as positive by only 26.7 % of the participants (n = 24/90). An invalid test is indicated by no lines or lack of a control line appearing on the test. Most invalid results were correctly interpreted (72.7 %, n = 258/355).

During results interpretation for FS1, which has three lines (control, HIV-1, and HIV-2), some participants perceived that a result of one line indicated a negative test result, two lines indicated a positive test result, and three meant the person was very ill.

"Firstly we all know that if there are two lines it means it is positive so here there are two lines and they say it is invalid, for a villager they cannot understand this, it doesn't matter where the lines are but as long as there are two lines to many people that is positive, so they better look into that." (Malawi, Female, FS2)

Another test did not explicitly label the locations of the test and control lines on the cassette (FS2). This test had the lowest frequency of correctly interpreted negative results (48.3 %, n = 14/29) and invalid results without a control line (41.4 %, n = 12/29). The highest percentage of participants (90.8 %) correctly identified the results of the flow-through test, which utilized different symbols for control and test.

Perception of Tests

Exit interview data from questionnaires revealed that more than half of participants using each test rated them as very easy or easy to use (Table 5). Over 80 % of users of all tests felt confident doing the tests and over 70 % felt confident reading results. More than 80 % of participants agreed that they were likely or very likely to use the test again if it were free and also indicated that they were willing to buy the test. There was a general sense of enthusiasm about the prospect of being able to undertake HIVST. Reasons provided in the qualitative interviews included time-savings, reduced fear of HIV status disclosure, and convenience.

"…sometimes there are queues at clinics. And also I am afraid that people will see me in that queue and know that I came for HIV test, whereas at home it is easy and everything you do is your secret. At least you will only have to go to the clinic if you have the disease." (South Africa, Female, FS2)

"It can be beneficial to your family and you as a person… you don't need to plan a journey to go to the hospital. You test yourself there and then." (Kenya, Male, FS4)

Discussion

The findings of this study help us better understand HIV self-test features that may promote or create barriers for correct use and interpretation and identify features that may need further development, contributing to a robust HIV self-test TPP. To our knowledge, this is the first study to compare usability of multiple HIV self-test prototypes (including both oral and fingerstick tests) in unsupervised contexts in the hands of naïve lay users in sub-Saharan Africa. Error rates are higher than were expected and higher than reported in other unsupervised HIVST studies. Several factors could account for the high error rates. The video monitoring methodology may have allowed the study staff to closely observe how participants were using the tests in a setting that more closely represents unsupervised testing. This increased monitoring may have captured errors that may not have manifested in other studies. Additionally, the use of prototype tests that may not yet be fully developed and commercially viable could account for user errors.

Our test instructions were developed and pre-tested with the on-site study staff. As with other rapid testing [19], both rapid tests using oral and fingerstick methods require significant adaptation of test instructions as well as clear labeling and numbering of all components to correspond with test instructions; instructions that specify more clearly how to handle the sample collection device and properly collect a sample may be helpful. In particular, instructions need to be developed and tested for understanding,

Table 5 Participants' perceptions of tests

	O1 (n = 33) n (%)	FS1 (n = 29) n (%)	FS2 (n = 29) n (%)	FS3 (n = 30) n (%)	FS4 (n = 29) n (%)
Test was easy to use					
Very easy/easy	23 (69.7)	16 (55.2)	17 (58.6)	18 (60.0)	19 (65.5)
Neither difficult nor easy	8 (24.2)	3 (10.3)	5 (17.2)	4 (13.3)	7 (24.1)
Very difficult/difficult	2 (6.1)	10 (34.5)	7 (24.1)	8 (26.7)	3 (10.3)
Felt confident doing the test					
Strongly agree/agree	28 (84.8)	26 (89.7)	26 (89.7)	26 (86.7)	26 (89.7)
Neither agree nor disagree	2 (6.1)	1 (3.4)	1 (3.4)	1 (3.3)	2 (6.9)
Strongly disagree/disagree	3 (9.1)	2 (6.9)	2 (6.9)	3 (10.0)	1 (3.4)
Felt confident reading the test					
Strongly agree/agree	24 (72.7)	28 (96.6)	24 (82.8)	29 (96.7)	25 (86.2)
Neither agree nor disagree	4 (12.1)	–	4 (13.8)	1 (3.3)	2 (6.9)
Strongly disagree/disagree	4 (12.1)	1 (3.4)	1 (3.4)	–	2 (6.9)
Missing	1 (3.0)	–	–	–	–
Would use if free					
Very/somewhat likely	31 (93.9)	26 (89.7)	29 (100.0)	28 (93.3)	27 (93.1)
Neutral	–	–	–	1 (3.3)	–
Not likely at all/not likely	2 (6.1)	3 (10.3)	–	1 (3.3)	2 (6.9)
Would buy if real test					
Very/somewhat likely	31 (93.9)	27 (93.1)	27 (93.1)	25 (83.3)	25 (86.2)
Neutral	1 (3.0)	2 (6.9)	1 (3.4)	1 (3.3)	1 (3.4)
Not likely at all/not likely	1 (3.0)	–	1 (3.4)	2 (6.7)	3 (10.3)

including populations with low literacy levels [20, 21]. Use of pictures is essential and, in some settings, additional local languages may be optimal. Other mechanisms for instruction should also be considered, such as community education sessions to demonstrate test steps (before distributing test kits) [8]. Designing a device and its presentation in a manner that is as intuitive and simple as possible to the intended user is critical [10], underscoring the need for piloting with the intended user population.

Our study has shown that fingerstick tests posed challenges due to the large proportion of participants who had not used a lancet before but improved their competency with a second single-use lancet, which demonstrated learning among lancet-naïve users. Therefore, it may be beneficial to include multiple single-use lancets in a test kit. Inclusion of a multiple-use lancet in a test kit may be less desirable from safety and waste-management perspectives. Our study has also shown that oral fluid sample collection posed a number of challenges to naïve users. Therefore, improved instructions with clear language and graphics, increased detail on proper swab collection, less complicated swab collection procedure, and color-coded or ergonomic devices to facilitate proper handling of the device could improve usability of oral fluid tests. Problems with blood sampling using a sample transfer device have been reported previously by health workers [22, 23] as well as lay users who otherwise found the kit easy to use and the instructions easy to understand [6].Due to this, one study site chose to modify the instructions so that participants were instructed to put the sample directly onto the test, and good results were obtained [7, 8]. This study corroborates these results and finds that technologies and test methods that integrate sample volume measurement with sample transfer into the test cassette increase the opportunity for the naïve test user to conduct the steps properly.

Both oral and fingerstick tests were considered acceptable by lay users and feasible to use as part of unsupervised HIVST. These findings are consistent with previous study findings that have found HIV testing to be acceptable among various populations, demonstrating the potential of HIVST as a strategy that may improve access to HIV testing [3].

HIV self-test kits need to minimize scope for error; integrating and/or decreasing the number of test components and test procedures appeared to enhance usability in our study. Even with prototypes that had fewer components, only a few users were able to conduct all steps correctly. While it is difficult to directly compare fingerstick and oral tests, the oral fluid tests had fewer steps and therefore fewer opportunities for error, indicating the importance of integration of steps into one device for HIVST. Quality control features such as sample adequacy,

correct test liquid volume indicator, or sequential procedure checks that provide feedback so the user has the potential to recover from errors can also contribute to a more usable test [13].

Interpretation of results can be problematic for trained testers in provider-assisted testing settings (Kenya National HIV Reference Laboratory, personal communication). We found that a non-trivial portion of users interpreted results incorrectly (29.6 %, n = 237/801). Additionally, weak positives prove difficult to interpret. Frequencies for correct results interpretation were lower than the frequencies of participants feeling confident that they had performed and interpreted the test correctly. This potentially could lead to self-testers believing erroneous results. Tests that use different symbols for test and control lines may mitigate this.

Limitations

Our study had a number of limitations. The sample size was selected in order to obtain in-depth information on the test features included in the study and should now be validated on a larger scale. Potential target populations for HIVST, such as men who have sex with men, were not explicitly captured in our study. Recruitment challenges were significant and bias may be present since immediate test results were not available, potentially excluding participants who were seeking immediate results, and including others who would otherwise not participate in our study. It is not clear how this may have influenced our results. Participants were asked to conduct one test; hence, user opinions in this study may have been limited by lack of exposure to other options. Test packaging, labeling, and instructions were developed to enable participants to use the prototypes for this study and were not necessarily as they would be for the final product. Again, further evaluations are needed, including verification of test performance. The quality of any of the samples collected was not assessed in this study because test results were not generated from the samples. Data have not yet been analyzed for potential cultural or context-specific influences that may occur among countries.

Conclusion

For HIVST programs to be successful, tests and instructions that are specific to the user and the context of use are needed. Manufacturers contend that it is reasonable that tests will fail if the instructions are not followed precisely. However, it is unreasonable to expect inexperienced lay users to properly use a consumer product that is complex, unintuitive, and has insufficient or unclear labeling with instructions targeted above comprehension levels. The high error rates and performance of the features integrated into these prototypes do not currently support their implementation in unsupervised HIVST programs or in settings where a demonstration of the test procedure is not available. Since the ideal HIV self-test is not yet available, we need to consider what factors can contribute to a more usable test. A thorough TPP informed with lessons learned from this study can inform development of the next generation of HIV self-tests. In particular, improved sample collection and transfer methods are needed. Instructions that are clear and easy to understand, integrated test components, fewer steps, and results that are easy to interpret will also contribute to increased test usability. Integrating these improvements in conjunction with further HIVST demonstration studies and education will help advance HIVST to fill the gap in access to universal HIV testing.

References

1. World Health Organization. Report on the first international symposium on self-testing for HIV: the legal, ethical, gender, human rights and public health implications of self-testing scale-up. Geneva: World Health Organization; 2013.
2. Pant Pai N, Sharma J, Shivkumar S, Pillay S, Vadnais C, Joseph L, et al. Supervised and unsupervised self-testing for HIV in high- and low-risk populations: a systematic review. PLoS Med. 2013;10(4):e1001414.
3. Krause J, Subklew-Sehume F, Kenyon C, Colebunders R. Acceptability of HIV self-testing: a systematic literature review. BMC Public Health. 2013;13:735.
4. Gaydos CA, Hsieh YH, Harvey L, Burah A, Won H, Jett-Goheen M, et al. Will patients "opt in" to perform their own rapid HIV test in the emergency department? Ann Emerg Med. 2011;58(1 Suppl 1):S74–8.
5. Ibitoye M, Frasca T, Giguere R, Carballo-Dieguez A. Home testing past, present and future: lessons learned and implications for HIV home tests. AIDS Behav. 2014;18(5):933–49.
6. Lee VJ, Tan SC, Earnest A, Seong PS, Tan HH, Leo YS. User acceptability and feasibility of self-testing with HIV rapid tests. J Acquir Immune Defic Syndr. 2007;45(4):449–53.
7. de la Fuente L, Rosales-Statkus ME, Hoyos J, Pulido J, Santos S, Bravo MJ, et al. Are participants in a street-based HIV testing program able to perform their own rapid test and interpret the results? PLoS One. 2012;7(10):e46555.
8. Choko AT, Desmond N, Webb EL, Chavula K, Napierala-Mavedzenge S, Gaydos CA, et al. The uptake and accuracy of oral kits for HIV self-testing in high HIV prevalence setting: a cross-sectional feasibility study in Blantyre, Malawi. PLoS Med. 2011;8(10):e1001102.
9. United States Food and Drug Administration. Testing for HIV. USFDA website 2012. http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/ApprovedProducts/LicensedProductsBLAs/BloodDonorScreening/InfectiousDisease/ucm080466.htm. Accessed 16 June 2014.
10. Draft Guidance for Industry and Food and Drug Administration Staff. Applying Human Factors and Usability Engineering to Optimize Medical Device Design. USFDA website 2011 June 22.

http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm259748.htm. Accessed 16 June 2014.
11. MaRS Entrepreneur Workbooks Defining your target product profile. MaRS 2011. http://www.marsdd.com/app/uploads/2011/03/Defining-Your-Target-Product-Profile_In-Vitro-Diagnostic-Products.pdf. Accessed 16 June 2014.
12. Mabey D, Peeling RW, Ustianowski A, Perkins MD. Diagnostics for the developing world. Nat Rev Microbiol. 2004;2(3):231–40.
13. Rubin J, Chisnell D. Handbook of usability testing. 2nd ed. Indianapolis: Wiley; 2008.
14. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. Behav Res Methods Instrum Comput. 2003;35(3):379–83.
15. Bahl R, Murphy DJ, Strachan B. Qualitative analysis by interviews and video recordings to establish the components of a skilled low-cavity non-rotational vacuum delivery. BJOG. 2009;116(2):319–26.
16. Schaeffer JH. Videotape: new techniques of observation and analysis in anthropology. In: Hockings P, editor. Principles of visual anthropology. 2nd ed. The Hague: Mouton de Gruyter; 1995. p. 255.
17. Sankar A, Golin C, Simoni JM, Luborsky M, Pearson C. How qualitative methods contribute to understanding combination antiretroviral therapy adherence. J Acquir Immune Defic Syndr. 2006;1(43 Suppl 1):S54–68.
18. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. London: Routeledge; 1996.
19. Gillet P, Scheirlinck A, Stokx J, De WA, Chauque HS, Canhanga OD, et al. Prozone in malaria rapid diagnostics tests: how many cases are missed? Malar J. 2011;10:166.
20. Anderson DA, Crowe SM, Garcia M. Point-of-care testing. Curr HIV/AIDS Rep. 2011;8(1):31–7.
21. Spielberg F, Levine RO, Weaver M. Self-testing for HIV: a new option for HIV prevention? Lancet Infect Dis. 2004;4(10):640–6.
22. Luchavez J, Lintag ME, Coll-Black M, Baik F, Bell D. An assessment of various blood collection and transfer methods used for malaria rapid diagnostic tests. Malar J. 2007;6:149.
23. Strategic Evaluation, Advisory and Development Consulting. Analysis of POCT/VCT Performed at South African Primary Health Care Clinics. Strategic, Evaluation, Advisory & Development (SEAD) Consulting website 2013. www.sead.co.za/downloads/POCT-clinics-2011.pdf. Accessed 16 June 2014.