Fraunhofer
PORTUGAL

FRAUNHOFER CENTER FOR ASSISTIVE INFORMATION AND COMMUNICATION SOLUTIONS – AICOS

REPORT

Android application usability testing
MalariaScope

By: Ana Vasconcelos, Tiago Devezas

Date: March 2014

Fraunhofer Center for Assistive Information and Communication Solutions – AICOS

Associação Fraunhofer Portugal Research

Rua Alfredo Allen 455/461, 4200-135 Porto, PORTUGAL

Table of Contents

1 Summary

System and objectives

MalariaScope (https://www.aicos.fraunhofer.pt/en/our_work/projects/malariascope.html) is a product developed by Fraunhofer Portugal AICOS (FhP) that consists of two components: an optical magnification system and an Android application. These two elements allow the analysis of blood samples and detection of the risk of malaria infection. MalariaScope is intended to be used by people with some medical training from developing countries in multiple contexts.

Method

We evaluated the usability of the Android application in two separate phases. Seven health technicians and researchers (one of them retired) from Portugal participated in the first phase. For the second phase of the evaluation six more health technicians and researchers from the National Health Institute Dr. Ricardo Jorge were recruited.
All of them were asked to complete the same eleven tasks (albeit in a different order on phase 2 due the implemented changes in the meanwhile). Tasks included common use cases such as creating a new patient, adding a new sample, adding views from the gallery and the camera, access the sample risk report, among others.

Results

Phase 1

Performance varied among the seven users that participated in the first phase of the evaluation. Four participants were able to complete all the eleven tasks with less than 10 errors, while three others made 20 or more errors. The average error rate was 12.86 (SD = 9.12). The minimum number of errors was 3 and the maximum was 24. On

average, little assistance was required from the facilitator in most cases (average assist rate was 5.43 (SD = 6.19; min = 0; max = 19). Six out of the seven participants requested 5 or less assistances. There was, however, a participant that needed 19 assistances to complete the tasks. It should be mentioned that this was a particular case: the participant was retired pharmaceutical researcher with 74 years of age and limited experience on touch devices. The global satisfaction measured with the SUS scale was 89.6 (SD = 5.7). This is well above the average 68 score calculated from the analysis of over 5000 users across 500 different evaluations[1], which is considered the threshold for an acceptable usability score.

Phase 2

The majority of the recommendations based on the findings of phase 1 were implemented before phase 2. Results from the six users that participated in the second phase show that the usability of the mobile application increased significantly with the changes introduced. Four participants were able to complete the eleven tasks with 5 or less errors, while the remaining two made less than 10 errors. The average error rate was 4.2 (SD = 3.13; min = 1; max = 9). This means that there was a 68% decrease in the average number of errors from the first to the second phase.

Very little assistance from the facilitator was required. Two participants required no assistances, and the remaining received 4 or less. The average assist rate was 1.5 (SD = 1.64; min = 0; max = 4). The average number of assists fell 72% from the first to the second phase.

[1] http://www.measuringusability.com/sus.php

One metric that increased in the second phase was the average of total deviations: in the first phase the average of total deviations was 0.71 while on the second was 1, representing an increase of 40%. However, we do not believe this to be a negative outcome. Rather, the alterations implemented provided the users alternative ways to navigate the screen of the application that, while not being the most efficient, allowed them to reach their goal.

The global satisfaction measured with the SUS scale was 83.75 (SD = 14.4). There was a slight decrease (7%) regarding the previous phase, but this value still remains well above the average 68 score.

2 Introduction

2.1 System description

MalariaScope is a project composed by two elements: an optical magnification system and a companion Android application. The product's goal is to simplify and make faster the first triage of blood samples potentially infected with Malaria. To achieve this goal, MalariaScope can be used by technical personnel without specialized knowledge in Malaria diagnosis. The user collects and prepares a blood sample of the patient, introducing it in a slot in the optical magnification prototype. Using the companion mobile application, installed in a smartphone that is coupled to the optical magnification prototype, the user can

take pictures of the sample using the smartphone's camera while using a button on the prototype to change the magnified views. The captured views can then be sent to analysis through the mobile app, which returns a report indicating the level of risk of each sample (and the individual views).

The results can then be analyzed by an expert, in order for the correct procedures and medication to be administered.

The product is intended for users in developing countries where Malaria constitutes a serious health problem. The users are expected to have experience working with blood samples and knowledge of how to operate a touch-based smartphone. Literacy is also required.

This usability evaluation report only encompasses the mobile application for Android smartphones, since the optical magnification prototype was not available at the time the tests were conducted.

2.2 Test objectives

This usability evaluation focused only on the MalariaScope mobile application.

The main goal of the tests was to iteratively detect and solve potential usability issues regarding the navigation flow between screens, the comprehensibility of the different actions and concepts and the adequacy of the information architecture (e.g., the relationship between patients, samples and views).

Since the process was iterative, the evaluation was split in two phases. After the first phase of tests, a set of recommendations to improve the general usability of the application was derived from the results and observations. The recommended changes were implemented before the tests in phase 2.

3 Method

3.1 Test facility

The first phase of the evaluation took place at two different settings: Fraunhofer AICOS's offices and a Family Health Unit. The locations were chosen for being the most convenient for the participants and not for simulating the intended context of use, which encompasses both urban and rural settings in developing countries, outdoors or in a laboratory. Due to the difficulty of mimicking outside conditions in a developing country, the best approach is to use an indoor, laboratory-like environment that can potentially be one of the regular contexts of use for the system. Still, some aspects like poor access to networks and electricity that can be part of the context of use are being left out of this evaluation.

The second phase of the evaluation took place at a single setting: the offices of a National Health Institute Dr. Ricardo Jorge in Porto. Six users participated. Similarly to the previous phase, the location was chosen for convenience purposes.

3.2 Equipment

All the participants used the same device to test the MalariaScope application both on the first and second phases: a Samsung Galaxy Note with a 5.3" WXGA (1280 x 800) screen, in standard color mode and automatic brightness.

3.3 Procedure

The sequence of events from greeting the participants until their dismissal was the following:

- Participants were greeted by the facilitators.

- They were given an informed consent to sign (available at section 0 of the Annex)
- One of the facilitators read aloud a script with information about the MalariaScope project and the two prototypes being tested (in this test only the mobile app was tested, so only information about it was provided) (available at section **Error! Reference s ource not found.** of the Annex)
- A short background questionnaire (age, profession, smarthphone/OS ownership and experience) was administered.
- Information about the test was read aloud from the script by one of the facilitators. It was provided a description of the concepts involved and permission to record the test was asked. Participants were asked to try and complete the task as if the facilitator was not present, but to ask for help if they felt they were stuck or did not understand the task description. The facilitator also tried to elicit some comments from participants during task execution to understand their though process.
- Participants were read the task instructions sequentially.
- Participants were given the SUS questionnaire (available at section 6.3 of the Annex).
- Questions and comments about the product were solicited.
- The facilitators thanked and dismissed the participants.

Participants were not compensated.

3.4 Tasks

Participants were asked to perform the following tasks:

1. Create a patient ("Joana", "30", "Female")
2. Edit the patient to add "Silva" to the name
3. Create a new sample for this patient with a picture from the gallery, and analyze it

4. Add three views from camera: take four pictures and before analyzing delete the first picture taken, choose to analyze them later
5. Change Patient and choose the patient with the name "Rui Almeida"
6. Open Sample 1
7. Change to Sample 2
8. View sample report
9. Analyze a view that is not analyzed
10. Open the view with the worst results and delete it
11. Delete sample

These tasks were devised after an expert evaluation of the application's usability and a conversation between the Human-Computer Interaction team and with the application developers (who worked together with experts from a National Health Institute while developing the app). In addition to representing the most frequent actions a user would have to execute in order to effectively use the application, they also allowed us to test potential issues with regard to the navigation flow, the comprehensibility of the different actions and concepts and the adequacy of the information architecture.

3.5 Usability metrics

3.5.1 Effectiveness

Effectiveness relates the goals of using the product to the accuracy and completeness with which these goals can be achieved. The following metrics were collected:

- **Completion rate**: The proposed task was considered completed when the user finished typing the given text on the smartphone.
- **Errors**: An error was counted every time the participant performed an action that did not contribute to task completion.

- **Deviations:** Deviations are defined as alternative flows to the completion of the task, that, while not being the ideal flow, still enable the participant to achieve task completion.
- **Assists**: An assist was considered every time the participant requested the assistance of the facilitator in order to perform the task. If the assistance was required because the task was not well explained it was not considered.

3.5.2 Efficiency

The efficiency metric was not considered for this evaluation due to the lack of metrics for comparison. This metric would be of better use on the evaluation of a following iteration. Nevertheless, the facilitators took into account the overall time spent using the application and the time spent in each task, and clear deviations from an appropriate time were noted.

3.5.3 Satisfaction

Satisfaction describes a user's subjective response when using the product. Satisfaction was measured through the SUS questionnaire administered after the test. Additionally, all the comments made by the participants during the test as well as in the informal conversation that followed were registered.

4 Results

4.1 Participants

4.1.1 Phase 1

In the first phase of the evaluation we recruited seven participants, including doctors, nurses and other professionals with laboratory work experience. The average participant age was 39.29 (SD = 17.53), one male and six females. Minimum age was 26 and maximum age 74. While cultural differences may occur regarding the target users (health workers in African countries), the participants are somewhat representative, since they, such as the intended audience, have a certain degree of knowledge on dealing with and analyzing blood samples.

Participant	Gender	Age	Education	Occupation/Role	Smartphone experience/OS used
P1	M	28	Higher	Medical doctor	iOS
P2	F	30	Higher	Medical doctor	Android
P3	F	29	Higher	Nurse	Feature phone daily / Android occasionally
P4	F	26	Higher	Diagnosis technician	iOS (previously Android)
P5	F	37	Higher	Nurse	Touch device with physical keyboard
P6	F	74	Higher	Retired pharmaceutical researcher	Feature phone
P7	F	51	Higher	Nursing teacher	Touch device (1 month experience)

Table 1. Participants (phase 1)

4.1.2 Phase 2

In the second phase of the evaluation we recruited six participants, which included researchers and lab technicians with laboratory work experience. The average participant age was 43 (SD = 8.75), two male and three female. Minimum age was 31 and maximum 52. While cultural differences may occur regarding the target users (health workers in African countries), the participants are somewhat representative, since they, such as the intended audience, have a certain degree of knowledge on dealing and analyzing blood samples.

Participant	Gender	Age	Education	Occupation/Role	Smartphone experience/OS used
P8	M	46	Higher	Researcher	Feature phone
P9	F	52	Higher	Lab technician	Feature phone
P10	F	37	Higher	Lab technician	Feature phone
P11	F	49	Higher	Pharmaceutical	Android
P12	M	31	Higher	Chemistry Grant student	Feature phone
P12	M	59	Higher	Researcher	Android

Table 2. Participants (phase 2)

4.2 Performance results

All the testing sessions were recorded using a video camera. The recorded videos were imported and coded using the Observer XT software. The coding scheme designed for this study had two groups. The Metrics group included Errors (point event), Deviations (point event) and Assists (state event). The Comments groups included Note (Point event), Task (state event) and Quote (state event).

The data were then analyzed in Observer XT, allowing the researchers to quantify the number of errors, deviations and assists for each

participant, as well as register the task (beginning and end), quotes (end and beginning) and additional notes.

An error was marked every time the participant performed an action that did not contribute to task completion. A deviation was marked every time the participant took an alternative path (i.e., different from the predefined ideal flow) to complete the task. An assistance was considered every time the facilitator intervened (at the participant's request or due to the facilitator's judgment) to help the participant in the completion of the task.

4.2.1 Phase 1

Participant	Unassisted Task Effectiveness [(%)Complete]	Assisted Task Effectiveness [(%)Complete]	Errors	Deviations	Assists
P1	73%	100%	24	0	3
P2	100%	100%	3	4	0
P3	82%	100%	23	0	4
P4	55%	100%	6	0	5
P5	82%	100%	9	0	3
P6	9%	100%	5	0	19
P7	64%	100%	20	1	4
Mean	**66%**	**100%**	**12,86**	**0,71**	**5,43**
St. Deviation	**29%**	**0%**	**9,12**	**1,50**	**6,19**
Min	**9%**	**100%**	**3**	**0**	**0**
Max	**100%**	**100%**	**24**	**4**	**19**

Table 3. Performance results by participant (phase 1)

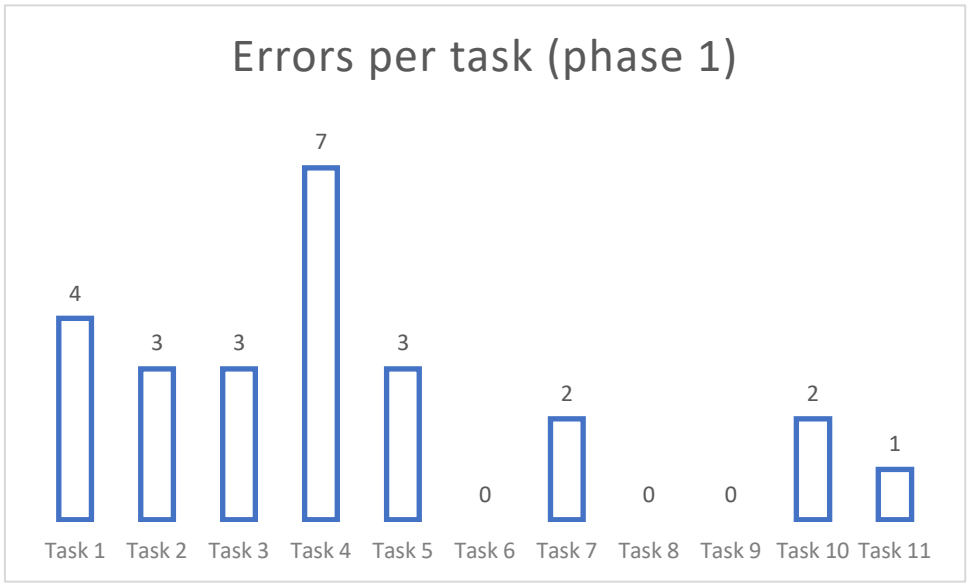

Figure 1. Errors per task (phase 1)

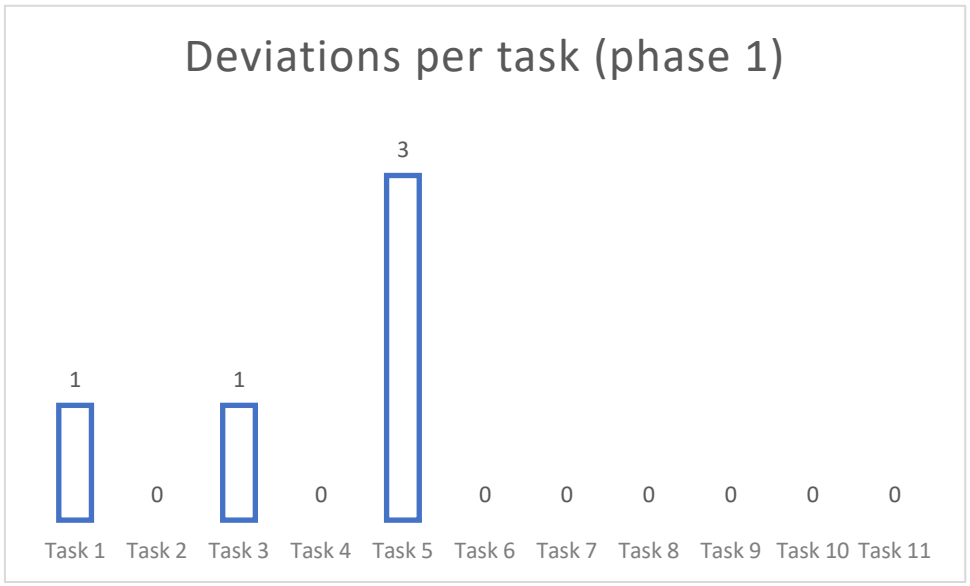

Figure 2. Deviations per task (phase 1)

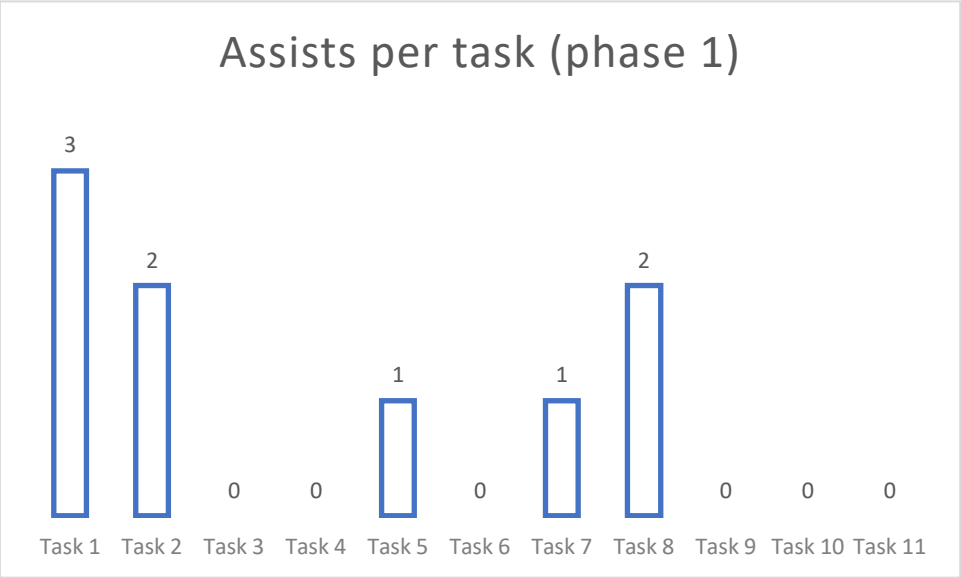

Figure 3. Assists per task (phase 1)

Performance varied among the seven users that participated in the first phase of the evaluation. Four participants were able to complete all the eleven tasks with less than 10 errors, while three others made 20 or more errors. The average error rate was 12.86 (SD = 9.12). The minimum number of errors was 3 and the maximum was 24.
On average, little assistance was required from the facilitator is most cases (average assist rate was 5.43 (SD = 6.19; min = 0; max = 19). Six out of the seven participants requested 5 or less assistances. There was, however, a participant that needed 19 assistances to complete the tasks. It should be mentioned that this was a particular case: the participant was retired pharmaceutical researcher with 74 years of age and limited experience on touch devices.

4.2.2 Phase 2

Participant	Unassisted Task Effectiveness [(%)Complete]	Assisted Task Effectiveness [(%)Complete]	Errors	Deviations	Assists
P8	82%	100%	5	1	1
P9	73%	100%	1	1	4
P10	91%	100%	1	2	1
P11	73%	100%	9	1	3
P12	100%	100%	3	0	0
P13	100%	100%	6	1	0
Mean	**86%**	**100%**	**4,17**	**1**	**1,5**
St. Deviation	**13%**	**0**	**3,13**	**0,63**	**1,64**
Min	**73%**	**100%**	**1**	**0**	**0**
Max	**100%**	**100%**	**9**	**2**	**4**

Table 4. Performance results (phase 2)

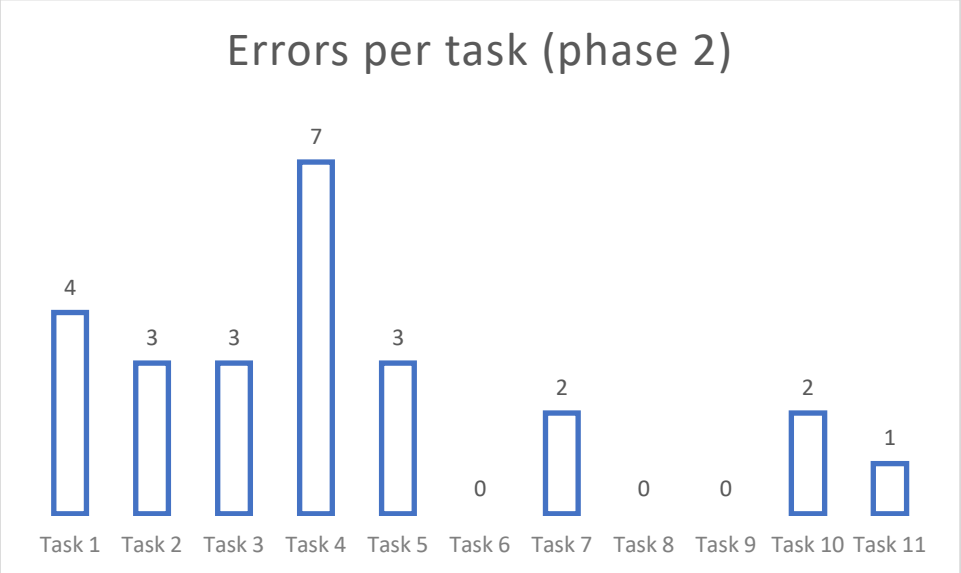

Figure 4. Errors per task (phase 2)

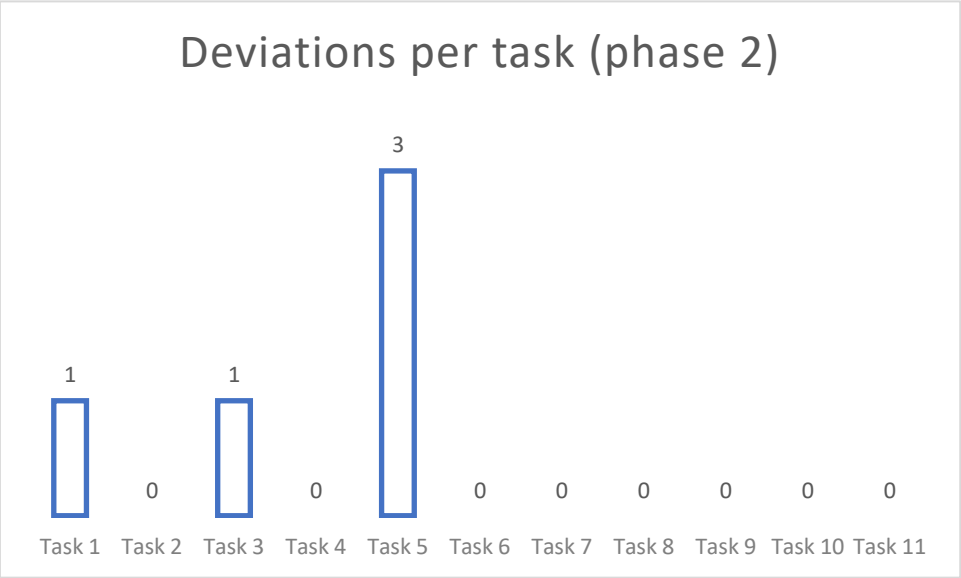

Figure 5. Deviations per task (phase 2)

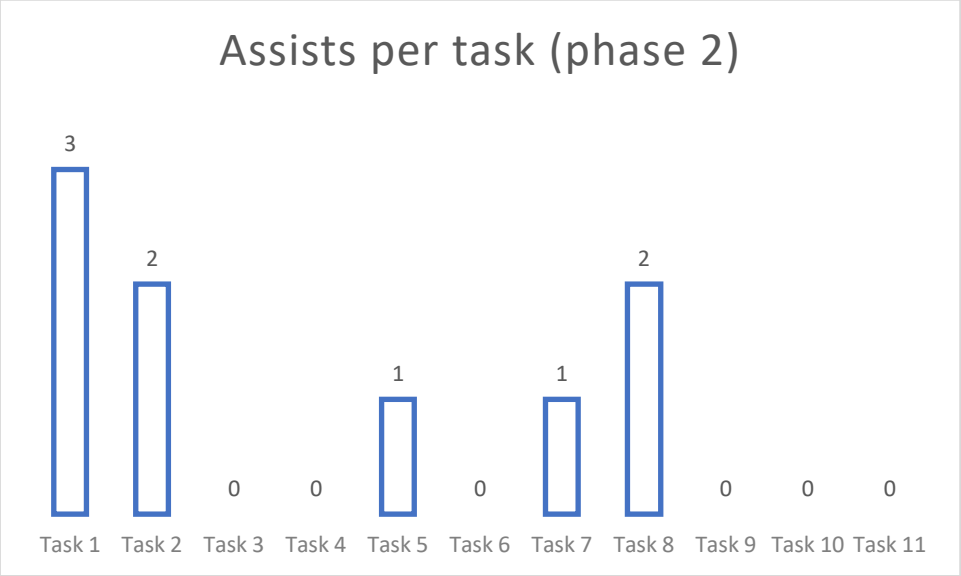

Figure 6. Assists per task (phase 2)

The majority of the recommendations based on the findings of phase 1 (cf. section 5.1) were implemented before phase 2. Results from the six users that participated in the second phase show that the usability of the mobile application increased significantly with the changes introduced. Four participants were able to complete the eleven tasks with 5 or less

errors, while the remaining two made less than 10 errors. The average error rate was 4.2 (SD = 3.13; min = 1; max = 9). This means that there was a 68% decrease in the average number of errors from the first to the second phase.

Very little assistance from the facilitator was required. Two participants required no assistances, and the remaining received four or less. The average assist rate was 1.5 (SD = 1.64; min = 0; max = 4). The average number of assists fell 72% from the first to the second phase.

One metric that increased in the second phase was the average of total deviations: in the first phase the average of total deviations was 0.71 while on the second was 1, representing an increase of 40%. However, we do not believe this to be a negative outcome. Rather, the alterations implemented provided the users alternative ways to navigate the screen of the application that, while not being the most efficient, allowed them to reach their goal.

It should also be mentioned that there were some potential constraints regarding the sample from phase 2 that give the results achieved greater relevance. First, four of the six participants reported using a feature phone daily and only one of these had some previous experience with touch devices. Additionally, two of the participants (P8 and P9) displayed a clear lack of interest and motivation for participating in the test, which might have influenced the results.

4.2.3 Comparison of performance results

Metric	Phase 1	Phase 2	Variation (%)
Average errors	12,86	4,17	-68%
Average deviations	0,71	1	40%
Average assists	5,43	1,5	-72%
Unassisted task completion rate	66%	86%	

Table 5. Comparison of performance results

The percent variation between the measured metrics in the first and second phase displays a clear increase in terms of effectiveness.
In terms of average number of total errors, there was a 68% decrease in phase 2 when compared to phase 1. While in phase 1 the average number of errors was 12.86, it fell to 4.17 in the second phase.
The number of assists also decreased significantly, from 5.43 to 1.50, less 72%. As for the average deviation, there was an increase of 40% in phase 2 (from 0.71 to 1). Based on or observations, we do not believe this to be a negative outcome. Rather, it mainly shows that the alterations implemented provided users with alternative ways to navigate the screens of the application that, while not being the most efficient, allowed them to reach their goal.

4.3 Satisfaction results

4.3.1 Phase 1 – SUS Scores

Participant	SUS Score (out of 100)	SUS Learnability Sub-score (out of 20)	SUS Usability Sub-score (out of 80)
P1	87,5	20	67,5
P2	90	17,5	72,5
P3	92,5	20	72,5
P4	100	20	80
P5	82,5	15	67,5
P6	85	10	75
P7	90	20	70
Mean	**89,6**	**17,5**	**70,8**
St. Deviation	**5,7**	**3,8**	**2,9**
Min	**82,5**	**10**	**67,5**
Max	**100**	**20**	**80**

Table 6. SUS Scores and sub-scores (phase 1)

4.3.2 Phase 2 – SUS Results

Participant	SUS Score (out of 100)	SUS Learnability Sub-score (out of 20)	SUS Usability Sub-score (out of 80)
P8	70	10	60
P9	67,5	15	52,5
P10	100	20	80
P11	100	20	80
P12	77,5	15	62,5
P13	87,5	17,5	70
Mean	**83,75**	**16,25**	**67,5**
St. Deviation	**14,4**	**3,8**	**11,2**
Min	**67,5**	**10**	**52,5**
Max	**100**	**20**	**80**

Table 7. SUS Scores and sub-scores (phase 2)

4.3.3 Comparison of SUS scores

Metric	Phase 1	Phase 2	Variation (%)
Average SUS Score	89,64	83,75	-7%

Table 8. Comparison of satisfaction results

On the second phase, the average SUS score decreased from 89.64 to 83.75. This represents a slight decrease (7%) regarding the previous phase, but this value still remains well above the average. This could be attributed to the lack of experience with smartphones of phase 2 participants, together with the lack of motivation demonstrated by two of the participants.

4.3.4 Phase 1 - Observations and participants' comments

Domain	Observations and comments
Samples/Views concepts	P1 and P3 expressed that the concepts and the relationship between samples and views caused initial confusion. P1 thought that he had deleted a sample when he had actually deleted a view.
Icons	The "Plus" button, to create samples/add views from the gallery/camera, was not intuitive to P1. He said that there should be an indication that it created a new sample. He also said the other icons were easily identifiable (P6 said the same thing about the icons on the right side).
Navigation	The list button to change samples caused problems to almost all participants. P2 used the device's back button instead of the ones on the header bar to go back because she did not notice that they were buttons (she thought it was the logo of the project). She eventually clicked the list button but through an elimination process. She also complained that she pressed the device's back button several times (thus exiting the app) because the system was slow to respond to the presses, and that a confirmation should appear before exiting the app. P2 was not expecting that, when clicking the back button inside a sample, the system took her to the patient's list. P3 had difficulty changing samples because she did not notice the list button. P5 did not understand that the list button was to change samples. She said that the only problems she had were related with the list/back button. P2, P5 and P6 said that the fact that the app logo and the back/list icons were grouped were confusing. When asked, P7 forgot what the list button was for, even though she had used it before. P5 and P7 said that the back icon was very small and not very visible when compared with the icons at the right side.
Color codes	P1 said that he did not understand the colors as an indicator of risk very well. He asked if there was a symbol to indicate that the view had not been analyzed and did not notice that the gray stripe was used to that end. However, the task was done when the color bug was present. After being shown the view list without the color bug, he said that it made sense. P2 asked what the view with worst results meant. After being told that it meant the most negative results to the patient, she clicked the red striped view. P3 identified the unanalyzed view by reading the text ("Already analysed: yes or no") one sample at a time and did not notice the gray stripe. She also did not look at the colors to identify the level of risk. However, she said that after knowing what the colors meant, they might be useful. P4 found the unanalyzed view through the text "Unavailable". She also mentioned the text that said "Already analysed: Yes or no". When told about the grey stripe, she said that she thought they referred to something else. She also referred to the red color as pink. P6, after reading the text of the views and looking at the colors, understood and inferred their meaning. She said that, after knowing their meaning, the colors would be useful for a quick identification.
General considerations	P1 said that he felt that he could do most of the tasks easily because he was used to smartphones. P3 said that the sample and view lists were clear. P4 said that her only trouble was with the gray stripe to identify the unanalyzed view. P6 and P7 said that the app was simple and intuitive but that some initial training (using it two or three times) might be needed.

Table 9. Observations and participants' comments (phase 1)

4.3.5 Phase 2 - Observations and participants' comments

Domain	Observations and comments
Samples/Views concepts	N.A
Icons	P9 commented that she did not understand the meaning of the three dots (more options) icon.
Navigation	N.A.
Color codes	P10 read the text to identify the view with the highest level of risk, rather than the color. She said that while the text indicates an objective measure (Few, Some, Plenty), the colors were subject to interpretation. In this regard, she commented that the colors might be labeled to help interpretation and that they should never be used as the only way to identify the level of risk. P11, on the other hand, identified the view with the highest risk through the red color. She explained that she associated each of the colors (green, yellow and red) to different levels of risk, with red being the highest.
General considerations	P10 commented that the only difficulty she faced was in setting the year using the date picker and attributed that fact to her lack of experience with smartphones. After completing the test, P10 told that she thought that the app was very intuitive even for people not used to smartphones. P11 commented that she made two errors because she was not paying proper attention to the task she was executing.

Table 10. Observations and participants' comments (phase 2)

5 Recomendations

The sources of the following recommendations are the results from the tests, the facilitators' notes and observations during the sessions, the user comments and the analysis of the videos coded using Observer XT.

5.1 Phase 1

The mockups illustrating some of the recommendations made in this section can be found in section 6.4 of the Annex.

Redesign the information architecture and hierarchy.

The application should be redesigned to help users understand the relationships between patients, samples and views. Tasks that required creating a new sample or adding views to a sample were the most troublesome to the users. Some of them specifically mentioned that the concepts and relationship between patients, samples and views caused initial confusion. Also, several assists were needed to clarify these concepts to some of the participants.

Center the screen's title, remove the app's logo and make the "Back" icon more visible.

Several users complained that the "Back" button was very small and not visible. This resulted in several errors in which the participants tried pressing the text of the screen title or the patient's info box rather than the "Back" button. The proximity between the app's logo and the "Back" button was also a source of confusion, leading the users to think that it did not perform any action.

Remove the "List" button and replace it with a "Back" button.

The "List" (navigation drawer) icon used to change samples, in addition to having the same visibility problems than the "Back" icon, also

confused some of the users. They did not understand the meaning of the icon, and some, despite having clicked there previously, did not go there immediately when asked to change samples. There were also users who clicked the "List" button when trying to change patient. Since at this stage they had used the "Back" button, they were presumably expecting to go back to the patients' list. By replacing the "List" icon with a (redesigned) "Back" button in the proposed information architecture, users will have the freedom to navigate back to the samples list as well as to the patients list.

Remove the gray stripe from views not yet analyzed.

Several users specifically referred that they did not use or associate the gray stripe with unanalyzed views and that they identified them through the text. Moreover, the color stripes should only be used to indicate the level of risk. Since no risk assessment has been made, removing the gray stripe helps to reinforce visually this separation between analyzed and unanalyzed views.

Review "Back" button behavior.

Pressing "Back" inside a Sample should return the user to the Sample list (as of right now, it sends the user to the Patients list);
Pressing "Back" inside the image list screen should return the user to the Sample page, not to the Camera;
Sometimes the app is slow to respond to "Back" presses. A dialog asking for exit confirmation should be displayed when the user presses the device's "Back" button sufficient times to exit the app.

Other recommendations.

Patients:

Create New: Use date of birth rather than age (so that the patient's age is updated automatically); Take the user to the Patient profile after saving.

Patient's profile:

Include the risk assessment in the patient's info box; Use alphabetical ordering on the patients list.

Samples:

Create New: New screen for creating samples (fields: Sample ID, Date); take the user to the Sample screen after saving.

Add an icon that reinforces the concept of sample (also helping distinguish between samples and views);

Add a "Delete" option to the dialog of the samples not yet analyzed;

Use reverse chronological order (most recent on top) in the Samples list and in the Views list;

Remove the "Exit" option from the "Options" menu.

Sample/view report:

Add the Patient's data to the screen.

5.2 Phase 2

Redesign and fix the behavior of the date picker.

Task 1 had the highest number of assistances, and all of them were related to the date picker. The results and our observations indicate that the date picker currently used is not intuitive to novice users. One user commented that the only difficulty she had during the test was setting the date using the date picker and attributed this to her lack of experience with touch devices. Many users tried to set the date manually because they did not know they could swipe to change the values. Additionally, there were some issues with manually setting the date. Thus, we recommend implementing the Android date picker (from a previous version of the OS) displayed in **Error! Reference source not found.**. Additional recommended fixes include setting the picker to the date set previously by the user when he/she edits it (not to today's date) and using a date picker in the date field of the Edit Patient screen.

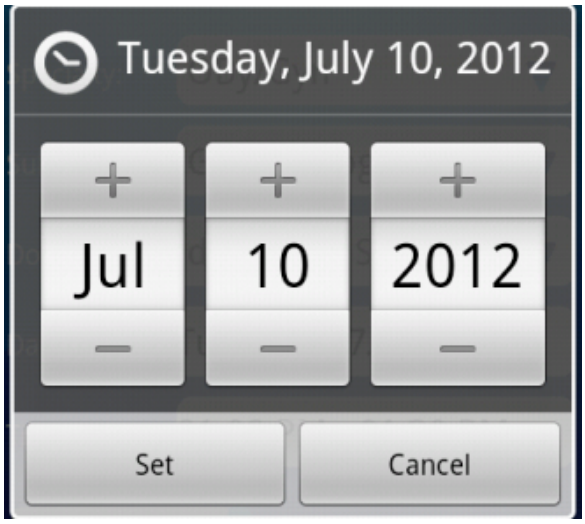

Figure 7. Recommended Android date picker

Other recommendations.

New Patient/Sample screen:

A confirmation dialog should appear when the user presses "Back", alerting him/her that all the data entered will be lost.

Captured Views screen:

When the user presses Back, replace the word "View" by "Sample" in the dialog's title and the second line of the body.

Edit Sample screen:

Similarly to Patients, Samples should have an Edit screen.

6 ANNEX

6.1 Instructions to participants

6.1.1 Introduction

Regarding the project

Malaria Scope is a project being developed under the scope of the ICT4D competence center that targets solutions for developing countries. Malaria is one of the most severe public health problems worldwide, being the leading cause of death disease in many developing countries. Malaria is conventionally diagnosed by microscopic examination, and this is the most widely accepted method by the medical community. However, the manual microscopy examination it's an exhaustive and time consuming activity, which requires considerable expertise and training of the healthcare workers. The proposed mobile-based system could work as a first triage framework for isolated laboratories, where a technician with no special skills in terms of malaria diagnosis collects blood from a patient, prepares the blood smear and uses the system to analyze the blood sample and shares results in order to provide the correct medication.

Regarding the prototypes

The system is composed of two prototypes: an optical magnification specifically designed for mobile devices and an android application that allows the user to capture, store and send analysis result. The user collects and prepares a blood smear of the patient and introduces it here [show correct place on the prototype]. On the application the user chooses or creates a new patient and creates a new sample for the current blood smear. Then, it uses the smartphone to capture different views that can be changed with this button on the prototype [show button]. This means that each sample will include several views. Do you have any question regarding the procedure?

Regarding the test

First of all I would like to ask for your permission to film the test under the compromise that these images are for research purposes only and will not be shown to anyone other than the researchers working on the results of this test. I would also like to ask you to read and please sign this informed consent form.

Our goal now is to evaluate the usability of this product and for that we will ask you to perform some tasks using the application and the optical magnification prototype. I will explain a task, you can ask me anything you don't understand about it and then you can try to accomplish the task. Try to do it as if I was not here but if you feel that you are stuck you can ask me for assistance. You can also voice your opinions regarding any aspect of the prototypes. Remember that we are testing the application and not the user, and that there is no right or wrong way to perform a task. Also, we are looking for both good or bad feedbacks so don't refrain from expressing a bad opinion or point out any errors that you may encounter. They are expected and we appreciate it if you let us know.

Regarding the application

This application allows creating patients. Different samples can be created for each patient (each one corresponding to a blood sample). Each sample can have several views (which are photographs of the sample captured on different angles).

6.1.2 After the test

Do you have any questions or comments? Thank you very much for your participation in this test, your opinion is very valuable to us.

6.2 Informed consent (Portuguese)

A *Associação Fraunhofer Portugal Research* faz trabalho de investigação destinado a encontrar soluções focadas na população sénior ou na população de países em desenvolvimento. O projecto Malaria Scope insere-se nesta última categoria, sendo um projecto que visa a criação de um protótipo de diagnóstico da Malária, a ser utilizado pela população de países em desenvolvimento.

A informação recolhida durante o teste realizado está relacionada com a usabilidade da aplicação e protótipo apresentados assim como alguns dados sociodemográficos. Esta informação será recolhida através da observação da interacção, dados de vídeo e som, assim como um questionário e entrevista.

Estes dados são depois usados para criar soluções mais fáceis e eficazes que permitam melhorar aspectos do sistema.

Gostaríamos de contar com a sua participação. A participação não envolve qualquer prejuízo ou dano material e não haverá lugar a qualquer pagamento. Os dados recolhidos são confidenciais. A *Associação Fraunhofer Portugal Research* tomará todas as medidas necessárias à salvaguarda e protecção dos dados recolhidos por forma a evitar que venham a ser acedidos por terceiros não autorizados.

A sua participação é voluntária, podendo em qualquer altura cessá-la sem qualquer tipo de consequência.

Agradecemos muito o seu contributo, fundamental para a nossa investigação!

<u>O participante</u>:

Declaro ter lido e compreendido este documento, bem como as informações verbais fornecidas e aceito participar nesta investigação. Permito a utilização dos dados que forneço de forma voluntária, confiando em que apenas serão utilizados para investigação e com as garantias de confidencialidade e anonimato que me são dadas pelo investigador. Autorizo a comunicação de dados de forma anónima a outras entidades que estabeleçam parceria com a Associação Fraunhofer Portugal Research para fins académicos e de investigação científica.

Nome: _____

Assinatura: _____. Data ___ / ___ / _____

<u>Investigador responsável pelo projecto "Malaria Scope"</u>:

Nome:

Telefone:

E-mail:

6.3 SUS Questionnaire (Portuguese)

Usando a escala abaixo, por favor coloque um círculo no número mais próximo da palavra que mais se aproxima aos seus sentimentos acerca do sistema avaliado.

1. Penso que gostaria de usar este sistema frequentemente

Discordo fortemente	1	2	3	4	5	Concordo fortemente

2. Achei o sistema desnecessariamente complexo

Discordo fortemente	1	2	3	4	5	Concordo fortemente

3. Achei o sistema fácil de usar

Discordo fortemente	1	2	3	4	5	Concordo fortemente

4. Penso que precisaria do apoio técnico para conseguir usar o sistema

Discordo fortemente	1	2	3	4	5	Concordo fortemente

5. Achei que as várias funções do sistema estavam bem integradas

Discordo fortemente	1	2	3	4	5	Concordo fortemente

6. Achei que havia demasiadas inconsistências neste sistema

Discordo fortemente	1	2	3	4	5	Concordo fortemente

7. Imagino que a maioria das pessoas consegue aprender a usar este sistema muito rapidamente

Discordo fortemente	1	2	3	4	5	Concordo fortemente

8. Achei o sistema muito incómodo de usar

Discordo fortemente	1	2	3	4	5	Concordo fortemente

9. Senti-me muito confiante ao usar o sistema

Discordo fortemente	1	2	3	4	5	Concordo fortemente

10. Precisei de aprender muitas coisas antes de conseguir começar a usar o sistema

Discordo fortemente	1	2	3	4	5	Concordo fortemente

6.4 Mockups

The following mockups were created after phase 1 to illustrate some of the changes of the system.

Figure 8. Mockups of the application (phase 1)