WhitePaper

Tobii® Technology

# Using Eye Tracking to Test Mobile Devices

## What to consider and how to set it up

# Using Eye Tracking to Test Mobile Devices

## What to consider and how to set it up

February 26, 2010

Tobii Technology AB

As mobile devices become more and more popular, interest in eye tracking studies on these devices grows accordingly. Carrying out eye tracking tests on such small screens or physical devices in general, is not, however, completely straightforward. In this paper we aim to provide information on the limitations of eye tracking when testing mobile devices, as well as share our experiences with three different set-ups using Tobii products.

File version: Tobii_Whitepaper_UsingEyeTrackingToTestMobileDevicesWhitepaper_270110.pdf

# Table of Contents

# 1 Executive Summary

With a worldwide increase in the number of mobile devices, interest in eye-tracking studies on these devices has also increased. Eye tracking can provide valuable information on a number of areas, such as how well a graphical interface works, where on a physical device the user looks for certain options (such as buttons) and how the user's attention shifts between the different parts of the phone when interacting with it. However, carrying out eye tracking tests on such small devices is not an uncomplicated task. Due to the way the human eye works, we are able to see a large part of the mobile device even if only one fixation has been registered by the eye tracker. Also the accuracy of the eye tracker (which is about 0.5° for all Tobii models) plays a more significant role the smaller the interface becomes. In addition, as the mobile device is three-dimensional, eye tracking interaction with it can result in data offsets and errors if not set up properly. In order to overcome these issues, several different set-ups have been developed. In our study we tested three of these which we judged to be the most practical and most suitable to provide most robust data.

## 1.1 The Different Set-ups

In our study we evaluated not only for which studies each set-up was the most suitable – we also tested the actual steps involved in setting up each setup. Below you will find a short summary of the key findings from this study.

### 1.1.1 The Emulator Set-up

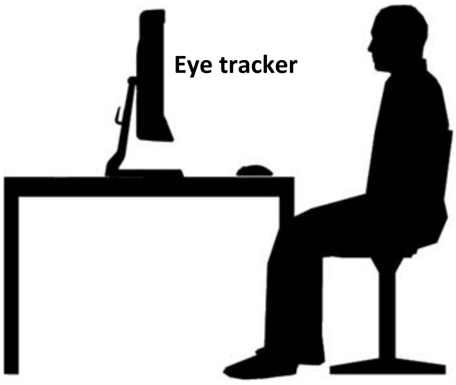



Figure 1. The Emulator setup

This set-up uses an emulator (a computer program that looks and works more or less like a mobile phone) which is run on a regular computer. The emulator is shown on a Tobii T-series Eye Tracker screen, or a regular screen if an X-series eye tracker is used, and is tested the same as any other piece of software in Tobii Studio. This is by far the simplest way of carrying out eye tracking tests on a mobile device or interface and provides very accurate eye tracking data as the interface is two-dimensional, more or less stationary and slightly larger than the actual device. Of course, this kind of testing cannot provide feedback on how participants would look at the actual, three-dimensional device, but does provide information on mobile interfaces as well as websites designed for a small screen. It can also be combined with a touchscreen overlay to provide a more natural interaction.

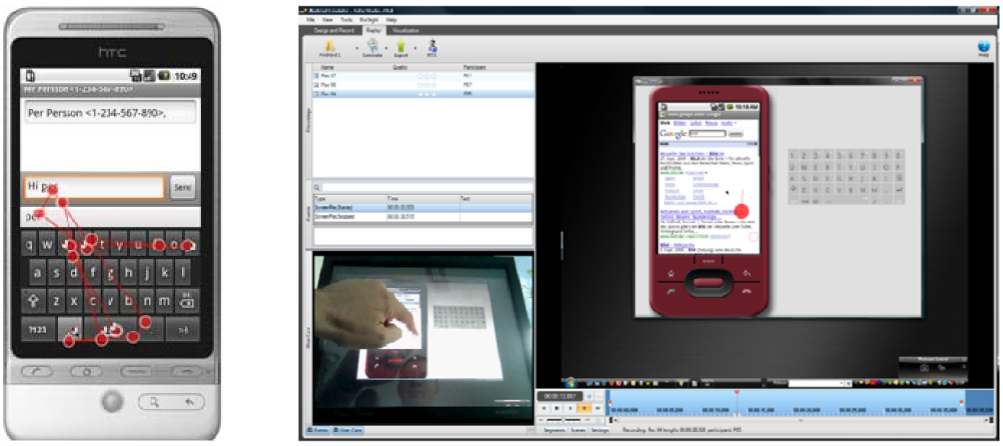


Figure 2. Example gaze plot from the Emulator set-up and illustration of using a touchscreen overlay.

### 1.1.2 The Below Table Set-up

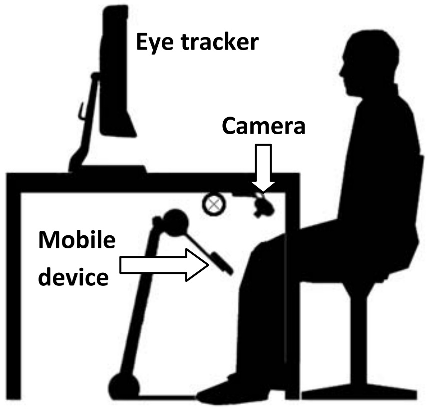


Figure 3. The Below the table setup

In this set-up we made use of Tobii Studio's video-capturing functionality by mounting the mobile device in a bracket on a stand under the table. We then filmed the device and the participant's interactions with it. The participant could not see the device directly, but was instead presented with the streamed video on the screen of a Tobii T-series Eye Tracker. As with the Emulator set-up, an X-series Eye Tracker combined with a regular screen can also be used.

This set-up was the only one tested that allowed for complete freedom of movement of the mobile device. As the image of the device was shown on the eye tracker's screen, tilting or moving the device did not introduce offsets in the data which is otherwise often the case when carrying out eye tracking tests on physical objects. The eye tracking data collected was also very precise as the image of the device was enlarged compared to the size of the actual device.


Figure 4. Gaze plots from when using the Below the table set-up.

### 1.1.3 The Stand-alone Eye Tracker Set-up

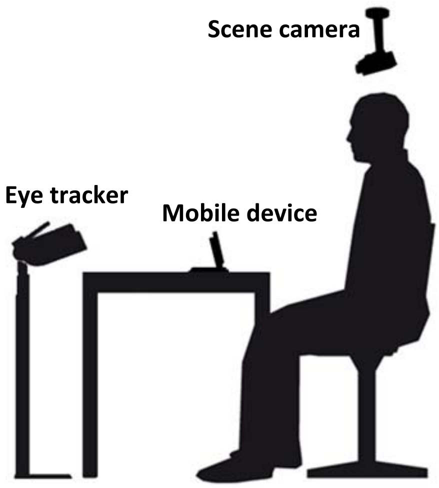


Figure 5. The Stand alone eye tracker setup

This set-up was the most challenging of those tested. While providing the most natural interaction with the device, i.e. the ability to see and touch it, it was the most sensitive in terms of movement of the test elements. In this set-up an X-series Eye Tracker was positioned upside down in front of the participant so that it was able to see straight into the participant's eyes. The mobile phone was placed in a holder on a table between the eye tracker and the participant. A scene camera was placed above (or beside) the head of the participant, i.e. a camera filming the devices so that Tobii Studio could superimpose the gaze data onto the video. As the eye tracker is optimized for collecting gaze data from items which are positioned above it, it has to be mounted upside down. The mobile phone needs to be positioned so that the eye tracker can 'see' the eyes of the participant above the mobile phone. Care hast to be taken not to position it too high as the participant's eyelids might obstruct the view of the eyes when the participant looks down at the phone. The device should also be kept stationary as even small movements can cause serious offsets in the data collected.

Even though this set-up is quite complicated, it gives the participants the opportunity to directly interact with the mobile device. It also shows the device in its real size which allows you to test how participants look at small interfaces.

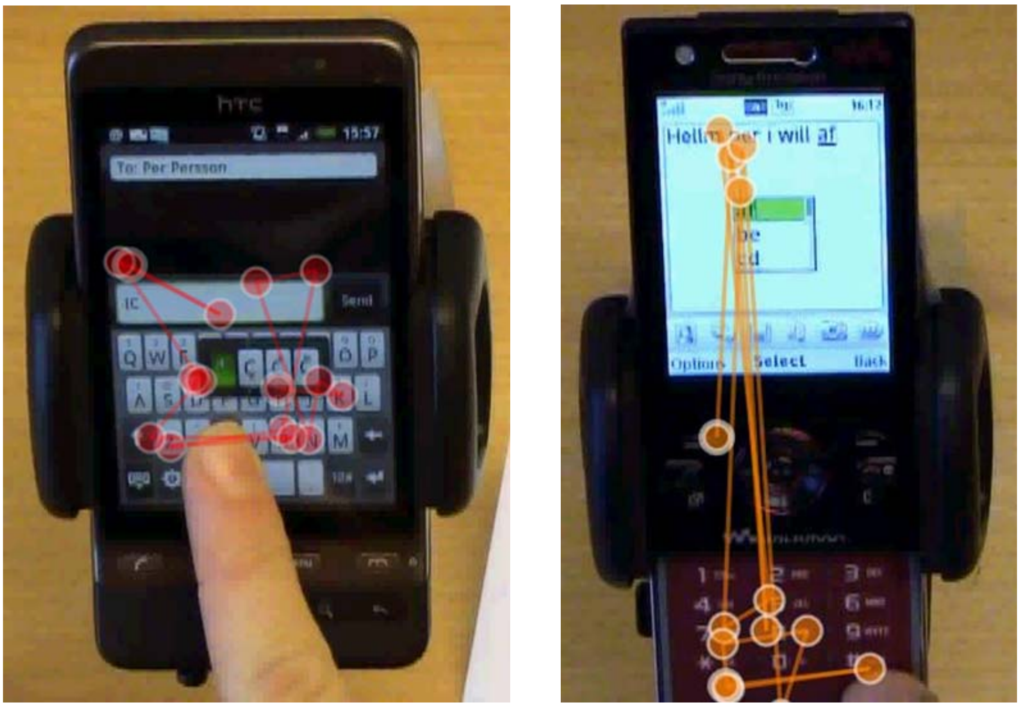


Figure 6. Gaze plots from when using the Stand alone eye tracker setup.

## 1.2 General Conclusion

There is not one particular set-up that is optimal for all kinds of testing. In this document we have given only a quick overview of the set-ups we have tested and the conclusions drawn from these tests. One of the most important conclusions is that the suitability of each set-up depends on what it is that you are testing. If you want to test the structure of a website designed for access through mobile devices, for example, then the emulator setup might well be the best choice. However, if you want to test where participants look when changing a battery or interacting with the physical device, the below table set-up may be more suitable. Also important to bear in mind is the difference in complexity of the different setups.

# 2 Introduction

According to a report presented in March 2009 by the International Telecommunications Union (ITU) (International Telecommunication Union, 2009), which is an organ within the United Nations, more than half of the global population now pays to use a mobile phone. This can be compared to approx. a quarter of the population worldwide having access to the Internet. Mobile phone use is primarily growing in developing countries where the lack of infrastructure limits the availability of other forms of communication. This means that people who might not be able to access information via a computer could use their mobile phones instead. With phones such as the iPhone and devices currently running Android, mobile Internet access has become a significant part of mobile phone functionality.

Plenty of research has been carried out on information display on small screens within the academic HCI community, but only now is this becoming a regular field of study within commercial usability testing. Testing the usability of mobile applications and/or devices is very different from conventional usability testing. Mobiles can be used almost anywhere at any time. Depending on which aspect of the mobile application and/or device is the focus of the testing the set-up of the usability test has to be designed accordingly. While testing of applications developed for regular computers can often be carried out in usability labs, this is not always possible when testing phone applications as the context of use plays a vital part in the interaction (Hagen, Robertson, Kan, & Sadler, 2005). The limited size of the display on a mobile phone also causes problems in recording and analyzing the usability test. These problems become even more prominent if the testing also aims to collect eye tracking data. To overcome some of these difficulties several different solutions for mobile device usability testing are currently used.

In this paper we aim to provide the reader with an insight into the possibilities of using Tobii eye trackers in mobile device usability testing. In order to fully understand the implications of using the different set-ups described in this paper, the first chapter describes the impact of the interaction between human vision and eye trackers when using eye tracking to study small devices. This is followed by a brief description of current research and methodology related to mobile interface usability testing, as well as research in which eye trackers has been used for this purpose. In the following chapter we provide our experiences of three different set-ups with Tobii products (the T-series and X-series Eye Trackers). Even though only Tobii products were used in this study, most of the issues identified would likely be encountered if using other suppliers' equipment in similar set-ups. The last chapter of this paper discusses and concludes our evaluation of the different set-ups. A set-up guide, which provides step-by-step guidance on setting up the tested scenarios, can also be requested from your sales representative. A video demonstrating the various set-ups can also be found on YouTube (search for Tobii Eye Tracking Mobile Devices or TobiiEyeTracking).

# 3 Human Vision, Eye Tracking and Mobile Devices

In order to understand how eye tracking can be used for mobile device testing some general information on how the eye works, as well as the eye tracker used, should be considered.

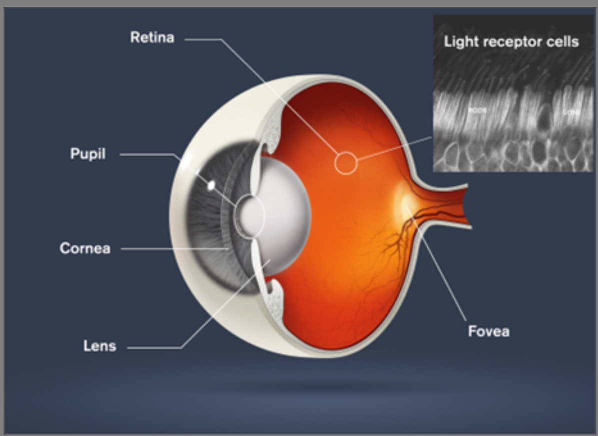
## 3.1 The Eye

The human visual field spans approx. 220° and is divided into three main regions: foveal, parafoveal and peripheral. We primarily register visual data through the foveal region (or fovea see Figure 7) which constitutes less than 8%, i.e. 1-2°, of the visual field. Our peripheral vision has very poor acuity and is mostly suited to picking up movements and contrasts. This means that, at a distance of approx. 50cm, we can clearly see an area which is roughly between 1.5 cm and 3.5 cm in diameter. When we move our eyes to fixate on a specific area of an image or object we are essentially placing the foveal region of the eye on top of it. By focussing our fovea on this area we are maximizing our visual-processing resources onto that particular area of the visual field. This means that since the fovea is able to collect and transmit more data to the brain than any other area of the retina, the more details and information the brain receives and uses to create the image we perceive we are seeing.

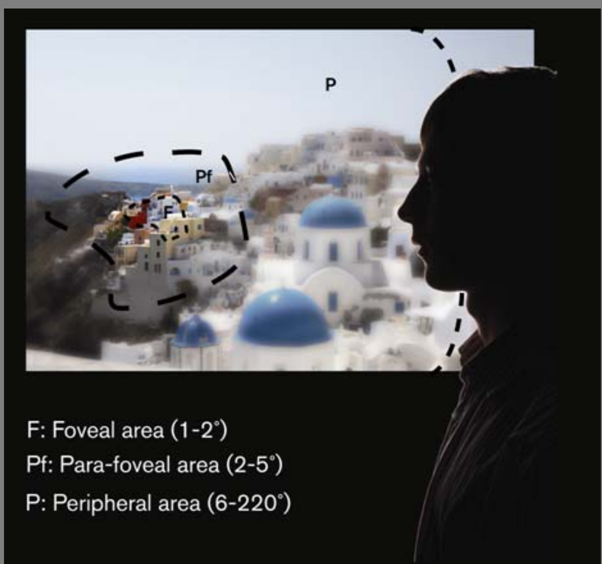
Whenever we look at the world, we consciously or unconsciously focus on only a fraction of the total information that we could potentially process. In other words we perform a perceptual selection process called attention. Visually this is most commonly carried out by moving our eyes from one place of the visual field to another - our gaze follows our attention shift. Even though we prefer to move our eyes to shift our attention, we are also capable of attending to the peripheral areas of our visual field without moving our eyes. This means that instead of focusing our attention on a specific gaze point, we take in what we see in the peripheral areas of our vision as well. Although we can use these two mechanisms separately we most often combine them. An example is when looking at a city landscape. We first detect a shape or movement in our visual field that appears to be interesting and use our peripheral vision to roughly identify what it is. We then direct our gaze to that location allowing our brain to access more detailed information.

Figure 7. The human eye and vision.

**Human vision**




The retina is a sensitive structure inside of the eye responsible for transforming light into signals, which are later converted into an image by the visual cortex in the brain. The fovea is a section of the retina that contains a high density of both kinds of light receptor cells found in the eye, i.e Cone and Rod cells. Rod cells, which are mostly located in the outer retina, have low spatial resolution, are able to support vision in low light conditions, do not discriminate colors, are sensitive to object movement and are responsible for the peripheral vision. Cone cells, which are densely packed within the central visual field, function best in bright light, process acute images and discriminate colors.




The figure to the left is a schematic representation of the human visual field. The main area that is in focus, F, corresponds to the area where we direct our gaze to – the foveal area., i.e. the area of the image registered by the fovea, As is illustrated in this image, the foveal area provides the sharpest image but is not circular. Hence, the area in focus will have a slightly irregular shape as well. Within the rest of the visual field (the para-foveal and peripheral areas) the image we perceive is blurry and thus harder to interpret and discriminate in high detail.

Eye movements have three main functions which are considered important when we process visual information: 1. to place the information that interests us on the fovea (fixations, saccades); 2. to keep the image stationary on the retina in spite of movements of the head or object (smooth pursuit); and 3. to prevent stationary objects from fading perceptually (microsaccades, tremors and drift). For eye tracking purposes, the most commonly mentioned terms are 'fixations' - a pause of the eye movement on a specific area of the visual field; and 'saccades' - rapid movements between fixations. During each saccade visual acuity is suppressed and, as a result, we are unable to see at all. We perceive the world visually only through fixations. The brain virtually integrates the visual images that we acquire through successive fixations into a visual scene or object. Furthermore we are only able to combine features into an accurate perception when we fixate and focus our attention on them. The more complicated, confusing or interesting those features are the longer we need to process them and, consequently, the more time is spent fixating on them. In most cases we can only perceive and interpret something clearly when we fixate on it or very close to it. This eye–mind relationship is what makes it possible to use eye movement measurements to learn something about human cognitive processes.

## 3.2 General Benefits and Drawbacks of Eye Tracking Studies on Mobile Devices

Most eye tracking studies aim to identify and analyze the visual attention patterns of individuals whilst they are performing specific tasks (e.g. reading, interacting with software, scanning an image, browsing the web, etc.). Similar tasks can also be analyzed when tracking participants who are interacting with a mobile device. Interaction with the physical device can also be tracked with an eye tracker and subjected to analysis. We have identified six different research areas for which eye tracking studies on mobile devices can be of particular interest; device-specific interface usability and design, mobile application interface usability and design, mobile web browsing, visual attractiveness and hedonic qualities, physical device interaction and mobile input methods.

The challenge of combining eye tracking with other methods of evaluation for mobile devices is that mobile device displays are generally very small. As mentioned previously, at a distance of 50 cm, i.e. the typical distance (from the eyes) at which we hold a mobile phone whilst using it, only one fixation is necessary for the brain to get an accurate and clear image of approx. a quarter of the display. Hence, using an eye tracker to analyze detailed reading or scanning patterns on a mobile device can be difficult no matter how accurately the eye tracker is able to determine the center of the fixation. However, eye tracking can provide useful information in mobile application or device testing if the chosen test set-up is the correct one and is carefully designed.

## 3.3 Mobile Phone Use

Millions of people using mobile phones in different situations and cultures across the globe have led to an almost infinite variety in the way that mobile phones are used. Different users require different characteristics from phones depending on their personal preferences as well as the context of use. Hence, there is a wide variety of phones available on the market today which offer differences in size, price, functionality, design, speed, interaction capabilities, networking capabilities etc. (Kiljander, 2004). While some users only use the phone to make calls others might need a phone which has almost the same functionality as a computer, e.g. e-mail, document browsing and editing features, and secure access to files on a remote server. Other users might want a phone designed for durability while using the phone as a mobile media center is of importance for another user group. When using a phone as an mp3-player, for example, the user is often mobile in the sense that he or she is walking or moving about (Pirhonen, Brewster, & Holguin, 2002) while users who use the phone to send e-mails are quite often stationary when doing so.

A function that a person uses whilst on the move should ideally be tested in as real a situation as possible. However, it might not be ideal to test the first prototype of the function in a real setting as such testing often requires a lot of resources (planning, set-up, participants, monitoring, analysis time etc). Instead, it is often more beneficial to test the application first in a usability lab (e.g. using an emulator) as this allows the most fundamental usability issues to be identified at an early stage and at a lower cost.

# 4 Current Methodologies

When testing the usability of mobile applications or devices, different methods are currently in use. By far the most commonly used methodology (71% according to a study by Kjeldskov, J. & Graham, C. (Kjeldskov & Connor, 2003)) is Laboratory Experiments. These tests generally take place in a usability laboratory or some sort of artificial and controlled environment. An important benefit of testing in a lab setting, as compared with various field study methodologies, is the control over different variables, which are replicable in a laboratory environment. Laboratory testing also requires less time and fewer resources (Intille, o.a., 2003) and offers the possibility to record the test sessions. Sometimes contextual or mobile tests take place in real-life settings to create a more natural testing environment, which then also provide the accuracy of laboratory equipment (Nielsen, 1994).

Another common methodology is field studies or observational fieldwork, which is a broad class of methods conducted in the participants´ natural environment. Field studies provide a more realistic environment for the test subjects. Field studies are often used to gain insight into and more understanding of *concepts* such as mobility (Kjeldskov & Connor, 2003), rather than for usability testing, even though there are exceptions (Isomursu, Kuutti, & Väinämö, 2004).

Comparing laboratory studies to field studies shows that field testing is significantly more time consuming and requires more resources than conventional laboratory testing. This is mainly due to more preparation being needed; the researcher needs to go to the participant instead of the other way around which means more travel, and, as the interaction that is going to be tested should be conducted naturally, the researcher needs to wait until the interaction happens. The data collected during field studies is also often very time-consuming to analyze. In some cases field testing does not reveal any additional usability issues which than could be captured via laboratory experiments (Kaikkonen, Kekäläinen, Cankar, Kallio, & Kankainen, 2005).

Some other methods used in the usability testing of mobile devices involve remote tracking and data or stream logging. This means that the participants´ behavior can be observed from a remote location through specialized software that tracks and records the users´ actions and transmits them directly from the mobile device. This type of testing is mainly useful to identify usability problems with content on the mobile device, such as web browsing. It does not allow for any testing of usability problems with the physical device itself (Waterson, Landay, & Matthews, 2002).

## 4.1 Mobile Devices and Eye Tracking

There are currently two eye tracking set-ups which are most commonly used for usability studies on mobile devices.

The first involves using eye tracking with a so-called emulator or simulator that simulates the device´s screen interface, and optionally also the keypad, on a monitor equipped with an eye tracker which accurately tracks the participants´ fixation points on the device. This method is especially useful for testing the usability of the interface in terms of navigation, as the size of the device is displayed larger than in reality which means that more fixations are needed in order to see the interface. However, it does not allow the user to actually physically handle the device so there are limitations in identifying usability issues with the device itself. The context of use, when using an emulator instead of a real mobile device, can also eliminate the possibility to identify usability issues which are context-dependent (Nielsen, Overgaard, Bach Pedersen, Stage, & Stenild, 2006).

The second method involves a head-mounted or mobile eye tracker which is sometimes combined with a small camera attached to the mobile device. The camera attached to the mobile device registers the screen and allows for monitoring of the participants´ interaction with the interface[1]. The head-mounted eye tracker registers the participants' gaze behavior in order to provide insight into the attention and perceptions of their conscious processes, e.g. which buttons are fixated and which elements of the mobile device attract the participants' attention. It also allows for usability testing in a real context, i.e. when the test participant is on the move. However, conventional head-mounted eye trackers can be quite obtrusive for the participants and limitations in precision and data capture reduce the usefulness of the data collected via these eye trackers (Duda, 2009).

### 4.1.1 Different Kinds of Eye Trackers

Those methods listed above are currently most commonly used. However, there are other possible methods of collecting gaze data from usability test participants whilst interacting with a mobile device. Which method to use depends on the type of study, e.g. what the goal of the study is, which aspect of the device or software is going to be explored, which eye tracker is available. In terms of the different types of eye tracker, each type offers its own set of benefits as well as drawbacks. Currently there are three categories of eye tracker available for usability testing of mobile devices:

- Head-mounted eye trackers
- Monitor-based eye trackers
- Standalone eye trackers

### 4.1.2 Head-mounted Eye Trackers

When considering mobile device testing in combination with eye tracking, head-mounted eye trackers seem to be the obvious solution. The reason being that they allow for data capture in a real context and enable the participant to manipulate the physical device freely. Many of the head-mounted eye trackers available today are obtrusive and have a low-quality scene camera. As we will discuss later on, when carrying out eye tracking studies on mobile devices, it is important to be able to see the different parts of the interface (e.g. buttons and icons) on the video feed. If the image is of too poor a quality, this can be difficult. Another issue is that the calibration procedure can be quite complicated and on some occasions the calibration is not valid if the participants look up from the device to look at something further away, as might be the case if the study takes place in a real-life setting. Also, as both the mobile device and the camera move during the study, it is often problematic to carry out a quantitative analysis of the data collected. Another aspect to be considered when using a head-mounted device is that the eye tracking will be conducted using the actual mobile device. This means that, as discussed earlier, if the device is held at arm's length (i.e. approx. 50 cm from the eyes) the participant will be able to see quite a large part of the mobile device in only one fixation.

### 4.1.3 Monitor-based Eye Trackers

Monitor-based eye trackers are very accurate but are currently only associated with screen- based interaction, e.g. using emulators or simulators for mobile device testing. This technology has the potential, however, to be linked to an actual device through an external camera connected to the eye tracker. An example of such a set-up is to let the eye tracker act as the participants´ eyes while the participant holds the device, with an illuminated camera attached, under the table on which the eye tracker is placed. This way the device is hidden from the participants' direct view and the participant is forced to use the live video feed shown on the eye tracker screen when interacting with the device. The obvious drawback of using monitor-based eye trackers when testing mobile devices is that the interaction with the device is less natural than with head-mounted eye trackers. However,

Figure 8. Tobii T-series Monitor based Eye Tracker

[1] Using a camera attached to the device is also a common method for studies in the field which do not include eye tracking.

as the eye tracking is carried out on an enlarged version of the device it allows for smaller parts of the mobile device and interface to be studied in the analysis.

#### 4.1.4 Standalone Eye Trackers

Standalone eye trackers provide more flexibility compared with monitor-based eye trackers when it comes to usability testing, as the item being investigated does not have to appear on a screen. However, there have not been many studies exploring the possibilities of using this technology for mobile device eye tracking studies. There are also some limitations involved in setting up the device, e.g. the participant must not obstruct the field of view of the eye tracker as gaze data will then be lost. Furthermore, careful measurements of the angles and distances in relation to the standalone eye tracker and the stimuli are crucial in order to guarantee accurate results.


Figure 9. Tobii X-series Standalone Eye Tracker.

# 5 Evaluation of Different Eye Tracking Methods for Mobile Device Usability Testing

Taking into account the limitations as described above, we decided to evaluate three of the set-ups currently possible using Tobii Eye Trackers. The different set-ups described and discussed in this white paper represent a number of realistic methods for usability testing mobile devices with the help of Tobii eye tracking equipment. When using eye tracking in usability tests of mobile devices the set-ups described in this paper are not the only set-ups possible. However, they were seen, at the time of writing, as the most practical and robust set-ups available. The selection of these set-ups was the result of careful consideration and testing of some of the most common set-ups used in the field of mobile device testing today. During this process of selection, methods that required too much time to set up or proved too difficult to use, for either the researcher or the participant, were discarded. This white paper provides an evaluation of three different set-ups and gives a comparison of the benefits and drawbacks of each.

## 5.1 Overview of the Tested Set-ups

The set-ups we tested had the following characteristics:

- **Tobii T-series Eye Tracker (T120) with the built-in TFT screen displaying an emulator as the stimulus.**
    - Participants used an emulated version of a mobile device with a mouse pointer acting as the participant's 'finger' when interacting with the device. The emulator was running on a PC and was displayed on a Tobii T-series Eye Tracker's built-in TFT screen.
    - The emulated mobile device image was enlarged (compared to the actual size of the device) in order to accurately get gaze data for the smaller interface details.
- **Tobii T-series Eye Tracker (T120) with the built-in TFT screen displaying a live feed from a small camera mounted under a table acting as the stimulus. Two different mobile devices were used.**
    - Participants interacted with a mobile device under a table while looking at the TFT screen of an eye tracker to guide their actions.
    - The mobile device was fixed in a holder standing on the floor for consistent eye tracking results that could be used in a quantitative way. (However, this could also be substituted by a camera directly attached to the device for more freedom of movement for the participant but this was not evaluated as part of this study.)
    - This set-up also offered the possibility of tracking the non-constrained use of a mobile device, e.g. opening the phone and replacing the SIM card. Hence, this set-up included a task involving interaction with the physical device which was not included in the other set-ups.

- **Standalone Tobii X-series Eye Tracker (X120) using an inverted set-up. Two different mobile devices were used as stimuli.**
    - o Participants used a mobile device positioned under an eye tracker which was mounted upside down.
    - o The mobile device was fixed in place in a mobile phone holder during testing in order to ensure consistent eye tracking results that can be examined in a quantitative way.

## 5.2 Mobile Devices Used for Testing

For the testing described in this paper we used two different mobile devices:

- Sony Ericsson W715, a modern but traditionally-designed phone with a tactile keypad and a medium-sized screen (2.4 inches) with a resolution of 240 x 320 pixels.
- HTC Hero, a modern mobile Smartphone with a relatively large capacitive touch screen (3.2 inches) with a resolution of 320 x 480 pixels.

These phones were mainly chosen as they represent the greatest differentiation in mobile phones today, i.e. one represents small, traditional phones while the other is an example of the new generation of 'smart phones' which have considerably larger screens and a touch screen interface. The similarities and differences between the devices are presented in Table 1.

Table 1. Characteristics of the tested devices.

| | Sony Ericsson W715 | HTC Hero |
|---|---|---|
| **Form Factor** | Compact, slide phone<br>95 x 47.5 x 14.3 mm | Touch screen phone<br>112 x 56.2 x 14.4 mm |
| **Display Type** | TFT regular screen<br>240 x 320 pixels, 2.4 inches | TFT capacitive touch screen<br>320 x 480 pixels, 3.2 inches |
| **Feedback type** | Tactile buttons | Haptic feedback vibrations |
| **Input methods** | Pressing buttons, keypad (with T9 text input support.), soft keys, call buttons and navigation button | Touching the screen, onscreen QWERTY keyboard, soft keys and call buttons, trackball |
| **Operating System (OS)** | Proprietary Sony Ericsson OS | Google Android OS |
| **Mobile Browser** | Proprietary and Opera mini | Proprietary and Opera Mini |

As the physical differences between the phones were eliminated when using an emulator and the focus of the study was to evaluate the eye tracking implications of using the different phones, we decided to test only one of the phones when evaluating the emulator set-up. The emulator that we used was able to accurately emulate the HTC Hero. This allowed us to compare the results of the emulator set-up with those of the other two set-ups.

## 5.3 Methodology

Our study was a qualitative comparison of the three set-ups. Even so, the study was set up in such a way that it also allowed us to compare our experiences and results when using the different set-ups systematically. To test the different set-ups consistently, a script was used with the same basic tasks for all set-ups. By using the same tasks, the results from the different set-ups were made comparable. The study was presented to the participants as a regular usability study for mobile devices, i.e. it was not presented as an evaluation of different set-ups for mobile device testing. Hence, the participants were made to behave as they normally would during a regular usability test of a mobile device or interface. The participants were asked to evaluate the usability of the interface of the two mobile phones. This included the phones' operating system as well as the usability of the different websites visited during the tests. Each participant was allocated only one phone and one set-up in order to prevent them from getting too familiar with the interface of the phone and the websites tested. The main focus of the study was to test the practicality of, and type of data produced by each set-up and not the actual usability problems of the mobile devices or interfaces. Hence, for most set-ups only four participants and two mobile devices were tested. If participants were found to be difficult to track due to reasons discussed later on in this white paper, extra participants were recruited as replacements. This also provided additional feedback on trackability issues with the different set-ups. After testing, the participants were asked some questions in order to gain their opinions on the set-up in question. These questions were used to evaluate whether the set-up could be considered as a realistic representation of regular interaction with the device, in the opinion of the participants. It also provided information on which aspects of mobile device testing the particular set-up was most useful for.

## 5.4 Emulator Set-up – With a Tobii T-Series Eye Tracker

This set-up requires the use of a stable mobile device emulator that will run on the Tobii Studio computer and will be displayed as a stimulus on a Tobii T-series Eye Tracker. Instead of using the actual mobile device, the participant will interact with a program (i.e. an emulator) that works and looks like the real phone but is running on a regular computer and is shown on the screen of the eye tracker. Instead of using the fingers to interact with the device, the user will use the mouse. In our case we made use of the Google Android OS emulator, emulating a smart phone very similar to the one used when testing the other set-ups, i.e. a HTC Hero. In our set-up the facilitator sat next to the participant to monitor and control the emulator between tests.

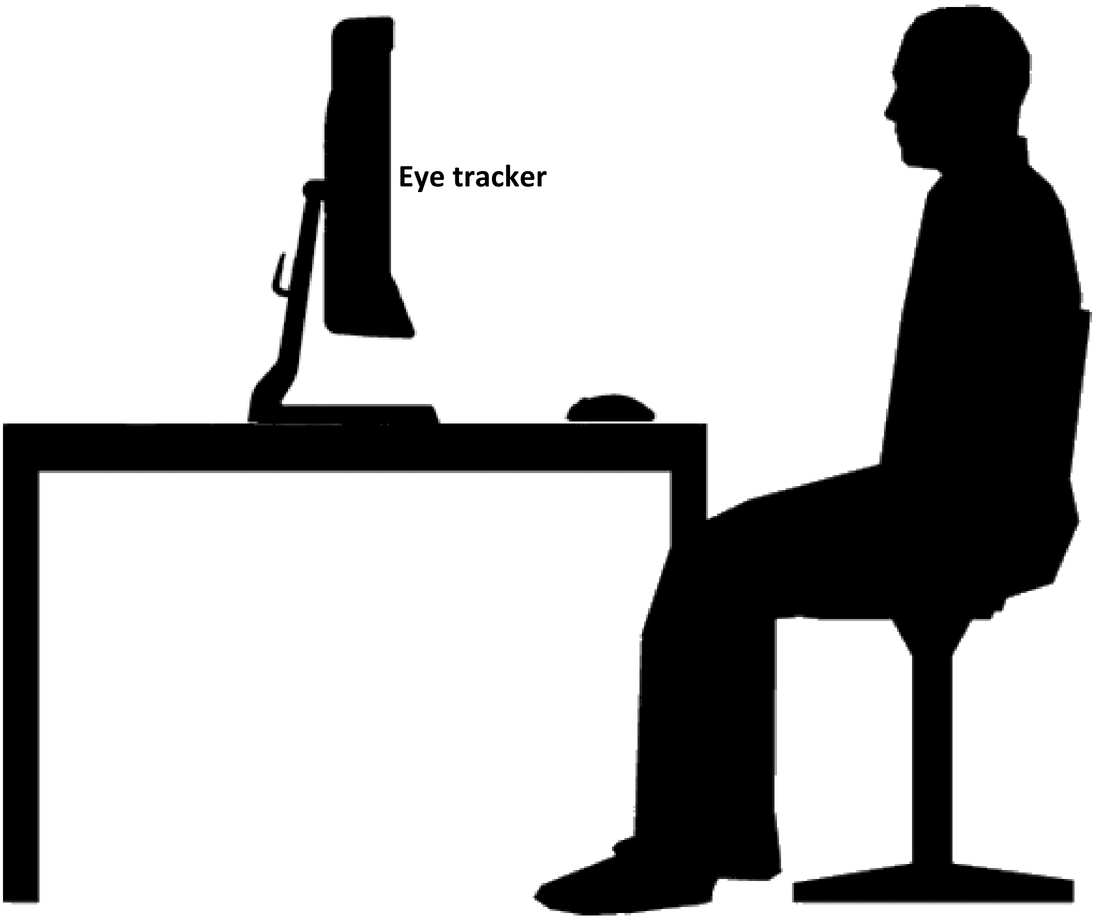



Figure 10. Emulator set-up.

### 5.4.1 Issues Encountered During Testing

The overall experience of the participants using the emulator was that the emulator didn´t really resemble an actual mobile device in terms of tactile interaction. The problem with interacting with the device using the emulator was twofold: As the device was emulated as a two-dimensional image, buttons and other interaction features on the device were hard to distinguish. The other problem was that interactions that were natural on a touch screen, such as scrolling by dragging the fingers over the screen or zooming by 'pinching' the image seen on the display, were difficult to achieve using a mouse. However, the software interface emulation was otherwise considered realistic by the participants.

Even though the number of participants was limited in the test, the collected data showed that similar usability problems related to the *software interface* were identified when using this set-up as with the other two set-ups. However, due to the fact that the testing did not include any interaction with the actual device, the participants were unable to identify usability problems related to the modes of interaction (e.g. buttons and the scroll ball) and feedback received when interacting with the device (e.g. vibrations or tactile feedback from pressing physical buttons).

Participants indicated that the emulator felt slower than they would expect from an actual mobile device. This is possibly due to the fact that they have come to expect more speed from a computer than from a mobile phone so when the computer worked at the speed of the mobile device it was perceived as being slower than it actually was.

Participants used the mouse as a regular mouse instead of the intended representation of their finger, e.g. they tried to use the right click button or double clicked on objects and made selections on the screen that would be too small to do on the actual device using the fingers. This led to the participants making fewer mistakes or encountering other usability problems when it came to device interaction than they did when using the actual devices.

Figure 11. T120 combined with a touch screen panel.

### 5.4.2 Alternative Method

One of the major problems when using an emulator is that the physical interaction with the device is eliminated. The usual input methods available to users, e.g. buttons and touch screen elements, were not available. This can partly be compensated for by using a set-up where the Tobii T-series Eye Tracker is combined with a touch screen panel, turning the screen of the eye tracker into a touch screen (see Figure 11). Testing of such a set-up has highlighted some challenges; e.g. the participants' arms sometimes obstruct the eye tracker's illumination modules, thereby disrupting the eye tracking. However, modifying the set-up slightly by tilting the eye tracker and placing the item to be tracked in a way which minimizes the arms' obstruction of the image sensors or illuminators, greatly improves the trackability.

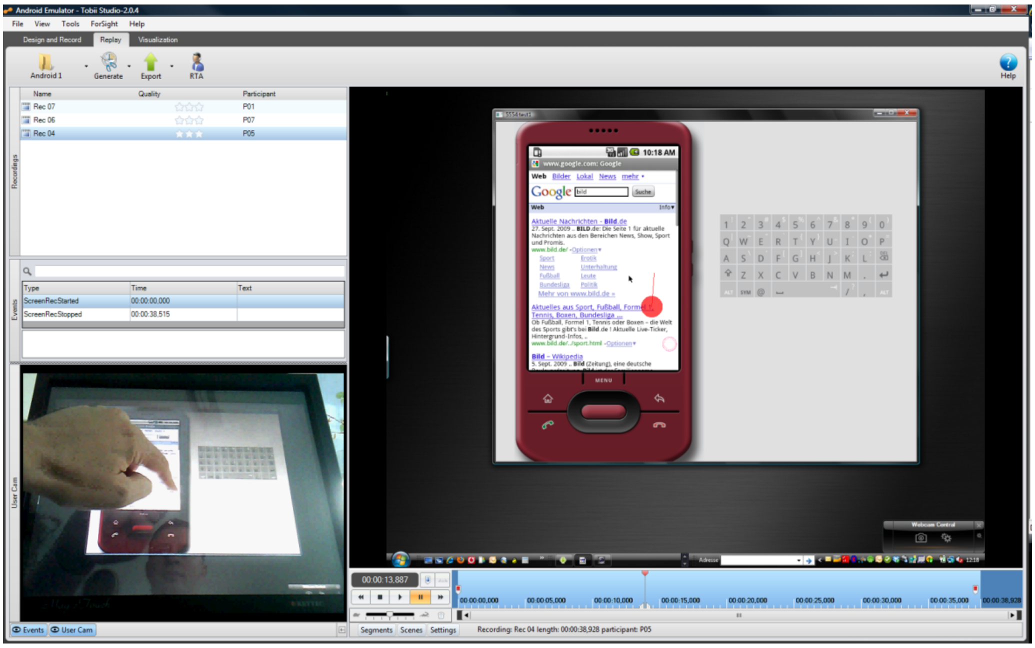


Figure 12. Screen shot from Tobii Studio showing data captured using the set-up. Photo: UsEye

Figure 13. Writing an SMS using the emulator.

### 5.4.3 Analysis of the Results

This is the simplest set-up for evaluating mobile software. However, it also highlights the problems with evaluating mobile interfaces, i.e. that the physical device introduces usability problems not necessarily connected to the interface shown on the display. While this method is suitable for evaluating interface-specific problems, e.g. site structure and the general perception of the design, it cannot provide feedback on what it is like to use the device and software in a real-life situation. However, this set-up has its benefits: It is very easy to set up and allows for very accurate eye tracking data as the device as shown on the screen is larger than in reality (see Figure 13). This means that less of the device's screen can be seen during only one fixation. This method also allows for testing software that is not yet fully functional or even for testing mock-ups in the early stages of the design cycle.

## 5.5 Below Table Set-up – With a Tobii T-Series Eye Tracker with built-in TFT Screen, Displaying a Live Feed From a Small Camera Mounted Under a Table

In this set-up the participants interacted with a mobile device mounted under a table whilst following their interactions on a Tobii T-series eye tracker through a live video feed from a small camera (which was filming the mobile device and the hands of the participants - see Figure 14). By using the mobile phone under the table, we avoided the participant getting distracted and looking at the device itself instead of the eye tracker. Furthermore, it provided the participants with a tactile interaction with the mobile device. Similar to when using the emulator, the projection of the mobile device on the screen can be made slightly larger than the real device which allows for more detailed eye tracking data. We tested this set-up in two different ways: first we fixed the phone to a holder and asked the participants to carry out the same tasks as in the previous set-ups. We then gave the participants a task unique to this particular set-up i.e. to handle the phone freely without any restrictions as long as the phone was kept under the table. More details can be found in the next section.

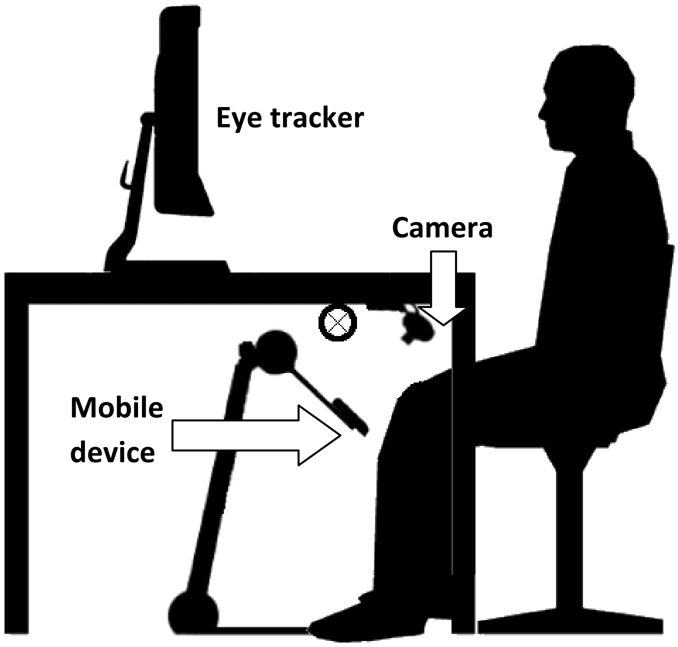


Figure 14. Illustration of the Below Table set-up where the mobile device is placed under a table.

### 5.5.1 Eye Tracking of 'Free Hand' Mobile Interaction

As the image shown to the participant was a two-dimensional representation of a three-dimensional device, we were able to track interactions which otherwise would have caused perspective errors (see Figure 17). To test this, we asked the participants to change the SIM card on the Sony Ericsson phone. To allow them to do this, we disconnected the phone from the holder and placed it on a flat surface under the table. This gave the participants more freedom of movement than was previously possible, although the freedom of movement was somewhat limited by the camera's field of view.

With the previous fixed set-up, it was possible to carry out a quantitative analysis using the 'Areas of interest' tool in Tobii Studio. However, when the device was removed from the holder and given to the participants to freely move around, this possibility was no longer available. Hence, this set-up is mainly suitable for qualitative analysis of user interaction.

### 5.5.2 Problems Encountered During Testing

During testing we noticed that participants experienced a learning curve when using this set-up. Initially, some participants had problems associating the movements they saw on the screen with the actual movements of their fingers. However, after some time, the co-ordination of their fingers according to the feedback from the camera became better and was perceived to be more natural.

Touch screen phones were proven to be harder to use naturally with this set-up, as compared with a traditional phone with a keypad. This was mainly due to the touch screen phone's lack of feedback on interactions, something that becomes more important since the perception of depth or perspective has been taken from the participant due to the fact that he/she is looking at a two-dimensional image on the eye tracker screen. While the participant can feel the buttons being pressed when using the traditional phone, the touch screen phone's only feedback is to vibrate. Furthermore, while the shape and layout of the buttons on the traditional phone can aid in identifying which button to press, with the flat screen of a touch phone we are only able to orientate ourselves by what we see.

The participants' fingers sometimes blocked the camera's view when interacting with the device. Also, the closer the fingers were positioned towards the camera, the larger they appeared on the eye tracker screen. Hence, the fingers sometimes appeared unproportionally large compared to the mobile device.

Small physical details on the mobile device itself or details in its graphic interface were hard to distinguish for some participants. This was mainly due to the relatively unclear, low-resolution and two-dimensional video stream they had to use for their interactions.

Configuring the set-up to prevent the light source from causing reflections on the surface of the mobile device was not uncomplicated. The camera settings also had to be set to account for changes in the brightness of the images shown on the screen of the device, e.g. if the device changed from showing a dark image to a black text on a white background, the text sometimes became unreadable as it blended into the background.

Figure 15. Writing an SMS.

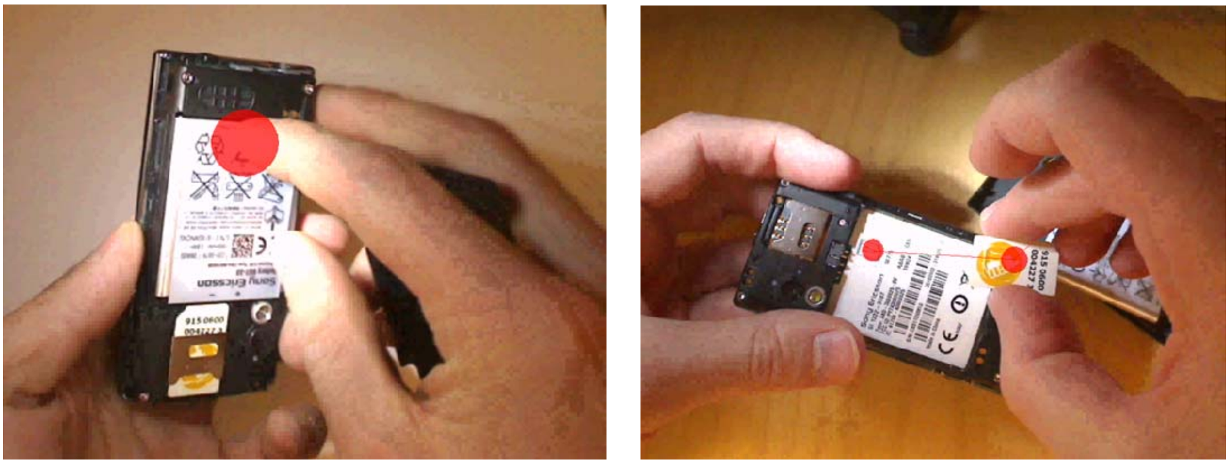


Figure 16. Replacing the SIM card.

### 5.5.3 Analysis Of The Results

Even though this method of testing is not as direct as the inverted eye tracker set-up, it does allow for fairly natural interaction with the mobile device. The participant is able to hold the device and, depending on how it is set up, can even move the device around. An alternative set-up, which we have not investigated here, is to attach the camera to the mobile device instead of to the table. Hence, the camera follows the movements of the device while still providing the participant with a stable image of the device's screen. Even though using a phone under a table is not something that people do under normal circumstances, our testing showed that the participants adapted fairly quickly to this mode of interaction. However, there was a distinct difference in the success rate for this type of interaction between the traditional phone and the touch screen device.

Based on our experiences with this set-up, we found it particularly useful for tracking interaction with the physical mobile device rather than the screen interface. This set-up is the only one tested that can accurately track participants when they freely handle the device, e.g. when changing a SIM card or otherwise turning the device over freely in their hands. It also provide highly accurate eye tracking data when testing the screen interface. However, the downside of the set-up is that the participants cannot look at the device directly.

## 5.6 Standalone Eye Tracker Set-up - With an Inverted Tobii X-Series Eye Tracker

In this set-up a Tobii X-Series Eye Tracker (X120) was used. As the Tobii X-series has been optimized to track objects situated above it, an inverted set-up was used. The participants used a mobile device positioned just below the eye tracker which is set up upside down (see Figure 17). The mobile device was placed in a mobile phone holder in order to ensure consistent eye tracking data that could be examined in a quantitative way. A high-quality video camera was used as a scene camera to record the screen of the mobile device. This scene camera was positioned above the participant to get a realistic view of what the participant was seeing. The inverted set-up provided the participant with an unobtrusive eye tracking experience and a natural way of using the mobile device. To register any physical behavior of the participant, a user camera[2] was also included in the set-up. To allow for participants of different heights the chair used was adjustable. In order to ensure consistent and accurate data, and to minimize set-up times between the different participants, all other parts of the set-up were made as stationary and fixed as possible.

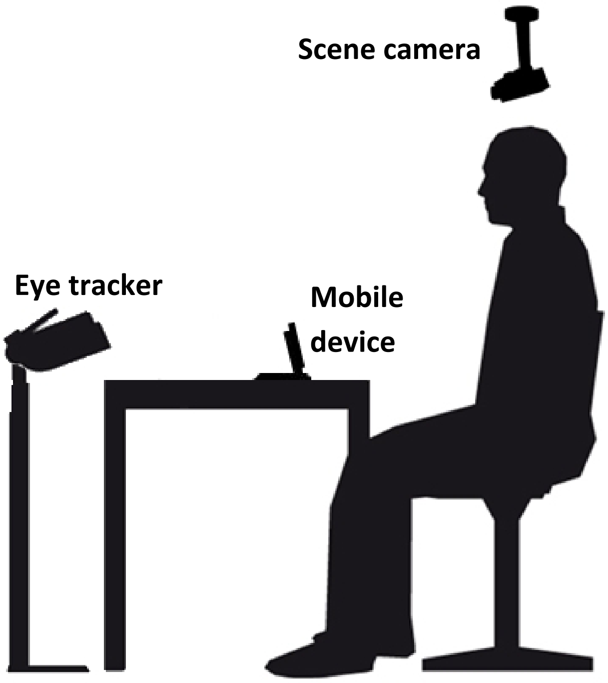



Figure 17. Inverted set-up using a Tobii X-Series Eye Tracker

### 5.6.1 Issues Encountered During Testing

If you are considering implementing this set-up it is important to note that not all participants will be trackable. Compared to tracking stimuli situated above the X120 where almost all participants can be tracked, this inverted set-up brings the number of trackable participants down considerably. The current testing did not involve enough participants for a definite number to be provided, but even with a handful of participants, one or two could not be calibrated correctly. Certain factors decreased the chance of a participant being trackable, e.g. glasses, thick eye lashes or extensive use of mascara or ‘droopy’ eyes. The general problem with

[2] A *user camera* is a camera that is positioned so that the participant’s face and reactions can be recorded. The video captured by this camera can be seen in Tobii Studio together with the eye tracking data and the user interaction. The data collected by the user camera can provide additional information that can be used in a qualitative analysis of the test results.

positioning the eye tracker above the stimuli is that the participant will be looking down which means that the eyelids and/or eye lashes will block the eye tracker's view of the eyes. Hence, it is important to place the eye tracker in a way that allows it to 'see' straight into the eyes of the participant. When glasses are used and the participant is looking down, the reflection of the IR light on the surface of the glass can be distorted in such a way that it eliminates the possibility of tracking the eye movements. To avoid this, the eye tracker can be positioned slightly offset from the centre of the stimuli and rotated so the IR light is not directed straight towards the glasses. This modification to the set-up needs to be compensated for by the eye tracker. Hence, it is very important that the numbers put into the X-series configuration tool[3] are correct. When running our study the set-up we used had a 40 mm offset with a 7° angle which significantly improved the trackability of the participants with glasses.

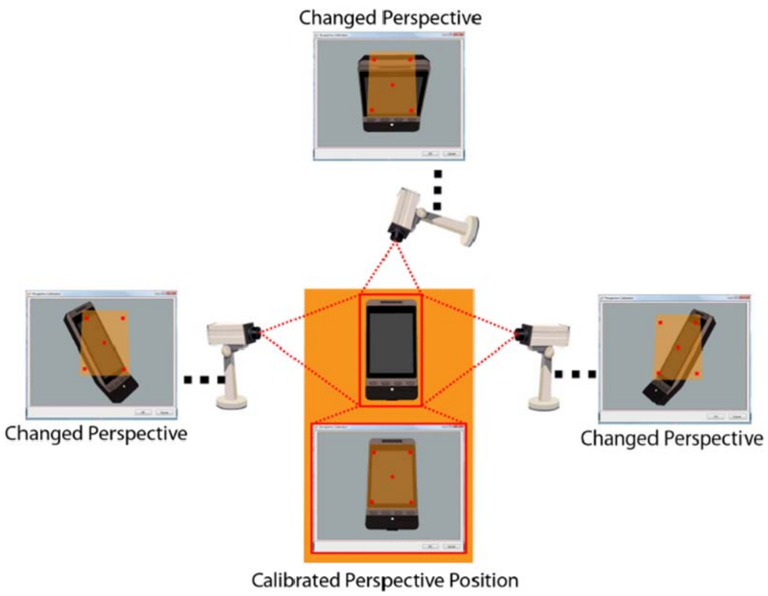



Figure 18. Perspective error.

If the calibration grid is not placed on the same plane as the stimuli, an error or offset might be introduced into the gaze data as presented in Tobii Studio. The reason being that the data collected is presented as two-dimensional points superimposed onto the video feed captured by the scene camera. If the distance between the eyes and the scene camera is large, the image seen by the eyes and by the camera will differ in perspective. What the calibration does is to compensate for this difference in perspective by mapping the gaze points from the eyes onto the image seen by the camera. However, if this perspective difference is altered after calibration is completed, this mapping is no longer accurate (see Figure 18).

Figure 19. Position of the scene camera.

Figure 19 shows the position of the scene camera. The participants sometimes obstructed the view of the scene camera by leaning forward too much. To compensate for this, the scene camera had to be moved closer to the eye tracker. We also noted that participants with bangs or other hair styles with a long fringe sometimes obstructed the view of the scene camera and/or the eye tracker. As the scene camera could not be repositioned during testing without carrying out a new image mapping and calibration, we compensated for this

[3] The X-series configuration tool is used to give the eye tracker essential information which enables it to create a model of the participant's eyes. The eye model is used to ensure robust eye tracking data even if the participant moves in relation to the eye tracker after calibration.

problem by asking the participants to try to keep the hair away from the face as much as possible and not to lean forwards too much.

Participants felt the mobile device holder used in our set-up did not allow for a comfortable or natural use of the mobile device because of its bulkiness. The holder was a traditional mobile phone holder for cars which was not designed for touch screen phones. Hence, the holder partly obstructed parts of the screen from being interacted with.

### 5.6.2 Analysis of the Results

Although this set-up provided some problems when it came to trackability and the physical set-up of the devices used, it did offer the most natural interaction with the phones used in terms of the participants actually being able to see and touch the mobile device. By using a table to support the phone and the arms of the participants, the physical strain of interacting with the mobile device was kept to a minimum. This also means that this set-up could potentially be used for tests of a longer duration, as compared with a set-up without arm support.

The data from this set-up showed some offsets or parallax issues which were mainly due to small movements in the set-up. This made it impossible to do any quantitative analysis of the data. A more fixed set-up is required to avoid these issues. Overall eye tracking accuracy was good as long as all the measurements and calibration stayed constant throughout the session. However, when analyzing the data resulting from a set-up such as this one, where the actual physical device is used, it has to be kept in mind that what the participant sees is not only the point of fixation – the fovea allows the participants to see areas around the fixation point as well. Depending on the distance between the participant and the device, what the participant is actually able to see clearly within one fixation can vary from only a small part of the device (or the device's screen) to the whole device. When deciding on the distance between the eye tracker, the mobile device and the participant, a tradeoff has to be made in terms of trackability versus precision. The closer the mobile device is positioned to the eyes of the participant, the smaller is the area which the participant can see with only one fixation. However, positioning the eye tracker closer to the participant and, hence, further away from the eye tracker also means that the participant will be looking down more and the risk of the eyelids or eye lashes obstructing the eye tracker's view of the eyes increases.

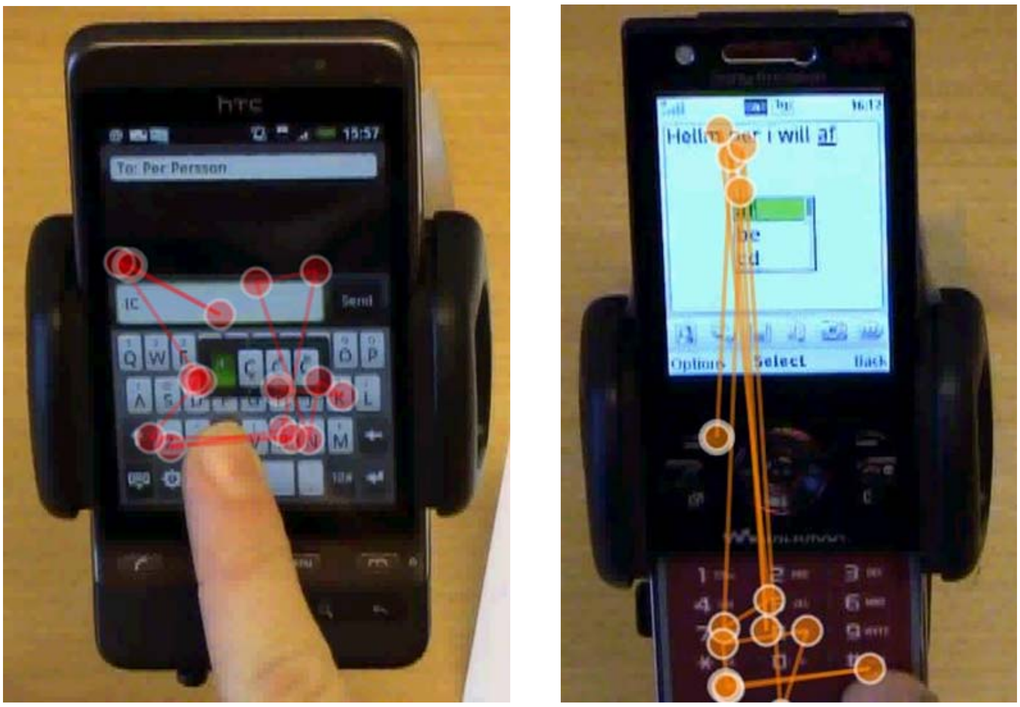



Figure 20. Writing an SMS.

The images from the video camera initially used during the test were not of a very high resolution. A high-definition camera is recommended for this set-up as it is important to see the details of what is happening on the mobile device's screen. Even though the mobile device was positioned in a static setting which greatly simplifies quantitative analysis of the data, the set-up did not allow for automatic recording of any participant-generated events, e.g. mouse clicks or key presses. This means that the generation of scenes which is necessary for a quantitative analysis had to be done manually. In order to generate these scenes, the events (e.g. selections and confirmations) had to be seen on the screen. This means that details such as clicks on the screen had to be seen on the video recording.

Some participants were, for various reasons, hard to track. Even though slight shifts in the set-up (such as changing the offset and viewing angle of the eye tracker) can improve trackability for some participants, this does not completely solve the problem. Furthermore, the Tobii X-series Eye Trackers have not been designed or optimized to work in this inverted way with lower tracking functionality as a result. PLEASE NOTE: It is highly recommended that this set-up is used only for studies where potentially large dropout rates in terms of untrackable participants are acceptable.

# 6 General Discussion

Each set-up tested had its own set of benefits and drawbacks. When selecting which set-up to use, it is important to first evaluate what the purpose of the test will be, e.g. whether it is to evaluate the actual physical device or rather the interaction with the screen content. You should also consider what equipment and resources are available. Based on our experiences with the different set-ups, we have summarized the benefits and drawbacks of each, as well as its suitability for the different types of usability testing , in easy-to-read tables.

## 6.1 Selecting the Right Set-up According to the Data Analysis

Table 2. Type of data analysis methodology.

| | Emulator on T-series | Below table T-series | Inverted X-series |
|---|---|---|---|
| **Qualitative data analysis** | ★★★ | ★★★★★ | ★★★★★ |
| **Quantitative data analysis** | ★★★★ | ★★★ | ★★★ |

In both set-ups with a physical mobile device, the device has been more or less fixed in place by a holder. The main purpose of this is to simplify quantitative analysis. Using the 'Scene' functionality in Tobii Studio, the recordings can be segmented so that all participant interactions can be analyzed together. Thus heat maps, gaze plots and areas of interest can be created. However, in order to create the scenes the researcher needs to go through each recording individually and select the parts that should be grouped into scenes. Alternatively, if using Tobii Studio Enterprise, events can be manually recorded by the moderator during the testing by selecting specific shortcut keys on a computer running the Tobii Remote Logger. Scenes also need to be created for the emulator set-up in order to be able to carry out quantitative analyses later on. The difference with the emulator set-up is that as the participants are interacting with an application on the Tobii Studio computer, Tobii Studio can automatically record the mouse and keyboard activity.

All the set-ups we tested are suitable for qualitative data collection. However, a more complete picture of the user interaction with the mobile device can be gained from the two set-ups with physical devices than from the emulator set-up. Therefore, if qualitative analysis is going to be performed and enough time end effort can be dedicated to it, we would recommend using either the Below Table or Inverted X-Series set-up.

## 6.2 Selecting the Right Set-up for the Context and Purpose of the Testing

Table 3. Type of usability test.

| Comparison of what to test and where | Emulator on T-series | Below table T-series | Inverted X-series |
|---|---|---|---|
| **Device-specific interface usability** | ★★★★★ | ★★★★ | ★★★★ |
| **Mobile application interface testing** | ★★★★★ | ★★★★ | ★★★★ |
| **Mobile web browsing** | ★★★★ | ★★★★ | ★★★★ |
| **Visual attractiveness & hedonic quality** | ★★ | ★★★★ | ★★★★ |
| **Physical device interaction** | ★ | ★★★★ | ★★★ |
| **Mobile input methods** | ★ | ★★★ | ★★★★ |

The suitability of the three set-ups evaluated in this paper depends on the type of usability test planned. In the table above we have graded the suitability of each set-up for common usability tests.

### 6.2.1 Device-specific Interface Usability

*Device-specific interface usability* refers to the testing of visual elements of an interface, such as the appearance of buttons and objects on a phone. This includes analysis of the visual interaction with the software (looking at the screen and the elements there presented e.g. text, logos, dynamic content etc.) as well as interaction with the interface (manipulating buttons, the screen or specific controls e.g. using menus, typing a message). Hence, the set-up needs to be able to accurately register what the participants are looking at, i.e. small details on the screen, and the participants must be able to interact with the mobile device as naturally as possible.

The area of the device's screen that can be seen in only one fixation varies with the inverted X-Series set-up depending on the distance between the participant and the device, as well as the distance between the participant and the eye tracker. The eye tracker also has a degree of inaccuracy of approx. 0.5°. Hence, it can be difficult to distinguish whether the participant was looking at a small object on the screen or an object located next to it. As the emulator shows an enlarged version of the device's interface on a Tobii T-series Eye Tracker, much smaller details can be analyzed as compared with the inverted X-Series set-up. Interferences, such as poor video quality or errors caused by changes in the set-up, are also eliminated when using an emulator.

The below table set-up introduces some visual limitations for the participants as the actual screen is not visible. Our testing highlighted that not all elements of the interface were perceived correctly just from looking at the video feed. Other issues such as lighting and contrast also played a big part in the visual quality. A higher quality camera might solve some of these issues, but the lack of a natural, visual interaction would still remain.

### 6.2.2 Mobile Application Interface Testing

Mobile application interface testing refers to the testing of content shown on the mobile device screen (games, browsers, applications, etc). This does not include, e.g., testing scenarios where the participants have to locate the correct device-specific hardware button in order to interact in a particular way with the screen content. The reason for this being that the same application can be produced for several devices, e.g. Opera Mini. Hence, the buttons to be used, and the location as well as design of these, will vary depending on the device used for the testing even if the actual interface seen on the screen remains the same.

The issues listed previously in the *Device-specific Interface Usability section* also apply to *mobile application interface testing*.

### 6.2.3 Mobile Web Browsing Testing

*Mobile web browsing testing* focuses only on the web browsing and, just as with browsing on a normal computer, this type of testing is only concerned with the content of the website or websites visited during the test which is consistent independent of the device (e.g. phone or computer) used to view the content.

The major difference with this type of testing compared to the other described here is that Internet browsing is something that we normally carry out on a computer. Hence, the websites viewed via the mobile browser are most often designed for a larger screen and the modes of interaction provided by a regular computer. In terms of eye tracking and the inverted X-Series and below table set-ups, this does not add to or remove from the issues as listed in the previous two sections. However, it does have some implications for testing with emulators.

Emulators are conventionally used with a mouse and participants therefore tend to interact with websites as they would when surfing with a regular web browser. Our testing showed that participants used the cursor to select and interact with small objects, such as links or scroll bars in the browser, whereas when interacting with the physical device they would normally zoom in first and then use their finger to scroll and select.

### 6.2.4 Visual Attractiveness and Hedonic Quality Testing

*Visual attractiveness and hedonic quality testing* focuses on the device as an object. The goal is to learn more about the general qualities of the device, such as what people find interesting to look at or feel, how they hold the phone (ergonomics), the quality of the screen (pixels, color etc.) and so on.

Due to the fact that the participant is unable to physically interact with the device, feedback on physical qualities is very difficult to collect with the emulator set-up. However, it is possible to use this set-up to gather feedback on visual design ideas. Hence, this set-up is mainly useful for 'quick and dirty' prototype testing at an early stage of the design process.

The below table set-up allows the participant to hold the phone and interact with it in almost any way they please. However, as the visual feedback to the participant comes via a camera feed shown on a display, the actual size and color of objects might not be exactly the same as on the actual device. Even so, a properly configured implementation of the below table set-up can potentially provide valuable feedback for this kind of testing.

The benefit of the inverted X-Series set-up for this type of testing is that the participants get to see and interact with the actual device. Hence, no misrepresentations caused by poor lighting, video quality, perspective etc. are introduced. However, the set-up is sensitive to movement so it is not suitable for testing the ergonomics of the device or how the participants look at the device when they hold it naturally without the support of a holder.

### 6.2.5 Physical Device Interaction

Physical device interaction refers to the physical handling of a device, i.e. holding it and manipulating parts or buttons on it. Examples of this are changing a SIM card on a phone, charging the phone and turning it on or off.

As the emulator only emulates the device on a regular computer screen, testing the physical device interaction is not possible with the emulator set-up.

The below table set-up allows for almost full freedom of movement and interaction with the device while still maintaining accurate eye tracking. The different elements of the device are also projected slightly larger than in real life when seen on the screen. Hence, when analyzing smaller details, the participants' gaze data is more accurately represented than when interacting with the device at its actual size.

The inverted X-Series set-up tested here was very sensitive to changes in the position of the scene camera, the device and the eye tracker. Even small changes caused errors in the mapping between the image captured by the scene camera and the eye gaze data collected. However, as long as everything was fixed in place, participants were able to naturally interact with the touch sensitive controls of the devices tested, i.e. the touch screen and the buttons.

### 6.2.6 Mobile Input Methods

Testing *mobile input methods* includes evaluating the different methods of interacting physically with the mobile device, e.g. touch-screen, stylus, touchpad, joystick, trackball, etc.

As testing mobile input methods involves physical interaction with the mobile devices, using an emulator is not suitable. Even though adding a touch screen panel to the TFT-screen of the eye tracker mimics the touch screen interface of the mobile device, it lacks the haptic feedback provided by the device itself during interaction.

The major problem when testing mobile input methods using the below table set-up is that three-dimensional interactions are recorded by the camera and presented on screen as a two-dimensional image. Since objects positioned closer to the camera will appear larger than objects further away, this sometimes causes situations where, for example, a finger appears to cover more than one button on the interface even though this is not actually the case.

To interact with the mobile input methods currently in use on most mobile devices (such as the methods mentioned above) does not require the participant to actually hold the device. Hence, even if the device is fixed in a holder, the user can interact with it. Even so, the inverted X-Series set-up, as discussed previously, is sensitive to changes in the set-up after the calibration has been completed. Of those tested this set-up offers the most natural interaction with the device as the participant can both see and touch the device. However, if holding the device is necessary for the input method to function properly then the below table set-up is better suited.

## 6.3 Selecting the Right Set-up for the Device Being Tested

Table 4. Type of phone to be tested. Red fields indicate the set-up is problematic, yellow indicates that it is moderately suitable and green highlights the most suitable options for the purpose of testing or type of device.

| Evaluation of different type of devices | Emulator T-series | Below Table T-series | Inverted X-series |
|---|---|---|---|
| **Standard phones (keypad, small screen)** | Currently difficult to find fully functional emulators for most devices | Suitable set-up. Needs a high quality camera | Suitable set-up with HD zoom camera for screen capture |
| **Smart phones (large touch screen)** | Emulator not available for all devices | Steep learning curve. Hard to interact with naturally | Suitable set-up when using HD zoom camera for screen capture |

Our set-up evaluation includes only two types of mobile phones, but these phones represent the most common types of mobile devices found on the market today in terms of screen size and modes of interaction. As discussed earlier, different problems were experienced with each set-up. A set-up which was suitable for one phone model proved to be less suitable for the other. While both phone models could be tested successfully with the inverted X-Series set-up, the below table set-up proved to be unexpectedly difficult to use with touch screen phones as the visual feedback was sometimes obstructed by the participants' fingers. With the standard phone with regular buttons, the participants received immediate feedback via the sensation of touching the buttons.

With the emulator set-up, we noted a difference in the availability of the correct emulator for the two types of phones. Emulators are currently mostly used by software developers for phones. Hence, they sometimes require a lot of effort to be found and installed properly. Mobile phone suppliers also do not always provide emulators for all models.

## 6.4 General Characteristics of the Set-ups

Table 5.Practical considerations for each set-up. Red fields indicate the set-up is problematic, yellow indicates that it is moderately suitable and green highlights the most suitable options for the purpose of testing or type of device.

| Evaluation of different ET set-ups | Emulator T-series | Below Table T-series | Inverted X-series |
|---|---|---|---|
| **Interaction with the device** | Not very realistic interaction, no haptic feedback | Reasonably realistic | Realistic |
| **ET Accuracy** | High accuracy | High accuracy | Some known issues |
| **Set-up time** | Fast set-up | Reasonable set-up time | Extensive set-up time |
| **Set-up difficulty** | Very easy | Reasonably easy set up procedure | Difficult |
| **ET Calibration** | Easy calibration | Easy calibration | Difficult calibration |
| **Freedom of movement** | Phone/emulator fixed | Almost full freedom possible | Phone fixed |
| **Intrusiveness** | Very unobtrusive | Sitting position and equipment placement might be intrusive | Unobtrusive |
| **ET robustness** | Very robust | Robust | Not robust in some cases |
| **Set-up robustness** | Very robust | Reasonably robust | Vulnerable set-up (to moving) |
| **Additional equipment** | Very little equipment needed | Some additional equipment needed | Requires considerable additional equipment |
| **Portability** | Very portable | Reasonably portable | Not portable |
| **Captures user behavior (physical)** | Doesn´t capture hands | Captures all physical interaction | Captures all physical interaction |
| **Logging user actions** | Automatic click logging | Manual logging | Manual logging |
| **Learning curve set-up** | Steep learning curve | Reasonable learning curve | No learning required |
| **Usability of the set-up** | Easy, using a mouse | Uncomfortable position | Easy to use, non-intrusive |
| **Flexibility of the set-up** | Switch between emulators easily | Easy to switch the position of the device or device to test | Manual calibration for each device to be tested |

The practical implications of each set-up differ significantly and needs to be considered when selecting which set-up to use. The emulator set-up is very easy to set up and requires only a computer with a Tobii T-series Eye Tracker. However, it does not allow any physical interaction with the device and the input methods are very different from those of the actual device.

The below table set-up provides a compromise between a realistic interaction with the device and a robust eye tracking set-up. It does need some time and effort to set up, but once complete it can be used almost without any risk of the eye tracking data being corrupted. It also allows the participant to handle the device freely. However, the major drawback is that the device has to be kept under the table during the entire testing session which provides a somewhat artificial use environment as well as potential discomfort for the participant.

Even though the inverted X-Series set-up provides a natural and realistic interaction with the device, it is fairly complicated to set up and needs quite an extensive amount of equipment. This makes it highly unsuitable for testing outside of a controlled laboratory environment.

# 7 Conclusion

There is no perfect set-up for eye tracking studies on mobile devices. Our research has evaluated a number of possible methods using currently available eye tracking equipment from Tobii and found that all methods tested have their own benefits and drawbacks. It IS possible to successfully carry out eye tracking on mobile devices, but it is very important to choose the set-up wisely and keep the limitations of the set-up, as well as of eye tracking in general, in mind. In the table below we have summarized the strengths and weaknesses of each set-up.

Table 6. Strengths and weaknesses of the set-ups.

| Strengths & Weaknesses | Emulator T-series | Below Table T-series | Inverted X120 Table |
|---|---|---|---|
| **Strengths** | Low cost, accurate solution | Physical interaction possible with the device | Natural use approach |
| | Easy to set up | Accurate and easy calibration | Shows the full user interaction |
| | Portable solution for on- location testing | Robust set-up and eye tracking data | Non-intrusive |
| | Good for quantitative studies (on location) | Very versatile and flexible set-up | Good for qualitative studies |
| **Weaknesses** | No physical device interaction possible | Steep learning curve, unnatural interaction | Difficult to set up and calibrate |
| | Unnatural interaction, fixed big screen | Slightly intrusive method and equipment | Low eye tracking robustness and subject to trackability issues |
| | Shortage of suitable emulators | Manual processing of data (logging) | Not portable or quick to set up |
| | Limited testing possibilities | Test results may vary in quality (e.g. fingers overlapping image) | A lot of additional equipment necessary |

Eye tracking studies on mobile devices is a fairly new and unexplored area of research. Our evaluation included only three possible set-ups. There are however others available and new ones will certainly be developed in the future. Many of the issues discussed in this paper are applicable to these other set-ups also. This paper should be treated as a guide. It provides advice on selecting the most suitable set-up for your study and is based on limited testing by Tobii.

# 8 Bibliography


Duda, S. (2009). My phone is my castle: Facing the Special Challenges of Mobile Usability Studies. *CHI 2009 Workshop: Mobile User Experience Research: Challenges, Methods & Tools (Unpublished)* .

Hagen, P., Robertson, T., Kan, M., & Sadler, K. (2005). Emerging research methods for understanding mobile technology use. *Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future. 122*, pp. 1-10. Canberra, Australia: Computer-Human Interaction Special Interest Group (CHISIG) of Australia.

International Telecommunication Union. (2009, March 16). *Measuring the Information Society - The ICT Development Index.* Retrieved January 7, 2010, from International Telecommunication Union: http://www.itu.int/ITU-D/ict/publications/idi/2009/material/IDI2009_w5.pdf

Intille, S. S., Tapia, E. M., Rondoni, J., Beaudin, J., Kukla, C., Agarwal, S., et al. (2003). Tools for Studying Behavior and Technology in Natural Settings. In *UbiComp 2003: Ubiquitous Computing* (Vol. 2864/2003, pp. 157-174). Berlin / Heidelberg: Springer Berlin / Heidelberg.

Isomursu, M., Kuutti, K., & Väinämö, S. (2004). Experience Clip: Method for User Participation and Evaluation of Mobile Concepts. *Proceedings of the eighth conference on Participatory design: Artful integration: interweaving media, materials and practices. 1*, pp. 83 - 92 . Toronto, Ontario, Canada : ACM.

Kaikkonen, A., Kekäläinen, A., Cankar, M., Kallio, T., & Kankainen, A. (2005). Usability Testing of Mobile Applications: A Comparison between Laboratory and Field Testing. *Journal of Usability Studies , 1* (1), 4-17.

Kiljander, H. (2004). *Evolution and usability of mobile phone interaction styles.* Helsinki: Unpublished Ph.D. Thesis at Helsinki University of Technology.

Kjeldskov, J., & Connor, G. (2003). A Review of Mobile HCI Research Methods. In *Human-Computer Interaction with Mobile Devices and Services* (Vol. 2795/2003, pp. 317-335). Berlin / Heidelberg: Springer Berlin / Heidelberg.

Nielsen, C. M., Overgaard, M., Bach Pedersen, M., Stage, J., & Stenild, S. (2006). It's worth the hassle!: the added value of evaluating the usability of mobile systems in the field. *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles* (pp. 272 - 280). Oslo, Norway: ACM.

Nielsen, J. (1994). *Usability Engineering.* New York: Morgan Kaufmann.

Pirhonen, A., Brewster, S., & Holguin, C. (2002). Gestural and audio metaphors as a means of control for mobile devices. *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves* (pp. 291 - 298). Minneapolis, Minnesota, USA: ACM.

Waterson, S., Landay, J. A., & Matthews, T. (2002). In the Lab and Out in the Wild: Remote Web Usability Testing for Mobile Devices. *CHI '02 extended abstracts on Human factors in computing systems* (pp. 796 - 797). Minneapolis, Minnesota, USA : ACM.