Bachelor thesis
Information Science

Radboud University

The applicability of Grinbath's EyeGuide Tracker to mobile devices for usability testing

Author:
Jonathan Seesink
s4143590

First supervisor/assessor:
Erik Barendsen
E.Barendsen@cs.ru.nl

Second supervisor:
Sjaak Smetsers
S.Smetsers@cs.ru.nl

March 19, 2015

ABSTRACT

In software engineering, usability testing is a method for getting insights about the usability of a software product. Often this method results in conclusions based on qualitative data. Eye tracking methodology can be used to support these data in a quantitative way.

Grinbath's *EyeGuide Tracker* is an affordable eye tracker suitable for usability testing on computer systems with specific operating systems and screen resolutions. Unfortunately, these systems do not include mobile devices.

In this design & creation study we investigated a possible workaround for this incompatibility. The idea is to capture the screen image of the mobile device on a conventional computer screen using a video camera. The user would then look at the computer screen (allowing the eye tracker to follow the user's gaze) while operating the mobile device. In our study, this basic idea was tested in practice and refined through stepwise improvements, resulting in a candidate setup. Finally, a case study involving four test users was conducted to evaluate the suitability of this setup for proper usability testing.

The study revealed several problems, including delayed output of the video capture, disturbances originating from incoming sunlight and artificial illumination, and multiple attempts to touch on buttons without achieving change in the interface. Although a setup matching predefined minimal criteria was obtained, it was found to be unsuitable for usability testing on mobile devices in our case study.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

ET Eye tracking

UT Usability testing

EGT EyeGuide Tracker

1 INTRODUCTION

In software engineering usability is an important quality aspect. Software products can have a high utility by implementing the right functionality. However, if developers do not care about designing the product in a way it can be used efficiently, effectively and without exceeding cognitive capacities, they miss a necessary precondition for practical acceptability. Of course usability is also relevant for designing apps in the ever-growing field of mobile devices.

Usability testing (in the following UT) is a way to obtain insights about the usability of software. This thesis is about constructing and evaluating a low cost UT method for software applications on mobile devices. During this study mobile devices are defined as the set of tablets and smartphones.

In the following part we describe basic concepts and the technology used. Subsequently the aim of this study is examined.

1.1 BACKGROUND

1.1.1 *Usability & Usability Testing*

Usability is a concept in the field of human-computer interaction. Nowadays people have to accomplish lots of different tasks by using computers with different applications which, on their turn, use different interfaces. It is not feasible to let the user read manuals for every software application before he or she is able to use it. While in the past computer power was very expensive, nowadays the concept of "ubiquitous computing" breaks through more than ever (Weiser 1999). Usable software products do not only satisfy users, but are also able to prevent input errors or wrong interpretations of data. Although usability is important it cannot compensate deficiencies in other quality aspects of software like utility or reliability as underlined in the graphic below (Figure 1) by Nielsen (1994).

In the last decade mobile devices achieved a huge market share in the sector of digital products. They have a diverse target group with lots of different requirements. Usability is also important for software made for these devices. Usability has multiple dimensions. Typical dimensions are learnability, efficiency, memorability and user satisfaction (Nielsen 1994). It is difficult to make valid conclusions about the usability of software because the concept of usability is so heavily complex. This is why UT sessions were conducted by lots of usability experts. There are also inspection methods like cogni-

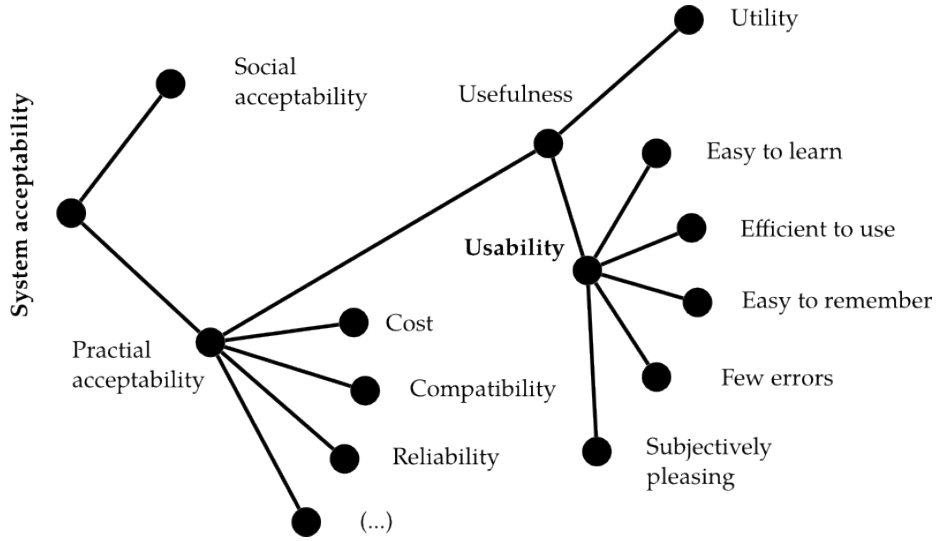


Figure 1: Model of attributes of system acceptability

tive walkthrough or heuristic evaluation to analyse the usability of software products theoretically. In contrast, UT methods involve real users using the computer software. Their interaction with the system is observed by practitioners assisting during the usability sessions, developers and other stakeholders. Often it is combined with think-aloud techniques, questionnaires and/or interviews.

Discount usability engineering aims to conduct UT sessions with less, but effective effort. It is not necessary to have lots of test users or absolutely representative test users to make conclusions about the usability of software (Nielsen 1989).

1.1.2 *Eye Tracking in Usability Testing*

Eye trackers are devices which are capable to register one's gaze and store this data for further analysis. Fixations (intervals in which a person's gaze is fixated on a specific point), smooth pursuits (eye movement while following a moving object) and saccades (rapid eye movements) are assumed to deliver insights in visual attention (Duchowski 2007). Eye tracking (in the following abbreviated with ET) is also used in the field of human-computer-interaction. Applications in this area are for example "Eye typing" (input methodology for handicapped people), Attentive User Interfaces that align their behaviour according to the user's behaviour and many more. These are examples of active ET. Passive ET does not have any influence on a system's behaviour.

Subjective utterances by test users are not always reliable (Siegenthaler, Wurtz & Groner 2010). Therefore reliable quantitative data can increase the value of usability evaluations. Passive ET is able to support UT with quantitative data by tracking the gaze position of the

test user. Practitioners use it as a tool to observe user's search strategies. Lots of research has been conducted approving the usefulness of ET for UT (including Benel, Ottens & Horst 1991, Cooke 2006). The new insights can be used to improve the alignment and selection of visual elements. An example is Çöltekin, Heil, Garlandini & Fabrikant's (2009) research about investigating the usability of a world atlas application. This study proofs that ET gives insights in search strategies of users which could not be obtained with solely traditional usability methods. The researchers used the SUS questionnaire (Brooke n.d.) and combined its results with the ET data.

1.1.3 *Cognitive Load Theory*

The cognitive load theory (CLT) has its origins in Sweller's (1988) article and is in general a learning theory. Task characteristics (task novelty, time pressure, reward systems), subject characteristics (which are mostly constant) and the dynamics between subject and the task/environment (motivation, state of arousal, noise) influence the cognitive load (Paas & Van Merriënboer 1994) and therefore someone's capacity to solve a problem. Figure 2 shows these causal factors schematically.

During UT sessions new interfaces must often be learned and problems have to be solved. Thus UT sessions can also be considered as problem solving tasks. Results of the UT sessions are also influenced by the causal factors we mentioned above. Consequently, the increase of cognitive load by applying a UT method must be as low as possible for valid results.

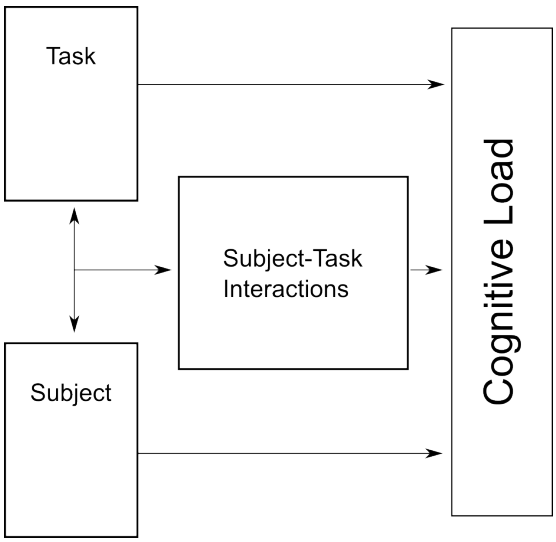


Figure 2: Causal factors of cognitive load

1.2 AIM OF STUDY

Generally speaking, eye trackers are expensive. The *EyeGuide Tracker* (in the following EGT) was developed by Grinbath, a company located in Texas. In 2011 they started to sell this product for an affordable price. Thus it is an interesting product for educational institutions or software projects with low budgets like free and open source software projects. Radboud University purchased some copies of this device for a software development course concerning usability.

The EGT is a head-mounted ET system which works with an infrared light. That light is reflected on the test user's eye. These reflections are recognized by the camera. Signals are sent wirelessly to a transmitter connected via the USB port of a Windows or Macintosh computer system. The software estimates the position of the pupil and relates it to a position on a computer screen based on a calibration procedure at the beginning of each UT session.

Grinbath's EGT system is not open source and therefore software applications or the driver cannot be changed by outsiders without permission.

The EGT system is only suitable for Windows and Macintosh systems. It is optimized for computer screens with a native resolution of 1280x1024 pixels or 1280x800 pixels. Consequently, it is not common to apply it to mobile devices with small screens and operating systems like Android or iOS.

The aim of this study is finding a test setup to use Grinbath's EGT on mobile devices which are run by other operating systems (e.g. Android or Blackberry OS) and which have alternative screen resolutions.

This research involves a design & creation study for finding a candidate setup for using the EGT on mobile devices. The research starts with a basic idea of the UT setup which is refined through stepwise improvements.

The idea is to capture the screen image of the mobile device on a conventional computer screen using a video camera. The camera's capture is displayed and live recorded by software running on a computer with working EGT software installed. The test user sees the display of the mobile device in full screen mode on the computer screen and has to interact with the mobile device by looking solely on the separate computer screen. Figure 3 illustrates this approach.

We choose this basic idea because it ensures that applications on an arbitrary operating system installed on the mobile device can be tested and because no new software has to be created.

The candidate setup is evaluated by a case study in which a UT session is conducted.

Results will add to the body of knowledge. We once saw a live video capture used as mediator for interaction on a conference, but

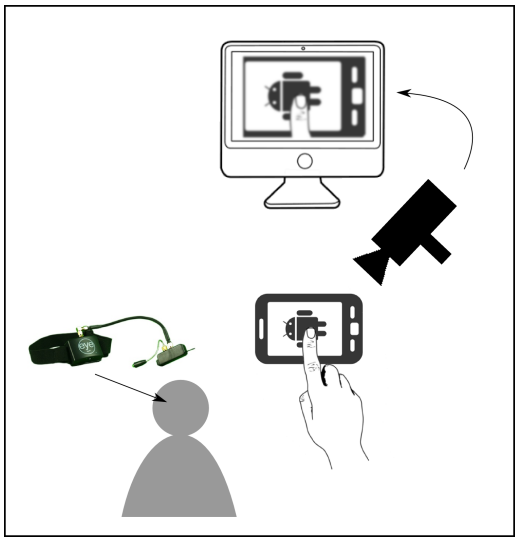


Figure 3: Functional schematic of the basic idea

we could not find any publication concerning this. The setup description obtained by this study can be used for educational purposes. Besides, the results could be valuable for software organizations which have less budget available for improving usability of their software projects like producers of free and open source software.

1.2.1 *Research question*

The research question is as follows: Can Grinbath's EGT be used for UT on mobile devices?

To answer this question it is divided into two subquestions.

1.2.1.1 *What is a workable UT setup with the EyeGuide Tracker for mobile devices?*

A test setup has to be found which can be used for UT sessions. Workable means that interaction with the mobile device is possible while looking solely on the separate computer screen, the mean of calibration values is appropriate and the setup is stable.

1.2.1.2 *Is the candidate setup suitable for UT on mobile device?*

Unquestionably it is necessary to evaluate the candidate UT setup. The evaluation focuses on the EGT's accuracy during UT and the quality of interaction. This involves the question whether the gathered ET data is accurate. Additionally the unnatural phenomenon of setting a camera as mediator for the interaction between test user and mobile device points to the following question: What are consequences of using a camera's screen capture as mediator for interaction?

2 METHOD

The research started with the basic idea as described in Section 1.2. During the design & creation study (according to Oates's (2006) terminology) this basic idea was refined through stepwise improvements, resulting in a candidate setup. Hardware, software, installation instructions and guidelines of the candidate setup were established in a setup description.

2.1 USED HARD- AND SOFTWARE

The following components were used during the study: A computer screen (resolution: 1280x1024, 17 inch), a camcorder (Sony DCR-HC27E), a FireWire card for connecting the camcorder to the PC, Grinbath's EyeGuide Tracker, a Nexus 7 tablet (Asus, model of 2012), a tripod (borrowed from chemistry department of Radboud University) Microsoft Windows 7 OS, VLC (for live video capturing) (version 2.1.4), the driver for the EGT transmitter (version 1.2.4.0, 08.04.2011) and other software by Grinbath for gathering and analysing EGT data.

(a) Camcorder

(b) FireWire card

(c) Grinbath's *EyeGuide Tracker*

(d) Tripod

Figure 4: Used hardware

2.2 STEPWISE IMPROVEMENTS

In total six steps were conducted which will be explained in this section.

2.2.1 *Criteria for adjustments on the setup*

Changes to the setup were made based on the following criteria.

- Visual elements (text and other objects) can be clearly seen through the camera's capture, touches on buttons cause actions linked to these buttons and the test user's finger can be clearly seen on the video capture.
- It must be possible to obtain calibration results in which the determined gaze positions were at least inside the red circle areas of the calibration result images produced by the EGT software. Figure 5 shows an example of such a calibration result image. Otherwise the determined gaze position differs with more than 40 pixels from the real screen position which would make conclusions about the point of attention impossible.
- The mobile device did not move from its initial position during usage.

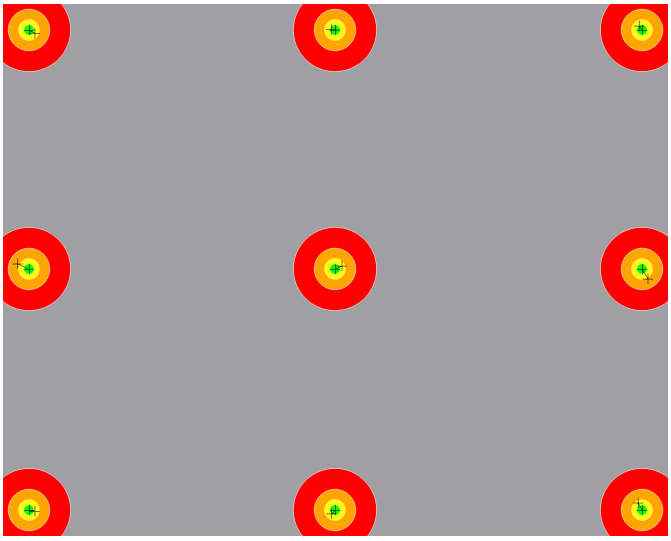


Figure 5: Example of a calibration result image

2.2.2 *Step #01: Video Camera*

The purpose of this step was selecting a proper video camera which captured the mobile device's output. Four camera's were tested. The first two runs were conducted with the webcam *Logitech E2500*, the

third run with the *Microsoft Lifecam Studio HD Webcam* (model Q2F-00015) and the last run with the DV camcorder *Sony DCR-HC27E*.

2.2.3 *Step #02: Application of EGT*

In this step the best positioning of the EGT on the test user's head was determined. In addition it was investigated what environmental illumination delivered best calibration results.

2.2.4 *Step #03: Drawing App*

This step was conducted to determine if apps with small buttons fulfilled the criteria. The drawing app "Autodesk Sketchbook" with small buttons was used during this step.

AUTODESK®SKETCHBOOK® An app for simple drawing of objects. A diversity of different shapes can be directly drawn. The app's version used in this study was 3.0.5 (last change on 27.11.2014). According to Google Play (*Page of Autodesk SketchBook on Google play platform* 2015) it was downloaded more than 5000000 times to the time of writing. It was rated with 4 of 5 points by 16.526 users.

2.2.5 *Step #04: Sloping board*

In this step it was determined whether the mobile device kept its physical position during usage by applying a sloping board as holder for the mobile device. Figure 6 shows a picture of the used sloping board. The string "This is a string for testing!:)" was typed into a text field of an arbitrary app.

2.2.6 *Step #05: Smartphone*

It was assumed that the setup would perform worse for smaller displays. The top of the test user's finger had to be moved with more concentration because movements were more sensitive when the camera had been placed closer to the test user's hand. In this step it was examined whether the setup matched the predefined criteria when applying it to smartphones with smaller displays than the Nexus 7 tablet. A respondent was selected by convenience to input a text string in her own smartphone. This way it was ensured that input errors and input speed did not depend on the test user's unfamiliarity with the device. She was prompted to type in the following string: "This is a string for testing!:)".

Figure 6: Picture of the used sloping board

2.2.7 *Step #06: LED light*

In this case a LED light was placed above the mobile device and it was investigated whether this adjustment improved the interaction.

2.3 CASE STUDY: USABILITY TESTING SESSION

This case involved four young adult students as test users of a UT session. The test users were invited to participate in the usability test (they were selected by convenience), they were put in a comfortable position, the monitor was aligned correctly (in such a way that the test user looked on the center of the computer screen while sitting in a relaxed position). The test users were asked to attach the eye tracker to their head. The camera and LED of the EGT were positioned based on findings in the design & creation study and the calibration process was performed. Then the UT session started. After testing, the test user filled in a questionnaire. The questions about the tested software were based on the "SUS" questionnaire. Beside the SUS, the questionnaires also contained questions about wearing the eye tracker and about the quality of the camera's capture.

The UT sessions were based on the research by Çöltekin et al. (2009) introduced in Section 1.1.2. An online map interface application was chosen for testing. For Android several applications in this category can be found. The app for this case was selected based on the following criteria: functionality to navigate through a map like in the research of Çöltekin et al. (2009) and buttons bigger than the estimated average size of a forefinger.

WORLD ATLAS An interactive online map interface by publisher "mindbeach". According to the Google play platform page (*Page of*

World Atlas on Google play platform 2015) the app was downloaded more than 1000000 times and has been rated with 4.1 of 5 points by 1860 users. The study used version 3.2.1 (last update on 2.10.2014).

We chose a question based task style. The user was asked to find information available somewhere in the software application. This information had to be found principally by the test user without external aid.

Çöltekin et al. (2009) designed two close-ended questions which "require an inference related to an attribute or a location" and one open-ended question that "requires the participant to compare two spatial distributions". We used this as a template to create questions that are applicable for the World Atlas app:

1. How often did Croatia score Olympic Gold?
2. Which country has the highest Death Rate?
3. What are the neighbouring countries of Argentina?

A test user got external help whenever he or she seemed to have difficulties finding the requested information within one minute.

2.3.1 *Definitions of terms used for analysis*

For the analysis of the case study two definitions were formulated. These definitions made it easier to compare results of different UT sessions.

TOUCH ERROR A situation where the test user's forefinger moved in a way that was interpreted by the researcher as attempt to touch an interactive button in the app but without causing any change in the interface.

The number of touch errors according to this definition were counted per task and participant. It is assumed that a touch error would not occur when the device was controlled directly. The results gave an indication for the difference between natural usage of the mobile device and the interaction via the camera's capture and the computer screen as mediators. The count of touch errors was used to quantify the quality of interaction.

CALIBRATION VALUE The sum of numeric values of the nine calibration points. The values were obtained by mapping the position of the pupil as determined by the EGT software to values according to Figure 7. Because there are nine calibration points the lowest calibration value can be 9 (representing highest precision) and the maximal calibration value is 45 (representing

lowest precision). A good calibration result is a necessary precondition for a high overall accuracy of the ET data. If the determined gaze position was very close to the border so that distinction was not clear, the mapping with the higher value was chosen.

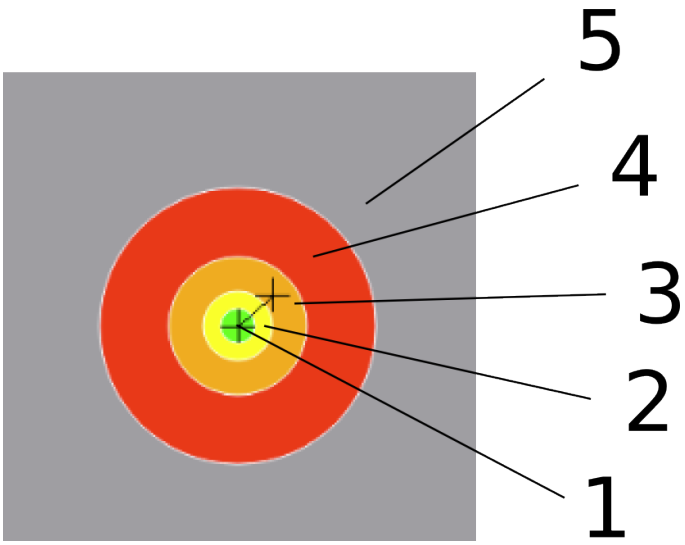


Figure 7: Mapping of the calibration image result to numeric values

3 RESULTS

3.1 STEPWISE IMPROVEMENTS

The interfaces on the mobile devices were not clearly visible when the *Logitech E2500* webcam was used for capturing the output of the mobile device. This was the case in both runs with the Samsung Star II and the Nexus 7 (see Figure 8 and Figure 9).

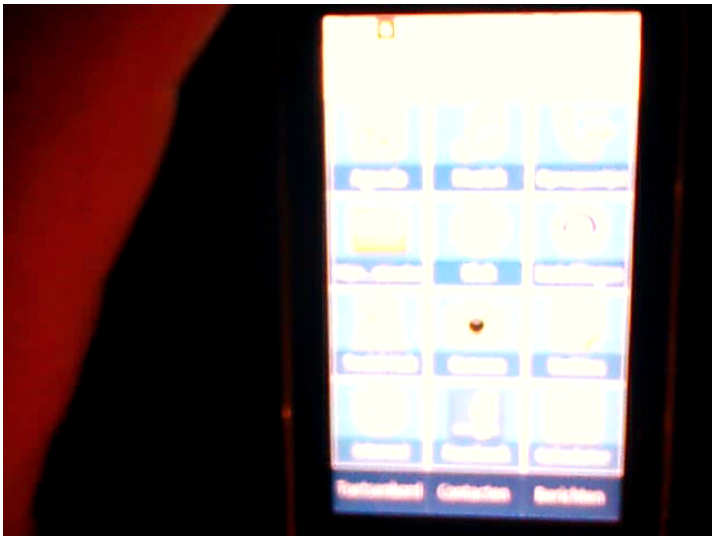


Figure 8: Unclear display during Step #1, Run #01

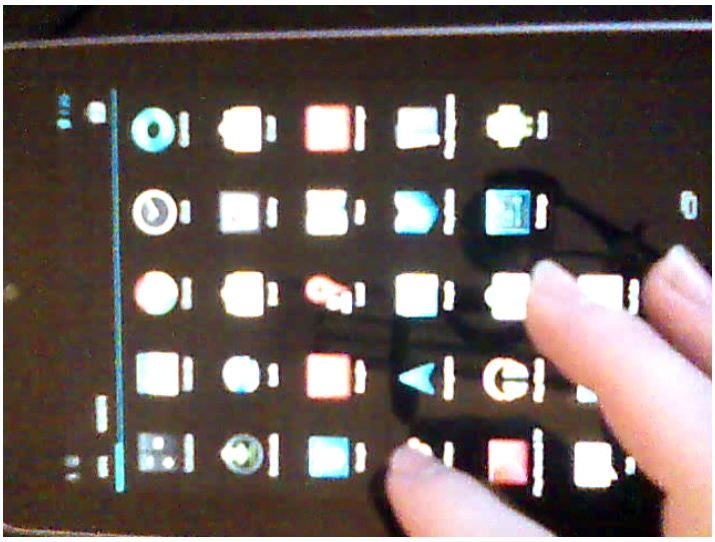


Figure 9: Unclear display during Step #1, Run #02

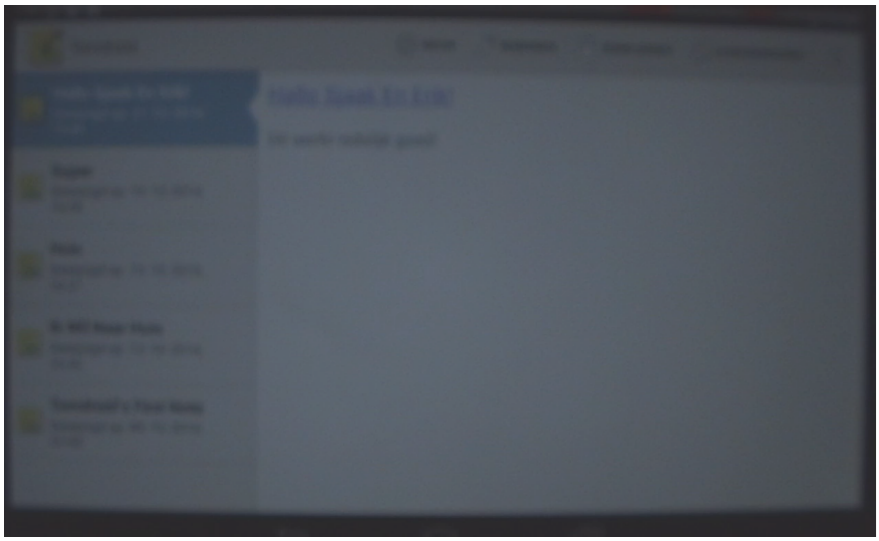


Figure 10: Blur effect during Step #1, Run #03

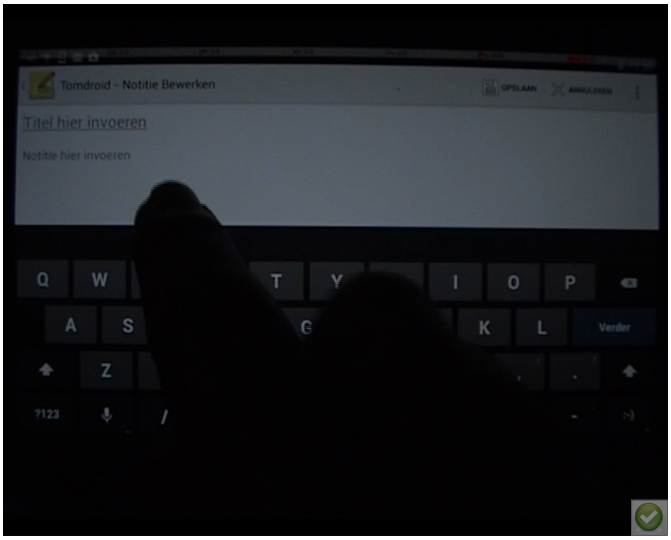
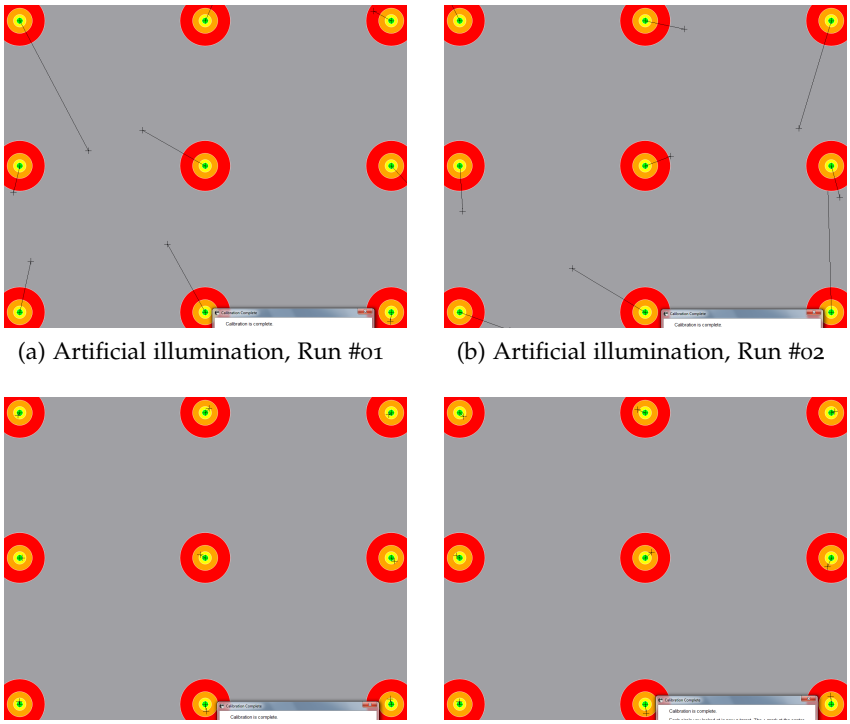


Figure 11: Result with camcorder during Step #01, Run #04

On the one hand it was possible to control the mobile device through the camcorder's recording. On the other hand there appeared to be a significant delay in the recording. This delay had a strong impact on the confidence of input in comparison with interacting naturally with the device. It was assumed that this delay was hardware related. When the whole screen changed to a overall bright condition the camcorder had to autofocus to keep the image sharp enough. This process resulted in a temporary blur effect where visual elements were not clear. An example of such a blur effect is provided in Figure 11. Visual elements were not clearly visible. Hence the cognitive load increased, which had serious impact on the validity of the UT method.

In step #02, the optimal position for the eye tracker was established: The camera was placed close to the test user's nose (inner side of the

eye) and was pointed to the left eye (under right corner) as close as possible.



(a) Artificial illumination, Run #01

(b) Artificial illumination, Run #02

(c) Room without illumination, Run #01

(d) Room without illumination, Run #01

Figure 12: Calibration result images in Step #02

Calibration values were higher when turning off the artificial light in the room. In a room in which sunlight was absent at the time of testing the calibration values were for the cases with artificial light turned on 42 respectively 45. The runs that followed resulted in calibration values of 17 respectively 22 by turning off the artificial light. Figure 12 shows the calibration result images created by the EGT software. The light (both natural and artificial) had huge influence on the calibration quality. Hence the UT setup was changed. We added the following restriction to the candidate setup: Artificial light had to be turned off and the UT setup had to be conducted in rooms protected against sunlight.

In step #03, the ET software could sometimes not locate the gaze position although the calibration result image was one of the best obtained during the design & creation part of the study (calibration result image see Figure 13). The gaze position disappeared or it jerked (with movements faster than any human eye could accomplish). Several times the ET environment registered completely wrong gaze positions. This was most obvious in the sixth minute of the replay recording when the respondent drew an object and the replay

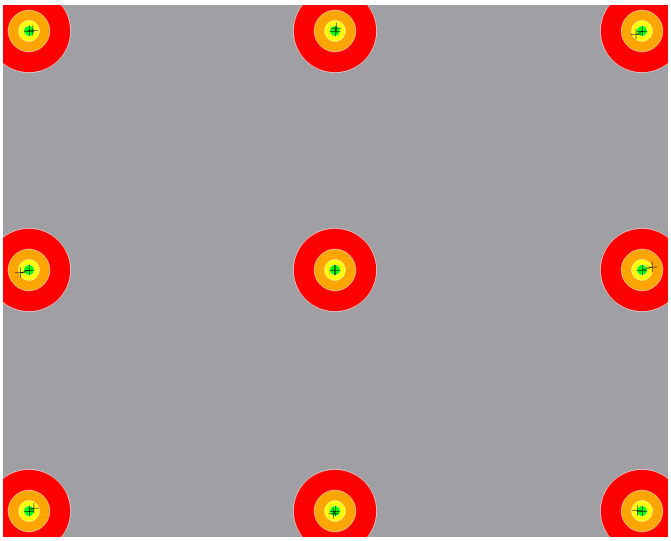


Figure 13: Calibration result image in Step #03

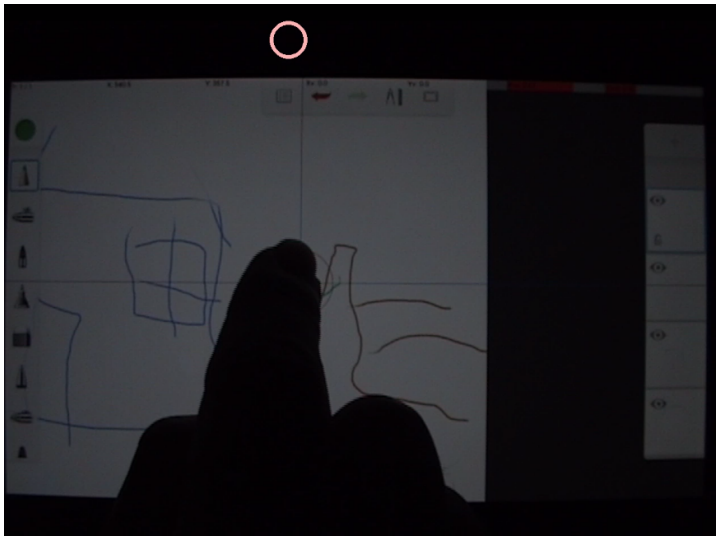


Figure 14: Gaze position not accurate

circle was almost continuously far away from the touch point of the respondent's forefinger. Certainly, this was not a point of attention. Figure 14 shows one frame of the replay video. The red circle represents the determined gaze position by the EGT.

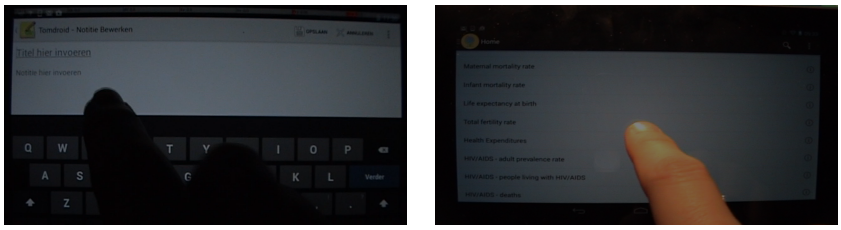
Identifying and touching the small buttons of the app was difficult. The researcher did often not succeed in touching an interactive button. Therefore apps with buttons smaller than the estimated size of the top of a forefinger were not recommended in the candidate setup.

In step #04 the researcher typed successfully the prescribed text string on the virtual keyboard of the mobile device. He needed 114 seconds for accomplishing this task. During the test the mobile device did not move from its initial physical position. As a result, the

sloping board was added to the candidate setup.

The participant struggled a lot with typing in the prescribed text string during step #05. For the string "This is s" the respondent needed 2:23 minutes. Several letters were typed after one touch and the respondent switched unintentionally to a wrong text field. These observations led to an exclusion of mobile devices smaller than 7 inches (size of the Nexus 7 tablet) in the candidate setup.

In step #06 the hand of the participant was clearly visible due to the LED placed above the mobile device. The difference can be seen in Figure 15. As a result the setup enabled more confident input by the user. The adjustment was added to the candidate setup.



(a) Setup without LED (b) Setup with LED

Figure 15: Difference of placing a LED above the mobile device

3.2 CASE STUDY: USABILITY TESTING SESSIONS

Three of the four respondents completed the calibration process with a calibration value less than 25. Respondent #06 got a calibration value of 30. He wore contact lenses during calibration and the UT session.

Some problems occurred during UT sessions.

During task #03 performed by respondents #07 and #08 the camcorder produced a grey picture on the computer screen either because of automatically switching to "demo mode" or because of a notification about lack of battery power when it was forgotten to connect the camcorder to the energy supply.

Respondent #06 closed the app by himself at the beginning of task #03 (touch errors were not counted during the period in which the test application was not active).

The on switch on the LED light was forgotten during the sessions with respondents #05 and #06.

Sometimes the software crashed (task #01, Respondent #06 and during task #03, Respondent #03).

In some cases it was forgotten to turn off software showing an interfering hardware info bar on top of the computer screen. This was

not recognized early enough. This resulted in an extra visual element, which increased the test user's cognitive load.

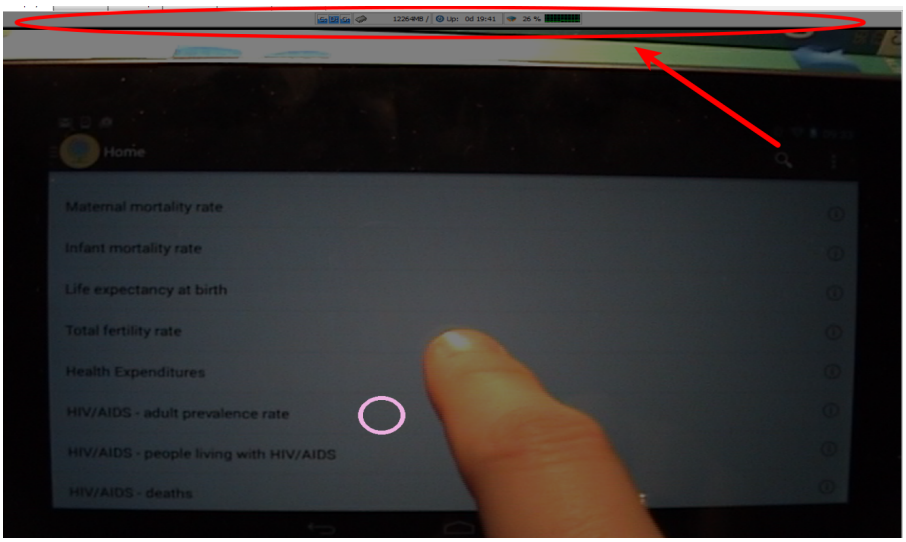


Figure 16: Hardware info bar by external software

In all UT sessions the EGT was placed on the test user's right eye instead of the left eye as determined in step #02 (section 3.1).

Sometimes interaction was difficult to observe because the test user's finger was not visible. This made counting of touch errors difficult.

Furthermore it was observed that some respondents looked directly on the screen of the mobile device. This happened e.g. when the camera switched automatically to the "demo mode" or when the app crashed.

When considering the replay video of respondent #07 (task #01) again the same effect as in section 3.1 arose. When touching on a button the data point the EGT registered was far away from the point where the user touched the screen. It was assumed that the user did not look on the back of the hand most of the time (like in Figure 17). Consequently, the eye tracker did not track the correct gaze position.

Table 1 shows the count of touch errors per respondent and task. The time is based on the original cut of the ET recording. The touch errors by respondent #06 were not counted because of lack of time. Touch errors were indicated upon all respondents. Respondent #07 exhibited most of the touch errors in task #01.

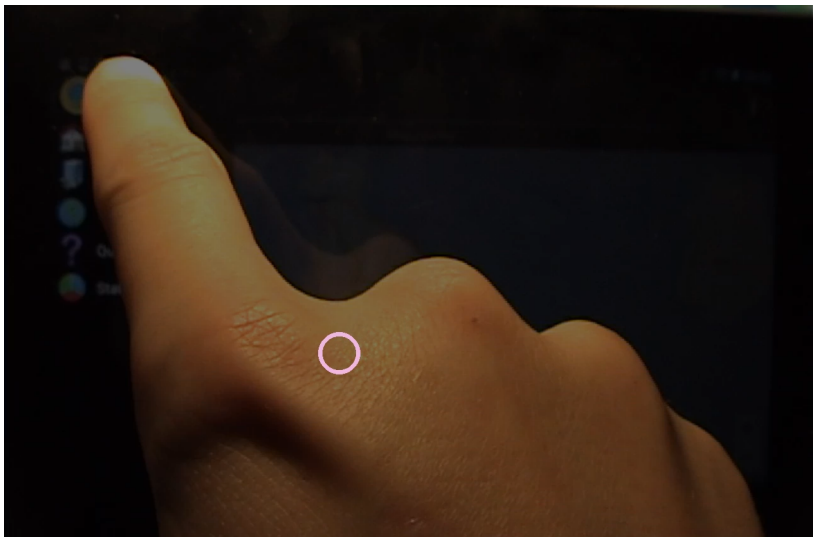


Figure 17: Case Study: Unlikely point of attention

	Respondent #05	Respondent #07	Respondent #08
task 01	2 (2:46) 1 (3:20)	1 (0:12) 1 (0:23) 2 (0:57) 1 (2:05) 1 (2:36) 1 (2:54) 1 (3:21) 1 (3:41)	2 (2:08)
task 02	none	1 (0:05) 1 (0:17)	1 (0:09)
task 03	1 (0:34) 1 (0:42)	1 (0:05) 1 (0:54) 1 (3:05) 1 (3:07)	none

Table 1: Number of touch errors in the case study

4 CONCLUSION & DISCUSSION

The research question was whether Grinbath's EGT can be used for UT on mobile devices. To answer this question two subquestions were formulated: First, what is a workable UT setup with the EyeGuide Tracker for mobile devices? Second, is the basic setup suitable for UT on mobile devices?

A candidate setup including a set of hardware, software and guidelines matching the predefined criteria was found (see Appendix A). Nevertheless, Grinbath's EGT was not suitable for UT on mobile devices with the approach used in this study.

The candidate setup can be used for educational purposes and further improvement.

The usage of a video camera as mediator for basic interaction proved to be possible according to the predefined criteria. Due to the many touch errors which were determined in the case study the approach was considered unnatural and therefore not suitable for UT. Moreover, some test users looked directly on the screen of the mobile device leading to partially invalid ET data.

The setup resulted in worse calibrations than those on videos or reports on the Internet. This gave rise to a separate case study of the EGT's accuracy itself. A picture with five visible objects was created to fill the computer screen. The task was to look at each object for five seconds in a specific order. 44% of all frames missed ET data (only runs with calibration points placed at least inside the red circle area in the calibration result image were considered for this calculation). Thus half of the time the EGT was not able to register ET data. Another remarkable observation was the absence of data points of the object placed in the upper right corner of the screen. Even with a maximal fixation radius of 99 and a minimal fixation duration of 0.02 seconds no fixation on that corner was identified (see Figure 18). These disappointing observations led to the suspicion that the used copy of the EGT was not accurate or a fundamental guideline in the setup was missed. Unfortunately, Grinbath's support via email did not lead to any new insights.

Point of discussion is also the influence of artificial light in the room on the calibration results. According to a mail from Grinbath support, only non-indoor light could interfere with the infra-red light of the camera. This conflicted with the results of step #02. The big differences between calibration results with and without artificial illumination suggest that the EGT software used more data than just the

reflected infra-red light for determining the position of the test user's pupil.

Disadvantages of the used definition of touch error are the following. Sometimes it was not clear if a user really touched the screen. Sometimes the test user accidentally hit a button. Actually this could also be considered as a touch error, but such a conclusion required knowledge about what a user was intended to do. Of course this could not be determined reliably by external observation.

While conducting the research only little budget was available. The research was not assigned to a specific location. Often the setup had to be dismantled and built up again. This made it difficult to ensure that results of different setups were comparable with each other.

Future research could investigate whether a better interaction is possible with a state of the art video camera with HDMI output and an advanced video capture card attached to the computer system. It was expected that such components could transmit the video capture faster to the computer screen.

The results showed that in some situations test users looked directly on the mobile device's screen. Research could be conducted to investigate in which situations a test user is tended to do this and how it can be prevented.

Alternative approaches for using the EGT on mobile devices could involve emulators on tablets where Windows is installed on or restricting the setup to apps made for Windows tablets. The EGT provides an API which can transmit raw data to a second computer system. A calibration procedure could be programmed for operating systems like Android or iOS which maps the data obtained by the EGT to positions of the mobile device's screen. All these approaches would not require a video camera as mediator for interaction and therefore could be more suitable for UT.

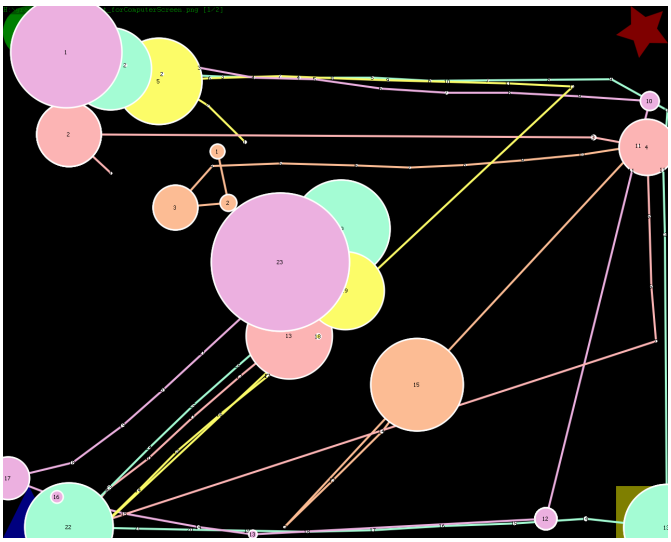


Figure 18: Separate case: Gaze in upper right corner could not be recognized by the eye tracker.

REFERENCES

Benel, D. C., Ottens, D. & Horst, R. (1991), Use of an eyetracking system in the usability laboratory, *in* 'Proceedings of the Human Factors and Ergonomics Society Annual Meeting', Vol. 35, SAGE Publications, pp. 461–465.

Brooke, J. (n.d.), 'Sus-a quick and dirty usability scale (1996)', *PW Jordan, B. Thomas, BA Weerdmeester and AL McClelland* .

Çöltekin, A., Heil, B., Garlandini, S. & Fabrikant, S. I. (2009), 'Evaluating the effectiveness of interactive map interface designs: a case study integrating usability metrics with eye-movement analysis', *Cartography and Geographic Information Science* **36**(1), 5–17.

Cooke, L. (2006), Is eye tracking the next step in usability testing?, *in* 'International Professional Communication Conference, 2006 IEEE', IEEE, pp. 236–242.

Duchowski, A. (2007), *Eye Tracking Methodology: Theory and Practice*, Methods in molecular biology, Springer.
URL: *http://books.google.nl/books?id=27jCNmafYU4C*

Nielsen, J. (1989), Usability engineering at a discount, *in* 'Proceedings of the Third International Conference on Human-computer Interaction on Designing and Using Human-computer Interfaces and Knowledge Based Systems (2Nd Ed.)', Elsevier Science Inc., New York, NY, USA, pp. 394–401.

Nielsen, J. (1994), *Usability Engineering*, Interactive Technologies Series, Morgan Kaufmann.
URL: *http://books.google.nl/books?id=95As2OF67f0C*

Oates, B. (2006), *Researching Information Systems and Computing*, SAGE Publications.
URL: *http://books.google.nl/books?id=ztrj8aph-4sC*

Paas, F. G. & Van Merriënboer, J. J. (1994), 'Instructional control of cognitive load in the training of complex cognitive tasks', *Educational Psychology Review* **6**(4), 351–371.

Page of Autodesk SketchBook on Google play platform (2015).
URL: *https://play.google.com/store/apps/details?id=com.adsk.sketchbook&hl=en*

Page of World Atlas on Google play platform (2015).
URL: *https://play.google.com/store/apps/details?id=com.mindbeach.android.worldatlas*

Siegenthaler, E., Wurtz, P. & Groner, R. (2010), 'Improving the usability of e-book readers', *Journal of Usability Studies* **6**(1), 25–38.

Sweller, J. (1988), 'Cognitive load during problem solving: Effects on learning', *Cognitive science* **12**(2), 257–285.

Weiser, M. (1999), 'The computer for the 21st century', *SIGMOBILE Mob. Comput. Commun. Rev.* **3**(3), 3–11.
URL: *http://doi.acm.org/10.1145/329124.329126*

A

APPENDIX: SETUP DESCRIPTION

Setup description for usability testing on mobile devices with Grinbath's EyeGuide Tracker

Introduction

This document is based on the bachelorthesis “The applicability of Grinbath's EyeGuide Tracker to Mobile Devices for usability Testing” by Jonathan Seesink (Radboud University, 2015). This setup makes it possible to use Grinbath's EyeGuide Tracker (EGT) for usability testing on mobile devices without having Windows or Macintosh installed on the mobile device. It summarizes the recommended usability testing setup found by a basic idea and several iterative adjustments. The basic idea involved a live video capture on a separate computer screen as mediator for the interaction with the mobile device. The test user has to interact with the mobile device by looking solely on the separate computer screen. Illustration 1 shows a functional schematic of the setup. The setup description provides information about hardware, software and guidelines.

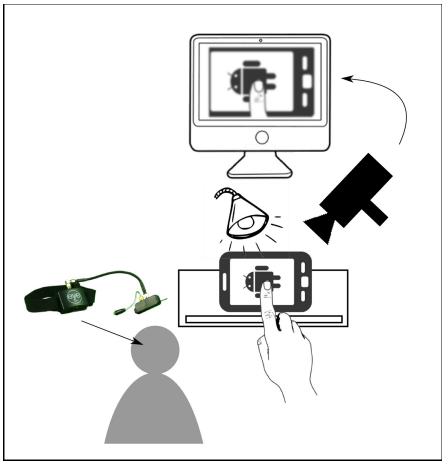


Illustration 1: Functional schematic of setup

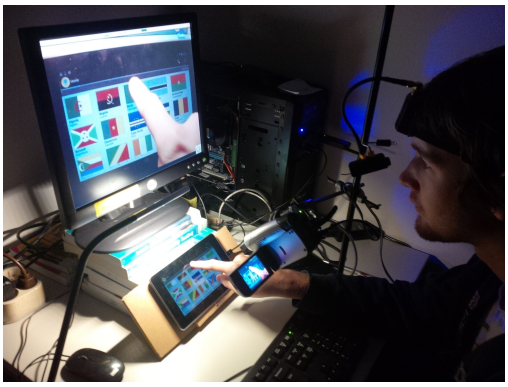


Illustration 2: Usability testing session with the proposed setup

Material

General

The setup requires the following materials:

- Table. It should provide space for computer screen, tablet and the tripod holding the camcorder. It is advised to place the computer with the transmitter as close as possible to the eye tracker. For this reason the computer was also placed on the table during the research.
- Chair (without casters). Test users should be able to sit straightly.
- Several plug connections.
- LED light for placing above the mobile device. (The research used the model “Harte” by Ikea).
- books or something comparable for levelling up the monitor to a proper height
- a mobile device with screen size bigger or equal to 7 inches

- a sloping board as holder for mobile device (e.g. as in Illustration 3)
- Room without sunlight and the option to turn off all artificial lights (of course this does not include lights coming from the computer screen, mobile device, the separate LED spotlight and the video camera).
- Camcorder with firewire slot
- Firewire cable suitable with the video camera's firewire slot and the PCI card of the computer system (e.g. a 4-conductor and 6-conductor firewire 400 alpha connector).

Illustration 3: Sloping board

Computer hardware

- Computer with Windows or Macintosh installed
- Computer screen with a native resolution of 1280x1024 pixels (ideal resolution according to this site[1])
- Firewire card if no firewire port is available on computer system (in most cases a PCI card slot is required for installation)
- Keyboard & mouse
- External HDD with disk space >= 1GB

The EyeGuide Tracker software Computer hardware should be chosen according to the system requirements by Grinbath (link[2]).

The external HDD is recommended due to the big files the EGT produces.

Software

Besides the Windows or Macintosh OS the following is needed:

- driver and data gathering software for the EyeGuide Tracker (delivered together with the EyeGuide Tracker) including the software products EyeGuide Analyse and EyeGuide Capture
- VLC for showing live capture of the video camera (version 2.1.4 used in this description)

Installation

1 http://support.grinbath.com/post/calibration_success.html
2 http://support.grinbath.com/post/system_requirements.html

Illustration 4 shows how to place the main components of the setup. This includes the computer screen, the LED light, the video camera, the sloping board, the computer (without any objects between transmitter and the EyeGuide Tracker), the computer screen (positioned in a proper height), the tripod and the mobile device. Ensure that test users can sit comfortable during usability testing sessions.

Illustration 4: Setup

Guidelines

- Before you start working with the EGT you should watch/read the following web pages. It is assumed that you read/watched the material below before you follow further instructions.
  - Step by step tutorial for getting started with the EGT[3]. Lots of useful information about setting up the EGT is embedded in the video.
  - Video about calibration with the EGT[4]
  - best practices by Grinbath (textual information on this web page[5] and a video[6])
  - suggested usability testing methodologies[7]
  - video about interpreting eye tracking data with EyeGuide Analyse[8]
- Make sure not to have any devices close to the transmitter of the EGT which can interfere with the wireless connection (mobile phones, wifi, etc.), see also this site[9].
- To reduce the amount of touch errors chose applications for usability testing that have buttons not smaller than the test user's forefinger.
- Use only mobile devices with screen sizes equal or bigger than 7 inches.

Setup headset and calibration

Please watch this video to get more information about the calibration process (link[10]). In that video Grinbath suggests to chose one of 4 headset positions:

- in the middle of the forehead
- on the side of the forehead just about the outer corner of the eye
- over the ear or

3 http://www.grinbath.com/tutorial/eyetracker/index.html
4 http://support.grinbath.com/post/optimal_lighting_with_the_led.html
5 http://support.grinbath.com/post/best_practices.html
6 http://support.grinbath.com/post/getting_started.html
7 http://support.grinbath.com/post/suggested_methodologies.html
8 http://support.grinbath.com/post/interpreting_data_with_eyeguide%E2%84%A2.html
9 http://support.grinbath.com/post/camera_interference.html
10 http://support.grinbath.com/post/calibrate.html

- on the side of the forehead with the camera housing inverted

During the research the following placement seemed to offer the best calibration results: The camera is placed close to the test user's nose (inner side of the eye) and points to the test user's left eye (under right corner) as close as possible.

Calibration

- Ask the test user to look to four different regions (inclusive the outer most corners of the computer screen) and determine if in all situations a constant green circle is rendered by the *EyeGuide Capture* software like in Illustration 5.
    - If this is the case continue with the calibration process.
    - If this is not the case move the camera with the IR light so that there is a larger distance between camera and the test user's eye.

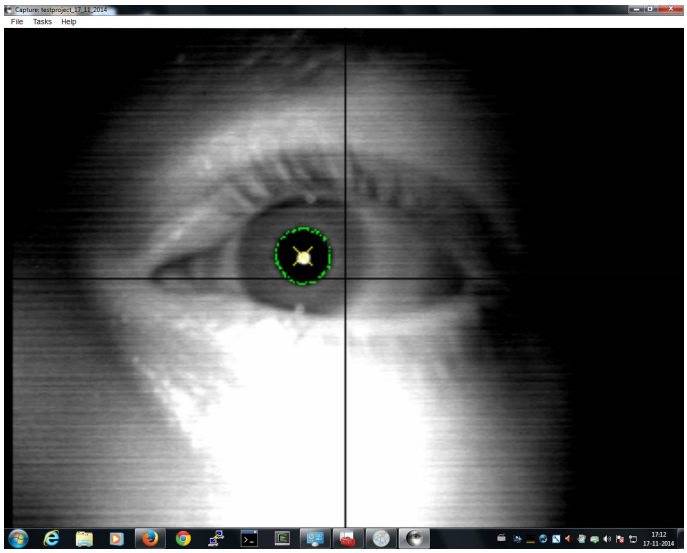


Illustration 5: Green circle represents a proper registration of the test user's pupil

Obtained from this site[11]

- If the participant is moving his/her head during the calibration, that will adversely impact the success of the calibration. Similarly, blinking during the calibration can also negatively affect the outcome of the calibration.

Show live capture with VLC

The video camera must be connected to a working fire wire port. For playing the live capture of the video

11 http://support.grinbath.com/post/calibration_success.html

camera just open VLC and select the option “Open Capture Device...” (see Illustration 6). In the following dialogue you do not need to change any parameters in most cases so you only have to click on the “Play” button. After the test user is ready and you started a task in EyeGuide Capture double click on the video pane of VLC to switch to full screen mode. Make sure that there are no other visual elements on screen visible except the capture of the mobile device's screen (e.g. set the option “autohide” for the Windows taskbar).

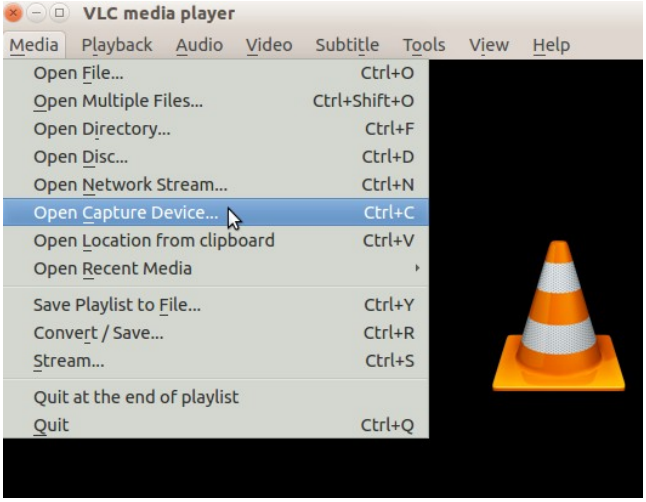


Illustration 6: Opening live capture in VLC

Recommendations

- This web page[12] suggests that you can set another delay for the DV-camera in VLC. Maybe this can reduce capture delay during interaction.
- Conduct usability testing sessions only for applications with buttons wider or longer than the average size of the top of the test user's finger.

12 https://www.videolan.org/doc/streaming-howto/en/ch11.html

B

APPENDIX: QUESTIONNAIRES

Test users of the case study filled in a questionnaire. These are provided in this section.

Questionnaire – Eye tracking usability test

Respondent-nr: 5

Date: 16-12-2014

Gender: male ☐ female ☒

Using lenses: yes ☐ no ☒

Handedness: Right ☒ Left ☐

		Strongly disagree				Strongly agree
Related to software		**1**	**2**	**3**	**4**	**5**
1	I think that I would like to use this app frequently		X			
2	I found the app unnecessarily complex		X			
3	I thought the app was easy to use				X	
4	I think that I would need the support of a technical person to be able to use this app		X			
5	I found the various functions in this app were well integrated			X		
6	I thought there was too much inconsistency in this app			X		
7	I would imagine that most people would learn to use this app very quickly				X	
8	I found the app very cumbersome to use		X			
9	felt very confident using the app		X			
10	I needed to learn a lot of things before I could get going with this app		X			
Related to eye tracker and camera						
11	Wearing the eye tracker during the test was annoying		X			
12	The eye tracker disturbed my work	X				
13	The text on the tablet was clearly readable through the camera's image				X	
14	The (other) visual elements on the tablet were clearly visible through the camera's image				X	
Related to hardware						
15	I was familiar with the mobile device (the tablet or smartphone) before the test.				X	

Questionnaire – Eye tracking usability test

Respondent-nr: 6

Date: 16-12-14

Gender: male ☒ female ☐

Using lenses: yes ☒ no ☐

Handedness: Right ☒ Left ☐

		Strongly disagree				Strongly agree
Related to software		**1**	**2**	**3**	**4**	**5**
1	I think that I would like to use this app frequently	X				
2	I found the app unnecessarily complex				X	
3	I thought the app was easy to use		X			
4	I think that I would need the support of a technical person to be able to use this app	X				
5	I found the various functions in this app were well integrated		X			
6	I thought there was too much inconsistency in this app				X	
7	I would imagine that most people would learn to use this app very quickly			X		
8	I found the app very cumbersome to use		X			
9	I felt very confident using the app			X		
10	I needed to learn a lot of things before I could get going with this app			X		
Related to eye tracker and camera						
11	Wearing the eye tracker during the test was annoying	X				
12	The eye tracker disturbed my work	X				
13	The text on the tablet was clearly readable through the camera's image		X			
14	The (other) visual elements on the tablet were clearly visible through the camera's image				X	
Related to hardware						
15	I was familiar with the mobile device (the tablet or smartphone) before the test.					X

Questionnaire – Eye tracking usability test

Respondent-nr: 7

Date: 17.12.2014

Gender: male ■ female □

Using lenses: yes □ no ■

Handedness: Right ■ Left □

		Strongly disagree				Strongly agree
Related to software		**1**	**2**	**3**	**4**	**5**
1	I think that I would like to use this app frequently		X			
2	I found the app unnecessarily complex				X	
3	I thought the app was easy to use		X			
4	I think that I would need the support of a technical person to be able to use this app		X			
5	I found the various functions in this app were well integrated		X			
6	I thought there was too much inconsistency in this app		X			
7	I would imagine that most people would learn to use this app very quickly				X	
8	I found the app very cumbersome to use			X		
9	I felt very confident using the app		X			
10	I needed to learn a lot of things before I could get going with this app			X		
Related to eye tracker and camera						
11	Wearing the eye tracker during the test was annoying			X		
12	The eye tracker disturbed my work		X			
13	The text on the tablet was clearly readable through the camera's image					X
14	The (other) visual elements on the tablet were clearly visible through the camera's image					X
Related to hardware						
15	I was familiar with the mobile device (the tablet or smartphone) before the test.					X

Questionnaire – Eye tracking usability test

Respondent-nr: 8

Date: 17.12.2014

Gender: male ☒ female ☐

Using lenses: yes ☐ no ☒

Handedness: Right ☒ Left ☐

		Strongly disagree				Strongly agree
Related to software		**1**	**2**	**3**	**4**	**5**
1	I think that I would like to use this app frequently				X	
2	I found the app unnecessarily complex		X			
3	I thought the app was easy to use					X
4	I think that I would need the support of a technical person to be able to use this app	X				
5	I found the various functions in this app were well integrated				X	
6	I thought there was too much inconsistency in this app			X		
7	I would imagine that most people would learn to use this app very quickly				X	
8	I found the app very cumbersome to use	X				
9	I felt very confident using the app			X		
10	I needed to learn a lot of things before I could get going with this app	X				
Related to eye tracker and camera						
11	Wearing the eye tracker during the test was annoying				X	
12	The eye tracker disturbed my work	X				
13	The text on the tablet was clearly readable through the camera's image					X
14	The (other) visual elements on the tablet were clearly visible through the camera's image					X
Related to hardware						
15	I was familiar with the mobile device (the tablet or smartphone) before the test.					X